Year : 2023

THREE ESSAYS IN DEVELOPMENT ECONOMICS:YOUTH EMPOWERMENT, PLAGUES, AND CONFLICT

Antunes Alice Rosenfeld

Antunes Alice Rosenfeld, 2023, THREE ESSAYS IN DEVELOPMENT ECONOMICS:YOUTH EMPOWERMENT, PLAGUES, AND CONFLICT

Originally published at : Thesis, University of Lausanne

Posted at the University of Lausanne Open Archive http://serval.unil.ch
Document URN : urn:nbn:ch:serval-BIB_CA6CFF5B2B671

UNIL | Université de Lausanne

FACULTÉ DES HAUTES ÉTUDES COMMERCIALES

DÉPARTEMENT D'ÉCONOMIE

THREE ESSAYS IN DEVELOPMENT ECONOMICS: YOUTH EMPOWERMENT, PLAGUES, AND CONFLICT

THÈSE DE DOCTORAT

présentée à la

Faculté des Hautes Études Commerciales
de l'Université de Lausanne

pour l'obtention du grade de
Docteure ès Sciences Économiques,
mention « Économie politique »

par

Alice Rosenfeld ANTUNES

Directeur de thèse
Prof. Rafael Lalive

Co-directrice de thèse
Dr. Elena Esposito

Jury

Prof. Christian Zehnder, président
Prof. Marius Brülhart, expert interne
Prof. Costanza Naguib, experte externe

LAUSANNE
2023

UNIL | Université de Lausanne

FACULTÉ DES HAUTES ÉTUDES COMMERCIALES

DÉPARTEMENT D'ÉCONOMIE

THREE ESSAYS IN DEVELOPMENT ECONOMICS: YOUTH EMPOWERMENT, PLAGUES, AND CONFLICT

THÈSE DE DOCTORAT

présentée à la

Faculté des Hautes Études Commerciales
de l'Université de Lausanne

pour l'obtention du grade de
Docteure ès Sciences Économiques,
mention « Économie politique »

par

Alice Rosenfeld ANTUNES

Directeur de thèse
Prof. Rafael Lalive

Co-directrice de thèse
Dr. Elena Esposito

Jury

Prof. Christian Zehnder, président
Prof. Marius Brülhart, expert interne
Prof. Costanza Naguib, experte externe

LAUSANNE
2023

UNIL | Université de Lausanne
HEC Lausanne
Le Décanat
Bâtiment Internef
CH-1015 Lausanne

I M P R I M A T U R

Sans se prononcer sur les opinions de l'autrice, la Faculté des Hautes Etudes Commerciales de l'Université de Lausanne autorise l'impression de la thèse de Madame Alice Rosenfeld ANTUNES, titulaire d'un bachelor en Science, Technology, and International Affairs de l'Université de Georgetown, titulaire d'un master en Environmental Policy and Regulation de la London School of Economics and Political Sciences, titulaire d'un master en Économie politique de l'Université de Lausanne, en vue de l'obtention du grade de docteure ès Sciences économiques, mention « économie politique ».

La thèse est intitulée :

THREE ESSAYS IN DEVELOPMENT ECONOMICS: YOUTH EMPOWERMENT, PLAGUES, AND CONFLICT

Lausanne, le 29 mars 2023

La Doyenne

Marianne SCHMID MAST

Members of the Thesis Committee

Prof. Elena Esposito
Assistant Professor, Collegio Carlo Alberto, University of Turin
Thesis co-supervisor

Prof. Rafael Lalive
Professor, University of Lausanne
Thesis co-supervisor

Prof. Marius Brülhart
Professor, University of Lausanne
Internal member of the thesis committee

Prof. Costanza Naguib
Assistant Professor, University of Bern
External member of the thesis committee

University of Lausanne
Faculty of Business and Economics

PhD in Economics
Subject area Political Economy

I hereby certify that I have examined the doctoral thesis of

Alice Rosenfeld ANTUNES

and have found it to meet the requirements for a doctoral thesis.
All revisions that I or committee members
made during the doctoral colloquium
have been addressed to my entire satisfaction.

Signature: _____ Date: _19.01.2023_____

Prof. Rafael LALIVE
Thesis supervisor

University of Lausanne
Faculty of Business and Economics

PhD in Economics
Subject area Political Economy

I hereby certify that I have examined the doctoral thesis of

Alice Rosenfeld ANTUNES

and have found it to meet the requirements for a doctoral thesis.
All revisions that I or committee members
made during the doctoral colloquium
have been addressed to my entire satisfaction.

Signature: _  _____ Date: _____19/1/2023_____

Dr. Elena ESPOSITO
Thesis co-supervisor

University of Lausanne
Faculty of Business and Economics

PhD in Economics
Subject area Political Economy

I hereby certify that I have examined the doctoral thesis of

Alice Rosenfeld ANTUNES

and have found it to meet the requirements for a doctoral thesis.
All revisions that I or committee members
made during the doctoral colloquium
have been addressed to my entire satisfaction.

Signature: _____ Date: _____19 Jan. 2023_____

Prof. Marius BRÜLHART
Internal member of the doctoral committee

University of Lausanne
Faculty of Business and Economics

PhD in Economics
Subject area Political Economy

I hereby certify that I have examined the doctoral thesis of

Alice Rosenfeld ANTUNES

and have found it to meet the requirements for a doctoral thesis.
All revisions that I or committee members
made during the doctoral colloquium
have been addressed to my entire satisfaction.

Signature: _____ Date: _____19.01.23_____

Acknowledgements

Acknowledgments are a difficult thing to write, and nobody ends up reading these anyway except for the people who are being Acknowledged. The whole point, though, is for those people who are in this part of the thesis to get an insight into how massively important they are to me. So in the hopes of being corny, I would like to thank the following humans and non-humans for their role in keeping me sane and happy in the process of writing this thesis:

First and foremost, I would like to thank my co-supervisors, Elena and Rafael, without whom I wouldn't be here today. Rafael was the person who helped me put together my application to the University of Lausanne, when I was just a sad civil servant in Washington, DC, trying to do something more interesting with my life. I remember calling him from my cubicle, seeing if there was some way an underqualified poli-sci major could get into a prestigious economics program in Lausanne. Elena was the person who inspired me to think about the Ph.D. program in Lausanne. I remember being elated when Elena, a young woman rising superstar in her field, agreed to supervise my thesis. Without their insight and encouragement, I would definitely not be writing these Acknowledgments today. Thank you also to my coauthors and mentors who helped me write and think of my thesis, Andrés Moya, Dominic Rohner, Ulrich Eberle, Camille Terrier, and Salvatore di Falco.

Just as importantly, I would like to thank my parents, who are the reason I am here today and who have opened up all the doors for me without ever asking for anything in return. They are the smartest and most inspiring people I know and I would be happy if I could one day grow into a fraction of the people they are. I love you very much and cannot ever begin to express how much your unconditional support means to me (except for here, where my dad wanted my mom to be acknowledged as Queen).

Thank you also to my brothers, Marc and Thomas, who provide me with endless entertainment, jokes, video games, and TV, even when we are an ocean apart. When we're not an ocean apart, our favorite thing to do is sit on the couch next to each other without talking to each other, and that's actually the activity that has kept me sane (question mark) all these years. Thank you also to my grandparents, Vovó and Vovô, who taught me to always look at life through a humorous and optimistic lens. Thank you to my uncles and aunts, Marcelo, Aldo, Claudinha, Paulinha, Thomas and Tina Dislich, and cousins, Juninho, Xandão, Christian, Gabriel, Solange, Duda, Alice, Lali, Julia & Christian Martins, and Marina & Oli who always helped me understand what I do from another perspective.

Thank you to my friends, my chosen family, who have always encouraged me to question the status quo and see life from a different perspective. My high school friends, Benjy and Hannah; my US friends, Cristy, Sara, Sarah, Steph, the Twins, Udayan, and Anjani; my master's friends, Sandra, Marie, and Olta; my friends with real adult lives (what even is that?) Céline, Justine, Naniko, Priya, and Chrystal; my current- and ex-office-#554-mates , Quentin, Uli, Andrea, Resuf, and Katja; all of y'all in the Lausanne department; all of y'all from Gerzensee and what used to be Women in Econ; and then there's my friends that just don't fit into any one category, but span most or all of them, Noëmi, Paola, Pauline, Lyndsey, Léo, Vaibhav, Albi, and Fernando.

Thank you to Francis, thank you for feeding me and loving me even when I haven't showered in days and have been wearing the same sweatpants for weeks. Thank you for broadening my horizons and making me laugh and letting me use your hoodie sleeve as a tissue when I cry. You are acknowledged.

Last but not least is Blue, who is not really a dog but more like a person, a soulmate, proof that Dog is indeed Man's best friend - even though I'm not a man. Du bist sehr brav! Das bist du! Let's go spazieren.

Contents

Introduction

It is hard to pinpoint what, exactly, development economics is, but it is often about the economics of improving standards of living. Big picture macroeconomic questions look at economic growth, GDP per capita, and why some countries are poor while others are rich. Microeconomic questions explore topics ranging from poverty, agriculture, environmental degradation, conflict, health, and education. Often, development economics is about evaluating policies and quantifying their impact on welfare.

In this thesis, I contribute to the development economics literature in that I look at how youth empowerment can help education have a more powerful impact on young adults that have been marginalized by urban poverty and conflict. I also contribute to the COVID-19 legacy by looking at the mental health effects of COVID-19 on these same marginalized young adults. Finally, I examine how a type of natural disaster, locust outbreaks, affects rural livelihoods and social conflict. I cover a breadth of topics in development economics, namely education, poverty, labor, and conflict economics. In addition, I use a variety of methods that have been at the forefront of improving our knowledge of human well-being. In particular, my thesis makes use of a randomized controlled trial, as well as more traditional econometric methods, including those that are currently being revolutionized by econometricians including Callaway and Sant'Anna (2021), de Chaisemartin and D'Haultfœuille (2020), and Sun and Abraham (2021).

Youth empowerment, vocational education, and labor market participation for marginalized young adults in Colombia. A variety of public policies have been piloted in an effort to better integrate vulnerable young adults into the labor market. A popular solution for underemployment has been vocational education - education that develops technical, social, and personal skills intending to integrate trainees in decent working conditions in the long term (Beerli et al., 2017).

There is some evidence of the effectiveness of vocational education in labor market integration for vulnerable young adults throughout lower and middle-income countries, but the results are far from conclusive. While some studies find positive results of vocational education on the economic success of vulnerable youth (in Colombia, Attanasio et al. (2011, 2017) have found positive, though limited impacts of vocational education), other studies find no effect at all (Card et al., 2011). It seems that vocational education has been most beneficial for the poorest quantiles of the population and for women (Ibarrarán and Rosas Shady, 2009). Yet it is costly (Blattman and Ralston, 2015) and not many people even take advantage of it. On top of that,

when people do start vocational education, a large share tends to not finish the program they are enrolled in (Cho et al., 2013).

The lack of a clear positive effect of vocational education on quality of life is puzzling. Some initial evidence suggests that bundling vocational education with additional treatments is more effective in providing access to a better future for participants. In particular, a new wave of projects is exploring the failed promise of vocational education programs by bundling them with complementary interventions. These include cash transfers (Blattman and Annan, 2015; Blattman et al., 2017) , reproductive education (Adoho et al., 2014; Bandiera et al., 2014), and cognitive-behavioral training (Blattman et al., 2017).

In this light, the first chapter of this thesis addresses the question of what the effect of adding a youth empowerment program to traditional vocational education for marginalized young adults is. This chapter aims to provide an answer to the vocational education puzzle by assessing the effect of training for disadvantaged young adults in Colombia using a randomized control trial where 300 young adults were sorted into three treatment groups: one receiving vocational education and youth empowerment training, one receiving vocational education only, and one untreated control group. The results of this chapter suggest that the youth empowerment program was initially successful in reducing informal employment, which is a source of lower incomes. Patterns in earnings suggest that the program also appears to have affected participants' earnings positively, though these earnings come from the informal labor market. This effect is clearer for women, who move away from the informal sector compared to the general population.

COVID-19 and mental health of marginalized young adults in Colombia. COVID-19 had a wide range of consequences on mental health, ranging from no impact, to either a positive or negative impact. A large stream of literature has documented the dramatic negative mental health implications of the COVID-19 pandemic, in terms of depression, distress, and trauma (see Serafini et al. (2020), Salari et al. (2020), Pierce et al. (2020), Giuntella et al. (2021), Brodeur et al. (2021), and Abbott (2021) among many others).

The main results of these studies are that on the one hand, health systems had fewer resources for providing adequate medical support. On the other hand, the policies implemented to contain the pandemic - such as lockdowns, closure of schools, shifting allocation of health resources, and curtailed livelihood opportunities - had particularly detrimental consequences in contexts with more vulnerable populations, weaker social safety nets, more informal employment, and less adequate economic resources to buffer against lost livelihoods (see Kola et al. (2021) for an insightful review). In countries where the vast majority of people are at a constant threat to their livelihood, the pandemic - together with the policy responses adopted to mitigate

it - represented a special burden for already vulnerable groups, such as women, women with children, and people with pre-existing physical and psychological vulnerabilities (see Silverio-Murillo et al. (2020), Becerra et al. (2020), Moya et al. (2021), and Parra-Saavedra and Miranda (2021)).

Addressing mental health issues was a priority for public institutions throughout the pandemic (Holmes et al., 2020; Moreno et al., 2020). The fear of another similar pandemic has pushed institutions to search for policies that can reduce the psychological consequences of similar shocks. We know people with better education had greater resilience, but there is limited knowledge on what kind of training helped people do better. There is still a limited understanding regarding what kinds of programs and policies can effectively increase the resilience of vulnerable populations in the face of similar shocks.

In this light, the second chapter of this thesis makes use of the previous randomized controlled trial to explore whether the same youth empowerment program, combined with vocational education, can help vulnerable young adults cope with shocks such as COVID-19. We had previously randomized 300 young adults in low-income suburbs of Bogotá, Colombia into a group receiving vocational education and a youth empowerment program, a group receiving only vocational education, and a non-treated control group. Using a survey we ran during the COVID-19 lockdown, we can explore if, and to what extent, the program shielded participants from the mental health shock of the pandemic. Young adults who received vocational education managed symptoms of depression and distress better during the lockdowns compared to participants in the control group. The effects are stronger for women. There is no effect on trauma, which is more pronounced for the entire sample. In the medium run, mental health bounces back to baseline values and there is no difference between groups. There seems to be an acute cushioning effect of treatment and vocational education seems to be the force behind the cushion. The main implication is that negative mental health effects of shocks such as those observed during the COVID-19 pandemic may be mitigated by programs like the one explored in this chapter of the thesis.

How locust outbreaks affect social conflict in the Horn of Africa. In political economics, it is well-established that shocks to agricultural production can cause social disruption and changes to social cohesion. While the potential disruption created by locust outbreaks has been a priority of development and aid agencies to help regions affected, not much effort has been put into understanding the impacts of locusts on the economy, and, more particularly, on conflict.

Locusts outbreaks can be very damaging to local economies. The most obvious consequence of locust outbreaks is on food systems. Their food of choice is key crops such as millet and sorghum. A locust swarm of one kilometer squared (which can comprise up to 150 million

locusts) can eat the same amount of food per day as 35,000 people and can travel up to 150 kilometers a day. To put this into perspective, a swarm the size of Paris eats the same amount of food in one day as half the population of France (Cressman, 2016). Locust outbreaks are expensive, both directly and indirectly. The 2003-2005 outbreak in the Horn of Africa resulted in crop loss worth $2.5 billion (Brader et al., 2006). Furthermore, the FAO estimates that household expenditures on food increased by three-to-five fold during that same outbreak, leading to negative consequences on health and nutrition.

There are a few studies that systematically estimate the effect of locust outbreaks on rural households, but most of the studies are on how locusts affect health. Conte et al. (2021) find that children exposed to a locust outbreak in Mali in 2004 are more stunted than children not affected by the outbreak. Similarly, Marending and Tripodi (2022) find a significant farm profit loss following a locust outbreak in Ethiopia in 2014, which results in stunted growth and body-mass-index scores for children affected by the outbreak. Vreyer et al. (2015) find that the locust outbreak in Mali from 1987-1988 reduced school enrollment of boys living in rural areas affected by the outbreak four years later. The effect of locusts on conflict remains unexplored.

Broadly speaking, there are two opposing forces explaining the impact of income shocks on conflict: the opportunity cost effect arises when an increase in prices of labor-intensive goods leads to a decrease in conflict, by increasing wages and thus increasing the potential loss someone may incur if they joined a rebellion, or by encouraging potential rebels to earn money outside of a rebellion. On the other hand, the rapacity effect occurs when prices of (less labor-intensive) goods increase and the gains from appropriation become greater than the gains from wages (thereby encouraging violence). These theories have been developed in the literature for over twenty years, with key works refining subtleties in the two forces (Collier et al., 2009; Bó and Bó, 2011; Dube and Vargas, 2013; McGuirk and Burke, 2020).

Previous literature has focused on how shocks to global food prices affect local conflict. Studies have used environmental sources of income shocks such as rainfall shocks (Miguel et al., 2004; Miguel and Satyanath, 2011) to proxy for local income variation. The third chapter of this thesis builds on this legacy by using an inherently local price shock to contribute to our understanding of the origins of conflict.

In addition, we know that natural disasters affect social structures. Following natural disasters (such as cyclones) or shocks to temperature or rainfall, some communities can experience more conflict (Hsiang et al., 2011, 2013; Hsiang, 2010; Dell et al., 2014). However, this is not the only social change that might occur. For example, Ager and Ciccone (2018a) find that communities with greater rainfall variation in the 19th-century United States had larger shares of their populations enrolled in religious organizations. Similarly, Bentzen (2021) finds that districts

of the world's neighboring districts hit by earthquakes are more likely to be religious. Finally, different coping strategies to risk can lead to social cohesion. Communities with less access to formal insurance systems tend to create their informal insurance systems in the face of risk, especially rainfall risk (see, for example, Rosenzweig (1988a,b); Rosenzweig and Stark (1989); Townsend (1995)).

In this light, the third chapter of this thesis studies the effects of locust outbreaks on conflict at the local level in the last twenty years. We use geolocalized data on locust sightings from the FAO, and match it to data on conflict. At the country-year level, we use a differences-in-differences specification with cell and country-year fixed effects to explore the yearly effect of locusts on conflict. We find that locusts have a significantly negative correlation with conflict. To explore the monthly effects of locust outbreaks, we run a two-way fixed effects event study. We find that locust outbreaks significantly reduce conflict one month later. To understand the underlying mechanisms that might be driving this surprising reduction in conflict, we match the locust data to household agricultural data in Ethiopia in 2014. We find that households having experienced a locust outbreak as early as one month before the survey are more likely to report a reduction in crop incomes, an increase in non-agricultural wage incomes, and an increase in food insecurity. The reduction in conflict could arise from three different channels: the locust outbreaks may be causing a reduction in appropriable resources - there are fewer things to steal. It is also possible that there are simply no resources left to compete over. Alternatively, the locust outbreaks might increase collective action and therefore increase social cohesion.

Structure of this thesis. The rest of this thesis proceeds as follows. Chapter 1 looks at the medium-run labor market impacts for marginalized young adults of combining vocational education with youth empowerment training. Chapter 2 uses evidence from the COVID-19 pandemic to find that the same program that added youth empowerment to vocational education provides greater resilience to shocks among vulnerable young adults. Finally, Chapter 3 explores how locust outbreaks have affected conflict in the Horn of Africa over the last twenty years.

Improving vocational education through youth empowerment: evidence from a randomized controlled trial in Colombia

Alice Antunes[1], [2]

April 5, 2023

Abstract

What is the effect of adding a youth empowerment program to traditional vocational education for marginalized young adults? Vocational education might be a way to increase the labor market attachment of young adults, but it has shown little to no effect in achieving results. This paper aims to provide an answer to this puzzle by assessing the effect of training for disadvantaged young adults in Colombia using a randomized control trial where 300 young adults were sorted into three treatment groups: one receiving vocational education and youth empowerment training, one receiving vocational education only, and one untreated control group. The results of the study suggest that the youth empowerment program was initially successful in reducing informal employment, which is a source of lower incomes. Patterns in earnings suggest that the program also appears to have affected participants' earnings positively, though these earnings come from the informal labor market. This effect is clearer for women, who move away from the informal sector compared to the general population.

Keywords: Vocational Education, Youth Empowerment, Colombia, Active Labor Market Policies, Formal Sector Employment

JEL Codes: O12, J24, D04, O17, E24

[1] Department of Economics, University of Lausanne.

[2] I would like to thank Elena Esposito, Rafael Lalive, and Andrés Moya. I would also like to thank Stefanie Enssle, Ira Amin, and Jan Weber of Vivamos Mejor; Edilberto Garcia and Andrés Clavijo of Fundación Apoyar; and iQuartil for the fast turnaround on the COVID-19 survey during the lockdown. This project is registered under RCT ID #AEARCTR- 0003869. The project's data collection and extensive participant follow-up work were funded by SDC-NADEL's Impact Award of 2020, the HEC Research Fund 2019, and Enterprise for Society (E4S). Research assistants who were instrumental in the data work are: Christian Arciniegas (University of Fribourg), Leonardo Pedretti (University of Lausanne), and Falone Moseka (University of Lausanne)

1 Introduction

A variety of public policies have been piloted in an effort to better integrate vulnerable youth into the labor market. A popular solution for underemployment has been vocational education - education that develops technical, social, and personal skills intending to integrate trainees in decent working conditions in the long term (Beerli et al., 2017).

There is some evidence of the effectiveness of vocational education in labor market integration for vulnerable young adults throughout lower and middle-income countries, but the results are far from conclusive. While some studies find positive results of vocational education on the economic success of vulnerable youth (in Colombia, Attanasio et al. (2011, 2017) have found positive, though limited impacts of vocational education), other studies find no effect at all (Card et al., 2011). It seems that vocational education has been most beneficial for the poorest quantiles of the population and for women (Ibarrarán and Rosas Shady, 2009). Yet it is costly (Blattman and Ralston, 2015) and not many people even take advantage of it. On top of that, when people do start vocational education, a large share tends to not finish the program they are enrolled in (Cho et al., 2013).

This study explores the effects of combining vocational education with a youth empowerment program on labor market integration and poverty alleviation for marginalized young adults. The youth empowerment program this study assesses combines soft skills training, psycho-social support, nation-building, and labor market integration. This is a long-running project provided by Fundación Apoyar, a local NGO in Bogotá. More specifically, the goal of the study is to find out how adding this youth empowerment program to vocational education affects participants' economic success (employment in the formal labor market, wages, socio-economic status) and their personal development in terms of mental health. Two additional questions can help clarify the main question: first, how do young adults with vocational education (whether they undergo the youth empowerment program or not) do compared to the general population? Second, how do young adults who also go through the youth empowerment program do compared to young adults who receive only vocational education? This latter would be the so-called added value of the bundled program of youth empowerment and vocational education.

To do so, I use the results of a randomized controlled trial[1] where 300 vulnerable young adults between the ages of 18 and 25 from Bosa (a neighborhood of Bogotá) and Soacha (a city in the neighboring state of Cundinamarca) were randomly allocated into equally-sized treatment groups: one receiving the vocational education and youth empowerment program offered by

[1] Registered under RCT ID #AEARCTR-0003869.

Fundación Apoyar (group VET+YE), one receiving vocational education only (group VET), and one control group with no additional support (Control group). Each group is composed of 100 young adults. Randomization was done via public lottery to reduce suspicion related to group allocation and to increase the transparency of the randomization process. The program ran throughout the academic year of 2019. The team collected data in December 2018 before the randomization, and again in August 2020, November 2020, and November 2021.

Participants following vocational education (groups VET+YE and VET) receive the nationally certified twelve-month classroom and training program given for the following professions: human resources, administration, early childhood care, logistics, and marketing. In addition to this, participants in the group that received youth empowerment on top of vocational education (VET+YE) follow a four-dimensional youth empowerment training program centered around: (1) socio-emotional skills development (group work and individual therapy), (2) development of a pacifist culture and citizen empowerment (to live with former paramilitaries and to learn to defend their rights), (3) business skills training (mock interviews, curriculum vitae development), and (4) labor placement support.

Colombia is a setting where a large proportion of the young adult population is vulnerable and where vocational education has been offered as a means of integrating young adults into the formal labor market. In Colombia, most young adults work without a contract: 63% of 18-25-year-olds hold informal jobs - that is, jobs without a proper work contract. Such jobs tend to be more dangerous, more seasonal, and do not provide any sort of social insurance or pension system. An additional 16% of this population are not in education, employment, or training (Inter-American Development Bank, 2019). The reasons for this are many, with anecdotal evidence highlighting discrimination and lack of skills - many of these young adults are lower-income young adults, are less well-educated, and are internally displaced from rural regions to more urban regions. Integrating these vulnerable youth into the labor market has been a policy priority in poverty alleviation and post-conflict nation-building efforts. Vocational education has been a hallmark of these efforts. The logic is that what is lacking in labor market integration is specific skills, and young adults are not getting these skills otherwise. Since vocational education is a way of providing young adults with marketable skillsets, it should follow that providing young adults with vocational education will give them access to the jobs they are trained for.

I analyze the data collected by this long process using the two-way fixed effects models elaborated on by the recent literature, which includes Sun and Abraham (2021). I begin by looking at a simple difference in means, which would theoretically provide the average treatment effect on the treated. I then elaborate on the difference in group means for two reasons. First, there are

slight imbalances between treatment groups that need to be controlled for in a baseline specification. And second, an event study method can exploit the various dates during which data was collected to measure the difference between groups in labor market outcomes, compared to the difference between groups before the program. Using alternative specifications, such as the one proposed by De Chaisemartin and d'Haultfoeuille (2020), does not change the results considerably.

The main results of the paper are first, that patterns suggest that the youth empowerment program affected participants' earnings. Three years after the start of the program, participants who participated in youth empowerment appear to be earning more than their peers who received only vocational education and those who received no further support. Participants in youth empowerment earn $36.15 more than their peers in the Control group during July 2021, compared to the difference before treatment.

Second, in terms of the source of earnings, youth empowerment initially significantly reduced earnings from the informal labor market compared to participants in the two other groups. However, there was an increase in informal labor market earnings eventually compared to their peers. Formal earnings increase for both vocational education treatment groups, though this increase is not statistically significant compared to participants in the control group. There is an initial lock-in effect of training in terms of earnings from the formal sector, but once participants have had enough time to find a job in the formal labor market, they tend to find one relatively quickly.

In addition, I run a heterogeneity analysis on gender and displacement status. The patterns in this analysis suggest that whereas women who went through VET+YE appear to have used the training to move away from the informal sector, men who received VET+YE (vocational education and youth empowerment) seem to have found the most benefits of the training in the informal sector. Men who received VET+YE seem to earn more compared to the Control group and compared to participants who only received VET. These earnings come mostly from informal earnings. For women, there is no significant difference in overall earnings between the three groups. That said, the initial drop in informal sector earnings in participants who followed VET+YE seems to come entirely from women - women who received VET+YE earn less in the informal sector, and more in the formal sector, compared to participants in the Control group and compared to participants in VET (vocational education) only in the first year after treatment. The magnitude of earnings is higher for men.

Participants who went through training who are internally displaced appear to earn more, compared to participants who are internally displaced (IDP) and who did not go through training, though this result is barely statistically significant, and only for two of the survey months

after the program ended. The initial dip in earnings in the year following treatment does not happen for internally displaced people who received VET+YE compared to internally displaced people in the Control group. Furthermore, the treatment effects on earnings are in general larger for IDPs (the difference in earnings in IDPs who went through treatment compared to the Control group are higher compared to the difference in earnings for non-IDPs in the different groups).

Finally, I run a simple cost-benefit analysis, where patterns suggest that participants in the youth empowerment program may be expected to earn between $2,158.59 and $2,683.73 more than participants in the Control group over the next ten years, ignoring statistical significance. The total cost of the VET+YE program per participant for 2019 was $1,395.67. Over the ten years following treatment, there is therefore an estimated net benefit between $762.92 and $1,288.06 per participant for following the VET+YE program. The corresponding rates of return would be 55% and 92%, respectively.

This study contributes to the literature on how to improve vocational education in labor market integration efforts for vulnerable young adults. In light of the limited impact of vocational programs among vulnerable youth, the question is what are these programs lacking, and what are the skill deficiencies that vulnerable youth have and that affect their performance in the labor market? Studies exploring bundling vocational education programs with complementary interventions (such as cash transfers (Blattman and Annan, 2015), reproductive education (Adoho et al., 2014; Bandiera et al., 2014), and cognitive-behavioral training (Blattman et al., 2017)), have found positive impacts of these projects on labor market integration and earnings of participants. This study assesses the effectiveness of adding a more comprehensive approach to vocational programs by comparing vocational education programs to vocational education and a youth empowerment training program, to see whether the combination of technical skills and youth empowerment provides positive results for participants. The project aims to see how vocational education can be made more consistently effective for well-being and labor market outcomes.

While the setting for this study is Colombia, some lessons can be extended to other post-conflict and urban poverty environments, in particular in Latin America but also potentially in other countries. The study sheds some light on how vocational education can be improved overall and contributes to the evolving literature on the role of youth empowerment in education and labor market integration.

The paper proceeds as follows: first, I explore the background literature that provides the context for the study design. I then explain the study design and examine balancedness and attrition. This clears the way for descriptive statistics and my results. I finish with an examina-

tion of the robustness of my results, then heterogeneous impacts of treatment, and finally with a cost-benefit analysis of the program.

2 Background

This section proceeds as follows. First, I show how the study fits into the larger literature. Second, I provide some context for the setting of the study, in particular the historical, social, and economic setting for Colombia, some background on the neighborhoods where the study was conducted, as well as an introduction to the NGO that developed and administered the youth empowerment program.

2.1 Literature review

2.1.1 Vocational education

In several countries across the world, access to good jobs in the formal sector is in general limited to more educated workers (Attanasio et al., 2011). Formal labor markets are even more inaccessible for less-educated young adults in lower-income contexts. For example, in Colombia, 63% of 18-25-year-olds hold informal jobs - that is, jobs without a proper work contract - and 16% are not in education, employment, or training (Inter-American Development Bank, 2019). This is such a common problem that this population is popularly known as NEETs (Not in Education, Employment, or Training), or NiNis in Spanish ("ni estudian, ni trabajan"). The reasons for this are many, with anecdotal evidence highlighting discrimination and lack of skills. Many public policies have been tried out to better integrate vulnerable youth into the formal labor market, both from the demand and from the supply side. A popular supply-side solution for underemployment has been vocational education - which develops technical, social, and personal skills to integrate trainees in decent working conditions in the long term (Beerli et al., 2017).

Vocational education has been a wildly popular policy tool to increase the employability of young and disadvantaged populations. In Latin America, Ibarrarán and Rosas Shady (2009) count no less than 8 countries that have instated national vocational education systems since the mid-1980s: Mexico's Probecat (1984-present), Chile's Chile Joven (1992-1997), Argentina's Proyecto Joven (1994-1998), Colombia's Jóvenes en Acción (2002-2005), Dominican Republic's Juventud y Empleo (1999-), Panama's ProCaJoven (2002-), Peru's Projoven (1996-).

Because vocational education has been such a popular policy tool, it has also garnered a lot of attention. Impact evaluations tend to find mixed results in vocational education programs. In a randomized evaluation, Attanasio et al. (2011) explored the Jóvenes en Acción program

from 2001-2005, which provided 3 months of in-classroom training followed by 3 months of on-the-job training to young adults between 18 and 25 in the lowest socioeconomic strata in Colombia. They find no effect of vocational education on employment rate, being in the formal labor market, or wages overall. They do, however, find that vocational education resulted in 22% higher wages for women. The results lasted in the long run, with a 13.6% increase in wages after 10 years for women. Similarly, also in a randomized evaluation, Card et al. (2011) explore the impact of the Juventud y Empleo program in the Dominican Republic, which ran from 2001 to 2006 and aimed to "increase the likelihood of employment for lowest-income members of the working-age population by facilitating access to the labor market through training and counseling". The study found no effect on employment rate or earnings, but conditional on working, there was a $7-10\%$ increase in hourly wages. It seems that vocational education has been most beneficial for the poorest quantiles of the population and for women (Ibarrarán and Rosas Shady, 2009).

Furthermore, vocational education seems to be costly (Ibarrarán and Rosas Shady, 2009). There is also low take-up of such programs - for example, the Pakistan's Skills for Employability pre-program survey found that 95% of interested households did not apply for the program (Blattman and Ralston, 2015); and there is a high drop-out of programs. Cho et al. (2013) did an analysis of drop-outs on an entrepreneurship and vocational training program for vulnerable youth in Malawi where one-third of their program participants dropped out, and found that women are most likely to drop out of these programs because of a variety of reasons: they have more dependents, spend more time on domestic chores, have family obligations and get married, and have to draw more on savings to participate in such programs.

2.1.2 Evidence on improving vocational education

The lack of a clear positive effect of vocational education on quality of life is puzzling. Some initial evidence suggests that bundling vocational education with additional treatments is more effective in providing access to a better future for participants. In particular, a new wave of projects is exploring the failed promise of vocational education programs by bundling them with complementary interventions. These include cash transfers (Blattman and Annan, 2015; Blattman et al., 2017) , reproductive education (Adoho et al., 2014; Bandiera et al., 2014), and cognitive-behavioral training (Blattman et al., 2017).

There has been a plethora of studies exploring the role of socio-emotional skills in education. In the psychology literature, what is called "socio-emotional learning", or "socio-emotional skills training" includes the process of acquiring skillsets to manage emotions, set and achieve goals, and handle interpersonal relationships through competencies that are considered impor-

tant for success in life such as self-awareness, self-management, social awareness, relationship skills, and responsible decision-making (Durlak et al., 2011; Taylor et al., 2017). In our project, we call this youth empowerment training.

In the child psychology literature, a few meta-analyses have found that school-based programs incorporating socio-emotional learning into their curriculum increase pro-social behavior, lower levels of problematic behavior, and improve academic performance (Sorrenti et al., 2020; Taylor et al., 2017; Blewitt et al., 2018). Effective mastery of social-emotional skills is associated with better well-being and school performance, as well as acting by internalized beliefs, caring for others, making good decisions, and taking responsibility for one's choices (Durlak et al., 2011). Other impact evaluations tend to find similar results for school-age children (Sorrenti et al., 2020).

This has been extended to other populations as well. In the development literature, some studies have explored the effect of different forms of socio-emotional learning (sometimes combined with vocational education) on well-being and labor market outcomes. Liberia's Sustainable Transformational Youth program found that employment and cognitive behavioral therapy increased self-control and noncriminal values, and lowered violence in ex-combatants (Blattman and Annan, 2015). A women's life skills program in Liberia resulted in more access to money and self-confidence, and less anxiety about circumstances (Adoho et al., 2014). A women's empowerment training program in Uganda, which combined vocational education with reproductive health education, found that four years post-intervention, women are more likely to engage in income-generating activities, are less likely to be pregnant, are less likely to get married early, and are less likely to have intercourse against their will (Bandiera et al., 2014).

There is thus extensive evidence that complementary interventions are important for the success of vocational education programs. What has not so far been explored is the value-added of bundling vocational education with the training of such skills. This study expands the current knowledge by exploring whether it is not only hard skills acquired through vocational education but also soft skills acquired through a youth empowerment program, that boosts an individual's ability to succeed after training.

To summarize, there have been studies on vocational education, which have not consistently found positive effects on life outcomes. There have been studies on vocational education programs bundled with complementary interventions, which have found mostly positive effects on life outcomes. Our study combines both approaches, comparing vocational education programs to vocational education in addition to other training programs, to see whether the combination of technical and youth empowerment provides positive results for participants.

2.1.3 Peace-building and mental health in a post-conflict setting

This study also contributes to the literature on peace-building in a post-conflict setting. A variety of studies explore the economic effects of conflict on social cohesion. For example, Cassar et al. (2013) show that trust significantly decreased as a result of the Tajik civil war; in terms of nation building, Depetris-Chauvin and Durante (2017) show that following a win for a country's national soccer team, individuals are less likely to report a strong sense of ethnic identity and more likely to trust people of other ethnicities. In the peace-building literature, Blattman and Annan (2016) shows how giving agricultural training and future economic incentives reduce the probability of rebellion for ex-fighters of the Liberian civil war, providing evidence that the opportunity cost of rebellion is a leading motivation for conflict.

There is also a literature on the mental health effects of violence. For example, Moya and Carter (2019) show how the psychological effects of violence can create its own kind of poverty trap by engendering hopelessness and underachievement. This builds on Moya (2018) who, collecting evidence from victims of violence in Colombia, finds that traumatic experiences can alter economic behavior by triggering severe anxiety.

The current study fits into the nation-building literature by exploring how both skills training and youth empowerment can help a group of young adults - affected by a conflict either directly via displacement or indirectly - improve their economic conditions.

2.2 The study context

This study took place in two lower-income suburbs of the capital of Colombia, Bogotá: Bosa (a neighborhood of Bogotá) and Soacha, the latter of which is in the neighboring state of Cundinamarca but borders the city. This section provides some details on the economic and historical background of Colombia, the two neighborhoods in which this study took place, and some information on the implementing NGO that developed the project this study aims to evaluate.

2.2.1 Colombia's economic and historical context

Colombia is one of Latin America's largest economies, and it suffers from high poverty rates and income inequality: with a GINI coefficient of 54.2 in 2020, Colombian income inequality was higher than that of its neighbors (Brazil's GINI was at 48.9, Ecuador's at 47.3, and Peru's at 43.8, for example)[2]. Rates of informal employment are high - 62.1% of employment was informal in 2019[3]. The economic context for young adults is also difficult, with 27.4% of young

[2]Source: World Bank Poverty and Inequality Platform. GINI Index: Colombia. Accessed 28 August, 2022. https://data.worldbank.org/indicator/SI.POV.GINI?locations=CO&view=map&year=2020

[3]Source: ILOSTAT. Statistics on the Informal Economy. Accessed on 28 August, 2022. https://ilostat.ilo.org/topics/informality/

adults in 2021 not employed, in education, or training[4].

Colombia has also suffered from a decades-long civil war that began in 1964 and has caused 220,000 deaths, 25,000 disappearances, and 6.7 million internally displaced peoples (IDPs). The conflict has implicated left-wing guerilla groups, the government military, and right-wing paramilitary groups. Briefly, the conflict was launched in the 1960s by communist guerillas; in the 1980s, right-wing paramilitaries originally formed by landowners responded; and in the 1990s, government forces joined the paramilitaries against guerilla groups. Parties negotiated an official peace in 2016 - though the conflict is ongoing (Bazzi et al., 2022).

This internal conflict has created one of the world's largest populations of IDPs, who often abandon their assets and their social and family network. The majority of IDPs move from the rural area to an urban environment. Economically, displacement results in the loss of assets, capital, and labor; furthermore, the emotional impact can have long-lasting repercussions, which include trauma. Furthermore, the probability of being unemployed as an IDP is three times higher than for any other population group in the country - in 2012, the unemployment rate of IDPs amounted to 35%, compared to a national average unemployment rate of 10% (UNDP, 2014).

The current economic landscape is further darkened by the large influx of Venezuelan immigrants which began in 2015 as a result of deteriorating economic conditions in Venezuela. Since then, Colombia has welcomed almost 2.5 million people from Venezuela - this is around one-third of all Venezuelan refugees in Latin America. This sudden and large influx of people has created important economic impacts that have affected mostly only a handful of departments, including Bogotá, which hosts 80% of all Venezuelan migrants. This influx of migrants has increased the number of people in vulnerable conditions in these locations, as well as more perceived insecurity, more costs for the health sector, increased incidence of diseases such as diphtheria, measles, and tuberculosis, increased burdens on the education sector (even if many of the children migrants are not enrolled in schools, which is another burden on the economy). Relating to employment, Venezuelan migrants seem to be crowding out workers from the formal labor market, and also given that they are willing to work for lower salaries, they are also more likely to be hired (Bank, 2018).

Within this context, it becomes important to understand which policies and programs can be implemented to support marginalized at-risk youth who may lack the technical skills required by the formal labor market and who may have suffered from exposure to the traumatic events of violence, conflict, and forced migration.

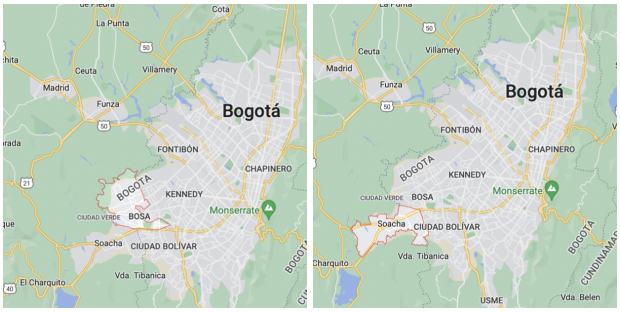
[4]Source: Departamento Administrativo Nacional de Estadística (DANE). Gran Encuesta Integrada de Hogares 2021.

2.2.2 Bosa and Soacha context

The neighborhoods in which this study takes place provide a unique opportunity to study urban marginalized young adults in Colombia. The study takes place in two lower-income suburbs of the capital of Colombia, Bogotá: Bosa (a neighborhood of Bogotá) and Soacha, the latter of which is in the neighboring state of Cundinamarca but borders the city. Many who live in these suburbs are lower-income and younger. In addition, a large percentage of the country's internally displaced people (IDPs) have settled in Soacha.

Within the study context, 70% of people living in Soacha and 35% of people in Bosa are eligible for social assistance, and 20% do not have health insurance. Most people stop studying after high school, and unemployment is higher than average (DANE, 2019). Anecdotal evidence supports the high criminality in Soacha in particular, where stories of 'social cleaning' by criminal gangs (murders perpetrated by gang members to clean the neighborhood of undesired members) are widespread and where young adults who do not participate in gang activities or suspected to be informers are especially targeted.

Figure 1: **Bosa and Soacha**



Notes: The figure highlights the two neighborhoods, Bosa, Bogotá (left) and Soacha, Cundinamarca (right), in which the study takes place. Source: Google Maps.

2.2.3 Project partners

This randomized controlled trial was undertaken in collaboration with Vivamos Mejor, a Swiss foundation that, together with its long-term local partners, e.g. Fundación Apoyar, is involved in various projects targeting vulnerable youth in disadvantaged areas of South America. In Colombia, Vivamos Mejor promotes programs to facilitate the acquisition of hard and soft skills and attitudes for transitioning toward formal and stable employment. Their local implementing partner, Fundación Apoyar, has been providing vocational education, along with youth empowerment training, to vulnerable young adults attempting to create a better future for themselves.

Fundación Apoyar has been working in these neighborhoods since 2012 as part of its young

adult labor market integration and youth empowerment project (*"Desarrollo laboral juvenil"*). The project targets young adults who fulfill certain vulnerability criteria, such as scoring low on the socioeconomic index SISBEN, being internally displaced, and having precarious housing situations. Fundación Apoyar works with the Association of Community Education, Development, and Labor (ASOCEDT) to provide these young adults with vocational education. On top of this, Apoyar provides these participants with soft skills training, psychosocial support, and labor market training and placement[5]. This is the project that the study aims to evaluate and that I describe in more detail below.

3 Study design

3.1 The intervention

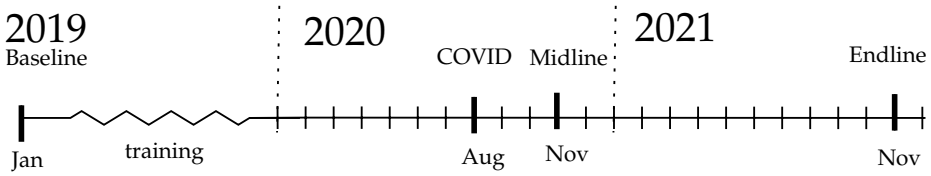
The study took place between 2019-2021 and aimed at evaluating the impact of Fundación Apoyar's youth empowerment program, which consists of soft skills training, psycho-social support, nation-building, and labor market integration. Residents of Bosa and Soacha were eligible if they were between the ages of 18-25, were finishing or had finished their high school education, and were economically disadvantaged (defined as being eligible for social assistance through the System of Identification of Social Program Beneficiaries (SISBEN)[6]). The study followed 300 vulnerable individuals that were randomized into three equally sized groups of 100 individuals each. The first group (group VET+YE) received vocational education with youth empowerment training. The second group (group VET) received vocational education. The last group (the Control Group) was left untreated.

Participants interested in the program could apply to one of five occupations: human resources, administration, early childhood care, logistics, and marketing. They were then randomized in early 2019 following a baseline data collection. Participants who were sorted into the control group did not receive support from the project to pursue their vocational education. The training programs ran throughout the academic year of 2019 (from February to December 2019). Participants received the standard nationally-verified vocational education certificate at the end of the program. Brochures with more information on each of the five occupations can be found in the appendix section A.1.1. The team then collected data on participants in August 2020, November 2020, and November 2021. Fundación Apoyar collected labor-related data every three months as of February 2020.

[5] https://www.fundacionapoyar.org.co/inicio/nuestra-historia/

[6] More details on vulnerability criteria, as well as on the overall selection of participants, can be found in section 3.2.

Figure 2: **Project timeline**



Notes: The figure illustrates the timeline of the intervention and the main stages of data collection.

Vocational education program Participants following vocational education received the nationally certified twelve-month classroom and training program for one of five professions during the calendar year of 2019: administration (*"auxiliar administrativo"*), early childhood care (*"atención a la primera infancia"*), sales and marketing (*"Mercado y ventas"*), human resources (*"recursos humanos"*), and supply chain and logistics (*"almacenamiento y aprovisionamiento"*).

Participants in the two treatment groups (VET+YE and VET) could select which of these vocational education training programs they would participate in. Participants in group VET received only vocational education, while those in group VET+YE complemented their vocational education training with the youth empowerment training program. The final numbers of participants per vocation and per group can be found in table 1 below. The administration program was only offered for participants who resided in Bosa, the child care, supply chain, and marketing programs were only offered for participants who resided in Soacha, and the human resources program was offered to participants in both Bosa and Soacha.

Table 1: **Vocational Education Training Intentions, by group**

Vocation	VET+YE	VET	Control	Total
Administration	14	14	14	42
Child Care	13	16	15	44
Supply chain	11	11	12	34
Marketing	21	17	21	59
Human Resources	41	41	38	120

Notes: The table illustrates how many participants, per group, were asking to be enrolled in each of the occupations offered by the program. Only VET+YE and VET participants were supported to participate in the program. Table 3 shows actual participation.

Each of the training programs consisted of morning and evening classes, Monday through Friday, for a total of 1,150-1,172 hours of coursework. This took place for one year, from February to December 2019. The first six months consisted of classroom-based lessons, while the

second six months included practical training. Participants received a certificate upon completing their training programs. More information on each of these programs can be found in the appendix section A.1.1.

Youth empowerment training program: In addition to this, participants in group VET+YE followed a youth empowerment training program which had the following four-dimensional structure: (1) socio-emotional skills and life plan development (done via group sessions and in individual psychotherapy), (2) political empowerment and post-conflict peacebuilding (to live with former paramilitaries and to learn to defend their rights), (3) business skills training (mock interviews, curriculum vitae development), and (4) labor placement support.

Sessions for the empowerment training took place two days a week, for a total of twelve hours per week throughout the entire academic year (February-December 2019), at each of the headquarters of Fundación Apoyar in Bosa and Soacha. What follows is a more thorough description of each of the dimensions of the Youth Empowerment program.

(1) Socio-emotional skills and life plan development/psycho-social support: The first dimension, socio-emotional skills development, focused on group-oriented soft skills training sessions that targeted eight areas: 1) the state of the self (self-esteem, self-image, self-care); 2) interpersonal relationships (empathy, communication, teamwork); 3) ambitions and aspirations (leadership, self-management); 4) conflict management; 5) assertiveness (active listening, emotions management); 6) intellectual capacity (attention span, concentration, critical thinking); 7) sensitivity; and 8) norms and values (responsibility, planning). These were in-person, group-level, training sessions that took place twice a week for two hours, throughout the entire academic year of 2019. They were provided by the in-house mental health specialist of each of the two headquarters of Fundación Apoyar in Bosa and Soacha.

In addition, participants had the option to see an in-house psychotherapist to help them work through trauma, adjust expectations, or deal with the stressors of difficult living situations such as early pregnancies and living in violent environments.

(2) Political empowerment and post-conflict peacebuilding : The second dimension contributes to the post-conflict peace-building efforts in Colombia via peace-building and political empowerment workshops. These took place twice a week, for two hours, in person. Around 10% of the young adults in the program are internally displaced due to the ongoing conflict. Part of their integration into the labor market includes being able to work alongside someone who might have been part of a group that was responsible for their internal displacement. Therefore, the peace-building dimension of the training program encourages young adults to learn to forgive people who may have been involved in their victimization. The intent is to contribute to the ongoing peace-building efforts in post-conflict Colombia. In addition, this dimension encourages

the political empowerment of these young adults.

The first half of this dimension of the youth empowerment training program consisted of a political empowerment module. This could be described as a political science course on the Colombian government structure and included lessons about Colombia's government, its history, some theory on political systems such as democracies, an analysis of the current Colombian government, as well as an overview of certain laws such as environmental, criminal, and military laws.

The second half of this dimension consisted of a peace-building module. This also included lessons on the theory of conflict, the history of the Colombian conflict and peace efforts, the people involved in the conflict and negotiation process, as well as courses on labor laws. The goal is to contribute to more sustainable peace in the community and the country.

(3) Business skills training and (4) labor placement support: The third and fourth dimensions could be labeled under the labor market umbrella. These courses took place twice a week, for two hours each time, throughout the entire academic year of 2019 (February-December 2019). The business skills training component of the training aimed to increase employability, by training on job interviews, CV development, and general presentation skills. Moreover, participants were offered labor placement support, through career fairs and network-building activities.

Each week, subjects participated in a workshop that aimed at creating soft skills related to the labor market, such as adaptation to an office environment, ability to get organized, management of time, management of people (being friendly, communication), conflict management, stress management, critical thinking, analytical thinking, storytelling, cloud computing, labor laws, creating a CV, and learning to interview. These workshops were organized by Fundación Apoyar and were sometimes supported by staff from the National Training Service (SENA, which is the national vocational education training institute), the ministry of labor, or other similar institutions.

3.2 Selection of participants and randomization

Selection of participants The intervention targeted disadvantaged young people in the municipality of Soacha (Cundinamarca) and the neighborhood of Bosa (Bogotá). The implementing partner, Fundación Apoyar, has been operating in both neighborhoods for many years and thus has created a reputation for itself as an organization helping locals create better lives for themselves. For the recruitment of program participants, Fundación Apoyar began canvassing in the late months of 2018 by talking to community leaders, advertising the programs on posters, and relying on mouth-to-mouth information thanks to an established network of contacts.

It was important to agree with Fundación Apoyar on a study design that would be accepted by the local community. Fundación Apoyar has a long track record of doing good in each neighborhood and local headquarters that have long been the seat of various community projects and are therefore trusted.

Candidates were accepted under the following additional conditions: they had to have completed or be completing, their secondary education[7], be between 18 and 25 years old, have a high vulnerability score, pass an interview with the in-house psychologist, and have a family in support of them following the program.

The recruitment process included a vulnerability screening of each participant, followed by an interview with one of Fundación Apoyar's psychotherapists. Only the 300 most vulnerable were accepted to the program, due to budgeting constraints. A participant was considered vulnerable if, first, they were eligible for social assistance (by receiving a score from the System of Identification of Social Program Beneficiaries (SISBEN)) of 2 or lower; second, if their living conditions were poor, as judged by Fundación Apoyar following a home visit; and, third, if they scored as vulnerable in a vulnerability score computed by Fundación Apoyar.

The first criterion of vulnerability relied on the System of Identification of Social Program Beneficiaries (SISBEN) vulnerability index. To give some context, SISBEN is a vulnerability index that is used to identify beneficiaries of social assistance in Colombia and is mandatory for all municipalities. Under its current iteration, it is an index between 0 (poorest) and 100 (richest) calculated on health, education, housing, and vulnerability (ILO, 2015). The SISBEN score that we use to measure vulnerability for our participants relies on the old score, which is a number between one (poorest) and six (richest). Young adults who hold a score between one and three were eligible for the project.

The second criterion for vulnerability relied on an in-house visit of a project worker from Fundación Apoyar to an applicant's home and lasted about one hour. These in-house visits assessed the living conditions of the applicants, in particular the type of home, who owns the home, how many people live in it, what kind of materials the floor, roof, and walls are made up of, whether there is water, light, gas, internet, or telephone access, whether the home is accessible by road or public transport, what kinds of appliances and how many of them there are (computer, television, microwave, washing machine, and so on), among other such criteria. Furthermore, these visits served to collect further information such as family incomes, lifestyle costs (including rent, food, and transport), whether the home provides all necessary living conditions, the education levels of household members, the occupation of family members, as well as their health conditions. An example of the household survey can be found in the appendix.

[7] Also known as the *Bachiller*.

The third criterion for vulnerability relied on Fundación Apoyar's vulnerability index, which ranged from 0 to 80. Anything above 40 was considered a high vulnerability.

Finally, the interview with Fundación Apoyar's therapists screened participants for their motivation to join the program and provided some background on their ability to complete the program, to avoid attrition. To reduce the risk of attrition, only subjects with families that approved of their studies were selected. For example, the goal was to avoid situations where participants could not finish the program because their families preferred that they work. Interview questions focused on the relationship with family, what applicants' academic performance is like, what their hobbies and plans for the future are, what their work experience and work style are, and their hopes for the program.

Stratification and randomization Overall, 361 candidates completed a baseline survey and were randomly assigned into one of three groups. Randomization was done via public lottery to reduce suspicion related to treatment allocation and to increase the transparency of the randomization process. This was done in light of the high mistrust of authorities in Colombia. A series of in-person gatherings were organized in Fundación Apoyar's headquarters between December 2018 and January 2019. There were several events to accommodate the working schedules of candidates.

Each participant was asked to fill out a survey and provide administrative data, and then picked out a lottery ball with A (Group VET+YE), B (Group VET), or C (Control Group) written on it, then went into a separate room with people who picked the same letter to receive a briefing on their treatment group. Randomization was stratified at the neighborhood and vocation levels. Participants received a cash incentive of approximately 10 USD to offset the opportunity cost of attending these sessions. For each event, the probability of being sorted into any of the three groups was always 1/3. This was ensured by removing the number of balls of each group available in the lottery bag according to how many people were present.

Overall, 361 individuals were randomized into the three different groups and completed the baseline survey. Up until April 2019, participants who dropped out were replaced by participants from a waitlist. The final sample is 300 participants who were followed throughout the four years of the experiment.

Table 2 shows whether the composition of the three treatment groups is balanced at baseline. Balancedness implies that the three groups are composed of roughly the same kinds of people, so that they may be compared. The idea is to create three groups that could be considered randomly allocated. Group means are presented in the first three rows, with standard deviations in parentheses (VET+YE in column 1, VET only in column 2, and control in column 3). Column 4 presents the difference in means between VET+YE and the control group, with standard

deviations in parentheses. Positive values in column 4 indicate that participants in VET+YE have higher levels of a variable compared to the Control group. Negative values in column 4 indicate that participants in VET+YE have lower levels of a variable compared to participants in the Control group. Column 5 presents the difference in means between VET and the control group, with standard deviations in parentheses. Positive values in column 5 indicate that participants in VET have higher levels of a variable compared to the Control group. Negative values in column 5 indicate that participants in VET have lower levels of a variable compared to participants in the Control group. Column 6 shows the difference between participants in VET+YE and VET. The number of observations is in the last row.

Column 4 indicates that compared to the Control group, participants in VET+YE are on average younger, have fewer children, are 11% less likely to be displaced, and score slightly higher on the socio-economic vulnerability index (SISBEN).

Column 5 indicates that compared to the Control group, participants in VET only are 14.8% more likely to be women, are on average one year younger, have fewer children, are 9% less likely to be displaced, are 9% more likely to work in the informal sector, and score on average 2 points lower on the distress test. Column 6, in turn, indicates that there are no significant differences between participants in VET+YE and VET.

Taken together, the slight unbalancedness between groups implies that when comparing group averages it is necessary to be cautious when it comes to the analysis. Imbalances among variables observed at baseline are natural and do not signify that the randomization failed; rather, such imbalances occur by pure chance. The question is whether they happen among variables that could confound the results, resulting in omitted variable bias. For instance, perhaps men do better in the labor market in general compared to women. If one of the treatment groups by chance has more men than women (which is the case - the Control group has more men than the VET-only group), then that treatment group may do better on average - mechanically - in the labor market compared to the other two treatment groups. Controlling for this in the regressions means that this potential bias is corrected for.

Potential design biases: It could be possible that the randomization process led to some disillusionment on the part of participants that were sorted into the Control group. The fear here is that this disillusionment would lead to an overall lack of investment into self-development and create a self-fulfilling prophecy whereby participants in the Control group do badly in life because they are negatively impacted by the outcome of the sorting. These concerns are valid. The study was designed with this fear in mind, and while it is not possible to eliminate this possibility, the team took great care to ensure participants in the Control group could see some benefit in participating in the study. Participants were briefed on the point of the study and

Table 2: **Balance at baseline**

Variable	(1) VET+YE	(2) VET	(3) Control	(4) Diff (1)-(3)	(5) Diff (2)-(3)	(6) Diff (1)-(2)
Female	0.710	0.808	0.660	0.050	0.148**	-0.098
	(0.456)	(0.396)	(0.476)	(0.449)	(0.018)	(0.107)
Age	20.140	19.980	20.980	-0.840*	-1.000**	0.160
	(3.269)	(3.574)	(3.423)	(0.077)	(0.045)	(0.742)
Number of Children	0.390	0.444	0.710	-0.320***	-0.266**	-0.054
	(0.650)	(0.772)	(0.868)	(0.004)	(0.024)	(0.591)
Displaced (1=yes)	0.080	0.091	0.190	-0.110**	-0.099**	-0.011
	(0.273)	(0.289)	(0.394)	(0.023)	(0.045)	(0.784)
SISBEN 1-6 scale	1.760	1.667	1.560	0.200**	0.107	0.093
	(0.553)	(0.535)	(0.538)	(0.010)	(0.162)	(0.227)
Earnings (USD/mo)	28.130	33.277	31.454	-3.324	1.823	-5.147
	(71.831)	(67.989)	(74.662)	(0.749)	(0.857)	(0.604)
Formal earnings (USD/mo)	5.220	5.126	14.764	-9.544	-9.638	0.094
	(40.255)	(35.897)	(60.706)	(0.192)	(0.175)	(0.986)
Informal earnings (USD/mo)	22.910	29.029	17.676	5.234	11.354	-6.119
	(61.488)	(61.277)	(51.565)	(0.515)	(0.159)	(0.483)
Employed (1=yes)	0.180	0.242	0.190	-0.010	0.052	-0.062
	(0.386)	(0.431)	(0.394)	(0.856)	(0.371)	(0.283)
Formal sector work (1=yes)	0.020	0.020	0.060	-0.040	-0.040	-0.000
	(0.141)	(0.141)	(0.239)	(0.150)	(0.155)	(0.992)
Work in the informal sector	0.160	0.222	0.130	0.030	0.092*	-0.062
	(0.368)	(0.418)	(0.338)	(0.549)	(0.088)	(0.266)
Work Experience	0.680	0.657	0.720	-0.040	-0.063	0.023
	(0.469)	(0.477)	(0.451)	(0.539)	(0.336)	(0.727)
PTSD severity index	30.010	30.788	31.705	-1.695	-0.917	-0.778
	(9.661)	(10.477)	(11.579)	(0.270)	(0.563)	(0.589)
PTSD DSM4 criterion	0.082	0.121	0.116	-0.034	0.005	-0.040
	(0.275)	(0.328)	(0.322)	(0.428)	(0.908)	(0.360)
Depression	4.227	4.323	4.500	-0.273	-0.177	-0.096
	(3.621)	(3.854)	(3.681)	(0.608)	(0.748)	(0.858)
Distress	10.062	8.583	10.703	-0.642	-2.120*	1.478
	(5.916)	(5.388)	(6.989)	(0.574)	(0.062)	(0.148)
Observations	100	99	100	200	199	199

Notes: The table shows whether the composition of the three treatment groups is balanced at baseline. Group means are presented in the first three rows, with standard deviations in parentheses (VET+YE in row 1, VET only in row 2, and control in row 3). Row 4 presents the difference in means between VET+YE and the control group, with standard deviations in parentheses, and row 5 presents the difference in means between VET and the control group, with standard deviations in parentheses. Row 6 shows the difference between participants in VET+YE and VET. The number of observations is in the last row. Stars represent significance levels, with one star representing 0.1 significance, two stars representing 0.05 significance, and three stars representing 0.01 significance.

encouraged to participate in the general development of the area. One participant, in the exit survey, even wrote "thank you for helping the vulnerable youth of Colombia." Furthermore, all participants received roughly one day's worth of pay (around 10 USD) for filling out data. Moreover, the events where data was collected from the Control group were designed to be fun

for participants. There, Fundación Apoyar organized lotteries to win certain prizes such as bicycles or radios, and there was entertainment (in terms of music and dancing) and refreshments provided.

Finally, exit surveys were conducted at the end of the randomization event, where participants were asked, (1) "How did you feel after finding out about your group?" and (2) "Why will you continue following the activities in your group?". I coded the answers provided by participants in the Control group as pessimistic if they contained words such as "sad," "disillusioned," or "demotivated". I coded the answers as neutral if they included words such as "normal" or "fine". Finally, I coded the answers as optimistic if they indicated motivation to continue studying, excitement at the idea of the impact evaluation, or encouragement at the idea of the economic support from the study.

For the first question, where participants share their feelings for being sorted into the Control group, 57% indicated pessimism such as sadness, 27% indicated neutral feelings, and 16% indicated they felt optimism (such as they felt better things would come from this). For the second question, about why they would continue participating in activities for the impact evaluation, 9% indicated a pessimistic answer (saying they would prefer not to participate), 39% indicated a neutral answer (such as having to move on from this), and 52% indicated an optimistic answer. These answers indicate that while it is not possible to rule out that feelings of discouragement would make participants in the Control group do worse in life, these answers indicate an overall acceptance of results as nothing more than one extra thing in life. Feelings of pessimism at the outcome (question 1) did not seem to discourage participants from wanting to participate in the study and look at the positive of receiving the benefits from being in the Control group (question 2).

In addition, in the medium run, the team ran a manipulation check to see whether participants think they learned something from their training, and to see what the "normal life" of someone in the Control group would look like. In November 2020, in addition to the midline survey, participants were also asked if they had done any sort of mental health training, if they had participated in cultural workshops, if they had done classes and what kind, and if they were currently working and in what. Looking at the Control group, it is possible to gather an idea of how they continued along their lives and can give some insight into whether participants in the Control group, in the medium run, felt cheated by not being in any of the treatment groups - by, perhaps, not continuing down any education path.

Looking at the manipulation check, it is possible to see that 23% of participants in the Control group are working in November 2020, 11% started studies (VET, technological, and high school), 5% finished their studies, and 20% have been invited to interviews. It does not seem

that they have been dramatically negatively impacted as a result of being sorted into a Control group for the study. Also worth mentioning is that several of the participants who did sign up for studies, particularly vocational education, ended up dropping out, citing a lack of time as the reason for dropping out. This is in line with the experience of the VET-only group, indicating that perhaps this is a common experience for young adults who choose to go into vocational education. The smaller dropout rates for VET+YE may be an extra benefit of the youth empowerment program that is being evaluated.

With this in mind, it could also be possible that the results are biased by the opposite effect - the Control group becomes a treatment group in and of itself and does better compared to the general population. This is more difficult to test for and to control for, though it would be ideal to have this information.

Attrition and noncompliance Here, I define attrition as people who completed the baseline survey but who did not complete the endline survey. I define noncompliance as people who completed the baseline survey, but who either dropped out of the treatment group they were assigned to (Groups VET+YE and Group VET), or people who were assigned to the untreated control group and followed through with vocational training regardless.

Table 3: **Breakdown of attritors and noncompliers by group**

Career	VET+YE	VET	Control	Total
Attrition				
Baseline	125	126	110	361
Endline	100	99	100	300
Noncompliance				
Finished VET	76	57	3	136
Did not finish VET	49	69	107	225

Notes: The table shows how many people were surveyed at baseline, per group, and how many people were surveyed at endline, per group (first two rows). The third and fourth rows show the number of people who responded they did not finish vocational education per group, reported during the midline survey, and the number of people who did finish vocational education per group also reported during the midline survey.

In terms of compliance, enumerators asked 299 participants during the November 2020 midline data collection to indicate whether they finished vocational education (and if not, why not). Of those participants, 136 finished vocational education: 76 in group VET+YE, 57 in group VET, and 3 in the control group finished vocational education (the three participants in the control group who finished vocational education are noncompliant). Most participants cited children,

pressure from work, and lack of resources as reasons for not finishing the program. Table 3 summarizes this information by group.

Table 4 shows the results of the regression of treatment on attrition, to explore whether attrition is random, compared to the control group. Effectively, I estimate:

$$y_{i,g,t} = \alpha + \theta_1[\text{VET + YE}]_g + \theta_2 \text{VET}_g + \gamma_1 X_{i,g} + \varepsilon_{i,g} \tag{1}$$

Where the dependent variable for attrition is equal to one if a participant is present in the baseline but not in the endline survey. The regression is run on the baseline and endline sample of the 300 final participants in the study. Also, $X_{i,g}$ indicates the control variables of gender, being internally displaced, and neighborhood. From the lack of statistical significance, it does not seem that being treated affects attrition. Being treated does not affect the probability of attrition.

Table 4: **Effects of treatment on attrition and noncompliance**

	In endline	In endline	Noncompliance
VET + YE	0.004	-0.003	-0.178***
	(0.004)	(0.003)	(0.064)
VET	0.006		
	(0.005)		
Control		-0.006	-0.401***
		(0.005)	(0.053)
Individual Controls	Yes	Yes	Yes
Observations	598	598	598
R-squared	.0000521	.0000521	.16

Notes: The dependent variable is attrition and noncompliance. VET + YE indicates the group of participants who received vocational education and youth empowerment, while VET indicates the group of participants that received only vocational education, the omitted category is the control group. Robust standard errors clustered at the individual level are reported in round brackets. ***, **, and * indicate significance at the 1-, 5-, and 10- % levels, respectively.

3.3 Data collected

The study seeks to find out whether bundling vocational education with youth empowerment training helps vulnerable young adults find better jobs following training in the medium run. To this end, the team collected a variety of data at various points in time: a baseline survey in January 2019 before the training; a COVID-19 lockdown survey in August 2020; a midline survey in November 2020; and an endline survey in November 2021. In addition, phone interviews were conducted every three months to collect information on each participant's address, earn-

ings, job status, and their family status (if they had moved in with their partner, were pregnant, and so on).

Earnings and labor market participation: participants self-reported their average hourly earnings from the last week as well as how many hours they worked in that past week. We use this information to approximate monthly earnings. Further, participants were asked whether these are from a formal job or not (i.e. if they have a contract). The team corroborated self-reported formal earnings by double-checking that the participant has current records contributing to the retirement pension called PILA, following Attanasio et al. (2017). These were collected either in person pre-randomization in December 2018, or via phone after the treatment was over in February 2020, August 2020, December 2020, March 2021, May 2021, and November 2021.

Mental illness variables: Variables included symptoms of **post-traumatic stress disorder (PTSD)**, which is a psychiatric disorder characterized by a failure to recover from a traumatic event; as well as by other mental health problems such as **depression** and **distress**; and, finally, whether training affected preferences that have always been assumed to affect economic outcomes (trust, negative and positive reciprocity, and so on, based on Falk et al. (2016)). These variables were collected at baseline (pre-randomization on December 2018), at the midline (November 2020), and at the endline (November 2021) via a paper survey.

Trauma was measured by the PCL-C, a standardized self-report rating scale of post-traumatic stress disorder. We compute a severity score index by summing up over the 17 questions. The severity score ranges between 17 and 85 where anything above 30 could be considered a potential for having symptoms of PTSD. We also follow the guidelines for constructing a 0-1 variable which indicates a presumptive diagnosis of PTSD.

The PCL-C is not a psychiatric evaluation and rather indicates whether participants may wish to seek additional support in diagnosing their trauma. Participants were asked to answer 1 when a statement indicated an answer of "Not at all", and 5 when a statement indicated an answer of "Extremely". We compute a severity score index by summing up over the 17 questions. The severity score ranges between 17 and 85. Anything above a score of 30 is indicated for exploring primary care for the treatment of PTSD. We also follow the guidelines for constructing a 0-1 variable which indicates a presumptive diagnosis of PTSD. A participant meets DSM-IV criteria for a presumptive diagnosis if they answer symptomatically ('Moderately' - 3 - or higher) to at least 1 question between questions 1-5, to at least 3 questions between questions 6-12, and at least 2 questions between questions 13-17. We tend to prefer the presumptive diagnosis variable given that it allows for more subtlety in the severity of questions.

Behavior variables were measured according to Blattman and Annan (2015), mainly to com-

pute mental health stability (distress, depression). These variables are created by asking a series of questions that are then summed up to give vectors of different lengths. These vectors are subsequently standardized for ease of interpretation. For example, in the case of depression symptoms, the respondent is asked to disagree/agree (giving a score from 0 to 3) with statements such as "I feel bad doing things I normally do" or "I am feeling sad or downhearted". All details for the construction of these variables can be found in the appendix section A.1.2.

Finally, the team also collected data on preferences that affect economic outcomes. The Global Preference Surveys created by (Falk et al., 2016) were administered, measuring trust, positive reciprocity, altruism (charity), negative reciprocity, risk preference, and time preference.

Demographic characteristics: include gender, age, marital status, number of children, being displaced by conflict (participants who are displaced have a registration card proving they are internally displaced due to the ongoing conflict), and socio-economic status (the (old) SISBEN index was a 1-6 index, given to people by the government, and is an indicator of whether or not someone is eligible for government aid). These were collected either in person pre-randomization in December 2018, or via phone after the treatment was over in February 2020, August 2020, December 2020, March 2021, May 2021, and November 2021.

4 Descriptive statistics at baseline

Table 5 provides some statistics at baseline on the participants in the sample.

Table 5: **Descriptive statistics of the sample at baseline**

	Total			VET+YE			VET			Control		
	mean	**sd**	**count**	**mean**	**sd**	**count**	**mean**	**sd**	**count**	**mean**	**sd**	**count**
Demographics at baseline												
Female	0.73	0.45	299	0.71	0.46	100	0.81	0.40	99	0.66	0.48	100
Age	20.37	3.44	299	20.14	3.27	100	19.98	3.57	99	20.98	3.42	100
Number of Children	0.52	0.78	299	0.39	0.65	100	0.44	0.77	99	0.71	0.87	100
Displaced	0.12	0.33	299	0.08	0.27	100	0.09	0.29	99	0.19	0.39	100
SISBEN	1.66	0.55	299	1.76	0.55	100	1.67	0.53	99	1.56	0.54	100
Labor market values at baseline												
Earnings (USD/mo)	30.95	71.35	299	28.13	71.83	100	33.28	67.99	99	31.45	74.66	100
Formal earnings (USD/mo)	8.38	46.98	299	5.22	40.26	100	5.13	35.90	99	14.76	60.71	100
Informal earnings (USD/mo)	23.19	58.27	299	22.91	61.49	100	29.03	61.28	99	17.68	51.57	100
Employed	0.20	0.40	299	0.18	0.39	100	0.24	0.43	99	0.19	0.39	100
Formal sector work	0.03	0.18	299	0.02	0.14	100	0.02	0.14	99	0.06	0.24	100
Work in the informal sector	0.17	0.38	299	0.16	0.37	100	0.22	0.42	99	0.13	0.34	100
Work Experience	0.69	0.47	299	0.68	0.47	100	0.66	0.48	99	0.72	0.45	100
Mental health at baseline												
PTSD severity index	30.83	10.58	292	30.01	9.66	98	30.79	10.48	99	31.71	11.58	95
PTSD DSM4 criterion	0.11	0.31	292	0.08	0.28	98	0.12	0.33	99	0.12	0.32	95
Depression	4.35	3.71	285	4.23	3.62	97	4.32	3.85	96	4.50	3.68	92
Distress	9.81	6.18	189	10.06	5.92	65	8.58	5.39	60	10.70	6.99	64

Notes: The table lists statistics for participants in overall (first four columns), for participants in VET+YE (columns 5-8), VET only (cols 9-12) and control (cols (13-16).

Demographically speaking, it is interesting to see the composition of the sample in terms of gender, displacement status, and socio-economic status as measured by the 1-6 SISBEN index as described in the data section, age, and the number of children. Most of the sample is made up of women: 217 out of the 299 participants are followed throughout the entirety of the study. There are 71 women in VET+YE, 80 in VET, and 66 in the control group. In addition, a non-negligible amount of people are displaced: 36 out of 299. There are 8 displaced people in VET+YE, 9 in VET, and 19 in the control group.

In addition, other variables of interest include labor variables including (conditional) earnings (and whether those are from an informal or formal source - all in USD per month), as well as whether someone is employed (and whether that is in the formal and informal sector), and whether they have work experience.

Finally, mental health variables are also of interest, such as PTSD severity, the potential diagnosis for PTSD according to the DSM4, the depression index, and the distress index, all as described in the data section above.

4.1 Vulnerability and earnings

Intuitively, one would expect that more vulnerable participants might be struggling more in terms of earnings at baseline. In our sample, it could be argued that women and displaced individuals could be considered the most vulnerable. Table 6 shows a histogram of women's and internally displaced people's *non-zero* earnings, compared to men and those who are not internally displaced, at baseline.

The first thing to note is that there are in general more men, and more not displaced people earning money. The first fact is surprising, given that our sample is made up mostly of women. The second is mechanical, as our sample is made up mostly of not displaced people. A large share of displaced people in our sample seems to be earning money.

Table 6: **Vulnerability and earnings**

	mean	min	max	median	sd	count
Earnings (USD/mo)						
men	175.84	43.5	348	181.25	85.12	20
women	139.90	7.25	377	130.5	78.23	41
not displaced	150.84	7.25	377	145	81.48	54
displaced	158.17	43.5	290	116	89.10	7
Formal earnings (USD/mo)						
men	274.39	226.49	348	261	47.20	5
women	226.78	145	377	226.78	95.02	5
not displaced	246.97	145	377	236.64	83.61	8
displaced	265.07	240.15	290	265.07	35.24	2
Informal earnings (USD/mo)						
men	149.56	43.5	236.64	174	68.38	15
women	130.25	7.25	232	116	67.80	36
not displaced	138.15	7.25	236.64	145	68.89	46
displaced	115.42	43.5	208.8	116	60.00	5

Notes: The table lists earnings statistics for women and internally displaced people at baseline. Earnings (based on positive earnings) are shown in the top rows; earnings from formal jobs are shown in the middle rows; earnings from informal jobs are shown in the last rows.

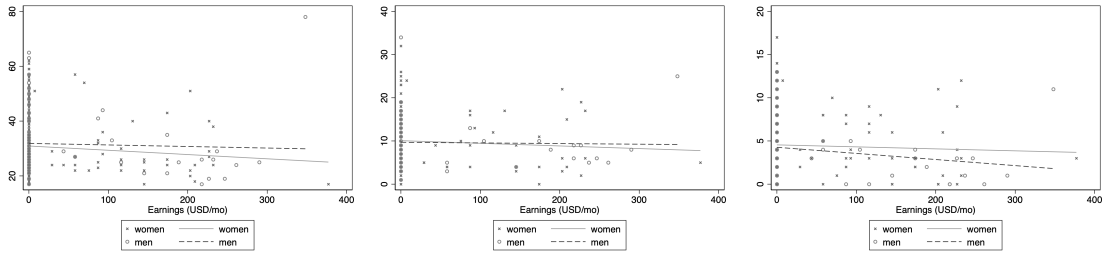
Second, men earn more than women at baseline. A simple comparison of group averages shows that the differences between men's and women's overall and formal sector earnings are statistically significant at baseline. Men earn on average $16.45 more than women at baseline, and when they earn formal sector earnings, they earn $11.50 more than women at baseline. All other differences between means are not statistically significant.

The difference in earnings between men and women is more strongly correlated when taking mental health into account. Figure 3 shows the correlation between (unconditional) earn-

ings and mental health, by gender, at baseline. It seems that, for women, having higher symptoms of trauma (panel 3a) and depression (panel 3c) is correlated with lower earnings. For men, on the other hand, having higher symptoms of trauma (panel 3a) and depression (panel 3c) does not seem to be correlated with earnings. However, having higher symptoms of distress (panel 3b) does seem to be correlated with lower earnings. It seems that women and people with higher symptoms of mental illness have lower earnings than men and people without mental illness at baseline.

Figure 3: **Gender, mental health, and (unconditional) earnings**

(a) Trauma and earnings (b) Distress and earnings (c) Depression and earnings



Notes: The graphs illustrate the relationship between mental health (trauma in panel a, distress in panel b, and depression in panel c), earnings (USD/month) and gender.

5 Methodology

In an ideal world, a simple difference in means from the randomized controlled trial design would effectively allow estimating the causal average effect of treatment on our outcomes. In theory, this would be the case because the randomized controlled trial creates a scenario in which the control group serves as a counterfactual group. By randomizing it is possible to create groups that have similar observerable and unobservable characteristics. This design, in theory, would imply that the difference in means between treatment and control groups is the causal average effect of treatment. However, several caveats are in order before establishing this.

Table 2 shows that there are some variables, such as gender and displacement status, which were not balanced across the three groups at the baseline. This implies that there could be some bias in the above numbers as a result - for instance, compared to both treatment groups, VET and VET+YE, participants in the control group have more children. If this implies that they have less time to work, they might mechanically earn less money than their peers who received training. It is for this potential omitted variable bias that we control for any unbalanced variables at baseline, in particular displacement status, gender, age, age squared, work experience at baseline, whether the participant worked in the informal sector at baseline, whether the par-

ticipant had any earnings at baseline, as well as whether the participant had two, three, four, or five children at baseline.

Furthermore, table 3 indicates that not everyone complied with the treatment - that is to say that some participants who were initially assigned to receive treatment dropped out for whatever reason. This complicates the exercise because participants who did not complete the training do not, in theory, benefit from the training as much as participants who did complete the training. There might be some effects of treatment assignment on treatment in and of itself. This can be mitigated by using the intent to treat it as an estimate. This effectively estimates the effect of treatment, weighted by the proportion of treatment each participant received.

Thirdly, it is possible that at each period when data was collected, particular issues are affecting participants who were in different groups differently. A clear example of such an event would be the COVID-19 pandemic-related lockdowns which were in effect throughout the summer of 2020. Such events affect all groups at the same time. Such unobserved time period-specific factors can be controlled by including time trends in the analysis. This averages out any time-varying unobserved variables that may be biasing the estimates.

Finally, the comparison exercise becomes more valuable by not only comparing treatment groups to the control group but also treatment groups before and after treatment. Differences-in-differences estimates the following numbers: (treated group after treatment at time t - treated group before treatment) minus (the untreated group after treatment at time t - the untreated group before treatment).

These caveats imply that to get a causal effect of the treatment, it is important to use a differences-in-differences analysis that also controls for omitted variables that may bias the estimates. This estimates the average effect of treatment on the treated participants (also known as ATT). I, therefore, follow Sun and Abraham (2021) in estimating a linear two-way fixed effects specification, using individual and time-specific fixed effects. To explore the treatment effect of VET+YE, I estimate equation 7:

$$Y_{i,t} = \alpha_i + \lambda_t + \sum_{t=E_1}^{\bar{T}} \mu_l \mathbf{1}\{t - E_1 = l\} \times [\text{VET+YE}_i] + \nu_{i,t} \tag{2}$$

Where Y is the outcome variable of interest in time period t for individual i. The outcomes of interest are overall earnings, informal earnings, and formal earnings. λ_t indicates time fixed effects, α_i indicates individual fixed effects, E_1 is the time of treatment (February 2019), $l = t - E_1$ are relative time periods since treatment, and $\bar{T}$ is the last survey period. In particular, $l = t - E_1 = 0$ if t is at baseline, and t is measured in quarters. The coefficient of interest is μ_l. VET+YE$_i$ is equal to one if a participant is in the VET+YE treatment group. When estimating the treatment effect of VET+YE compared to the Control group, I run equation 7 on the sample of

participants in VET+YE and the Control group so that the reference group is the Control group. When estimating the added value of VET+YE, I run equation 7 on the sample of participants in VET+YE and VET-only.

To explore the treatment effect of VET-only, I run equation 8 on the sample of participants in VET and the Control group, where VET_i is equal to one if a participant is in the VET treatment group:

$$Y_{i,t} = \alpha_i + \lambda_t + \sum_{t=E_1}^{\bar{T}} \mu_l \mathbf{1}\{t - E_1 = l\} \times [\text{VET}_i] + \nu_{i,t} \tag{3}$$

Given the relatively strong assumption of no heterogeneous effects on treatment in Sun and Abraham (2021), I check for robustness of results by estimating the regression following De Chaisemartin and d'Haultfoeuille (2020) and also by hand. Changing specifications does not significantly change results. These are reported in section 7.

6 Main Results

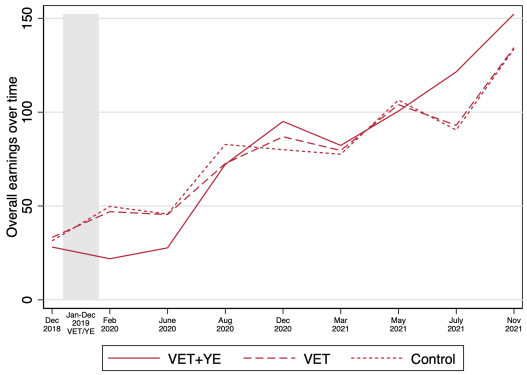
The purpose of this section is to explore how treatment has affected participants in the two years since their training ended. I will show results for labor market participation and earnings. For a detailed description of the effects of the program on mental health, please refer to Chapter 2 of this thesis.

6.0.1 Descriptive Evidence

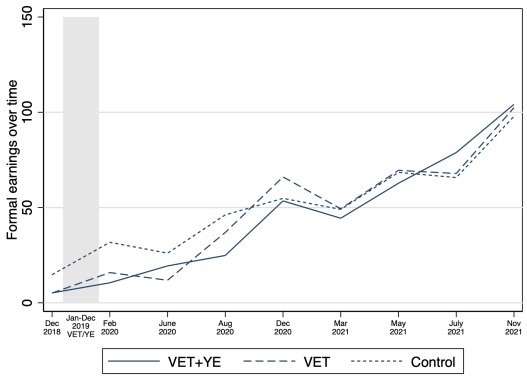
This section shows the raw outcome means for each group. The main outcome variables are earnings, which are set to zero in periods when a person does not generate any earnings. Reported below are also results for labor market participation, which is a binary variable that takes the value one if a person has any earnings, and zero otherwise. Earnings for the employed is the third outcome measure, which is like earnings but excludes people who do not generate positive earnings.

Figure 4: **Earnings over time**

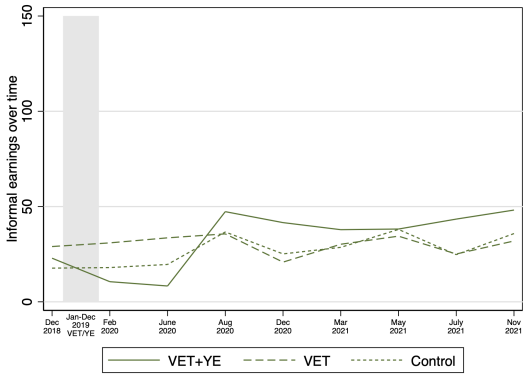
(a) Overall earnings



(b) Formal earnings



(c) Informal earnings



Notes: The figure illustrates average earnings, for the groups VET+YE (solid lines), VET (long-dashed lines), and the control group (short-dashed lines), from December 2018 to November 2021, unconditional on whether the individual is employed. That is to say, the earnings averages include the participants who also earn 0 USD/month, which pushes the averages down below the minimum wage (around 250 USD/month). The first panel shows overall earnings, the second formal sector earnings, and the third informal sector earnings.

Figure 4 displays average earnings for the three groups of participants, and by the source of income, e.g. overall, formal sector, or informal sector. First, group averages show that partici-

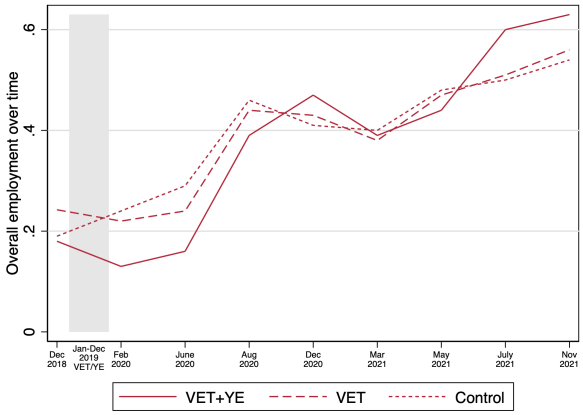
pants' overall earnings increase substantially throughout the study. While participants earned the equivalent of around 30 USD in December 2018, earnings increased more than five-fold to about 150 USD in November 2021. This growth in earnings is mainly driven by formal sector earnings, which are close to zero in December 2018 and reach almost 100 USD by the end of the study. In contrast, informal earnings are initially higher, around 30 USD, but remain at a level of no more than 50 USD throughout the period.

There are no major differences between the control group and the group following VET training. Income trajectories are fairly similar suggesting that VET training in itself has not had a major effect on earnings.

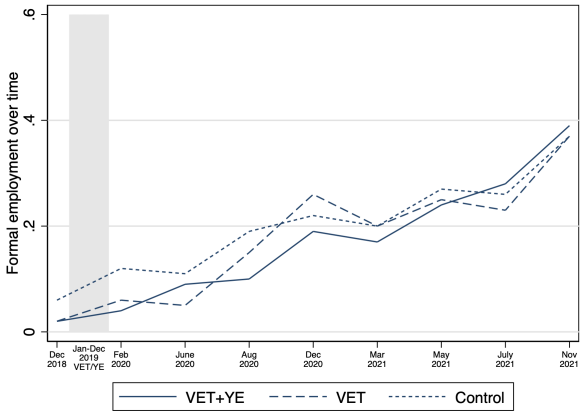
In contrast, a few months after vocational education, participants who followed the VET+YE treatment are earning less than the control group, but more than the control group towards the end of the observation periods, in July 2021 and November 2021. On closer inspection, this decrease appears driven by informal earnings, which are lower for VET+YE in the period December 2018 to June 2020, but then increase again and remain quite high throughout the period. Formal sector earnings contribute to this pattern, they are initially lower in the VET+YE group compared to the Control group, then higher towards the end of the sample period.

Figure 5: **Labor market participation over time**

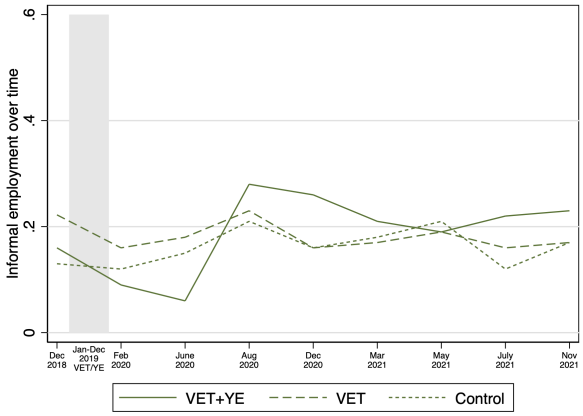
(a) Overall employment



(b) Formal employment



(c) Informal employment



Notes: The figure illustrates average employment rates, for the VET+YE treatment group (solid lines), VET treatment group (long-dashed lines), and the control group (short-dashed lines), from December 2018 to November 2021. The first panel shows overall employment, the second shows formal sector employment, and the third shows informal employment.

Figure 5 shows that labor market participation increases substantially over the three years,

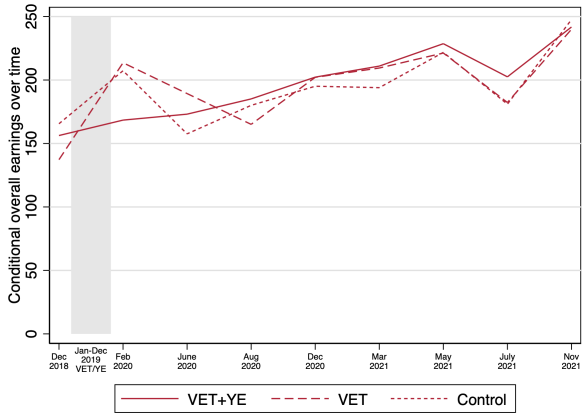
from December 2018 to November 2021. While initially around 20% of all individuals generate positive earnings, this proportion increases to around 60% in November 2021. Informal sector participation, at around 20%, is quite flat, which means that most of the gains in labor market participation come from the formal sector. Participation in the formal sector is close to zero initially and increases to around 40% over the three years that we observe.

Patterns in labor market participation are quite similar for the VET and the Control groups, except for when it comes to participation in formal employment. VET participants do not have jobs in the formal sector initially, in December 2018, whereas 10% of participants in the Control group individuals already participate in the formal sector. VET participants then catch up to the Control group and increase formal employment participation at a similar rate as participants in the Control group.

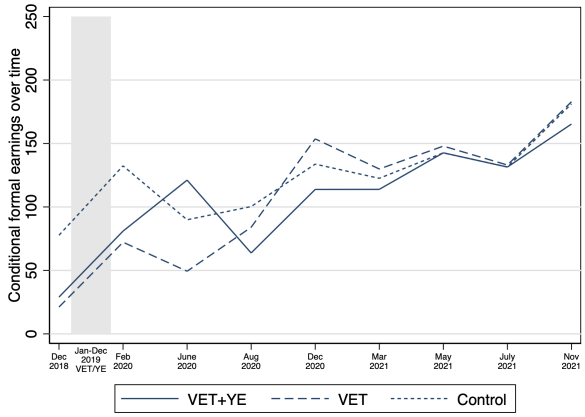
In contrast, Youth Empowerment shapes employment rates importantly, especially as far as informal participation is concerned. During the program, VET+YE participants have informal employment participation of less than 10%, in June 2020, which is lower than the Control group, almost 20% in June 2020. VET+YE however catches up to the control group and even overtakes the control group in terms of the informal employment rate from August 2020 throughout the end. Overall, this higher informal employment rate contributes to the slightly higher overall employment rate of VET+YE participants, compared to both VET and Control group, in November 2021.

Figure 6: **Conditional earnings over time**

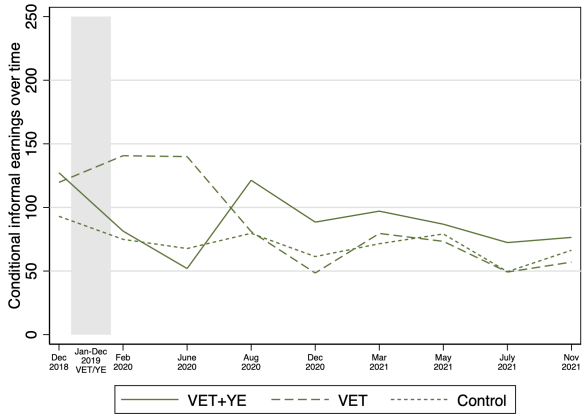
(a) Overall earnings



(b) Formal earnings



(c) Informal earnings



Notes: The figure illustrates average earnings, for the groups VET+YE (solid lines), VET (long-dashed lines), and the control group (short-dashed lines), from December 2018 to November 2021, conditional on whether participants are employed. The first panel shows overall earnings, the second shows formal sector earnings, and the third is informal sector earnings.

Figure 6 shows average earnings for the employed. Earnings for the employed increase from around 150 USD to 250 USD. This increase is strong, but less sizeable than the increase

in unconditional earnings from figure 4, which reflects changes in earnings for the employed as well as in the employment rate. Going back to figure 6, the increase in average earnings for the employed is driven by increases in formal earnings, which are around 50 USD at the beginning, and increase three-fold to 150 USD at the end of the sample period. Informal earnings, however, decrease over time from around 100 USD to around 60 USD. These movements in earnings for the employed reflect changes in hours worked, with individuals, on average moving out of informal employment and taking up formal employment. Another possible channel is increases in hourly wage rates, as individuals become more productive, their wage per hour work increases in the formal sector.

Again, there are no salient differences in earnings for the employed between the VET and Control groups. VET+YE individuals have lower average earnings from informal employment in the first part of the observation period, but higher informal earnings later on. This suggests that YE lowered informal employment both at the extensive margin (fewer individuals worked in that sector) and at the intensive margin (individuals worked fewer hours or for lower pay in the informal sector).

These findings suggest the formal sector is a better source of earnings in the long run. The strong persistence of informal sector earnings can be explained by the importance of informality in the context of the Colombian labor market.

6.0.2 Average Treatment Effects

Comparison of Group Averages: This section presents estimates of the average effect of VET, or VET+YE, on the treated, over time, by group. This implies, for each time period $t \in 1, ..., 9$ where 1 is December 2018 and 9 is November 2021, creating group averages of the outcomes of interest (e.g. earnings, formal sector earnings, informal sector earnings), and seeing if these averages are statistically significantly different from each other; that is to say, if the difference in averages is zero, there is no effect of treatment on that variable. This translates graphically as a flat line.

The graphs in figure 7 show the difference in means for earnings between VET+YE and the control group in solid lines, and VET and the control group in dashed lines, to explore the effect of each treatment on earnings compared to the control group. Simply put, we subtract the average earnings for the control group from the average earnings for group VET+YE, to find the difference between average earnings between the two groups. We do the same for the difference between VET and the control group. In words, if a line is close to zero, it means there is no difference between the treatment and the control group and the treatment does not have an effect on earnings

In each of the three panels, the levels of earnings over time for VET only are similar to the control group. Graphically, this translates as the VET-only line (the dashed line) mostly following the zero line (it is mostly flat). VET alone does not seem to have an effect on earnings.

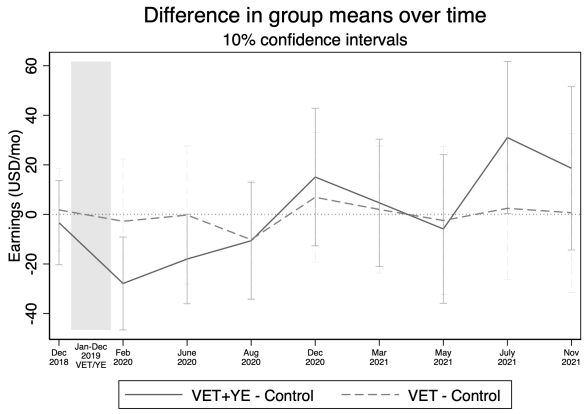
The levels of earnings over time for the VET+YE group (the solid line) vary compared to the control group, in each of the three panels. Participants in the VET+YE group started off earning less (there are significantly lower earnings immediately after the program), but then quickly earned more than participants in the control group. This implies that VET+YE seems to tilt the earnings path – the rate of growth in earnings is greater for VET+YE. There is an initial cost to the program but it seems to pay off over the long run.

The mostly insignificant estimates could be related to the low number of participants, but some of the values that are significant, combined with the trends, paint a consistent picture.

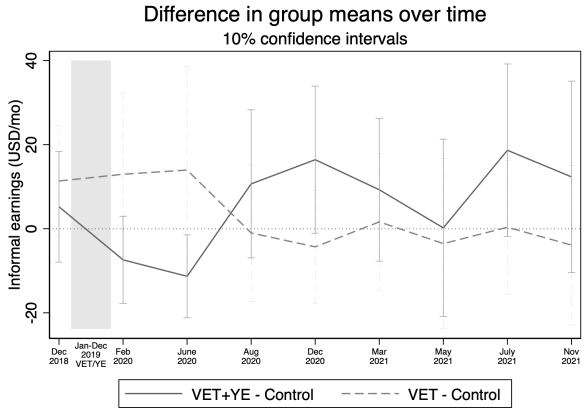
This is a path seen in several training programs - participants end the training program without a job, spend a couple of months looking for jobs, and then when they do find a job, it pays off more quickly.

Figure 7: **Difference in group means in earnings over time**

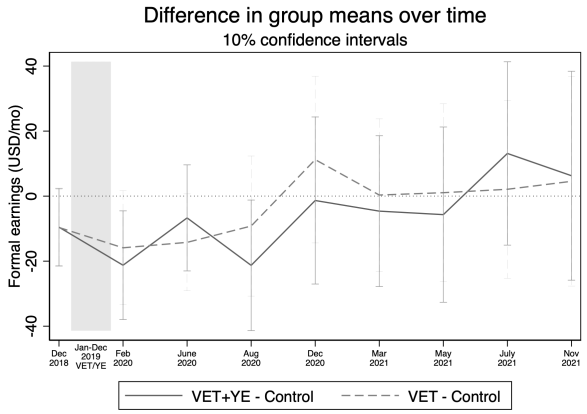
(a) Overall earnings



(b) Informal earnings



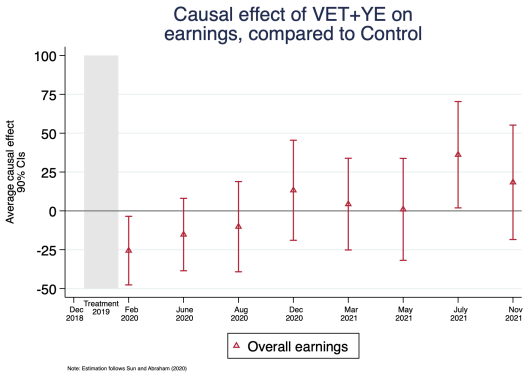
(c) Formal earnings



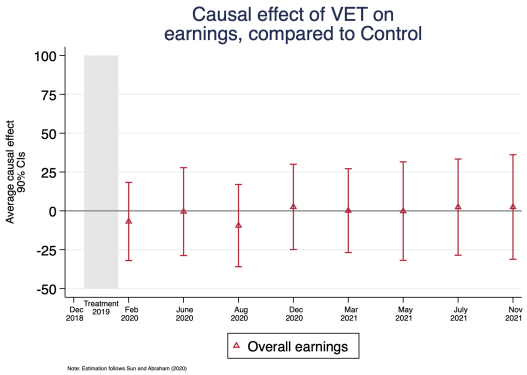
Notes: The graphs illustrate the difference in means in overall earnings (panel a), informal sector earnings (panel b), and formal sector earnings (panel c), between VET+YE and the control group in solid lines, and the difference in means in earnings between VET only and the control group in dashed lines. The vertical lines show the 90% confidence intervals

Figure 8: **Average effect of treatment on the treated on earnings over time**

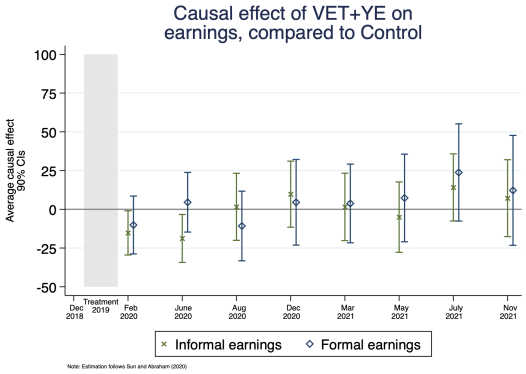
(a) Effect of VET+YE on overall earnings



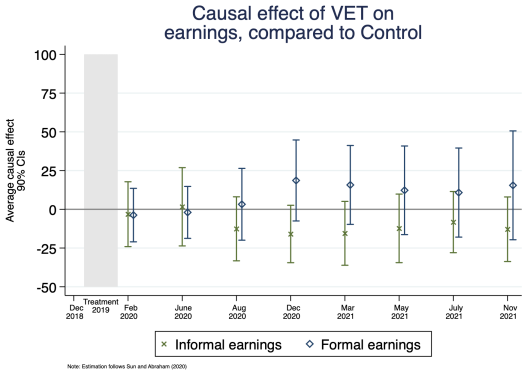
(b) Effect of VET only on overall earnings



(c) Effect of VET+YE on formal and informal earnings



(d) Effect of VET only on formal and informal earnings



Notes: The graphs illustrate the average causal effect of treatment on earnings over time. The figures on the right-hand side show the effect of VET only on earnings, and the figures on the left-hand side show the effect of VET+YE on earnings. The red triangles (panel a, panel b) show the effect of treatment on overall earnings, the green exes in panels c and d show the effect on formal sector earnings, and the blue diamonds in panels c and d show the effect on informal sector earnings, all in USD/month. The capped lines show the 90% confidence intervals.

The figures in 8 show the results of the two-way fixed effects regression estimated according to Sun and Abraham (2020). Results are similar compared to the ones found in figure 7, which indicates that the average in means is close to the actual causal average treatment effect.

Firstly, looking at the figure on the top, the patterns for earnings for VET+YE do seem to be towards a positive trend (panel 8a), especially after the COVID-19 pandemic. Immediately after treatment, VET+YE affects earnings significantly negatively, meaning that participants earn $25.58 per month less, compared to participants in the control group, in February 2020. This could be because participants who received VET+YE looked for jobs immediately after training and held out for better quality or better-paid jobs. However, in July 2021, participants in VET+YE earn $36.15 per month more, compared to participants in the control group. These

values can be seen in table 7. In terms of earnings from the formal labor market (blue diamonds, panel 8c), there is no significant difference between participants who received VET+YE and the control group, though the trend is positive. In terms of earnings from the informal labor market (green exes, panel 8c), we see some evidence that participants from VET+YE might have been holding out for better jobs right after training, as they earn $15.18/month in February of 2020 and $18.78/month less in informal earnings compared to the control group. However, the trend changes, and after a while, there is no significant effect of treatment on informal earnings.

On the other hand, VET only seems to have little effect on earnings over time (panel 8b). The lines in the graphs on the right-hand side are mostly flat and show that the effect of VET only on earnings does not differ significantly from zero. There also does not seem to be any indication that this trend might be changing. This remains true when we decompose earnings into earnings from jobs that are formal (blue diamonds, panel 8b) and earnings from informal jobs (green exes, panel 8b).

It should be noted that: 1) VET+YE, while not always cleanly estimated, seems to have a positive trend on earnings overall, and results suggest a similar pattern, though statistically insignificant, for formal and informal earnings (that is to say, participants in VET+YE appear to be making more money no matter the sector they are in); and 2) the estimates for the effect of treatment (both of VET+YE and VET only) on formal sector earnings are systematically bigger than the effect of treatment on informal sector earnings, which might imply that people who are in the formal sector are earning more money. The table 7 shows these in numbers.

6.0.3 Added value of youth empowerment training

Figure 9 shows the added value of VET+YE on earnings, compared to participants who did VET only. This is computed by comparing the earnings of participants in group VET+YE to the earnings of participants in group VET only. Figure 9 suggests that, compared to participants who only did vocational education, participants who did vocational education + youth empowerment earn less right after the training program, and earn more later, at a faster rate. This is mainly through informal earnings. This should be interpreted as a reading of patterns, given that the coefficients - when they are significant - are only significant at the 10% level. The summer of 2021, when the effects of COVID-19 were mostly dissipated, seems to be when earnings take off for participants in the VET+YE treatment. Going through youth empowerment training is associated with higher earnings of $33.74 in July 2021, compared to going through vocational education only, compared to the difference before treatment. Table 8 shows these values in numbers.

The drop in higher earnings in November of 2021 might be related to seasonal changes

Table 7: **Causal effect of treatment on earnings, compared to Control**

	All earnings		Formal earnings		Informal earnings	
	VET+YE	VET	VET+YE	VET	VET+YE	VET
Feb 2020	-25.58*	-6.836	-10.13	-3.737	-15.23*	-3.147
	(13.42)	(15.28)	(11.35)	(10.49)	(8.672)	(12.72)
June 2020	-15.25	-0.488	4.556	-2.001	-18.78**	1.634
	(14.17)	(17.20)	(11.70)	(10.18)	(9.403)	(15.36)
Aug 2020	-10.18	-9.479	-10.77	3.233	1.614	-12.59
	(17.66)	(16.10)	(13.66)	(14.09)	(13.16)	(12.56)
Dec 2020	13.26	2.547	4.514	18.61	9.768	-15.94
	(19.58)	(16.69)	(16.79)	(15.89)	(12.97)	(11.26)
Mar 2021	4.338	0.124	3.773	15.72	1.561	-15.51
	(17.95)	(16.42)	(15.43)	(15.48)	(13.22)	(12.53)
May 2021	0.956	-0.170	7.313	12.26	-5.034	-12.31
	(19.95)	(19.28)	(17.19)	(17.38)	(13.80)	(13.45)
July 2021	36.15*	2.412	23.79	10.80	14.12	-8.266
	(20.81)	(18.82)	(19.08)	(17.48)	(13.17)	(12.00)
Nov 2021	18.38	2.485	12.22	15.47	7.192	-12.87
	(22.38)	(20.45)	(21.55)	(21.35)	(15.10)	(12.67)
Observations	1750	1728	1749	1728	1749	1728

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

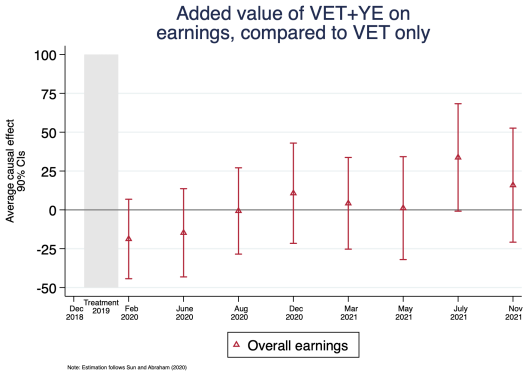
Notes: The dependent variable is overall earnings (col 2-3), formal earnings (col 4-5), and informal earnings (col 6-7). Columns 2, 4, and 6 show the difference between VET + YE (the group of participants who received vocational education and youth empowerment) and the control group, while VET (the group of participants who only received vocational education) compared to the control group (the omitted category is the control group). Robust standard errors clustered at the individual level are reported in round brackets. ***, **, and * indicate significance at the 1-, 5-, and 10-% levels, respectively.

in earnings and employment: as the Christmas season approaches, more people, in general, might be hired, so the difference in earnings between all participants might disappear seasonally. However, this might also be driven by the dissipation of the shock of COVID on the general labor market. It is possible that as the labor market opened in August of 2020, people with more skills were in greater demand, and as such found jobs much more quickly than participants who had fewer skills (e.g. VET only). The latter possibility is more concerning than the former as it would imply general equilibrium effects might be biasing our results; however, it might also be a promising result, as it could imply the youth empowerment program might be providing faster recovery from devastating shocks such as COVID-19.

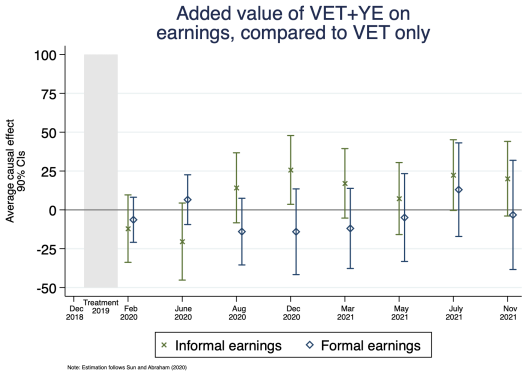
It would be possible to test for the latter possibility by running another survey in November of 2023. However, given that participants who were in the control group throughout the three years of the impact evaluation were given spots in the VET+YE program of 2023, the results of such a survey would not shed light on the matter.

Figure 9: **Added value of VET+YE on earnings over time**

(a) Added value of VET+YE on overall earnings



(b) Added value of VET+YE on informal and formal earnings



Notes: The graph illustrates the average causal effect of treatment on earnings over time, in particular, the added value of VET+YE compared to VET only. The first row shows the effect of treatment on overall earnings, and the second row the effect on formal and informal sector earnings, all in USD/month. The shaded area shows the 90% confidence intervals.

Table 8: **Added value of VET+YE on earnings over time**

	All earnings	Formal earnings	Informal earnings
Feb 2020	-18.75	-6.392	-12.08
	(15.55)	(8.827)	(13.20)
June 2020	-14.76	6.558	-20.41
	(17.26)	(9.743)	(15.08)
Aug 2020	-0.701	-14.00	14.21
	(16.89)	(13.04)	(13.70)
Dec 2020	10.71	-14.09	25.71*
	(19.62)	(16.77)	(13.47)
Mar 2021	4.214	-11.95	17.07
	(17.94)	(15.67)	(13.60)
May 2021	1.126	-4.949	7.277
	(20.13)	(17.20)	(14.09)
July 2021	33.74	12.99	22.39
	(21.03)	(18.32)	(13.84)
Nov 2021	15.90	-3.254	20.06
	(22.32)	(21.37)	(14.60)
Observations	1750	1749	1749

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The dependent variable is overall earnings (col 1), formal earnings (col 2), and informal earnings (col 3). VET + YE indicates the group of participants who received vocational education and youth empowerment, while VET indicates the group of participants that received only vocational education, the omitted category is the control group. Robust standard errors clustered at the individual level are reported in round brackets. ***, **, and * indicate significance at the 1-, 5-, and 10-% levels, respectively.

7 Heterogeneity and robustness of results

7.1 Heterogeneity

7.1.1 By gender

The figures below show the gendered effects of treatment on earnings. I do a split sample censor by gender and estimate the two-way fixed effects according to Sun and Abrahams (2020). Effects of treatment on women are shown in figure 10 below. First, women who undergo VET only do not have significantly different earnings compared to the control (see panel 10c). This is also the case for women who go through VET+YE (see panel 10a). Women who undergo VET+YE (panel 10b) seem to move away from informal earnings compared to the control group right after training, an effect that is not seen for women who only went through VET (panel 10d). One can see this pattern in informal employment rates for women who went through VET+YE right after training - they are 11% less likely to work in the informal sector compared to the control group in February 2020, and 20% less likely to work in the informal sector compared to the control group in August 2020. This difference disappears after COVID.

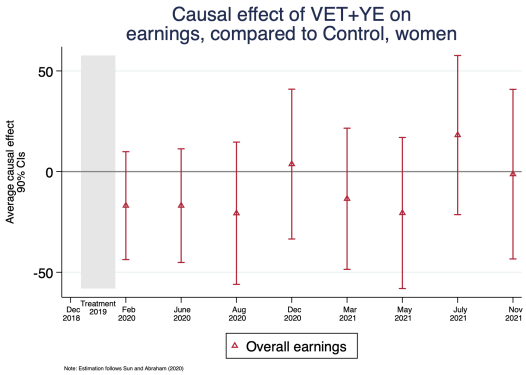
Figure 11 shows the effect of training on men. Unlike for women, the gains in earnings for men in VET+YE are more consistently significantly positive compared to men in the control group (11a). Like for women, men who undergo VET only do not have significantly different earnings compared to the control (see panel 11c). Men who undergo VET+YE (panel 11b) also initially move away from informal earnings compared to the control group right after training. Men in VET+YE tend to earn more from informal sources compared to men in VET only after COVID-10 (11f). It does seem the displacement effect for informal sector work is stronger for women than for men.

The most striking thing about the effect of training for men is the sheer magnitude of the difference, which is at least twice as high as the difference in earnings for women. For men, VET+YE is associated with higher earnings compared to men in the control group by December 2020 (which continues to be true until the end of the study) - see panel 11a. This starts at around \$50/month higher earnings and climbs up to nearly \$100/month in higher earnings. Men in VET+YE also earned \$100/month more than men in VET only 18 months after training. These higher earnings seem to come from informal earnings.

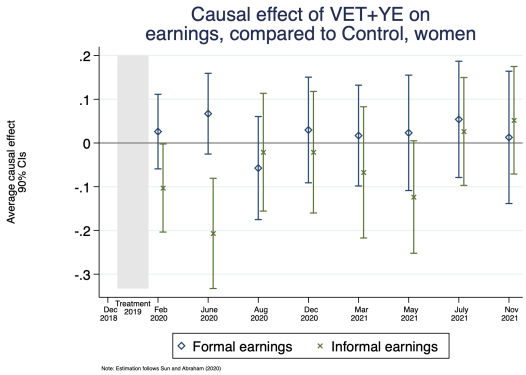
Whereas women who went through VET+YE seem to have used the training to move away from the informal sector, men who received VET+YE seem to have found the most benefits of the training in the informal sector.

Figure 10: **Average effect of treatment on the treated on earnings over time for women**

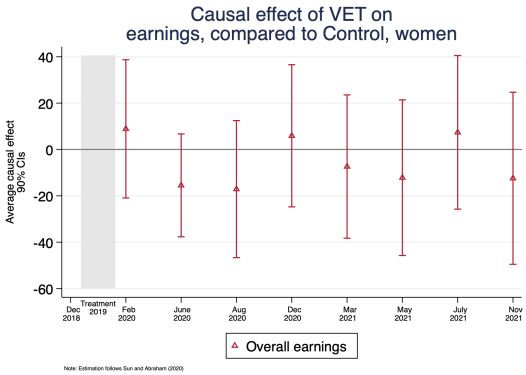
(a) Effect of VET+YE on earnings



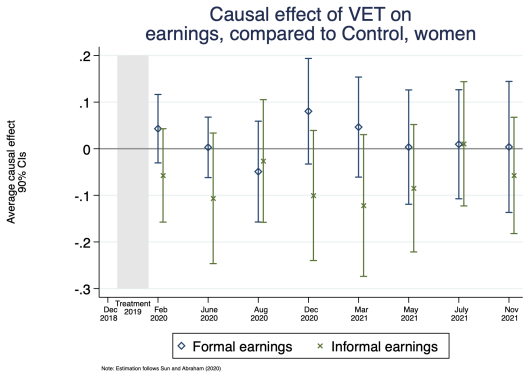
(b) Effect of VET+YE on earnings



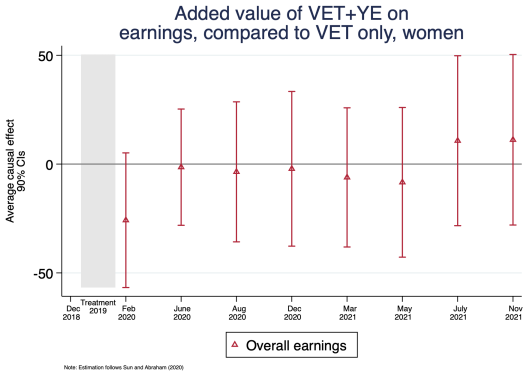
(c) Effect of VET only on earnings



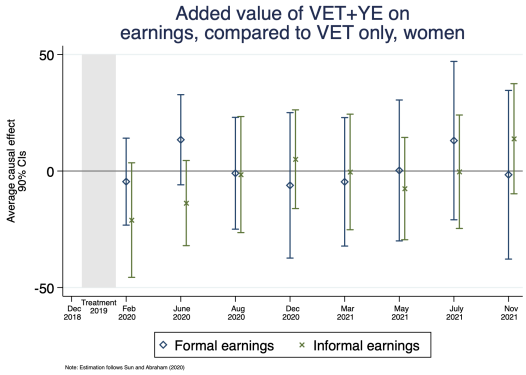
(d) Effect of VET only on earnings



(e) Added value of VET+YE on earnings



(f) Added value of VET+YE on earnings



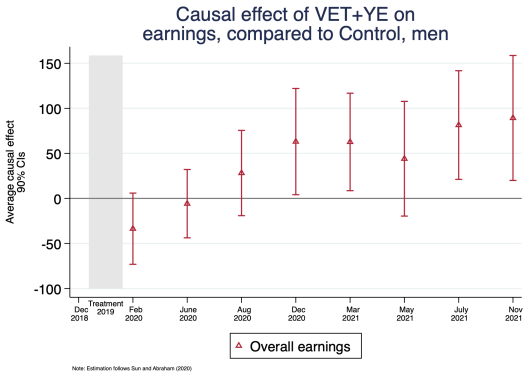
Notes: The graphs illustrate the average causal effect of treatment on earnings over time for women: overall earnings in triangles, informal earnings in "x"es, and formal earnings in diamonds. Panels a) and b) plot the effect of VET+YE on these three variables compared to the control group over time; panels c) and d) plot the effect of VET only on these three variables compared to the control group over time, and panel e) and f) plot the effect of VET+YE compared to VET only over time. The capped lines show the 90% confidence intervals.

7.1.2 By traumatic events: being displaced

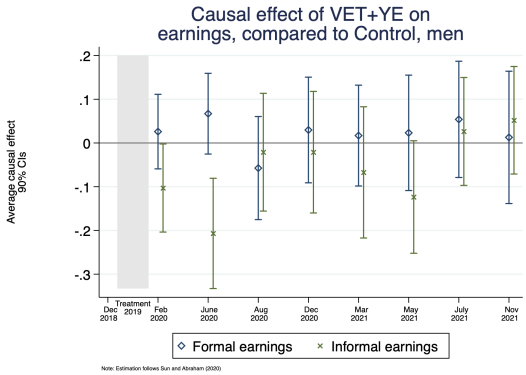
Figure 12 shows the effect of treatment on internally displaced people. Because the number of internally displaced people is low (around 30), these numbers should be taken with a grain of

Figure 11: **Average effect of treatment on the treated on earnings over time for men**

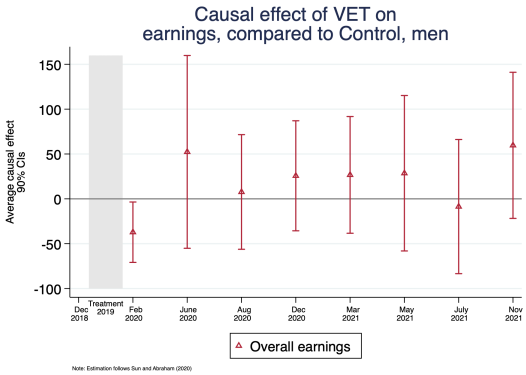
(a) Effect of VET+YE on earnings



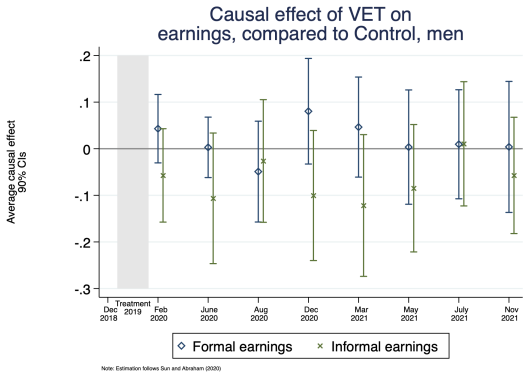
(b) Effect of VET+YE on earnings



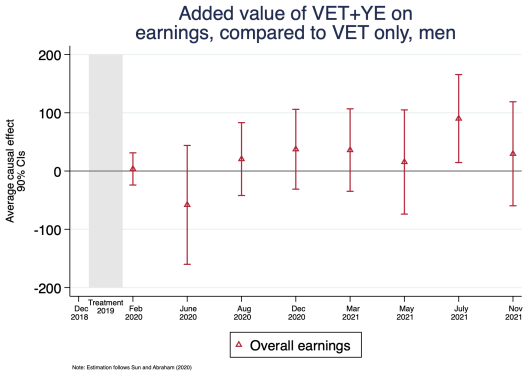
(c) Effect of VET only on earnings



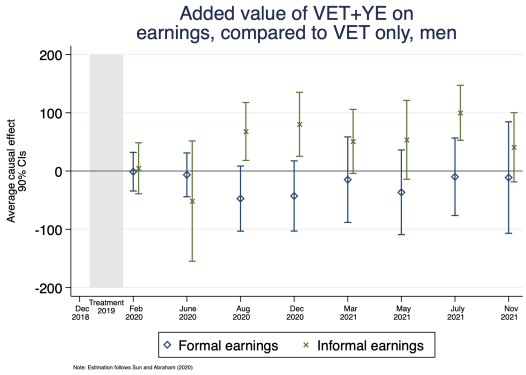
(d) Effect of VET only on earnings



(e) Added value of VET+YE on earnings



(f) Added value of VET+YE on earnings



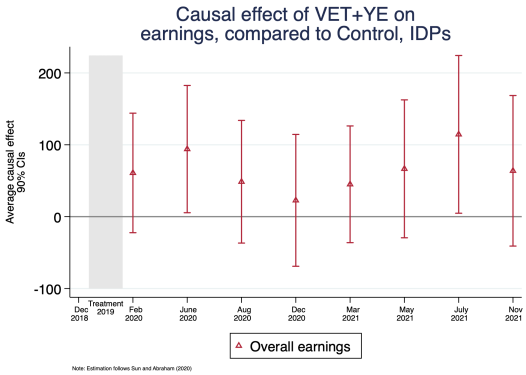
Notes: The graphs illustrate the average causal effect of treatment on earnings over time for men: overall earnings in triangles, informal earnings in "x"es, and formal earnings in diamonds. Panels a) and b) plot the effect of VET+YE on these three variables compared to the control group over time; panels c) and d) plot the effect of VET only on these three variables compared to the control group over time; panel e) and f) plot the effect of VET+YE compared to VET only over time. The capped lines show the 90% confidence intervals.

salt. Participants who went through training who are internally displaced earn more, compared to participants who are internally displaced and who did not go through training. These earnings seem to mostly come from formal sector earnings. The initial dip in earnings in the year

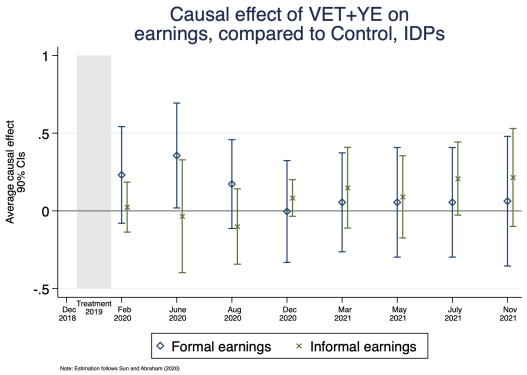
following treatment does not happen for internally displaced people who received VET+YE compared to internally displaced people in the Control group. Furthermore, the treatment effects on earnings are in general larger for IDPs (the difference in earnings in IDPs who went through treatment compared to the Control group are higher compared to the difference in earnings for non-IDPs in the different groups). Treatment effects for non-internally displaced people are shown in figure 13.

Figure 12: **Average effect of treatment on the treated on earnings over time for internally displaced people (IDPs)**

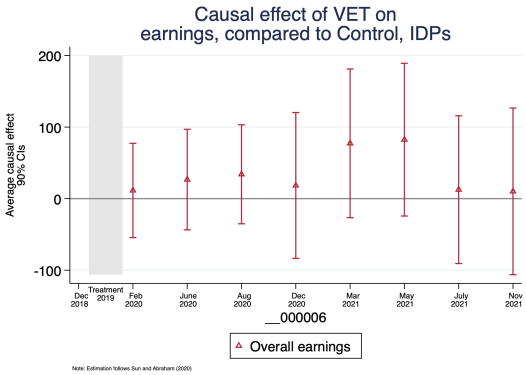
(a) Effect of VET+YE on earnings



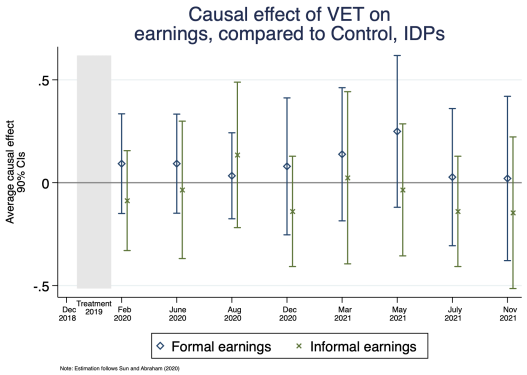
(b) Effect of VET+YE on earnings



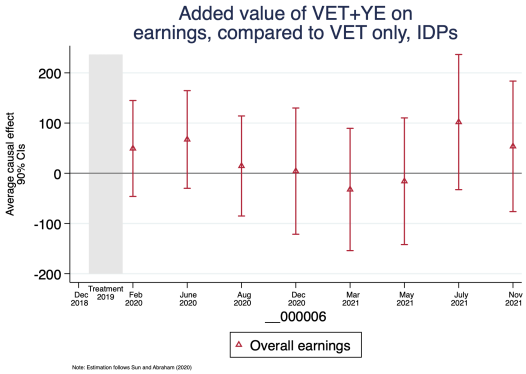
(c) Effect of VET only on earnings



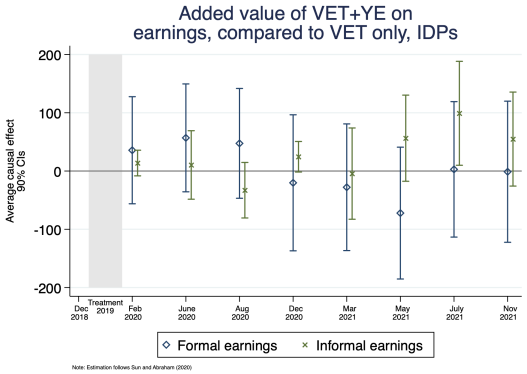
(d) Effect of VET only on earnings



(e) Added value of VET+YE on earnings



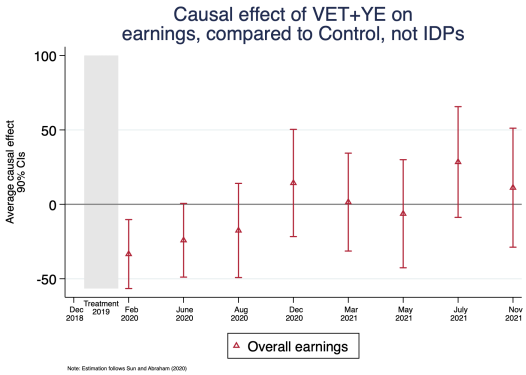
(f) Added value of VET+YE on earnings



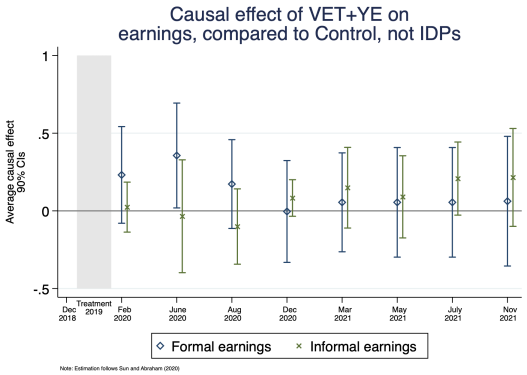
Notes: The graphs illustrate the average causal effect of treatment on earnings over time for IDPs: overall earnings in triangles, informal earnings in "x"es, and formal earnings in diamonds. Panels a) and b) plot the effect of VET+YE on these three variables compared to the control group over time; panels c) and d) plot the effect of VET only on these three variables compared to the control group over time; panel e) and f) plot the effect of VET+YE compared to VET only over time. The capped lines show the 90% confidence intervals.

Figure 13: **Average effect of treatment on the treated on earnings over time for non internally displaced people (IDPs)**

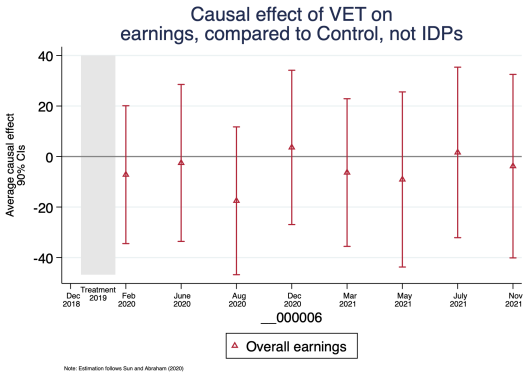
(a) Effect of VET+YE on earnings



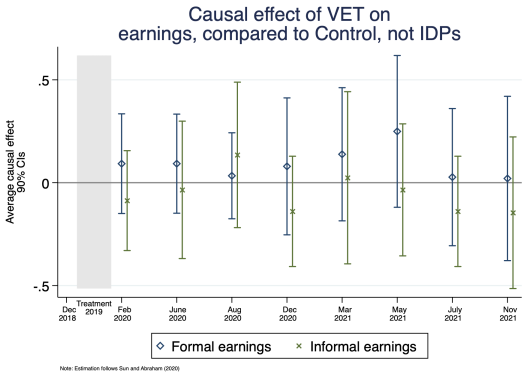
(b) Effect of VET+YE on earnings



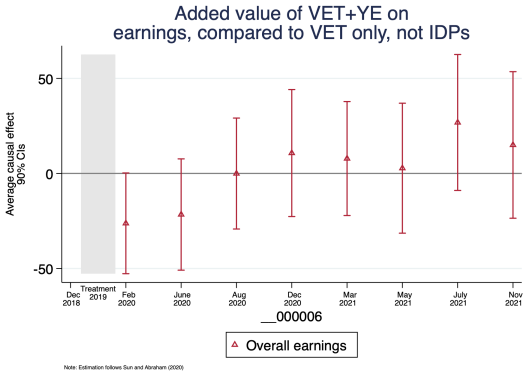
(c) Effect of VET only on earnings



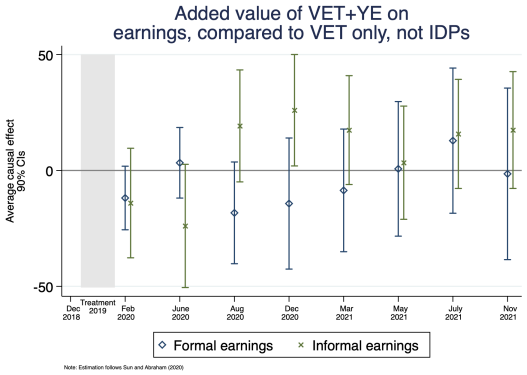
(d) Effect of VET only on earnings



(e) Added value of VET+YE on earnings



(f) Added value of VET+YE on earnings



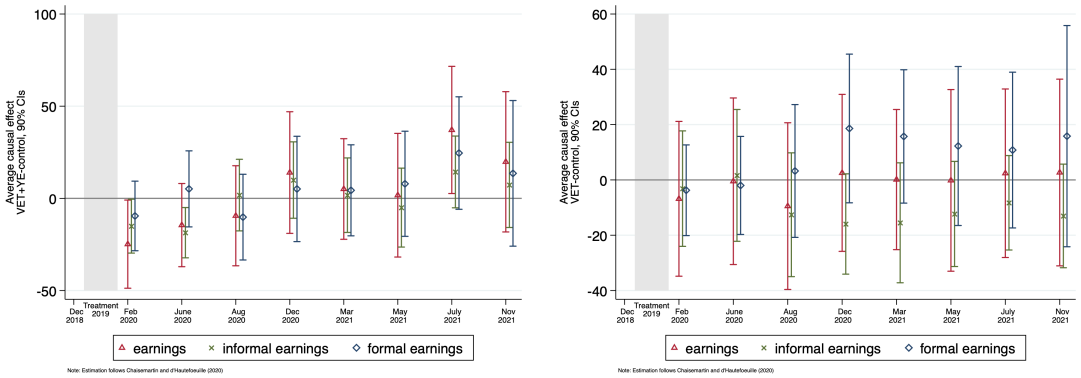
Notes: The graphs illustrate the average causal effect of treatment on earnings over time for not IDPs: overall earnings in triangles, informal earnings in "x"es, and formal earnings in diamonds. Panels a) and b) plot the effect of VET+YE on these three variables compared to the control group over time; panels c) and d) plot the effect of VET only on these three variables compared to the control group over time; panel e) and f) plot the effect of VET+YE compared to VET only over time. The capped lines show the 90% confidence intervals.

7.2 Robustness

Alternative specifications The event study graphs were created following Sun and Abrahams (2020), which hold only the assumptions of parallel trends and no anticipatory behavior (which are standard for differences-in-differences), but also for the more strict assumption of treatment effect homogeneity - or, more precisely, that "each cohort experiences the same path of treatment effects". For this reason, I check the results using the more flexible specification of De Chaisemartin and d'Haultfoeuille (2020). Figure 11 runs the decomposition proposed by De Chaisemartin and d'Haultfoeuille (2020) for labor market outcomes. The results do not change significantly.

Figure 14: **Average effect of treatment on the treated on mental health over time**

(a) Effect of VET+YE on earnings (b) Effect of VET only on earnings



Notes: The graphs illustrate the average causal effect of treatment on earnings over time: overall earnings in triangles, formal earnings index in triangles, and informal earnings in "x"es. Panel a) plots the effect of VET+YE on these three variables compared to the control group over time; panel b) plots the effect of VET only on these three variables compared to the control group over time. The capped lines show the 90% confidence intervals.

8 Cost-benefit analysis

8.1 Summary

The above analysis shows that participants in the VET+YE program seem to earn more than participants in the Control group, but not consistently in a statistically significant manner. Looking at patterns in earnings, participants in VET+YE may be expected to earn up to between $2,158.59 and $2,683.73 more than participants in the Control group over the next ten years, ignoring statistical significance. The total cost of the VET+YE program per participant for 2019 was $1,395.67. Over the ten years following treatment, there could therefore be an estimated net benefit between $762.92 and $1,288.06 per participant for following the VET+YE program.

The corresponding rates of return would be 55% and 92%, respectively. Table 9 provides a summary of these figures and adds to it the benefits of following the VET-only program compared to the Control group.

Table 9: **Gains compared to Control group for overall earnings (in USD)**

	over 24 months	over 10 years	rate of return
VET+YE gains			
gains that don't depreciate	$66.22	$2,683.73	92%
gains that depreciate (5% depreciation/yr)	$66.22	$2,158.59	55%
VET only gains			
gains that don't depreciate	-$28.05	$202.35	NA
gains that depreciate (5% depreciation/yr)	-$28.05	$156.13	NA

Notes: The table shows the expected gains of following the VET+YE program and of following the VET-only program, compared to the Control Group. The costs of the VET+YE program are estimated at $1,395.67 per participant.

8.2 In-depth cost-benefit analysis

The simplest way to explore the cost/benefit tradeoff of the VET+YE program would be to look at the gains in earnings of VET+YE compared to the Control group and VET compared to the Control group in the medium run. The cost-benefit analysis first describes the costs and benefits of the program in terms of overall earnings, regardless of source (whether it is formal or informal), and then by breaking down by source of earnings. It should be noted that the below analysis does not take statistical significance into account.

8.2.1 Costs and benefits of treatment on overall earnings

Figure 8 showed that there were roughly three phases of the effect VET+YE on earnings throughout the two years, compared to Control participants: a learning phase between February 2020 and August 2020 where participants lost on average $17 compared to Control participants[8]; a stabilizing phase between December 2020 and May 2021 where participants earned on average $6.18 more than Control participants; and a reaping phase as of July 2021 where participants earned on average $27.26 more than Control participants. Over the first roughly 24 months, then, VET+YE participants earned $66.22 more than participants in the Control group compared to the difference before treatment. Assuming the "reaping" figure of $27.26 in gains per month is the effect of VET+YE on earnings for the next ten years, one could expect that participants in VET+YE would earn $2,683.73 more than participants in the Control group over the

[8] It is possible to think of this as the hidden cost, or opportunity cost of joining the program.

following ten years. Assuming, as in Attanasio et al (2011), that there is an annual discount rate of 5%, this figure could drop to $2,158.59 in higher earnings over the next ten years for participants in VET+YE compared to the Control group.

The total costs of the VET+YE program for 2019 were COP 457,781,841, which, in 2019[9] was equivalent to $139,567.63. Per participant (there were 100 of them), this program, therefore, costs $1,395.67.

In the short run, the costs of the program are not recouped. However, over the ten years following treatment, there may be a net benefit of $762.92 for VET+YE participants under the conservative scenario of depreciation of gains - this is the lower bound for the net return for the program - and a net benefit of $1,288.06 under the less conservative scenario where gains remain constant for the next 10 years. The corresponding rates of return would be 55% and 92%, respectively [10].

Extrapolating for the same period, participants in VET earned on average $2.44 (both in July and November 2021) more than the Control group 18 months after treatment relative to the difference before treatment. Using the same logic as before, the learning phase is worth $\$ - 5.53$, the stabilizing phase is worth $0.87 and the reaping is $2.4. This translates to a loss of $28 compared to the Control group over the first two years; and a gain of $202.35 over the ten years after treatment without depreciation of gains. With a depreciation of gains, this would be $156.13. Table 9 provides a summary of these figures.

8.2.2 Cost-benefit analysis in terms of formal and informal sector earnings

It is possible to run the same analysis as above, but differentiate between earnings from the formal and informal sectors. This analysis breaks down the benefits in terms of earnings from the formal sector and earnings from the informal sector and aims to see the source from which overall benefits come. Again, this analysis does not take statistical significance into account.

Costs and benefits of treatment on formal earnings: First, looking at the benefits of the VET+YE program in terms of formal sector earnings, one can see the three phases we saw before: the learning phase gives an average loss of $7.91 per month in formal earnings, the stabilizing phase saw a gain in $5.66 per month and the reaping phase saw a gain in $19.21 per month compared to the Control group. In the short run, over the two years after training ended, this is equivalent to a gain in formal earnings of $104.47 compared to the Control group. Over the next ten years, this is equivalent to a gain in formal earnings of $1,948.84 without depreciation of gains. With program costs of $1,395.67, this is equivalent to net benefits of $553.17 and

[9] Using the average 2019 exchange rate of COP3,280.67 per USD.

[10] That is, $\frac{netreturn=1,288.06}{cost=1,395.67} * 100 < x < \frac{netreturn=762.92}{cost=1,395.67} * 100$

a net return of 40% in formal sector earnings. With a depreciation of gains of 5%, the expected gain in formal earnings would be $1,578.82 ten years from now, giving net benefits of $183.15 and a net return of 13% in formal sector earnings.

For participants in VET the learning phase is worth -$4.19, the stabilizing phase is worth $13.12 and the reaping is $12.25. This translates to a gain of $153.84 compared to the Control over the first two years; and a gain of $1,329.58 over the ten years after treatment without depreciation of gains.

Costs and benefits of treatment on informal earnings: First, looking at the benefits of the VET+YE program in terms of informal sector earnings, one can see the three phases we saw before: the learning phase gives an average loss of $7.91 per month in formal earnings, the stabilizing phase saw a gain in $3.37 per month and the reaping phase saw a gain in $10.27 per month compared to the Control group. In the short run, over the two years after training ended, this is equivalent to a gain in informal earnings of $20.70 compared to the Control group. Over the next ten years, this is equivalent to a gain in informal earnings of $1,006.453 without depreciation of gains. With program costs of $1,395.67, this is equivalent to a loss of $389.22 in informal earnings.

For participants in VET the learning phase is worth -$2.83, the stabilizing phase is worth -$13.54 and the reaping is -$13.24. This translates to a loss of $226.811 compared to the Control over the first two years; and a loss of $1,497.841 over the ten years after treatment without depreciation of gains.

9 Conclusion

This study sought to understand the impact of adding youth empowerment to vocational education on the labor market outcomes for marginalized young adults in two lower-income suburbs of Bogotá, Colombia in the medium run. This is a long-running program by the Colombian NGO Fundación Apoyar, supported by Vivamos Mejor in Switzerland. I compare these results to those of marginalized young adults who received access to traditional vocational education, and to young adults who received no support.

I use the results of a randomized controlled trial where 300 vulnerable young adults between the ages of 18 and 25 from Bosa (a neighborhood of Bogotà) and Soacha (a city in the neighboring state of Cundinamarca) were randomly allocated into equally-sized treatment groups: one receiving the vocational education and youth empowerment program offered by Fundación Apoyar (group VET+YE), one receiving vocational education only (group VET), and one control group with no additional support (Control group). The program ran throughout the academic

year of 2019. The team collected data in December 2018 before the randomization, and again in August 2020, November 2020, and November 2021.

The study was particularly interested in the effects of training on earnings and labor market participation, whether these were in the formal or informal sector.

The main results are first, that the youth empowerment program appears to have affected participants' earnings. Three years after the start of the program, participants who participated in youth empowerment appear to be earning more than their peers who received only vocational education and those who received no further support.

Second, in terms of the source of earnings, youth empowerment initially reduced earnings from the informal labor market. However, there was an increase in informal labor market earnings eventually. Formal earnings increase for both vocational education treatment groups, though not consistently statistically significantly different from the control group. This indicates that there may be an initial lock-in effect of training in terms of earnings from the formal sector, but once participants have had enough time to find a job in the formal labor market, they tend to find one relatively quickly.

Moreover, the cost-benefit analysis indicates that over the ten years following treatment, there could be an estimated net benefit between $762.92 and $1,288.06 per participant in the empowerment program, with corresponding rates of return between 55% and 92%. This does not take into account statistical significance. This is a very high rate of return, especially in light of the COVID-19 crisis, which led to a sizeable increase in unemployment and mental health deterioration.

Why should we care? First, vocational education has not always been a reliable source of labor market integration for vulnerable young adults. Adding a youth empowerment component to such training programs seems to help young adults earn more in the medium term. Taken together, the positive results from the event study and the cost-benefit analysis indicate the importance of designing programs that properly target vulnerable youth in different contexts. Here, the longstanding program of Fundación Apoyar is well-designed to help urban and suburban conflict-affected young adults who are seeking a better life for themselves.

This research can be complemented in several ways. For more evidence on the added value of youth empowerment training, researchers could do a similar impact evaluation with a youth empowerment-only treatment branch. Furthermore, disentangling the effects of the different youth empowerment branches could provide further evidence on what, exactly, is most effective for young adults who are seeking vocational education. Vandewalle et al (TBD) are examining such a program that disentangles the effects of different soft skills training programs in Vietnam that could provide further evidence in this direction.

Concretely, in the Colombian context, providing a program such as Fundación Apoyar's could improve the usual vocational education system. There has been a concerted effort to make vocational education pay off, and this program is an example of a successful attempt to do so. In this situation, adding youth empowerment, coupled with a nation-building curriculum and job search support, ensures the success of these particular vocational degrees. The youth empowerment program is particularly tailored to help vulnerable young adults in a setting that combines a strong informal labor market, urban poverty and crime, and a nation that is actively working on peace-building. The anecdotal legacy of success, as well as the results of this policy brief, indicate that while there may be some costs to such a program, the benefits are worth it.

A.1 Supplementary material

This Appendix accompanies the paper Improving vocational education through youth empowerment: evidence from a randomized controlled trial in Colombia.

A.1.1 Curricula and household survey

The pages that follow provide an introduction, in Spanish, for each of the five vocational education programs offered. These were provided by Fundación Apoyar. In addition, the pages that follow provide an example of the household surveys conducted by Fundación Apoyar to categorize the living conditions and meet the families of each applicant.

Recursos Humanos en convenio con la Institución LA CONCORDIA

EL EGRESADO DE LA INSTITUCIÓN LA CONCORDIA EN CONVENIO CON CEDEP, ESTÁ EN LA CAPACIDAD DE : APOYAR ADMINISTRATIVAMENTE EMPRESAS Y ORGANIZACIONES DE CUALQUIER SECTOR; EN PROCESOS DE SELECCIÓN DE PERSONAL, RECLUTAMIENTO, BIENESTAR LABORAL, CONTRATACIÓN, NÓMINA, LIQUIDACIÓN DE CONTRATOS, SEGURIDAD SOCIAL , DESARROLLO ORGANIZACIONAL, MANEJO Y TRAMITES DE DOCUMENTOS; CON CAPACIDAD DE ANÁLISIS Y TRABAJO EN EQUIPO, DESEMPEÑAN UN PAPEL IMPORTANTE EN LAS EMPRESAS, SON LOS LIDERES ENCARGADOS DE LA COMUNICACIÓN ENTRE LOS TRABAJADORES Y LOS GERENTES DE LA EMPRESA.


CEDEP
Centro de Educación para el Trabajo y el Desarrollo Humano
"EL MUNDO AVANZA Y NOSOTROS CON ÉL"
Av. Caracas No. 45-00
Teléfonos: 3205900-3205884
www.cedep.edu.co CEDEP3@
MERCADEO Y VENTAS
Resolución No. 02-0075 04/12/2009 Y 02030 DE 11/09/2014.
Denominación del programa: TÉCNICO LABORAL POR COMPETENCIAS
Certificación: AUXILIAR DE MERCADEO
Duración: 1150 horas (1 año) (12 meses clase)
Jornadas de Lunes a Viernes
Mañana: 8 a.m. a 11 a.m.
Tarde: 2 p.m. a 5 p.m.
Noche: 6 p.m. a 9 p.m.
Sábados: 8 a.m. a 4 p.m.
PENSUM ACEDEMICO
CONOCIMIENTOS BÁSICOS
GESTIÓN ADMINISTRATIVA
GESTIÓN CONTABLE
DERECHO COMERCIAL
ÁREA MERCADERÍA
MERCHANDISING
INVESTIGACIÓN DE MERCADOS
TÉCNICAS MARKETING
ESTRATEGIA DE VENTAS
SERVICIO AL CLIENTE
ESTRATEGIAS DE NEGOCIACIÓN
FUNDAMENTOS PUBLICIDAD
ÁREA CONSTITUCIONAL
ÉTICA LABORAL
OFIMÁTICA
TÉCNICAS DE COMUNICACIÓN
INGLES TÉCNICO
SEMINARIO DE EMPRENDIMIENTO
ÁREA APLICATIVA
PRACTICAS LABORES
TRABAJO DE GRADO
REQUISITOS:
Copia Doc. identidad al 150%
Copia diploma Acta de grado o Certificado 9° Grado
2 fotos 3x4 fondo azul
MATRICULAS PERMANENTES...!

CEDEP
Centro de Educación para el Trabajo y el Desarrollo Humano

"EL MUNDO AVANZA Y NOSOTROS CON ÉL"

Av. Caracas No. 45-00
Teléfonos: 3205900-3205884
www.cedep.edu.co CEDEP3@

ATENCION A LA PRIMERA INFANCIA

Resolución No. 02-0075 04/12/2009 Y 02030 DE 11/09/2014.

Denominación del programa: TÉCNICO LABORAL
Certificación: EN ATENCION A LA PRIMERA INFANCIA
Duración: 1172 horas (1 año) (12 meses clase)

Jornadas de Lunes a Viernes

Mañana: 8 a.m. a 11 a.m.
Tarde: 2 p.m. a 6 p.m.
Nocturna: 6 p.m. a 9 p.m.
Sábados: 8 a.m. a 4 p.m.

PENSUM ACADEMICO

SEMESTRE 1

- FUNDAMENTOS PEDAGOGICOS
- MEDICINA E INFANCIA
- POLITICAS EDUCTIVAS
- COMPETENCIAS COMUNICATIVAS
- DIDACTICA DEL INGLES
- EXPLORACION ACADEMICA I
- GENEALOGIA DE LA ESCUELA –PRACTICA 1

SEMESTRE 2

- PROYECTO EDUCATIVO
- PSICOLOGIA DEL DESRROLLO INFANIL
- RECREACION Y DEPORTE
- PEDAGOGIA DE LA LECTO ESCRITURA
- DIDACTICA D ELA INFORMATICA
- EXPLORACION ACADEMICA II
- CULTURA CURRICULAR –PRACTICA 2

SEMESTRE 3

- TENDNECIAS Y MODELOS PEDAGOGICOS
- PSICOLOGIA INFANTIL
- EXPRESION CORPORAL
- PENSAMIENTO MATEMATICO
- RECURSOS DIDACTICOS
- EXPLORACION ACADEMICA 3
- CULTURA CURRICULAR 3

SEMESTRE 4

- SOCIOLOGIA DE LA EDUCACION
- PSICOLOGIA DEL APRENDIZAJE
- APRESTAMIENTO
- ETICA Y VALORES
- PUERICULTURA
- GESTION EMPRESARIAL
- EXPLORACION ACADEMICA 4
- PRACTICA DOCENTE 4

REQUISITOS:

- Copia Doc. identidad al 150%
- Copia Acta de grado o Diploma, Certificado 9° Grado
- 2 fotos 3x4 fondo azul
- Fotocopia de icfes
- Certificacion afiliación a EpS

CEDEP
Centro de Educación para el Trabajo y el Desarrollo Humano

"EL MUNDO AVANZA Y NOSOTROS CON ÉL"

Av. Caracas No. 45-00
Teléfonos: 3205900-3205884
www.cedep.edu.co CEDEP3@

ATENCION A LA PRIMERA INFANCIA EN CONVENIO CON "ESCUELA IBEROAMERICANA EICT"

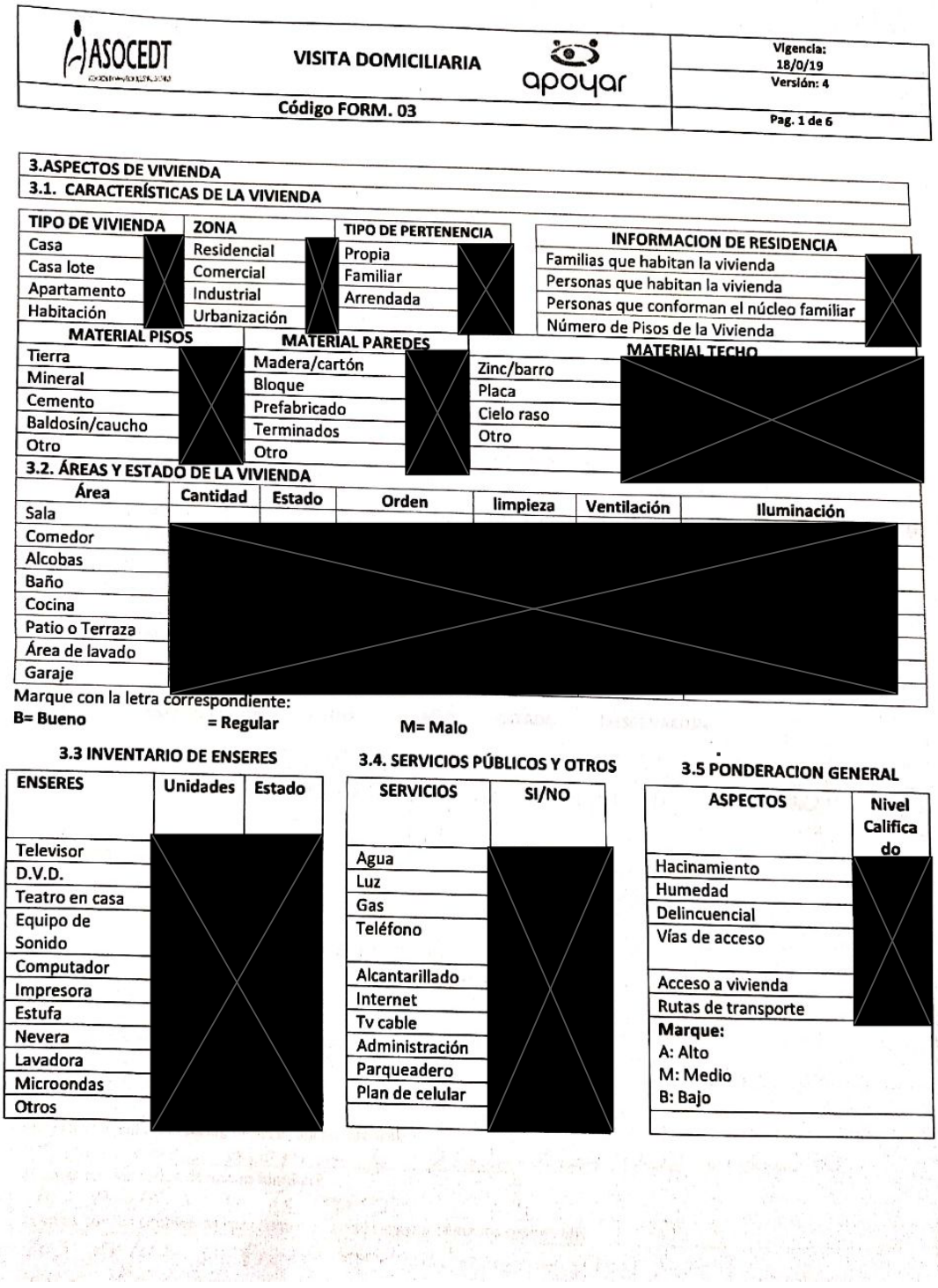
EL TÉCNICO EN ATENCIÓN A LA PRIMERA INFANCIA, ESTÁ EN LA CAPACIDAD DE CONOCER E IDENTIFICAR EL DESARROLLO INTEGRAL DE LOS NIÑOS Y NIÑAS, FORTALECER EL CONJUNTO DE HABILIDADES Y ACTITUDES ÉTICAS, LÚDICAS, PEDAGÓGICAS CON EL FIN DE DASARROLLAR UN PROYECTO EDUCATIVO INSTITUCIONAL DE ACUERDO CON LOS ESTANDARES DE CALIDAD ESTABLECIDOS POR EL MINISTERIO DE EDUCACION NACIONAL QUE ATIENDE A LAS NECESIDADES DE LA COMUNIDAD Y ASI CONTRIBUIR A LA SOCIEDAD.

Auxiliar Administrativo

EL EGRESADO ADMINISTRATIVO, SE CARACTERIZA POR SER ALTAMENTE COMPETENTE EN LA GESTION DE EMPRESAS Y CONOCEDOR DE LAS TECNICAS ADMINSITRATIVAS EN CADA UNA DE LAS AREAS FUNCIONALES DE OFICINA, LO CUAL LE PERMITIARÁ PARTICIPAR PROACTIVAMENTE EN LA PROYECCION, CRECIMIENTO, Y DESARROLLO DE LAS ORGANIZACIONES DE LA CUAL FORME PARTE; ASI MISMO SON POSEEDORES DE ACTITUDES EMPRENDEDORAS QUE LES PERMITEN PROYECTAR Y EJECUTAR POSIBLES ESCENERIOS DE CREACION DE EMPRESAS CON BASE EN LOS CONOCIMIENTOS ADQUIRIDOS EN EL TRANSCUROS DE SU FORMACION EN LIDERAZGO Y RESPONSBILIDAD SOCIAL

PÉNSUM DEL PROGRAMA TÉCNICO LABORAL AUXILIAR EN ALMACENAMIENTO Y APROVISIONAMIENTO

SEMESTRE	MÓDULO No.	CURSO	NOMBRE CURSO	HORAS PRESENCIALES	HORAS AUTÓNOMAS
PRIMER SEMESTRE ACADÉMICO	1	1	Procesos Administrativos	24	24
		2	Legislación comercial y laboral	24	24
	2	1	Informática (Windows, Word, Excel)	24	24
		2	Manejo de Tecnologías de la Información (Tic's)	24	24
	3	1	Ética para el Trabajo	24	24
		2	Desarrollo Humano Integral	24	24
SEGUNDO SEMESTRE ACADÉMICO	4	1	Técnicas de Almacenamiento de Mercancías	24	24
		2	Recibo y Despacho de Mercancías	24	24
	5	1	Fundamentos de Seguridad y Salud en el Trabajo	24	24
		2	Prácticas de Trabajo Seguras y Saludables	24	24
	6	1	Inglés Nivel A1.1	24	24
		2	Inglés Nivel A1.2	24	24
TERCER SEMESTRE ACADÉMICO	7	1	Técnicas de Control de Inventarios	24	24
		2	Técnicas de Valoración de Inventarios	24	24
	8	1	Procesos y Procedimientos para la Gestión Humana	24	24
		2	Técnicas para la Elaboración de Instructivos y Formatos Organizacionales	24	24
	9	1	Mentalidad Emprendedora	24	24
		2	Creación de Empresa	24	24
			Pasantía Empresarial	192	



ASOCEDT

VISITA DOMICILIARIA

apoyar

Vigencia: 18/0/19
Versión: 4
Pag. 1 de 6

Código FORM. 03

3.ASPECTOS DE VIVIENDA

3.1. CARACTERÍSTICAS DE LA VIVIENDA

TIPO DE VIVIENDA	ZONA	TIPO DE PERTENENCIA
Casa	Residencial	Propia
Casa lote	Comercial	Familiar
Apartamento	Industrial	Arrendada
Habitación	Urbanización	

INFORMACION DE RESIDENCIA
Familias que habitan la vivienda
Personas que habitan la vivienda
Personas que conforman el núcleo familiar
Número de Pisos de la Vivienda

MATERIAL PISOS	MATERIAL PAREDES	MATERIAL TECHO
Tierra	Madera/cartón	Zinc/barro
Mineral	Bloque	Placa
Cemento	Prefabricado	Cielo raso
Baldosín/caucho	Terminados	Otro
Otro	Otro	

3.2. ÁREAS Y ESTADO DE LA VIVIENDA

Área	Cantidad	Estado	Orden	limpieza	Ventilación	Iluminación
Sala						
Comedor						
Alcobas						
Baño						
Cocina						
Patio o Terraza						
Área de lavado						
Garaje						

Marque con la letra correspondiente:

B= Bueno **= Regular** **M= Malo**

3.3 INVENTARIO DE ENSERES

ENSERES	Unidades	Estado
Televisor		
D.V.D.		
Teatro en casa		
Equipo de Sonido		
Computador		
Impresora		
Estufa		
Nevera		
Lavadora		
Microondas		
Otros		

3.4. SERVICIOS PÚBLICOS Y OTROS

SERVICIOS	SI/NO
Agua	
Luz	
Gas	
Teléfono	
Alcantarillado	
Internet	
Tv cable	
Administración	
Parqueadero	
Plan de celular	

3.5 PONDERACION GENERAL

ASPECTOS	Nivel Calificado
Hacinamiento	
Humedad	
Delincuencial	
Vías de acceso	
Acceso a vivienda	
Rutas de transporte	

Marque:
A: Alto
M: Medio
B: Bajo

A.1.2 Data construction

The variables for depression and distress were created by participants filling in questionnaires, which were adapted from Blattman and Ralston (2015).

Table A1: **Questions to create depression and distress indices**

	Variable	Range	Spanish question	English question
B63	Depression	0-3	Me siento mal haciendo cosas que normalmente hace	Feel bad doing things you normally do
B64	Depression	0-3	La paso mal escuchando gente cuando estoy pensando en malas cosas	Hard time hearing people when thinking about bad things
B65	Depression	0-3	Me siento triste o "bajoneado"	Felt sad or down hearted
B66	Depression	0-3	Me siento cansado aún sin hacer nada	Feel tired when even not doing anything
B67	Depression	0-3	Pierdo el apetito (ganas de comer) cuando me siento mal	Lose appetite from feeling bad
B68	Depression	0-3	Siento que no soy importante para nadie	Felt unimportant to everyone
B69	Distress	0-3	Sento y pienso malas cosas	Sit and think bad things
B70	Distress	0-3	Tengo malos sueños / pesadillas	Have bad dreams
B71	Distress	0-3	Siento que las cosas malas se repiten	Seems like bad things happening again
B72	Distress	0-3	Transpiro o siento calor cuando pienso en cosas malas	Sweat, feel warm, etc when thinking about bad thigns
B73	Distress	0-3	Prefiero evitar ciertos lugares	Stay away from places
B74	Distress	0-3	Hablo con mi mismo(a)	Talk to yourself
B75	Distress	0-3	Siento que mi corazón está acongojado	Feel that your heart is spoiled
B76	Distress	0-3	Me siento frustrado(a)	Feel frustrated
B77	Distress	0-3	Siento que el espíritu de alguien me está perturbando / obstaculizando	Think someone's spirit is hampering you
B78	Distress	0-3	Mi cuerpo está "seco" (cansado) por las preocupaciones	Body is dry from worrying
B79	Distress	0-3	Mi corazón no está satisfecho	Heart does not feel satisfied
B80	Distress	0-3	Siento que el corazón arde	Feel heart burning

The trauma severity index was created from the PCL-C which is a self-administered test for post-traumatic stress disorder (PTSD):

Table A2: **Questions in PCL-C, in Spanish**

PCL-C question	Question	Min	Max	Item class
1	Recuerdos, pensamientos o imágenes reiterados y perturbadores	1	5	B
2	Sueños repetidos e inquietantes de una experiencia estresante del pasado?	1	5	B
3	Actuando repentinamente o sintiendo como si una experiencia estresante volviera a suceder (como si la estuvieras reviviendo)?	1	5	B
4	Te sentiste muy alterado(a) cuando algo te recordó una experiencia estresante del pasado?	1	5	B
5	Tienes reacciones físicas (por ejemplo, latidos cardíacos, dificultad para respirar, o transpiración) cuando algo te recuerda una experiencia estresante del pasado?	1	5	B
6	Evitas pensar o hablar sobre una experiencia estresante del pasado o evitas tener sentimientos relacionados con ella?	1	5	C
7	Evitas actividades o situaciones porque te recuerdan una experiencia estresante del pasado?	1	5	C
8	Tienes problemas para recordar partes importantes de una experiencia estresante del pasado?	1	5	C
9	Sientes una pérdida de interés en las cosas que solías disfrutar?	1	5	C
10	Sientes distancia o rechazo de otras personas?	1	5	C
11	Te sientes emocionalmente adormecido(a) o incapaz de tener sentimientos amorosos hacia las personas cercanas a ti?	1	5	C
12	Sientes como si tu futuro se cortara de alguna manera?	1	5	C
13	Tienes problemas para quedarte dormido?	1	5	D
14	Te sientes irritable o tienes ataques de ira?	1	5	D
15	Tienes dificultades para concentrarte?	1	5	D
16	Estás "súper alerto(a)" o vigilante y en guardia?	1	5	D
17	Te sientes nervioso(a) o fácilmente asustado?	1	5	D

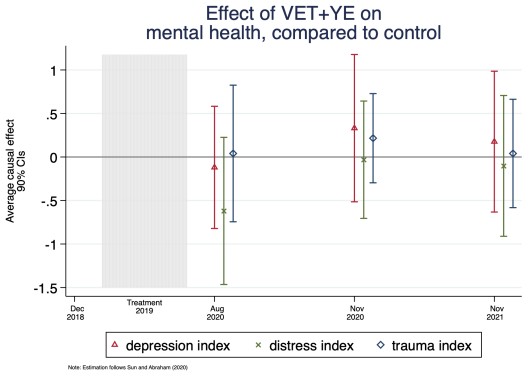
The variables are built by adding up the answers for each question, where 1 signifies "not at all" and 5 signifies "a lot". A variable of over 30 is considered severe. An individual should consider seeking professional treatment if their answers indicate a diagnosis for PTSD is possible, which entails that they answer between 3-5 to at least one 'B' item (Questions 1–5), at least 3 'C' items (Questions 6–12), and at least 2 'D' items (Questions 13–17). The 0-1 variable of a potential diagnosis is what we call "DSM-IV diagnosis", since it follows these guidelines, by the Diagnostic and Statistical Manual of Mental Disorders - fourth edition (colloquially known as the DSM-IV), which is the manual for the classification of mental disorders used by mental health professionals to diagnose patients.

A.1.3 Heterogeneity field of study

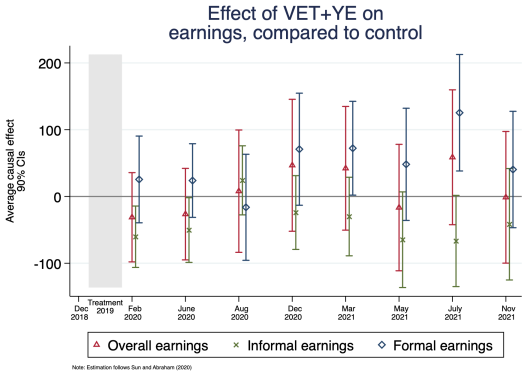
This section documents heterogeneity of the treatment in terms of earnings and mental health, following Sun and Abrahams (2020) and truncating the sample based on vocation.

Figure A1: **Average effect of treatment based on field of study**

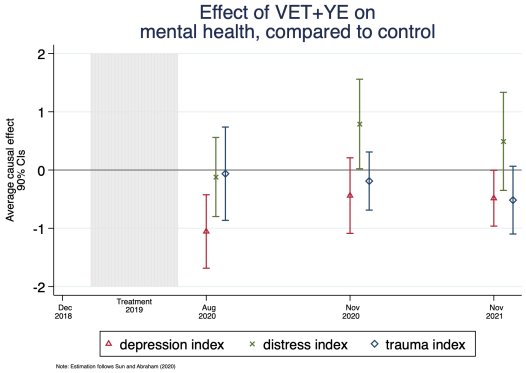
(a) Effect of VET+YE on mental health, administration



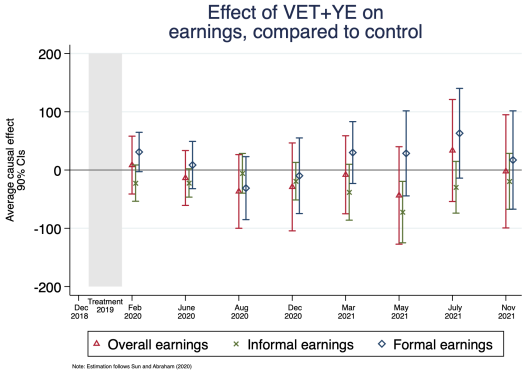
(b) Effect of VET+YE on earnings, administration



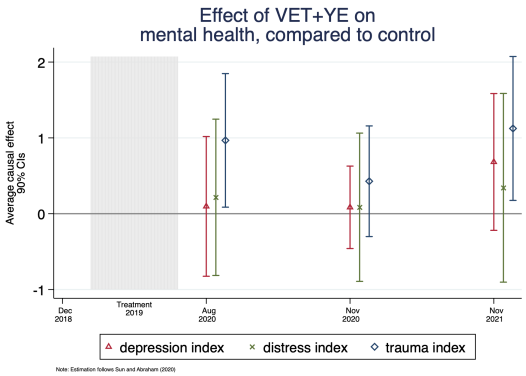
(c) Effect of VET+YE on mental health, child care



(d) Effect of VET+YE on earnings, child care



(e) Effect of VET+YE on mental health, logistics



(f) Effect of VET+YE on earnings, logistics

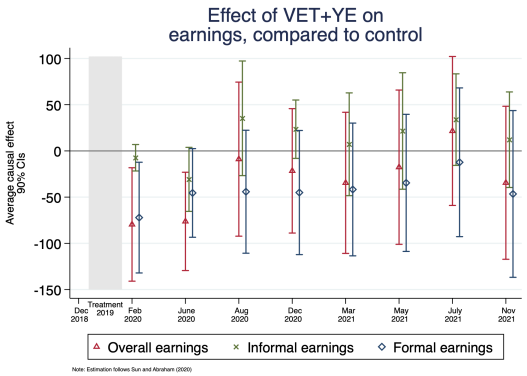
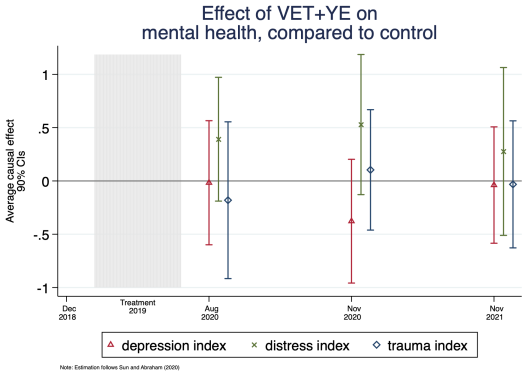
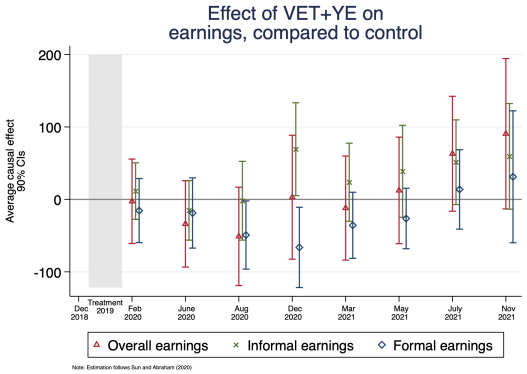


Figure A2: **Average effect of treatment based on field of study**

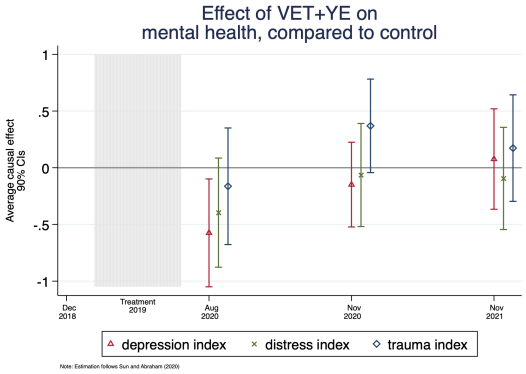
(a) Effect of VET+YE on mental health, marketing



(b) Effect of VET+YE on earnings, marketing



(c) Effect of VET+YE on mental health, HR



(d) Effect of VET+YE on earnings, HR

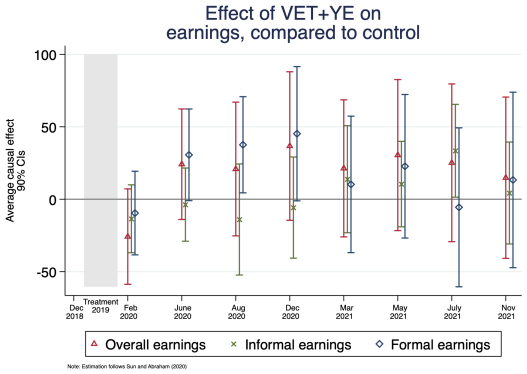
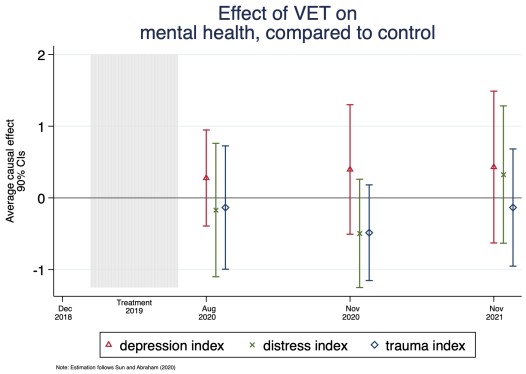
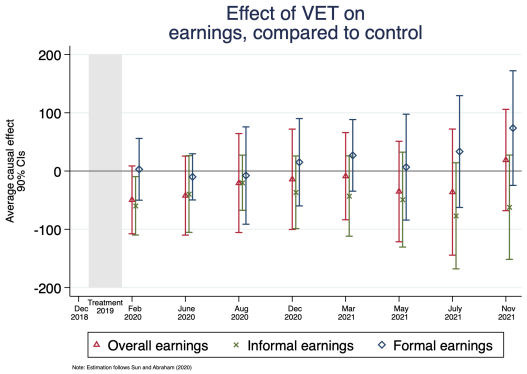


Figure A3: **Average effect of treatment based on field of study**

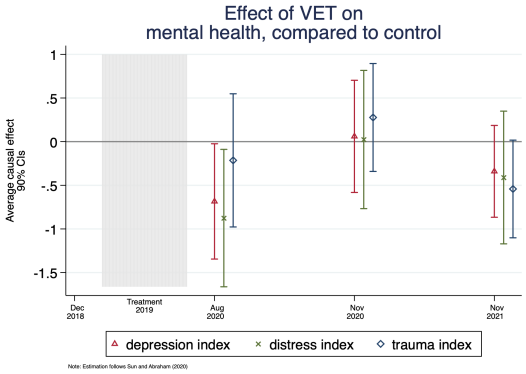
(a) Effect of VET on mental health, administration



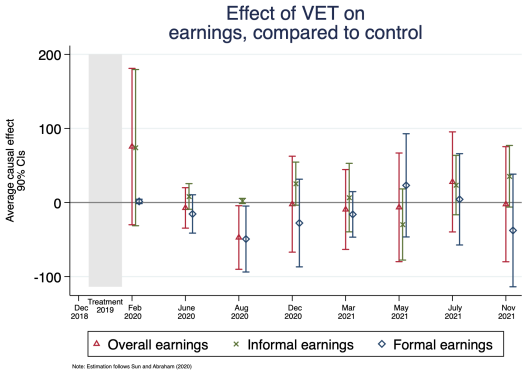
(b) Effect of VET on earnings, administration



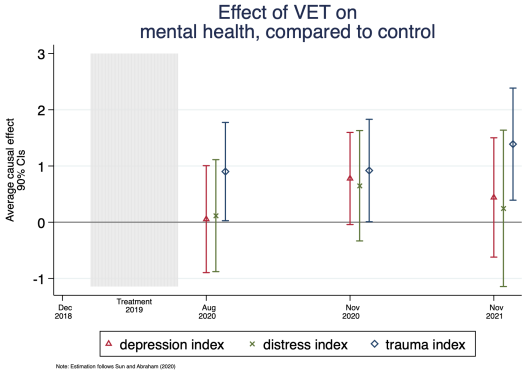
(c) Effect of VET on mental health, child care



(d) Effect of VET on earnings, child care



(e) Effect of VET on mental health, logistics



(f) Effect of VET on earnings, logistics

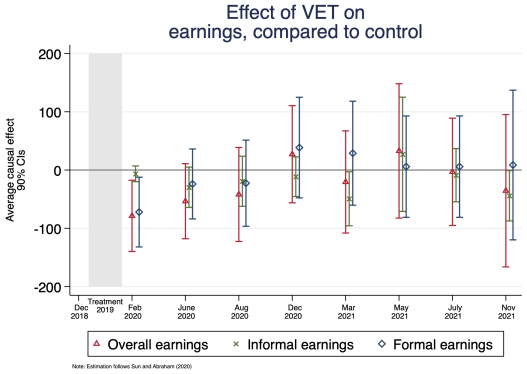
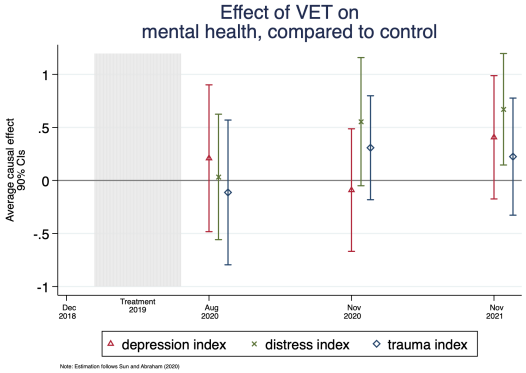
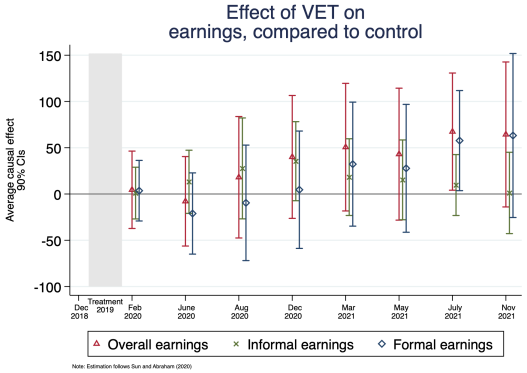


Figure A4: **Average effect of treatment based on field of study**

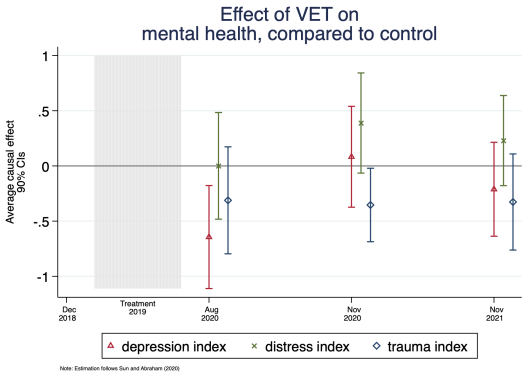
(a) Effect of VET on mental health, marketing



(b) Effect of VET on earnings, marketing



(c) Effect of VET on mental health, HR



(d) Effect of VET on earnings, HR

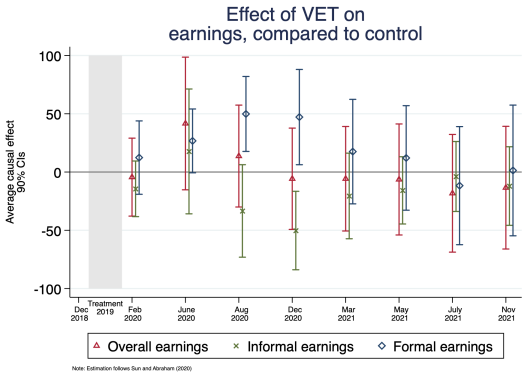
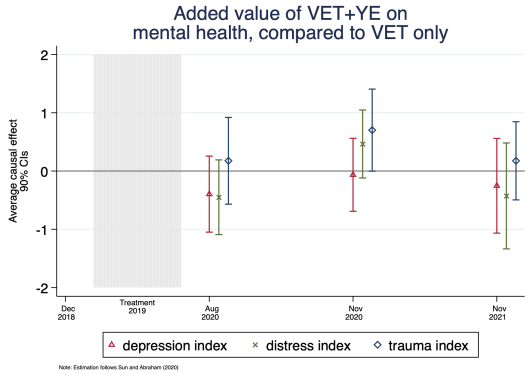
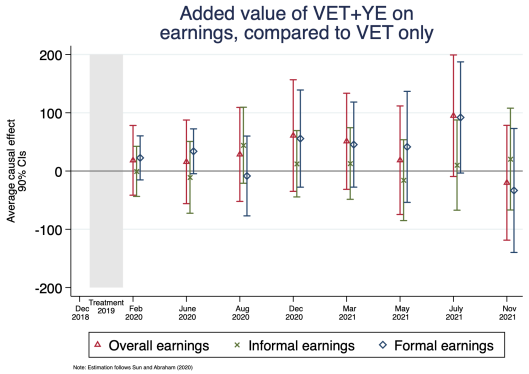


Figure A5: **Average effect of treatment based on field of study**

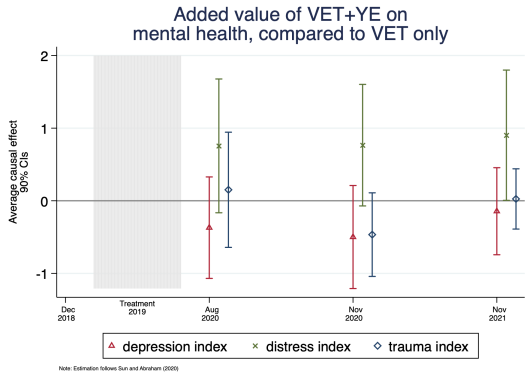
(a) Added value of VET+YE on mental health, administration



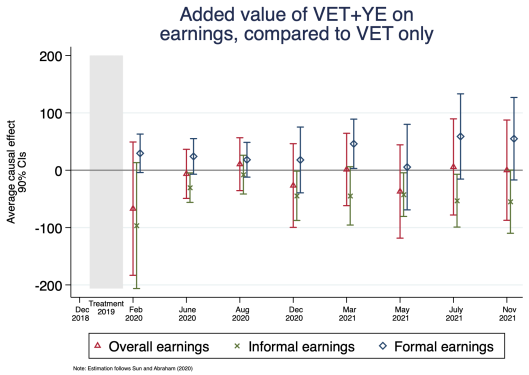
(b) Added value of VET+YE on earnings, administration



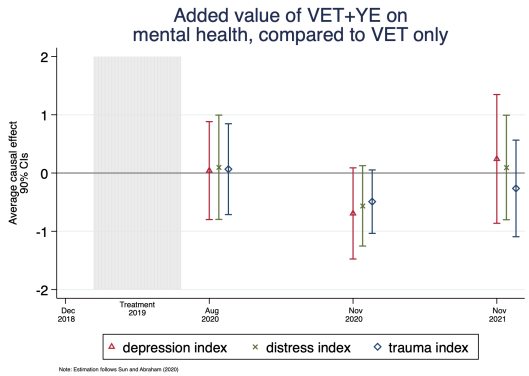
(c) Added value of VET+YE on mental health, child care



(d) Added value of VET+YE on earnings, child care



(e) Added value of VET+YE on mental health, logistics



(f) Added value of VET+YE on earnings, logistics

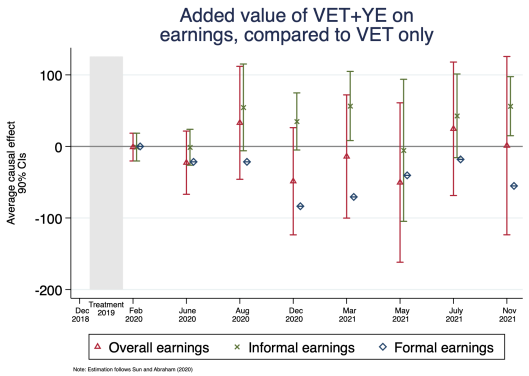
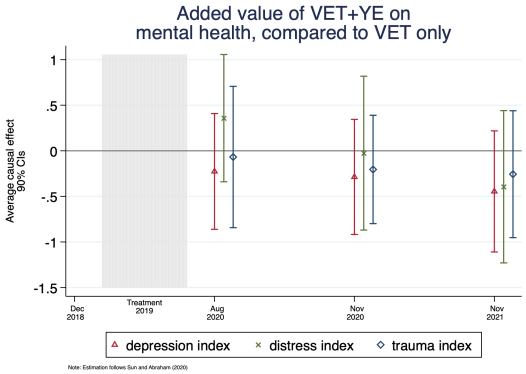
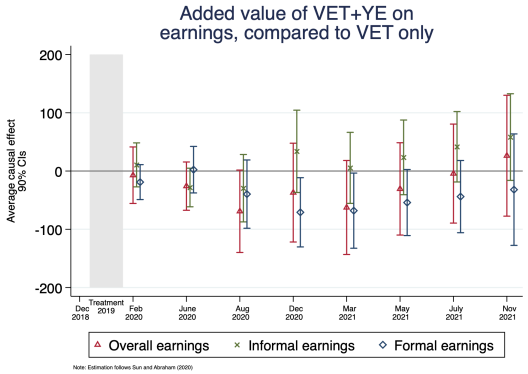


Figure A6: **Average effect of treatment based on field of study**

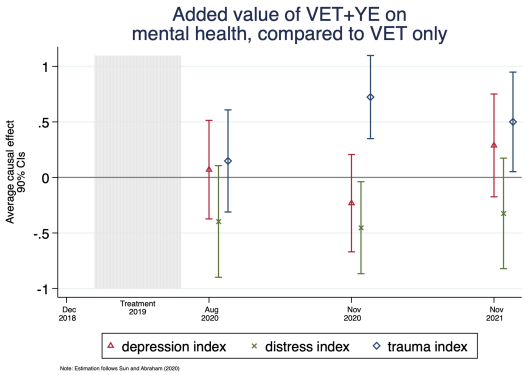
(a) Added value of VET+YE on mental health, marketing



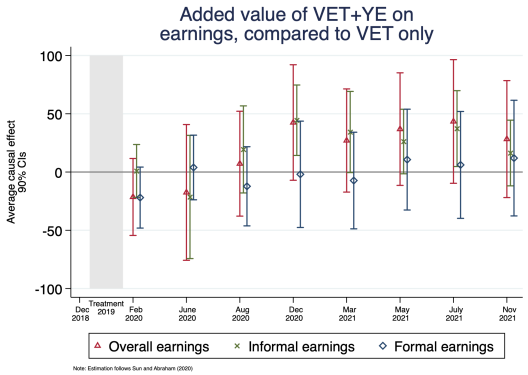
(b) Added value of VET+YE on earnings, marketing



(c) Added value of VET+YE on mental health, HR



(d) Added value of VET+YE on earnings, HR



Adding youth empowerment to vocational education provides greater resilience to shocks among vulnerable young adults: Evidence from the COVID-19 pandemic in Colombia

Alice Antunes[1] Elena Esposito[2] Rafael Lalive[1] Andres Moya[3], [4]

April 5, 2023

Abstract

We use the results of a randomized controlled trial to explore whether a youth empowerment program, combined with vocational education, can help vulnerable young adults cope with shocks such as COVID-19. We had previously randomized 300 young adults in low-income suburbs of Bogotá, Colombia into a group receiving vocational education and a youth empowerment program, a group receiving only vocational education, and a non-treated control group. Using a survey we ran during the COVID-19 lockdown, we can explore if, and to what extent, the program shielded participants from the mental health shock of the pandemic. Young adults who received vocational education managed symptoms of depression and distress better during the lockdowns compared to participants in the control group. The effects are stronger for women. There is no effect on trauma, which is more pronounced for the entire sample. In the medium run, mental health bounces back to baseline values. There seems to be an acute cushioning effect of treatment and vocational education seems to be the force behind the cushion. The main implication is that negative mental health effects of shocks such as those observed during the COVID-19 pandemic may be mitigated by programs like the one implemented in our study.

Keywords: COVID-19, Vocational Education, Youth empowerment, Psycho-social support, Mental Health, Active Labor Market Policies

JEL Codes: J24, I12, I26, I31, J15

[1] Department of Economics, University of Lausanne.

[2] ESOMAS Department, Collegio Carlo Alberto, University of Torino

[3] Department of Economics, University of Los Andes.

[4] We would like to thank Stefanie Enssle, Ira Amin, and Jan Weber of Vivamos Mejor; Edilberto Garcia and Andrés Clavijo of Fundación Apoyar; and iQuartil for the fast turnaround on the COVID-19 survey during the lockdown. This project is registered under RCT ID #AEARCTR- 0003869. The project's data collection and extensive participant follow-up work were funded by SDC-NADEL's Impact Award of 2020, the HEC Research Fund 2019, and Enterprise for Society (E4S). Research assistants who were instrumental in the data work are Christian Arciniegas (University of Fribourg), Leonardo Pedretti (University of Lausanne), and Falone Moseka (University of Lausanne)

1 Introduction

During the COVID-19 pandemic and ensuing lockdowns, many people experienced increased uncertainty regarding sources of income, food insecurity, the death of loved ones, and fear for their health. These were justifiably sources of anguish with psychological implications that are not yet understood. According to US data, 42% of people surveyed by the US Census Bureau in December 2020 reported symptoms of anxiety or depression, compared to 11% in the preceding year. Similarly, in the UK, the incidence of individuals suffering from depression jumped from 10% in March 2020 to 19% in June 2020. This also holds for lower- and middle-income countries. In Bangladesh, a nationwide study of more than 10,000 individuals documented a prevalence of depression of 33% in the population during the hardest months of the pandemic (Mamun et al. (2021)). These consequences were worse for more vulnerable populations, such as women, lower-income individuals, and people affected by conflict (see Kola et al. (2021), Mamun et al. (2021), Hamadani et al. (2020), and Campo-Arias and De Mendieta (2021)). In Colombia, for example, a study of young mothers in a rural, conflict-affected area found that mental health deteriorated significantly more for women who were internally displaced, had lower education, and had pre-existing mental health conditions (Moya et al., 2021). Moreover, experts say that the effects on mental health are likely to persist beyond the end of the pandemic (Abbott, 2021).

Why are some people better able to deal with the psychological consequences of the pandemic than others? We use the results of a randomized controlled trial to explore whether vulnerable young adults who had undergone vocational education with a youth empowerment program during the previous year were better able to weather the shocks engendered by the COVID-19 pandemic. This study is interested in the short and medium-run mental health effects of the pandemic on the study participants. In this study, like in others (see Moya et al. (2021)), we refer to the effects of the pandemic as the effects of COVID-19 itself combined with the effects of the policy responses and the resulting socio-economic consequences. We do this because it is not possible to untangle the different factors.

We hypothesize that adding a youth empowerment program to the traditional Colombian vocational training program can provide vulnerable young adults with a set of tools that can increase their resilience in the face of extreme shocks. At first glance, it may seem counter-intuitive that a hard skills training program such as the one we study could generate resilience during a time when the unemployment rate soared: after all, the program was designed to improve labor outcomes. We have several ideas about why this might be but can only answer in part with the data we have. It is possible that the training program provided a cushion on mental health because participants with treatment knew they had the skills that the labor

market would need once it recovered. Alternatively, it provided a cushion on mental health, period, because participants had had psycho-social support and had learned to deal with their mental health during the program. Finally, participants may have been employed more than their counterparts who had not received treatment. Of these hypotheses, only the last can be disentangled from our data and we do not find it to be the case.

Using answers to a survey conducted during the COVID-19 lockdowns in Colombia on participants of the randomized controlled trial, we can explore how the pandemic affected the mental health of vulnerable young adults who had gone through training, compared to vulnerable young adults who did not. The randomized controlled trial randomly allocated 300 young adults in low-income suburbs of Bogotá, Colombia into a treatment group receiving vocational education and youth empowerment (Group VET + YE), a treatment group receiving vocational education (Group VET), and a non-treated control group (Control Group). The training programs (Group VET + YE and Group VET) ran from February to December of 2019. We ran a baseline survey right before the program took place and a follow-up survey during the COVID-19 lockdown in August 2020. The COVID-19 survey repeated many of the questions of the baseline survey, including questions on participants' mental health[11] - in particular symptoms of post-traumatic stress disorder (PTSD), depression, and distress. We complimented the questions from the baseline survey with questions from a questionnaire designed by the Innovations for Poverty Action (IPA) for COVID-19 specifically, the RECOVR survey[12]. The RECOVR survey had been designed to track how people's lives were affected by the pandemic, and the Colombian questionnaire had been formulated with the Colombian pandemic context in mind. In the medium run, we collected mental health data in November 2020, and again in November 2021.

We are interested in exploring the effects of the COVID-19 pandemic on urban, vulnerable young adults. Our study targeted vulnerable young adults that live in two low-income suburbs of Bogotá, Colombia: Bosa, a neighborhood of Bogotá, and Soacha, a neighboring town with a bare minimum of infrastructure. Many of the residents of these neighborhoods are internally displaced from the long-running conflict in Colombia. In addition, there is a high prevalence of employment in the informal labor market in these neighborhoods. The informal labor market was particularly affected by the pandemic. Young adults that had completed the minimum required education were eligible to join the program if they belonged to the lowest socio-economic levels as measured by the System of Identification of Social Program Beneficiaries (SISBEN).

Our findings reveal that in the short run, all participants had higher incidences of trauma,

[11] We also collect extensive details on employment and earnings, but this study is interested in the mental health effects of the pandemic.

[12] More information on the survey can be found here: https://www.poverty-action.org/questionnaire/recovr-colombia-survey-round-2.

depression, and distress. Young adults who went through the two training programs, however, had smaller increases in depression and distress compared to participants in the Control group. On average, we find that being exposed to the youth empowerment program (Group VET+YE) is associated with an increase in depression that is 0.32 standard deviations lower than the increase in depression for participants in the Control group. The increase in symptoms of distress for participants in VET+YE is 0.2 standard deviations lower than the increase in distress for participants in the Control group, and the increase in PTSD symptoms is 0.08 standard deviations lower. These latter two results are not statistically significant.

Vocational education in and of itself seems to be the force behind the cushion for excess stress during the crisis. There is a greater resilience of the participants in both treatment groups (VET+YE and VET) - they are experiencing symptoms similar to others but these tend to be less severe. There is no significant difference in the mental health values between participants in VET+YE and VET, though the cushion is of a higher magnitude for participants in VET+YE compared to participants in VET only. This finding is in line with the literature that finds that people with more education suffered less during the pandemic. We do not find any differences in labor outcomes during the pandemic between the three groups. Our results indicate that vocational education was particularly beneficial for vulnerable sub-groups, women, displaced individuals, participants with little prior attachment to the labor market, and with worse psychological health before the intervention.

In the medium run, mental health bounces back to baseline values. There seems to be a strong, acute, cushioning effect of treatment - in particular for symptoms of depression -, though the effect dissipates as everyone's levels of mental health go back to their initial values. In the medium run, there is no significant difference between participants of different groups on mental health.

The study underlines the importance of updating traditional education systems with new training programs that target soft skills. There has been a discussion in policy circles about the benefits of so-called transversal skills in education (Inter-American Development Bank, 2019). Anecdotal evidence indicates that it is not only hard skills but also soft skills that help young adults have a fulfilling life. This study adds to the current state of knowledge by providing evidence for how adding a youth empowerment program to traditional vocational education can increase resilience to shocks such as COVID-19. Our results indicate that soft skills played an important role in the resilience of participants during the COVID-19 pandemic.

This paper has several contributions to the literature. First, we contribute to the increasing understanding of COVID-19 lockdowns on vulnerable populations in low- and medium-income countries. We know that there has been an increase in distress and depression overall

(Kola et al., 2021), and more particularly for vulnerable sub-populations where social safety nets are weaker and informal employment is eminent (Moya et al., 2021). Furthermore, our study contributes to the understanding of how non-tangible skills training programs can help vulnerable populations improve their well-being. Liberia's Sustainable Transformational Youth program found that employment and cognitive behavioral therapy increased self-control and non-criminal values, and lowered violence in ex-combatants (Blattman and Annan, 2015). A women's life skills program in Liberia resulted in more access to money and self-confidence, as well as less anxiety (Adoho et al., 2014). A women's empowerment training program in Uganda, which combined vocational education with reproductive health education, found that four years post-intervention, women are more likely to engage in income-generating activities, are less likely to be pregnant, and are less likely to get married early (Bandiera et al., 2014).

The paper proceeds as follows. Section 2 provides some background on the literature that already exists in the field. Section 3 gives information about the study design. Section 4 presents short-term results, and section 5 presents medium-run results. Section 6 presents some exercises on split samples to check for heterogeneous effects and presents results from alternative specifications.

2 Related Literature

2.1 Evidence before this study

A large stream of literature has documented the dramatic mental health implications of the COVID-19 pandemic, in terms of depression, distress, and trauma (see Serafini et al. (2020), Salari et al. (2020), Pierce et al. (2020), Giuntella et al. (2021), Brodeur et al. (2021), and Abbott (2021) among many others). In low- and medium-income countries, several studies have shown that the consequences on the psychological well-being of the population have been just as severe as in wealthier countries (see Mamun et al. (2021), Hamadani et al. (2020), and Campo-Arias and De Mendieta (2021)).

The main results are that on the one hand, health systems had fewer resources for providing adequate medical support. On the other hand, the policies implemented to contain the pandemic - such as lockdowns, closure of schools, shifting allocation of health resources, and curtailed livelihood opportunities - had particularly detrimental consequences in contexts with more vulnerable populations, weaker social safety nets, more informal employment, and less adequate economic resources to buffer against lost livelihoods (see Kola et al. (2021) for an insightful review). In countries where the vast majority of people are at a constant threat to their livelihood, the pandemic - together with the policy responses adopted to mitigate it - repre-

sented a special burden for already vulnerable groups, such as women, women with children, and people with pre-existing physical and psychological vulnerabilities (see Silverio-Murillo et al. (2020), Becerra et al. (2020), Moya et al. (2021), and Parra-Saavedra and Miranda (2021)).

2.2 Added value of this study

Addressing mental health issues was a priority for public institutions throughout the pandemic - they will be dealing with the repercussions for years to come (Holmes et al., 2020; Moreno et al., 2020). The fear of another similar pandemic has pushed institutions to search for policies that can reduce the psychological consequences of similar shocks. We know people with better education had greater resilience, but there is limited knowledge on what kind of training helped people do better. There is still a limited understanding regarding what kind of programs and policies can effectively increase the resilience of vulnerable populations in the face of similar shocks. To the best of our knowledge, this study is the first to evaluate whether vocational education, combined with youth empowerment, reduces the negative consequences of the pandemic on mental health.

Furthermore, our study design allows us to disentangle the value-added of combining vocational education with non-tangible skills training programs. Vocational education is a popular policy tool promising to improve the well-being of vulnerable populations that have not consistently fulfilled its promises (Attanasio et al., 2011; Card et al., 2011). Such experiments have compared treatment groups of vocational education against untreated control groups, with mostly null effects. Vocational education programs combined with other kinds of training, on the other hand, have shown more promise (Blattman and Annan, 2015; Bandiera et al., 2014; Adoho et al., 2014). These experiments in general compare a treatment group of vocational education and some non-tangible skills training programs against untreated control groups. Our study combines both approaches to explore the value-added of combining vocational education with a youth empowerment program. We, therefore, compare a treatment group receiving vocational education and youth empowerment (group VET+YE), to a treatment group receiving vocational education only (group VET) and to an untreated control group.

The group receiving vocational education and youth empowerment (group VET+YE) was trained to be more resilient to shocks to have a higher quality of life. While our design did not set up to answer the question of whether group VET+YE would be better off as a result of the COVID-19 crisis, the setting does provide a good opportunity to test whether the training program was effective.

Implications of all the available evidence: Vulnerable groups are particularly at risk in the face of extreme shocks such as the COVID-19 pandemic. Targeted labor market programs

mixing vocational education with youth empowerment can offer important tools for fostering resilience, with long-term positive effects going well beyond labor market gains.

3 Methods

The study was designed to study how vulnerable young adults fare in a highly volatile economic and social environment. For this reason, COVID-19 provided a unique opportunity to study the effects of a training program such as ours in light of intense economic pressure. In this study, we are interested in finding out whether vocational education combined with youth empowerment helps vulnerable young adults weather the mental health impacts of the COVID-19 pandemic. To this end, we explored the impacts of the COVID-19 pandemic on a cohort of participants in a randomized controlled trial offering vocational education and youth empowerment.

3.1 COVID-19 in Colombia

The training programs (Groups VET + YE and VET) ran from February to December 2019. A few months later, in March 2020, COVID-19 reached Colombia, and after a set of preliminary measures, the government mandated a national quarantine which lasted until the end of August 2020. All in all, COVID-19 was the primary cause of death in Colombia in 2020, with about 50,000 fatalities. Beyond mortality, the COVID-19 pandemic and the related policies adopted to mitigate the health burden of the disease had large negative impacts on Colombian society, with important disruptions in terms of mental health and economic well-being.

According to the IPA's RECOVR survey, 43% of individuals reported that somebody in the household had skipped necessary healthcare during the pandemic. More than half of the respondents declared that they had to reduce their meals in the week and that they were not able to find $270 to pay for emergency expenses. Respondents in lower socioeconomic levels were more greatly affected, even though they could more easily receive subsidies from the government, given that they were already in the system as vulnerable (for instance, beneficiaries that were already part of existing programs such as *Colombia Mayor*, *Jóvenes en Acción*, *Familias en Acción*) (Garcia et al., 2021). In terms of mental health and psychological well-being, more than 40% of the children have reported additional anxieties and 7% of the respondents declared worrying about sexual violence from their partners.

In terms of employment, only half of the respondents who were employed on February 2020 were still employed a few months later; 60% of the respondents saw their debts increase during the quarantine. The formal sector saw fewer losses of jobs compared to the informal sector but was slower to recover the jobs lost. Workers in the informal sector saw greater losses

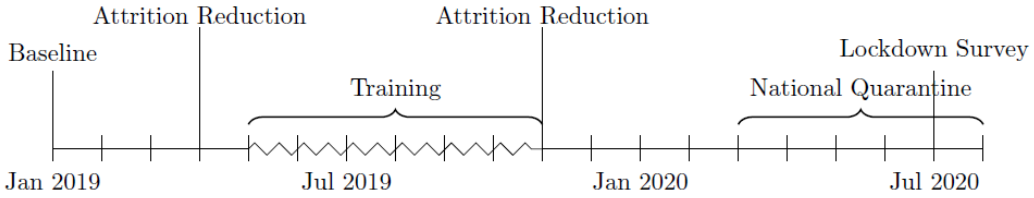
in earnings. Finally, women were more likely to stop working compared to men (Becerra et al., 2021).

3.2 The intervention

The randomized controlled trial targeted disadvantaged young adults in low-income suburbs of Bogotá, Colombia: Bosa and Soacha (see the Supplementary Material Section B.1.2.1 for further information regarding the study area). For the experiment, we randomly allocated 300 participants into a treatment group receiving vocational education and a youth empowerment program (Group VET + YE), a treatment group receiving only vocational education (Group VET), and a non-treated control group (Control Group).

Our randomized controlled trial proceeded as follows: our implementing partner in Bogotá recruited participants throughout October-December 2018 in Bosa and Soacha by canvassing and networking with important households in each neighborhood. Between December 2018 and January 2019, participants filled out baseline surveys in exchange for a cash incentive of approximately 10 USD and were subsequently randomized into one of three groups (VET+YE, VET, control). The participants that had been sorted into one of our two treatment groups (VET+YE and VET) and then followed the nationally certified vocational education programs from February-December 2019. In December 2019, participants from the control group were brought into our first series of "attrition reduction" events - which we created to keep in touch with participants in the control group - where they provided information about where they were in life (what they had done the previous year, whether they had jobs and how much they were making), in return for a cash incentive of approximately 10USD and for a chance to win a little gift (such as a bicycle) in a lottery. We organized a series of events to accommodate different working schedules. This was an important part of our design because our target population is highly mobile. Finally, after the first series of COVID-19-related lockdowns were instated, we ran a phone-based survey in August 2020 to understand the effects of COVID-19 on our study participants.

Figure 1: Project Timeline



Notes: The figure illustrates the general timeline of the intervention, of the data collection main stages, and of the national quarantine adopted to mitigate the effect of the COVID-19 pandemic.

The timeline of our randomized controlled trial – see Figure 1 – allows us to assess the impact of the COVID-19 pandemic on the mental well-being of our study participants, comparing individuals in the group that had just received vocational education and youth empowerment, or only vocational education, with the impact of the COVID-19 pandemic on the control group. It is important to note that while our study was not designed with COVID-19 in mind (such a thing would require omniscient foresight), it was designed to study how marginalized youth fare in a highly volatile economic environment. For this reason, COVID-19 provided a unique opportunity to study the short-term effects of a training program such as ours in light of intense economic pressure.

Participants following vocational education (groups VET + YE and VET) received the nationally certified classroom and training program for the following professions: human resources, administration, child care, logistics, and marketing. Participants in the control group applied to be eligible for the study under specific vocations. Participants who were subsequently sorted into the control group did not end up following the vocation they applied to.

In addition to this, participants in group VET + YE received the nationally certified vocational education program, complemented by a youth empowerment program targeting marginalized young adults from Bosa and Soacha. This consisted of once-a-week training that took place in the same locations as the vocational courses. The curriculum was four-dimensional, centering around: (1) socio-emotional skills development (group work and individual therapy), (2) the development of a pacifist culture and citizen empowerment (to live with former paramilitaries and to learn to defend their rights), (3) business skills training (mock interviews, curriculum vitae development), and (4) labor placement support.

The first dimension, socio-emotional skills development, focused on group-oriented soft skills training sessions that target self-esteem and confidence. For example, some weekly train-

ing included theater sessions, to help participants learn how to speak in public. In addition, participants had the option to see an in-house psychotherapist to help them work through traumas, adjust expectations, or deal with the stressors of difficult living situations such as early pregnancies and living in violent environments.

The second dimension, the development of a pacifist culture and citizen empowerment was specifically designed for the population. More than 10% of the young adults in the program are internally displaced due to the ongoing conflict. Part of their integration into the labor market includes being able to work alongside someone who might have been part of a group that was responsible for their internal displacement. Therefore, the pacifist culture dimension of the training program encourages young adults to learn to forgive people who may have been involved in their victimization. The intent is to contribute to the ongoing peace-building efforts in post-conflict Colombia. In addition, this dimension encourages the political empowerment of these young adults. The idea is to teach them an awareness of their political rights to encourage them to take ownership of their futures.

The third and fourth dimensions could be labeled under the labor market umbrella. Business skills training targeted skills such as how to act in a work environment, how to succeed in job interviews, and how to write up one's curriculum vitae. Labor placement support included encouraging participants to attend career fairs.

3.3 Experimental design

Our study ran from late 2018 through the end of 2021. It followed 300 young adults who were recruited by Fundación Apoyar, our local implementing partner, in the suburbs of Bosa and Soacha, soliciting motivated applicants through posters, house visits, and interactions with community leaders. Participants in the study were recruited via Fundación Apoyar, which made announcements about their program throughout the neighborhoods of Soacha and Bosa. Because Fundación Apoyar had been working in the neighborhoods for several years, their program was already known in the neighborhoods.

The recruitment process included a vulnerability screening of each participant. Only the 300 most vulnerable were accepted to the program, due to budgeting constraints. A participant was considered vulnerable if they had a SISBEN ranking of 1 or 2[13]; if their living conditions were poor, as judged by the NGO following a home visit; and if they scored as vulnerable in a vulnerability score computed by the NGO. Also, participants were accepted under the following additional conditions: they had to have completed or be completing their mandatory school

[13]SISBEN, or System of Identification of Social Program Beneficiaries, is a household vulnerability index that is used to identify the beneficiaries of social assistance programs in Colombia. Households received a score of 1 if they were considered most vulnerable, and 6 if they were considered most privileged. The scoring system has been updated into an index between 0 (vulnerable) to 100 (privileged). This study uses the old 1-6 scoring system.

requirement, be between 18 and 25 years old, pass an interview with the in-house psychologist, and have a family in support of them following the program. To reduce the risk of attrition, only subjects with families that approved participation in the program were selected.

The 300 selected participants were then randomly allocated to three groups: a treatment group receiving vocational education and soft skills training/psycho-social support (Group VET+YE); a treatment group receiving vocational education (Group VET); and a non-treated control group (Control Group). Each group was composed of roughly 100 young adults. Randomization was done via public lottery at the neighborhood headquarters of the managing NGO, to reduce suspicion related to treatment allocation and to increase the transparency of the randomization process. This was especially important also in light of the high mistrust of authorities in certain parts of Colombia. Randomization was stratified at the neighborhood and vocation levels.

3.4 Outcome variables

We are interested in finding out whether bundling vocational education with a youth empowerment program helps vulnerable young adults cope with the psychological consequences of the COVID-19 pandemic. We had previously administered a baseline survey before the training, before randomization, in December 2018. To understand how the different training groups were affected differentially during COVID-19, we administered an additional survey in August 2020, during the last weeks of the COVID-19 lockdown. This was done via phone surveys due to the lockdowns. Finally, a midterm survey was administered in November 2020, and an endline survey in November 2021.

Baseline Survey. The baseline survey was conducted right before randomization and included also an interview to collect data on demographics, income, and employment data. Psychological well-being was elicited, using written responses, through three types of measures:

- **Indexes of Depression and Distress:** To elicit depression and distress we followed Blattman et al. (2017). In particular, we asked an array of questions aiming to elicit symptoms of distress and symptoms of depression. For instance, in the case of depression symptoms, the respondent is asked to disagree/agree (giving a score from 0 to 3) with statements such as "I feel bad doing things I normally do" or "I am feeling sad or downhearted". The full set of depression symptoms assessed is summarized in the Supplementary Material, Section B.1.2.4. The index of depression and distress is then built summing up the answers to each question. Because each variable ends up being an index of different lengths, we also compute a version of the variable after transforming it into a z-score, with a mean

of 0 and a standard deviation of 1.

- **PTSD checklist:** To assess symptoms of post-traumatic stress disorder (PTSD), we also relied on the so-called PTSD checklist - civilian version (PCL-C), which is self-administered. This questionnaire allowed us to gauge whether individuals express any symptoms of trauma. Participants were asked to answer 1 when a statement indicated an answer of "Not at all", and 5 when a statement indicated an answer of "Extremely". We compute a severity score index by summing up over the 17 questions. The severity score ranges between 17 and 85. We also follow the standard psychological guidelines for constructing a 0-1 variable which indicates a presumptive diagnosis of PTSD (according to the Diagnostic and Statistical Manual of Mental Disorders, which is considered best practice in psychology). A participant meets the criteria for a presumptive diagnosis if they answer symptomatically ('Moderately' - 3 - or higher) to at least 1 question between questions 1-5, to at least 3 questions between questions 6-12, and at least 2 questions between questions 13-17. We tend to prefer the presumptive diagnosis variable given that it allows for more subtlety in the severity of questions.

- **Other Variables** The baseline survey included an extensive set of questions to gather demographic, income data, and information on preferences (patience, altruism...) and other personal traits. A full list can be found in the Supplementary Materials Section B.1.2.4.

COVID-19 Survey The COVID-19 survey was conducted via telephone interviews due to sanitary restrictions and to protect participants' health. The telephone interviews (each lasting 20-30 minutes) were assigned to the company iQuartil and were implemented between July and August 2020. Participants were compensated with a small sum of money equivalent to USD 10. The survey included: the questions to compute the Indexes of Depression and Distress following Blattman et al. (2017) described above and the PTSD checklist, also described above. On top of these variables, we also asked questions on the lockdown based on the IPA RECOVR surveys, a standardized survey designed to track how people's lives are affected by the COVID-19[14].

Demographic Characteristics Demographic variables include: gender, age, marital status, number of children, being displaced or not (participants who are displaced have a registration card proving they are internally displaced due to the ongoing conflict), and socio-economic status (the (old) SISBEN index was a 1-6 index, given to people by the government, and is an indicator of whether or not someone is eligible for government aid).

[14] see $https://www.poverty-action.org/recovr/recovr-survey$

We also collect administrative labor data, including formal/informal employment (both self-reported and verifying for formal employment by double-checking that the participant has current records contributing to the retirement pension system called PILA, which is the method followed by Attanasio et al. (2017), hourly wages and monthly wages (both formal and informal). In this case, informal employment is any employment without a formal contract, which then also precludes the employee from social safety nets and benefits related to income.

3.5 Randomization and balance

Table 1 describes the sample at baseline, comparing demographic characteristics, labor characteristics, and mental health status across the three different groups. Women are the majority in all groups, with a slightly lower percentage in the Control group compared to the VET-only group. This raises the question as to whether women are over-represented in the sample. Anecdotal evidence from Fundación Apoyar indicates more women sign up for their program. That said, women tend to be more represented in technical education in Colombia: a descriptive study from 2013 found that of the people enrolled in technical studies, 61% were women (Bornacelly, 2013). There are several reasons this might be the case, some of them including such hypothetical ones as: within socio-economic stratus, women tend to seek higher education; men might be more likely to be involved in criminal activity than women; or women might be more likely to be of lower socio-economic status than men, making them eligible for the program. We do not test these. Individuals in the Control group tend to be slightly older than participants in the other two groups. In addition, they tend to have more children. Last, individuals in the Control group are significantly more likely to be displaced. Given the significance of these differences, our baseline regressions will account for all these demographic characteristics. Reassuringly, labor market characteristics at baseline are homogeneous across the different groups, except informal sector employment, which is slightly higher in participants in group VET only. This is also the case for mental health conditions, with the sole exception of distress at baseline, which tends to be lower among participants allocated to the VET-only group.

Table 1: **Balanced Table at Baseline**

	(1)	(2)	(3)	(4)	(5)	(6)
Variable	VET+YE	VET	Control	Diff (1)-(3)	Diff (2)-(3)	Diff (1)-(2)
Female	0.710	0.808	0.660	0.050	0.148**	-0.098
	(0.456)	(0.396)	(0.476)	(0.449)	(0.018)	(0.107)
Age	20.140	19.980	20.980	-0.840*	-1.000**	0.160
	(3.269)	(3.574)	(3.423)	(0.077)	(0.045)	(0.742)
Number of Children	0.390	0.444	0.710	-0.320***	-0.266**	-0.054
	(0.650)	(0.772)	(0.868)	(0.004)	(0.024)	(0.591)
Displaced (1=yes)	0.080	0.091	0.190	-0.110**	-0.099**	-0.011
	(0.273)	(0.289)	(0.394)	(0.023)	(0.045)	(0.784)
SISBEN 1-6 scale	1.760	1.667	1.560	0.200**	0.107	0.093
	(0.553)	(0.535)	(0.538)	(0.010)	(0.162)	(0.227)
Earnings (USD/mo)	28.130	33.277	31.454	-3.324	1.823	-5.147
	(71.831)	(67.989)	(74.662)	(0.749)	(0.857)	(0.604)
Formal earnings (USD/mo)	5.220	5.126	14.764	-9.544	-9.638	0.094
	(40.255)	(35.897)	(60.706)	(0.192)	(0.175)	(0.986)
Informal earnings (USD/mo)	22.910	29.029	17.676	5.234	11.354	-6.119
	(61.488)	(61.277)	(51.565)	(0.515)	(0.159)	(0.483)
Employed (1=yes)	0.180	0.242	0.190	-0.010	0.052	-0.062
	(0.386)	(0.431)	(0.394)	(0.856)	(0.371)	(0.283)
Formal sector work (1=yes)	0.020	0.020	0.060	-0.040	-0.040	-0.000
	(0.141)	(0.141)	(0.239)	(0.150)	(0.155)	(0.992)
Work in the informal sector	0.160	0.222	0.130	0.030	0.092*	-0.062
	(0.368)	(0.418)	(0.338)	(0.549)	(0.088)	(0.266)
Work Experience	0.680	0.657	0.720	-0.040	-0.063	0.023
	(0.469)	(0.477)	(0.451)	(0.539)	(0.336)	(0.727)
PTSD severity index	30.010	30.788	31.705	-1.695	-0.917	-0.778
	(9.661)	(10.477)	(11.579)	(0.270)	(0.563)	(0.589)
PTSD DSM4 criterion	0.082	0.121	0.116	-0.034	0.005	-0.040
	(0.275)	(0.328)	(0.322)	(0.428)	(0.908)	(0.360)
Depression	4.227	4.323	4.500	-0.273	-0.177	-0.096
	(3.621)	(3.854)	(3.681)	(0.608)	(0.748)	(0.858)
Distress	10.062	8.583	10.703	-0.642	-2.120*	1.478
	(5.916)	(5.388)	(6.989)	(0.574)	(0.062)	(0.148)
Observations	100	99	100	200	199	199

NOTE: The table displays the demographic characteristics and psychological status of project participants, measured at baseline (January 2019), across the three groups. The fourth column presents the difference in the mean value of each variable between the first and the third column and the second and the third column respectively. P-values are reported in parentheses.

3.6 Attrition and compliance

Overall, 361 individuals were randomized into the three different groups and filled out baseline surveys. Up until April 2019, participants who dropped out were replaced by participants from

a waitlist. Participants who were on the waitlist were first sorted into groups and replaced participants that dropped out of their respective groups and vocations. Three hundred final participants were followed throughout the three years of the study.

Here, we define attrition as people who completed the baseline survey but who did not complete the COVID-19 survey. We define noncompliance as people who completed the baseline survey, but who either dropped out of the treatment group they were assigned to (Groups VET+YE and Group VET), or people who were assigned to the untreated control group and followed through with vocational training regardless.

Overall, 78 people who completed the baseline survey did not complete the lockdown survey. Of these, 28 were in group VET+YE, 37 were in group VET, and 13 were in the control group. One participant from group VET, unfortunately, passed away. This is an effective attrition rate of 21.6%. However, we followed 300 participants officially as opposed to all 361 who completed the baseline and were able to track down 283 for the COVID-19 lockdown survey (this would make the attrition rate 5.6%).

Table 1 below shows the balanced characteristics of those attritors. There does not seem to be any statistically significant difference between attritors in the three different groups, except attritors from the Control group tend to have fewer children than attritors from groups VET+YE and VET only.

Table 2: Balanced Table for Attritors

	(1)	(2)	(3)	(4)	(5)
Variable	VET + YE	VET	Control	Diff (1)-(3)	Diff (2)-(3)
Displaced	0.107	0.091	0.154	-0.047	-0.063
	(0.315)	(0.292)	(0.376)	(0.680)	(0.547)
Gender	0.679	0.818	0.846	-0.168	-0.028
	(0.476)	(0.392)	(0.376)	(0.271)	(0.826)
Age	20.714	20.242	20.615	0.099	-0.373
	(3.547)	(3.260)	(3.572)	(0.934)	(0.735)
Age squared	441.214	420.061	436.769	4.445	-16.709
	(156.392)	(139.641)	(160.548)	(0.933)	(0.728)
Work experience (1=yes)	0.821	0.545	0.769	0.052	-0.224
	(0.390)	(0.506)	(0.439)	(0.703)	(0.169)
Informal sector work	0.286	0.242	0.154	0.132	0.089
	(0.460)	(0.435)	(0.376)	(0.373)	(0.523)
Positive earnings	0.286	0.333	0.231	0.055	0.103
	(0.460)	(0.479)	(0.439)	(0.720)	(0.507)
One child	0.357	0.273	0.000	0.357**	0.273**
	(0.488)	(0.452)	(0.000)	(0.012)	(0.036)
Two children	0.036	0.061	0.154	-0.118	-0.093
	(0.189)	(0.242)	(0.376)	(0.185)	(0.323)
Three children	0.036	0.000	0.000	0.036	0.000
	(0.189)	(0.000)	(0.000)	(0.503)	()
Four children	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	()	()
Observations	28	33	13	41	46

NOTE: The table displays the demographic characteristics and psychological status of project participants who were measured at baseline but who did not follow up with the COVID survey, across the three groups. The fourth column presents the difference in the mean value of each variable between the first and the third column and the second and the third column respectively. P-values are reported in parentheses.

In terms of compliance, we asked 299 participants during the November 2020 midline data collection to indicate whether they finished vocational education (and if not, why not). Of those participants, 136 finished vocational education: 76 in group VET+YE, 57 in group VET, and 3 in the control group finished vocational education (the three participants in the control group who finished vocational education are noncompliant). Most participants cited children, pressure from work, and lack of resources as reasons for not finishing the program. Table 3 summarizes this information by group. Importantly, the drop-out rates for participants in VET

only are dramatically higher than dropout rates for participants in VET+YE. This seems to be similar to other studies that find large drop-out rates of participants who sign up for vocational education programs (Cho et al., 2013).

Table 3: **Breakdown of attritors and noncompliers by group**

Career	VET+YE	VET	Control	Total
Attrition				
Baseline	125	126	110	361
COVID-19 survey	97	89	97	283
Noncompliance				
Finished VET	76	57	3	136

Notes: This table shows attritors and noncompliers, by group.

Table 4 shows the results of the regression of treatment on attrition and noncompliance. Effectively, we estimate:

$$y_{i,g,t} = \alpha + \theta_1[\text{VET + YE}]_g + \theta_2 \text{VET}_g + \gamma_1 X_{i,g} + \varepsilon_{i,g} \tag{4}$$

Where the dependent variable for attrition is equal to one if a participant is present in the baseline but not in the COVID-19 survey; and where the dependent variable for noncompliance is equal to one for groups VET+YE and VET if the participant did not finish vocational education, and for the control group if the participant did do vocational education. Also, $X_{i,g}$ indicates the control variables of gender, age, age squared, number of children, being internally displaced, work experience, work in the informal sector, earnings, and socio-economic class as identified by the old SISBEN category.

Table 4, column 1, shows that, compared to participants in the control group, participants in the VET+YE program are 10% less likely to be present in the COVID-19 survey; and participants in the VET-only program are 17% less likely to be in the COVID-19 survey. This is likely highly mechanical, as participants in the Control group did not drop out of the study, while some participants who were sorted into the two treatment groups did. Column 2 shows that participants in VET+YE were 19% more likely to be noncompliant compared to participants in the Control group, and participants in VET only were 32.9% more likely to be noncompliant compared to participants in the Control group. This is also likely mechanical, as there are almost no participants in the Control group who underwent vocational education regardless of being sorted into a non-treated group. The coefficient on noncompliance for VET only is enormous - one-third of

Table 4: **Effects of treatment on attrition and noncompliance**

	Attrition	Noncompliance
VET + YE	-0.104**	0.192***
	(0.051)	(0.048)
VET	-0.172***	0.329***
	(0.056)	(0.058)
Individual Controls	Yes	Yes
Observations	361	283
R-squared	.0425	.13

Notes: The dependent variables are attrition and noncompliance. VET + YE indicates the group of participants who received vocational education and soft skill training, while VET indicates the group of participants that received only vocational education, the omitted category is the control group. Noncompliance is estimated only on participants who completed the COVID-19 survey. Robust standard errors clustered at the individual level are reported in round brackets. ***, **, and * indicate significance at the 1-, 5-, and 10-% levels, respectively.

participants who started the program and who were surveyed for COVID-19 did not complete their vocational education program the year before.

4 Results in the short run

4.1 COVID-19 and mental health in the short run: preliminary evidence

The COVID-19 lockdown in Colombia was particularly strict, with important repercussions for all the participants in the study. Participants in the vocational training groups (Groups VET + YE and VET) had finished their training programs a few months before the lockdowns were imposed. Regardless, all participants were affected. Almost two-thirds report having had their employment affected by the lockdown, and are currently receiving aid to help them during the lockdown. About one-quarter of participants report having to move during the pandemic, and more than half have developed anxiety, trouble sleeping, aggression, or sadness due to the pandemic. Overall averages, illustrated in figure 2, show that, while very few program participants were directly affected by the health consequences of the epidemic, their job status and the composition of their family changed because of the pandemic. Many of these outcomes do not vary greatly by group, which is why only overall averages are reported. Graphs breaking down these variables by group can be found in Supplementary Material Section B.1.2.3.

Figure 3 presents preliminary evidence on the differential psychological consequences of the COVID-19 pandemic across our three groups. We track, in particular, the variables associated

Figure 2: **COVID-19 effects - overall averages**

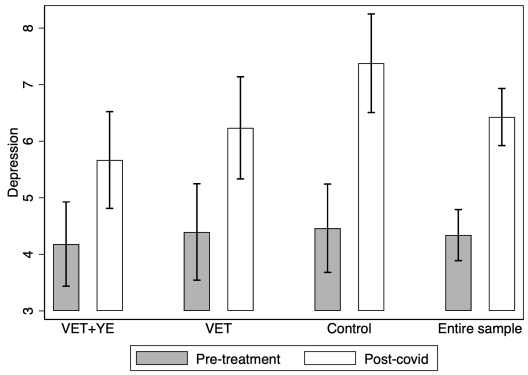
Note: This figure shows overall averages for reported effects of COVID-19 on our study participants. Variables are sorted by: demographic effects of the pandemic; job-related effects of the pandemic; and mental health effects of the pandemic.

with depression, distress, and trauma. As a reminder, the indices on depression and distress are created from a series of questions where participants filled in values between 0 and 3 when answering questions. Questions for depression include questions such as "I feel bad doing things I normally do" or "I feel unimportant to everyone". Questions for distress include questions like "I have bad dreams" or "I talk to myself".

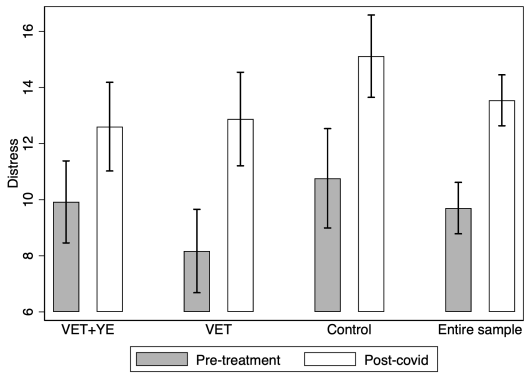
The trauma severity index was created from the PCL-C which is a self-administered test for post-traumatic stress disorder (PTSD). The variable is built by adding up the answers for each question, where 1 signifies "not at all" and 5 signifies "a lot". A variable of over 30 is considered severe. An individual should consider seeking professional treatment if their answers indicate a diagnosis of PTSD is possible, which entails that they answer between 3-5 to at least one 'B' item (Questions 1–5), at least 3 'C' items (Questions 6–12), and at least 2 'D' items (Questions 13–17). The 0-1 variable of a potential diagnosis is what we call "DSM-IV diagnosis", since it follows these guidelines, by the Diagnostic and Statistical Manual of Mental Disorders - fourth edition (colloquially known as the DSM-IV), which is the manual for the classification of mental disorders used by mental health professionals to diagnose patients. Calculating a simple difference in pre and post-averages by group, we can see that average levels of depression, distress, and trauma increased significantly across all groups, with all indexes increasing by around a half.

Figure 3: **Mean effects on depression, distress and trauma across treated and control groups**

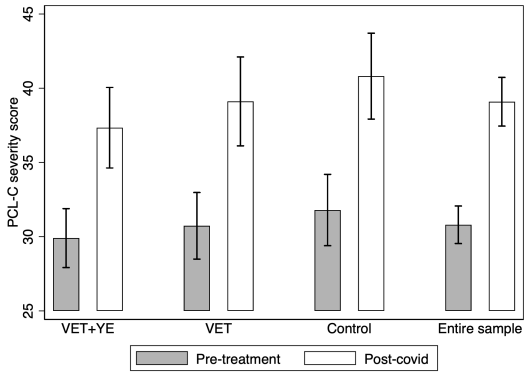
(a) Depression



(b) Distress



(c) PTSD Severity



Notes: This figure illustrates the average levels of Depression (panel a), Distress (panel b), and PTSD severity (panel C) for participants that received vocational education and youth empowerment (VET + YE), for participants that received only vocational education (VET), and for participants in the control group (Control), before the intervention and after the intervention, the COVID pandemic and national quarantine. Bar heights indicate the mean levels of each index. Gray bars indicate pre-COVID-19 levels, and white bars indicate post-COVID-19 levels. The capped lines indicate confidence intervals at the 0.05 significance level.

The different bars in panels a) and b) indicate that the increase in depression and distress was more pronounced in the Control group. Out of the two treated groups, the increase in de-

pression and distress is smaller in the group that received vocational education and youth empowerment, where the increase tends to not be statistically significant. In other words, while all participants were negatively affected by the pandemic, those exposed to a mix of soft and hard skills training are much less negatively affected. It seems at first glance that the share of participants that were only offered vocational education is not as resilient as the ones in the $VET + YE$ group, though levels at follow-up between VET + YE are not statistically different than those for VET only. Results on PTSD severity are qualitatively similar; however, the magnitudes of differences across the three groups are smaller and not statistically significant.

The figure shows that there is a big spike in symptoms of trauma severity, distress, and depression during the COVID-19 pandemic. The spikes between December 2018 and August 2020 are highest for participants in the control group: trauma (panel c) increases by almost 10 points from 31.705 to 40.814, depression (panel a) increases by 3 points from 4.5 to 7.4, and distress (panel b) increases by 4.44 points from 10.7 to 15.14.

In comparison, for participants who received VET only, trauma severity (panel a) increased by 8.32 points from 30.78 to 39.1, depression (panel b) increased by 1.9 points from 4.3 to 6.2, and distress (panel c) by 4.2 points from 8.6 to 12.8. For participants who received VET+YE, on average, trauma severity (panel a) increased by 7.3 points from 30.01 to 37.3, depression (panel b) increased by 1.45 points from 4.22 to 5.67, and distress (panel c) by 2.5 points from 10.06 to 12.5. These numbers are expressed clearly in table 5.

The increase in mental illness was higher during the COVID-19 pandemic for participants who did not receive youth empowerment, but most importantly for participants in the control group. Table 5 shows these numbers and also explores the difference in group means. In particular: participants in the control group scored, on average, 3.47 points higher in the trauma severity index compared to participants in VET+YE, 1.7 points higher in the depression index, and 2.6 points higher in the distress index. These are all statistically significant. In addition, participants in the control group score, on average, 1.7 points higher in the trauma severity index compared to participants in VET, 1.17 points higher in the depression index, and 2.26 points higher in the distress index. These values are also statistically significant, when not controlling for anything. The differences in group means are not statistically significant when comparing VET+YE to VET only. These trends indicate that vocational education overall, and youth empowerment in particular, seems to have helped provide some resilience to the COVID-19 pandemic.

Are there any potential reporting biases that could lead to Hawthorne effects or other such experimental design biases? For instance, could it be that participants who received VET+YE would feel like they should report smaller symptoms of depression, distress, and trauma be-

Table 5: **Mean baseline and post-COVID-19 values of mental health indices, and differences in means**

	VET+YE	VET	Control	Diff (1) - (2)	Diff (2) - (3)	Diff (1) - (3)
Depression						
Before	4.227	4.323	4.500	-0.096	-0.177	-0.273
				(0.538)	(0.550)	(5.313)
After	5.670	6.236	7.402	-0.566	-1.166*	-1.732***
				(0.622)	(0.630)	(0.610)
Distress						
Before	10.062	8.583	10.703	1.478	-2.120*	-0.642
				(1.015)	(1.126)	(1.139)
After	12.505	12.876	15.144	-0.371	-2.268**	-2.639**
				(1.155)	(1.105)	(1.076)
Trauma						
Before	30.010	30.788	31.705	-0.778	-0.917	-1.695
				(1.436)	(1.584)	(1.533)
After	37.340	39.112	40.814	-1.772	-1.702	-3.474*
				(2.029)	(2.096)	(1.997)

Notes: The table shows the mean values of depression (rows 2-3), distress (rows 5-6), and trauma (rows 8-9) at baseline (top rows) and post-COVID-19 (bottom rows), for groups VET+YE (column 2), VET (column 3), and Control (column 4). Column 5 shows the difference in means between VET+YE and VET, column 6 shows the difference in means between VET and Control, and column 7 shows the difference in means between VET+YE and Control. Standard errors are in parentheses. ***, **, and * indicate significance at the 1-, 5-, and 10-% levels, respectively, for the two-sample t-test results.

cause they received the youth empowerment training and feel like they are less entitled to stress compared to their peers (e.g. they feel privileged, in a way)? A just-as-valid concern is that participants in VET+YE would be expressing greater symptoms of depression, distress, and trauma precisely because of their training - they might be looking for greater sympathy for these distressing conditions.

These concerns are slightly mitigated by the fact that the mental health variables are indices constructed from answers given to several questions (see section 3 for a more in-depth discussion of the construction of our variables). The indices were designed to be more objective. Also, participants are less likely to lie on several questions. They are also not, to the best of our knowledge, aware of how the questions will be used in the study. Finally, it is just as likely that any reporting bias goes in the same direction for all groups. For instance, it could be that participants in the Control and VET-only groups exaggerate their mental health variables to receive extra aid. If participants in VET+YE also exaggerated their mental health variables, then reporting bias may be less of a concern. Overall, it is difficult to test for such biases ex-post, but

the above reasons should assuage some concerns.

4.2 COVID-19 and mental health in the short run: regression results

This section presents the main results of the analysis using regression estimates. Regression analysis allows us to study the effect of the pandemic across the three different groups while i) controlling for baseline characteristics increasing the precision of the estimates; ii) accounting for individual fixed effects, thus keeping fixed all characteristics that might vary at the individual level.

We propose two baseline specifications. Both account for demographic characteristics that are unbalanced at baseline. The first specification also accounts for the lagged value of the dependent variable - that is to say, it controls for psychological well-being measured at baseline. More formally, our first baseline specification is the following:

$$y_{i,g,t} = \alpha + \theta_1[VET + YE]_g + \theta_2 VET_g + \gamma_1 \times y_{i,g,t-1} + \gamma_2 \mathbf{X_{i,g}} + \varepsilon_{i,g} \tag{5}$$

where $y_{i,g,t}$ measures the level of depression, distress, and PTSD severity for individual i at the time of the lock-down survey. $[VET + YE]_g$ takes on the value of 1 for all participants belonging to the group that received both vocational education and youth empowerment, $[VET]_g$ takes on the value of 1 for all participants belonging to the group that received only vocational education, and the omitted category refers to the control group. $y_{i,g,t-1}$ is the lagged value of the dependent variable, depression, distress, and PTSD severity, for individual i measured during the baseline survey. $\mathbf{X_{i,g}}$ includes a set of demographic controls, namely age, dummies for gender, displaced individuals, number of children, work experience, earnings, and formal employment at baseline.

The second baseline specification accounts for individual fixed effects by looking at the change in the outcome variable between the two survey periods. More concretely, we estimate:

$$\Delta y_{i,g} = \alpha + \theta_1[VET + YE]_g + \theta_2 VET_g + \gamma_2 \mathbf{X_{i,g}} + \varepsilon_{i,g} \tag{6}$$

where $\Delta y_{i,g}$ measures the change in the level of depression, distress, and PTSD severity for individual i that materialized between the lockdown survey and the baseline survey. It can be shown that this specification, when ignoring the individual level controls, corresponds to a two-way fixed effects (individual and time) specification, where we interact a post-COVID variable with the treatment group $[VET + YE]_g$ and VET_g. The formulation we propose in equation 6, including the set of demographic controls $\mathbf{X_{i,g}}$, allows us to also control for the possible time-varying effect of baseline demographic characteristics. Robust standard errors are clustered at

the individual level.

Table 6 presents baseline results. The dependent variable is the depression z-score index in the first two columns, the distress z-score index in the third and fourth columns, and the PTSD severity index in the fifth and sixth columns. In odd columns, we report results for specification 5, and in even columns for specification 6.

Table 6: **Mental health impacts of the COVID-19 Pandemic across different intervention groups**

	Mental health impacts of the COVID-19 pandemic					
	Depression		Distress		PTSD Severity	
	Post-COVID19	Δ Post-Pre	Post-COVID19	Δ Post-Pre	Post-COVID19	Δ Post-Pre
VET + YE	-0.400***	-0.321*	-0.261*	-0.201	-0.168	-0.089
	(0.138)	(0.176)	(0.156)	(0.210)	(0.141)	(0.152)
VET	-0.325**	-0.291	-0.370**	-0.138	-0.147	-0.101
	(0.153)	(0.191)	(0.173)	(0.230)	(0.150)	(0.157)
Lagged Dependent Variable	Yes	NA	Yes	NA	Yes	NA
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	269	269	180	180	276	276
R-squared	.194	.0633	.219	.0407	.254	.0699

Notes: The dependent variables are the indexes of Depression, Distress and PTSD severity. In even columns the index is measured in July 2020, at the end of the national lockdown. In odd columns, the dependent variable measures the difference between the index measured in July 2020 and January 2019. VET + YE indicates the group of participants who received vocational education and soft skill training, while VET indicates the group of participants that received only vocational education, the omitted category is the control group. The Lagged Dependent Variable is the index measured at baseline, on January 2019. Individual controls include age, dummies for gender, displaced individuals, number of children, work experience, earnings and formal employment at baseline. Robust standard errors clustered at the individual level are reported in round brackets. ***, **, and * indicate significance at the 1-, 5-, and 10-% levels, respectively.

All estimates of the coefficients $[VET + YE]_g$ and VET_g are negative for all the indexes examined. Coefficients for distress and PTSD severity tend to be less precisely estimated. In terms of magnitude, participants that received vocational education and youth empowerment before the pandemic tended to report levels of depression 0.32 standard deviations lower than individuals in the control group (column 2). The coefficients for distress and PTSD severity provide qualitatively similar insights, as they are negative and similar in magnitude, but not precisely estimated. Interestingly, also individuals that received only vocational education before the pandemic tend to have a lower level of depression, distress, and PTSD severity. While the coefficients for VET_g only tend to be smaller than the coefficients for $[VET + YE]_g$, the difference is not statistically significant, as shown in table 7.

Robustness. Robustness tables are reported in the Supplementary Material. In Table B3 we compare individuals in the group that received $[VET + YE]_g$ with individuals in the control group via a split sample, while in Table B4 we compare individuals in the group that received $[VET]_g$ with individuals in the control group via a split sample. Results are robust to these

Table 7: **Mental Health Impacts of COVID-19 Pandemic across the two Treated Groups**

	Vocational Education versus Vocational Education with Youth Empowerment					
	Depression		Distress		PTSD Severity	
	Post-COVID19	Δ Post-Pre	Post-COVID19	Δ Post-Pre	Post-COVID19	Δ Post-Pre
VET + YE	-0.063	-0.024	0.103	-0.052	-0.054	-0.028
	(0.144)	(0.177)	(0.193)	(0.228)	(0.147)	(0.152)
Lagged Dependent Variable	Yes	NA	Yes	NA	Yes	NA
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	180	180	117	117	184	184
R-squared	.222	.0732	.211	.0373	.257	.0383

Notes: The sample include participants who received vocational education and soft skill training, and participants that received only vocational education. The dependent variables are the indexes of Depression, Distress and PTSD severity. In even columns the index is measured in July 2020, at the end of the national lockdown. In odd columns, the dependent variable measures the difference between the index measured in July 2020 and January 2019. VET + YE indicates the group of participants who received vocational education and soft skill training, the omitted category is the group that received only vocational education. The Lagged Dependent Variable is the index measured at baseline, on January 2019. Individual controls include age, dummies for gender, displaced individuals, number of children, work experience, earnings and formal employment at baseline. Robust standard errors clustered at the individual level are reported in round brackets. ***, **, and * indicate significance at the 1-, 5-, and 10-% levels, respectively.

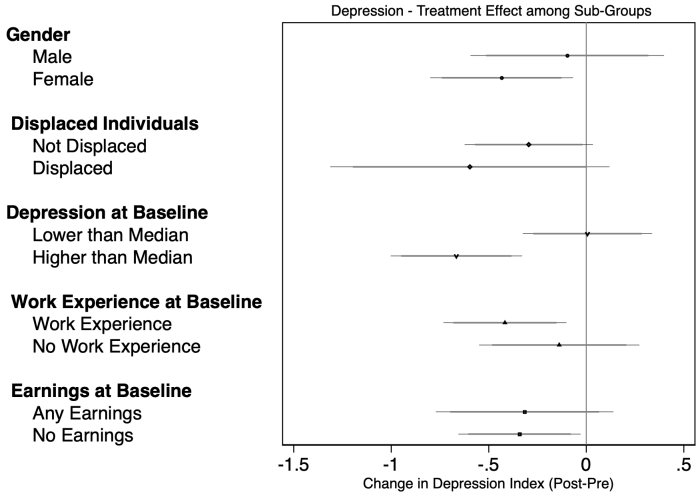
alternative specifications.

Heterogenous Effects. Next, we explore whether the effect of treatment was larger for more vulnerable sub-groups of participants. We explore several dimensions of marginalization and vulnerability: i) demographic vulnerability, focusing on females and displaced individuals; ii) psychological vulnerability, prior mental health conditions measured at baseline; and iii) labor market marginalization, in terms of lack of work experience and earnings at baseline. For simplicity, we bundle individuals in the group $[VET + YE]_g$ and group $[VET]_g$, considering both as treated.

Figure 4 summarizes heterogeneous effects on depression, while we present results on distress and PTSD Severity, which are qualitatively providing the same insights, in the Supplementary Material, section B.1.4 (see tables B9 and B9). Further heterogeneity analyses, by gender and median decomposition, can be found in the same section.

The effect of treatment in mitigating negative mental health consequences of the pandemic is stronger when looking at all dimensions of vulnerability. First, female and displaced individuals tend to benefit more from vocational education. Similarly, individuals with little or no attachment to the labor market tend to benefit more from vocational education. Importantly, individuals that benefit more from the programs are the ones with a higher baseline level of depression: the negative effect on their post-COVID levels of depression is entirely driven by this group.

Figure 4: **Heterogeneous Effect of Treatment Among Sub-Groups of Participants**



Notes: The figure illustrates the effect of treatment among sub-groups of participants on depression levels during COVID-19. Treatment is defined as both VET + YE and VET, the omitted category is the control group. The dependent variable measures the difference between the depression index measured in August 2020 and January 2019. Regressions account for individual controls, including age, dummies for gender, displaced individuals, number of children, work experience, earnings, and formal employment at baseline. The center indicates the coefficient value. Lines indicate 90% (thick lines) and 95% (thin lines) confidence intervals.

5 Main results in the medium run

In the medium run, we need to expand our model to allow for dynamic effects. In an ideal world, a simple difference in means from the randomized controlled trial design would effectively allow us to estimate the causal average effect of treatment on our outcomes over time. In theory, this would be the case because the randomized controlled trial creates a scenario in which the control group serves as a counterfactual group. This arises because participants who are similar to each other - they come from a similar background, live in the same neighborhood, and have similar levels of education - are randomly assigned to one of each treatment group. This design, in theory, would imply that the difference in means between treatment and control groups is the causal average effect of treatment. However, several caveats are in order before we can establish the actual causal effect of treatment.

However, there are some variables, such as gender and displacement status, which were not balanced across the three groups at the baseline. This implies that there could be some bias in the above numbers as a result. Furthermore, not everyone complied with the treatment - that is to say that some participants who were initially assigned to receive treatment dropped out for whatever reason. This can be mitigated by using the intent to treat it as an estimate. This effectively estimates the effect of treatment, weighted by the proportion of treatment each participant received. Thirdly, it is possible that at each period when data was collected, particular

issues are affecting participants who were in different groups different. Such unobserved time period-specific factors can be controlled by including time trends in our analysis. This averages out any time-varying unobserved variables that may be biasing the estimates.

We, therefore, follow Sun and Abraham (2021) in estimating a linear two-way fixed effects specification, using individual and time fixed effects. To explore the treatment effect of VET+YE, we estimate equation 7:

$$Y_{i,t} = \alpha_i + \lambda_t + \sum_{t=E_1}^{\bar{T}} \mu_l \mathbf{1}\{t - E_1 = l\} \times [\text{VET+YE}_i] + \nu_{i,t} \tag{7}$$

Where Y is the outcome variable of interest in time period t for individual i. The outcomes of interest are overall earnings, informal earnings, and formal earnings. λ_t indicates time fixed effects, α_i indicates individual fixed effects, E_1 is the time of treatment (February 2019), $l = t - E_1$ are relative time periods since treatment, and $\bar{T}$ is the last survey period. In particular, $l = t - E_1 = 0$ if t is at baseline, and t is measured in quarters. The coefficient of interest is μ_l. VET+YE$_i$ is equal to one if a participant is in the VET+YE treatment group. When estimating the treatment effect of VET+YE compared to the Control group, we run equation 7 on the sample of participants in VET+YE and the Control group so that the reference group is the Control group. When estimating the added value of VET+YE, we run equation 7 on the sample of participants in VET+YE and VET-only.

To explore the treatment effect of VET-only, we run equation 8 on the sample of participants in VET and the Control group, where VET$_i$ is equal to one if a participant is in the VET treatment group:

$$Y_{i,t} = \alpha_i + \lambda_t + \sum_{t=E_1}^{\bar{T}} \mu_l \mathbf{1}\{t - E_1 = l\} \times [\text{VET}_i] + \nu_{i,t} \tag{8}$$

Given the relatively strong assumption of no heterogeneous effects on treatment in Sun and Abraham (2021), we check for robustness of results by estimating the regression following other specifications. Changing specifications does not significantly change results. These are reported in section 7.

5.1 Descriptive evidence

Figure 5 shows how mental health has evolved throughout the period of study, from December 2018 to August 2020 when the COVID-19 surveys were conducted via phone, to November 2020 when the midline survey was conducted in person, and finally to November 2021 when the endline surveys were administered in person. We track the variables associated with depression, distress, and trauma.

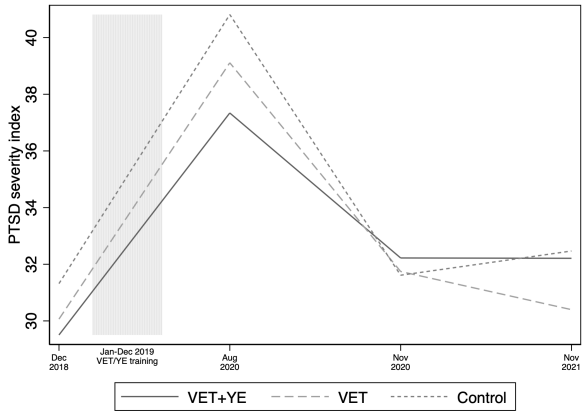
First, figure 5 shows that while there are some differences at baseline between each group, on average these differences seem to be insignificant, statistically speaking. This is confirmed with a simple t-test of the difference in means between each group in December 2018. The average value of the trauma severity is around 30.25, which is above the threshold of 30, which is considered severe. Depression, which is a variable between 0 and 18, is on average 4.25, which is low, and distress, which is a variable between 0 and 36, is on average 9.75.

Second, figure 5 shows that the indexes of mental illness dissipate after the lockdown is lifted. By November 2020, the values for trauma severity (panel a), depression (panel b), and distress (panel c) are similar to, though still slightly higher than, what they were at baseline. In particular, the average value of the trauma index is 31.8, average depression is 4.6, and average distress is 10.02.

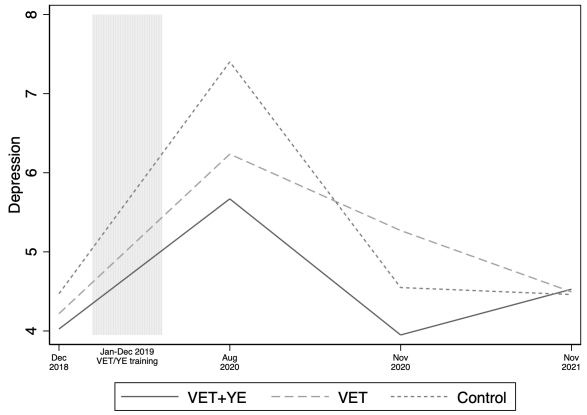
Curiously, the values for depression and distress are highest for participants who only received VET (the long dashed line in each panel). Participants who receive VET only score, on average, 1.3 points higher in the depression index than participants who received VET+YE, statistically significantly so. Statistical significance is confirmed by a simple t-test of the difference in means. Trauma severity, depression, and distress seem to hover around similar values one year later, in November of 2021, though depression is no longer statistically significantly different between groups.

Figure 5: **Mental health symptoms over time**

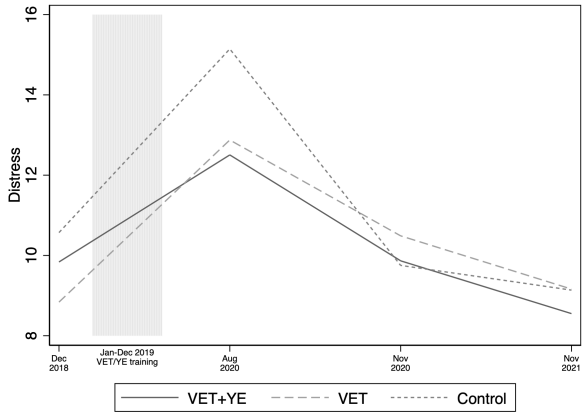
(a) PTSD severity



(b) Depression



(c) Distress



Notes: The figure illustrates average indices, by group, of the PCL-C severity index (the indicator of trauma that takes values between 17-85, panel a), depression (panel b), and distress (panel c), the latter two measured according to Blattman and Annan (2015), over time. Solid lines show values for VET+YE, long dashes are for VET only, and short dashes are for the control group.

5.2 Average treatment effects

The graphs in figures 6, 7, and 8 plot the effect of VET+YE and VET on indices of mental illness over time, in particular trauma, depression, and distress.

5.2.1 Interpreting the coefficients:

To interpret the coefficients, it is important to first understand how each variable is computed. The indices of depression, distress, and trauma are each vectors between 0 and a number of a different length, which makes it difficult to interpret differences between the effects of training on these indices. To be able to compare depression, distress, and trauma, we create indices of each of these three variables via standardization, which is a way of putting everything onto the same scale. A standard score represents how many standard deviations over or below the mean a person falls.

Let us illustrate this with an example: if a person is associated with a standardized depression value of 1.5, this indicates that they are more depressed than the average person in the sample by 1.5 standard deviations.

This is also the logic used to interpret the coefficients in figures 6, 7, and 8 below. The coefficients in panels 6a, 7a, and 8a stand for the effect of VET+YE on mental health, compared to the control group. The coefficients in panels 6b, 7b, and 8b stand for the effect of VET on mental health, compared to the control group. The coefficients in panels 6c, 7c, and 8c stand for the effect of VET+YE on mental health, compared to VET only (the so-called value-added of VET+YE).

In this light, interpreting the effects of treatment on each mental health index becomes simpler. Seeing a coefficient of x at time t for a treatment g would therefore imply the following interpretation: at time t, receiving treatment g is associated with a change of x standard deviations of the index of depression/distress/trauma compared to the control group. A negative coefficient would imply that the treatment is associated with less mental illness (and therefore greater resilience).

5.2.2 Depression

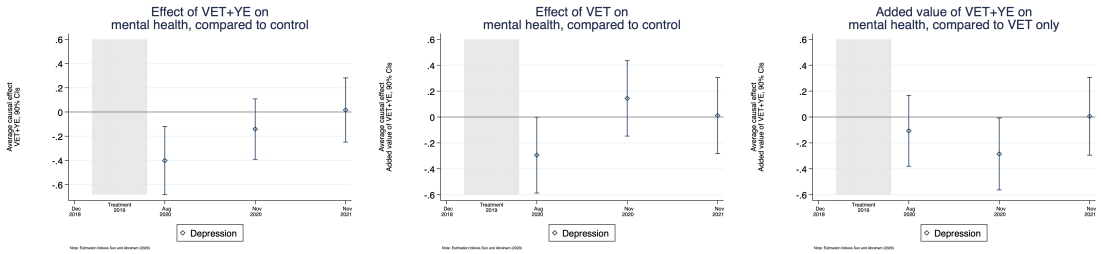
Figure 6, panels 6a and 6b, shows that vocational education and youth empowerment reduced depression compared to participants in the control group during the COVID-19 lockdowns. Specifically, VET+YE reduced depression by 0.4 standard deviations compared to the control group during the COVID-19 lockdowns; and VET only reduced depression by 0.29 standard deviations compared to the control group during the COVID-19 lockdowns. This is consistent

with the short run results. After the pandemic, the lines bounce back to the zero line. The cushioning effect of vocational education does not seem to be long-lasting.

Panel 6c confirms the intuition that the VET+YE training provides a stronger, or perhaps longer-lasting, cushion against depression from the pandemic. In November 2020, it appears that participants in VET+YE have almost 0.28 standard deviations lower depression compared to participants in VET-only. We know from figure 5 that by November 2020 symptoms of depression had dramatically reduced, indicating a recovery of mental health throughout the entire sample, though it seems it stayed higher for participants in VET-only.

Figure 6: **Average effect of treatment on the treated on depression over time**

(a) Effect of VET+YE (b) Effect of VET only (c) Added value of VET+YE compared to VET-only



Notes: The graphs illustrate the average causal effect of treatment on depression over time. Panel a) plots the effect of VET+YE on depression compared to the control group over time; panel b) plots the effect of VET only on depression compared to the control group over time; and panel c) plots the effect of VET+YE compared to VET only over time. Vertical lines show the 90% confidence intervals.

5.2.3 Distress

Figure 7, panels 7a and 7b, show that vocational education and youth empowerment reduced distress compared to participants in the control group during the COVID-19 lockdowns, though not statistically significantly so.

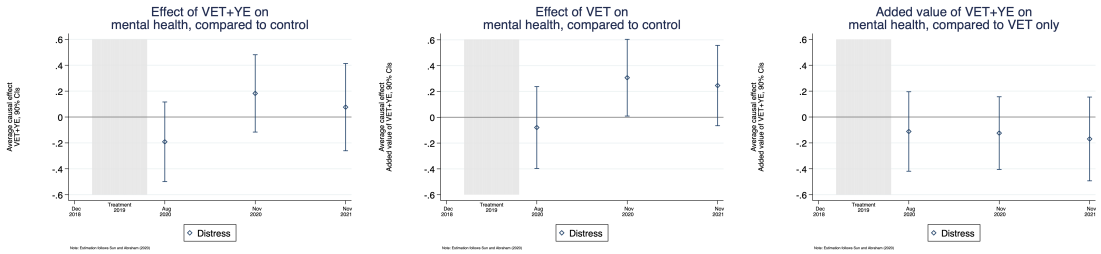
After the pandemic, there are some differences in terms of how participants in VET-only did compared to Control participants (panel 7b). Participants who received only vocational education are associated with an increase of distress equivalent to 0.3 standard deviations compared to the control group in November 2020: they seem to have higher symptoms of distress compared to the control group. This value remains almost the same until the next year, though it is estimated less precisely in November 2021. It is unclear why participants who receive only vocational education might experience higher distress compared to participants in the control group, though it is possible that vocational education only does not necessarily always provide cushioning against mental illness in normal times. This pattern is not present when comparing

VET+YE to the control group (panel 7a).

Second, panel 7c shows participants who received vocational education and psycho-social training are associated with lower distress than vocational education only, even though this is not a statistically significant result. It is, however, consistent in November 2020 and November 2021. Combined with the previous story, it is possible that psycho-social/soft skills training provides young adults with some longer protection against mental illness in the medium term. However, this statement should be taken with a degree of caution since the variables are not statistically significant

Figure 7: **Average effect of treatment on the treated on distress over time**

(a) Effect of VET+YE (b) Effect of VET only (c) Added value of VET+YE compared to VET-only



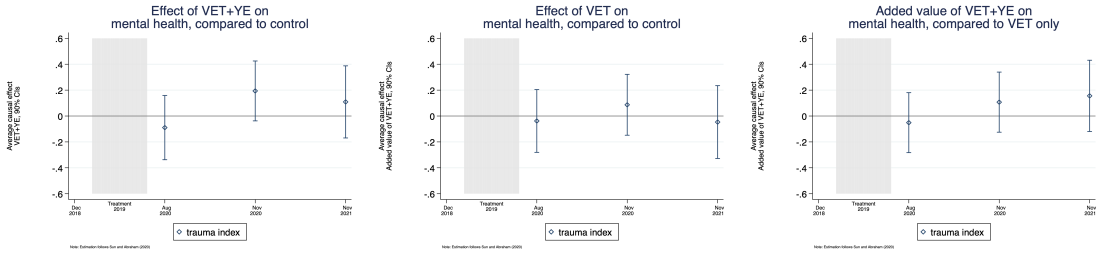
Notes: The graphs illustrate the average causal effect of treatment on distress over time. Panel a) plots the effect of VET+YE on distress compared to the control group over time; panel b) plots the effect of VET only on distress compared to the control group over time; and panel c) plots the effect of VET+YE compared to VET only over time. Vertical lines show the 10% confidence intervals. The capped lines show the 90% confidence intervals.

5.2.4 Trauma

We saw in figure 5 that trauma symptoms increased for everyone in the sample, and figure 8 confirms that the increase in trauma is similar for participants in all groups. Neither vocational education nor youth empowerment training cushions participants against increases in trauma symptoms during the pandemic.

Figure 8: **Average effect of treatment on the treated on trauma over time**

(a) Effect of VET+YE (b) Effect of VET only (c) Added value of VET+YE compared to VET-only



Notes: The graphs illustrate the average causal effect of treatment on trauma over time. Panel a) plots the effect of VET+YE on trauma compared to the control group over time; panel b) plots the effect of VET only on trauma compared to the control group over time; and panel c) plots the effect of VET+YE compared to VET only over time. Vertical lines show the 10% confidence intervals. The capped lines show the 90% confidence intervals.

5.2.5 General remarks

Table 8: **Mental health impacts of the COVID-19 pandemic across different intervention groups in the medium run**

	Depression		Distress		Trauma	
	VET+YE	VET	VET+YE	VET	VET+YE	VET
Aug 2020	-0.401**	-0.295*	-0.191	-0.0795	0.0119	-0.0903
	(0.171)	(0.178)	(0.187)	(0.193)	(0.193)	(0.186)
Nov 2020	-0.141	0.144	0.182	0.307*	0.215	-0.0300
	(0.152)	(0.177)	(0.182)	(0.181)	(0.147)	(0.147)
Nov 2021	0.0162	0.0110	0.0764	0.246	0.107	-0.0655
	(0.161)	(0.179)	(0.205)	(0.189)	(0.166)	(0.167)
Observations	738	705	671	632	751	725

Notes: The dependent variables are the indexes of depression, distress, and trauma severity. In odd columns are the estimates for VET-only compared to Control. VET + YE indicates the group of participants who received vocational education and youth empowerment, while VET indicates the group of participants that received only vocational education, the omitted category is the control group in both cases. Individual controls include age, dummies for gender, displaced individuals, number of children, work experience, earnings, and formal employment at baseline. Robust standard errors clustered at the individual level are reported in round brackets. ***, **, and * indicate significance at the 1-, 5-, and 10-% levels, respectively.

Taken together, the overall conclusion is that vocational education seems to have effects in protection against mental illness during times of crisis. Tables 8 and 9 provide a summary of the above figures: table 8 shows the mental health effects of the COVID-19 pandemic of VET+YE and VET participants compared to Control group participants, and table 9 shows the

added value of VET+YE compared to VET only. More striking, perhaps, is what happened after the COVID-19 lockdowns were lifted. Figure 5 shows that by November 2020 symptoms of depression, distress, and trauma had dramatically reduced, indicating a recovery of mental health throughout the entire sample.

What might be causing these results? First, and most cynically, participants in the VET+YE and VET groups might have planned for a transition period during the end of training and finding a first job. It is possible that they are less concerned by the lockdowns as they may have planned for a period of unemployment regardless of the crisis.

Second, vocational education in and of itself may have provided people with the skills to handle crises. The youth empowerment program might just have stronger cushioning effects compared to the general population, but it does not provide quantifiable differences compared to only vocational education. Both VET-only and VET+YE groups seem to be doing similarly, and thus it is something about vocational education itself that helped cushion mental health stress during the COVID-19 pandemic.

Table 9: **Added value of VET+YE: Mental health impacts of the COVID-19 pandemic in the medium run**

	Depression	Distress	Trauma
Aug 2020	-0.107	-0.112	0.102
	(0.166)	(0.187)	(0.177)
Nov 2020	-0.285*	-0.124	0.245
	(0.169)	(0.171)	(0.151)
Nov 2021	0.00527	-0.169	0.173
	(0.183)	(0.197)	(0.162)
Observations	715	639	732

Notes: The dependent variables are the indexes of depression, distress, and trauma severity. The treatment is the group of participants who received vocational education and youth empowerment. The omitted category is the VET-only group. Individual controls include age, dummies for gender, displaced individuals, number of children, work experience, earnings, and formal employment at baseline. Robust standard errors clustered at the individual level are reported in round brackets. ***, **, and * indicate significance at the 1-, 5-, and 10-% levels, respectively.

6 Heterogeneity and robustness of medium run results

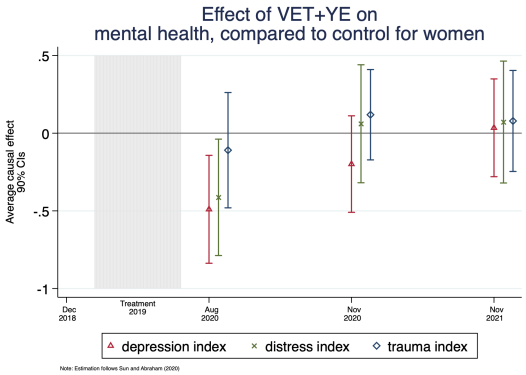
6.1 Heterogeneity in the medium run

6.1.1 By gender

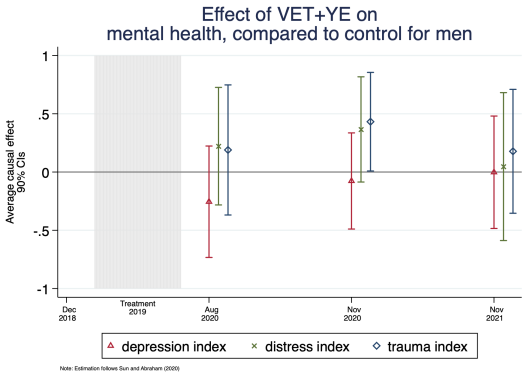
The figures in 9 show the heterogeneous effects of treatment by gender on mental health in the medium run. We do a sample censor by gender and estimate the two-way fixed effects according to Sun and Abrahams (2020). This corroborates the short term results we saw earlier. During the COVID-19 lockdowns in August 2020, women who went through vocational education (both treatment groups) were more resilient to the shocks to mental health that the lockdowns generated (see panels 9a and 9c). For men, this cushioning effect disappears (see panels 9b and 9d). When it comes to symptoms of trauma, what is interesting is that women who went through youth empowerment have a spike in trauma symptoms compared to participants who went through vocational education only 9f in November, 2020, shortly after the lockdown ended. Men in VET+YE do not have such a spike compared to men who only went through VET.

Figure 9: **Average effect of treatment on the treated on mental health over time for women and men**

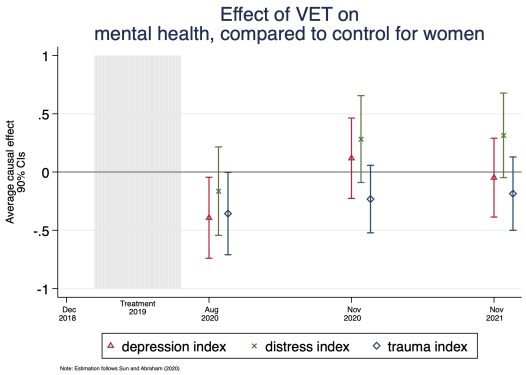
(a) Effect of VET+YE on mental health



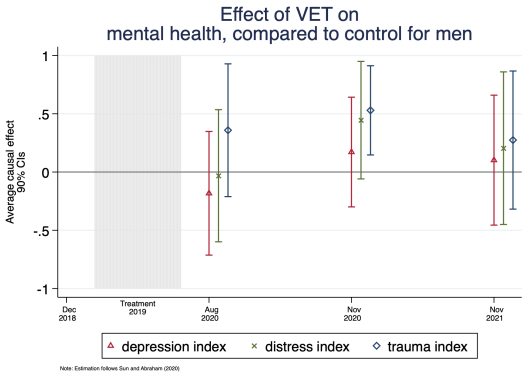
(b) Effect of VET+YE on mental health



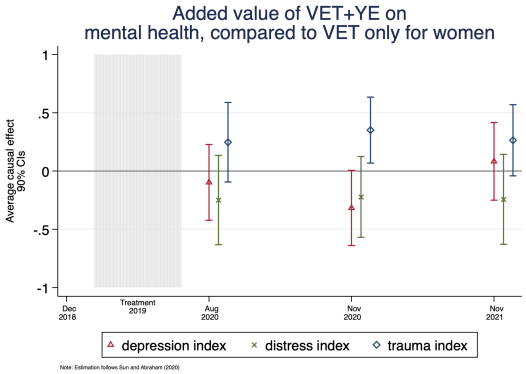
(c) Effect of VET only on mental health



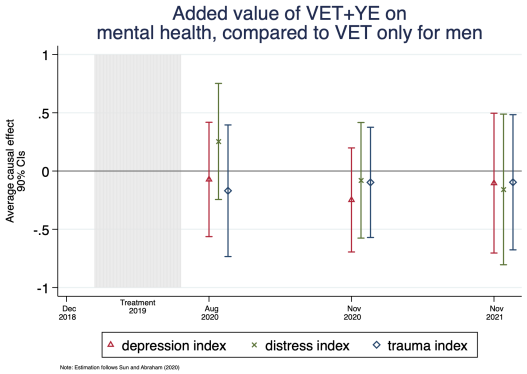
(d) Effect of VET only on mental health



(e) Added value of VET+YE on mental health



(f) Added value of VET+YE on mental health



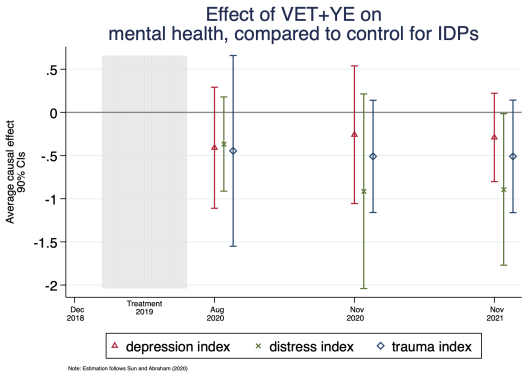
Notes: The graphs illustrate the average causal effect of treatment on mental health over time for women: the trauma index in diamonds, depression index in triangles, and distress in "x"es. Panels a) and b) plot the effect of VET+YE on these three variables compared to the control group over time; panels c) and d) plot the effect of VET only on these three variables compared to the control group over time; and panel e) and f) plot the effect of VET+YE compared to VET only over time. Panels a), c), and e) plot the effect on women; panels b), d), and f) plot the effects on men. The capped lines show the 90% confidence intervals.

6.1.2 By traumatic events: being displaced

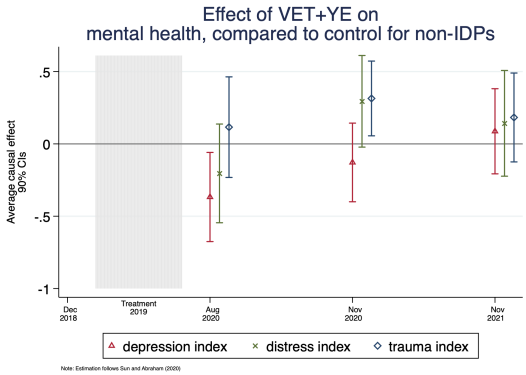
Figure 10 shows the effect of treatment on internally displaced people. Because the number of internally displaced people is low (around 30), these numbers should be taken with a grain of salt. However, it seems there is not a huge effect difference in terms of the effect of treatment on being internally displaced. There does seem to be some indication that people who are not internally displaced have less resilience to the opening of the economy post-COVID in November 2021 - see panels 10b, 10d, and 10f. It is possible that people who have been displaced have some resilience that help them benefit from shocks to mental health.

Figure 10: **Average effect of treatment on the treated on mental health over time for IDPs**

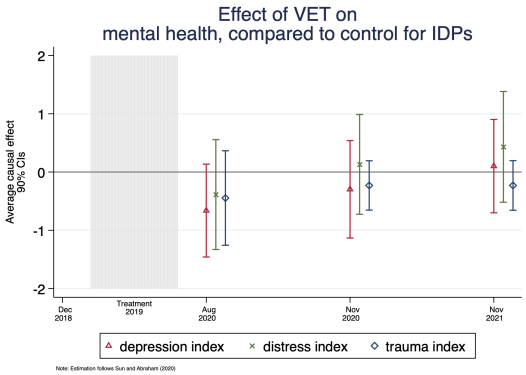
(a) Effect of VET+YE on mental health



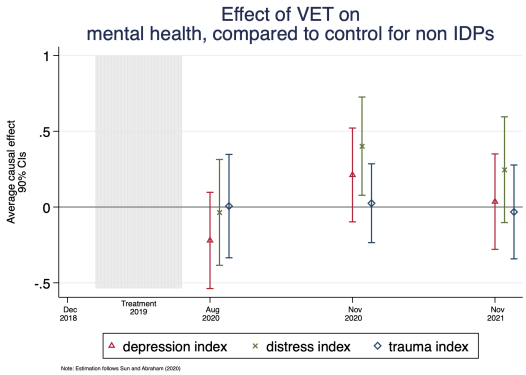
(b) Effect of VET+YE on mental health



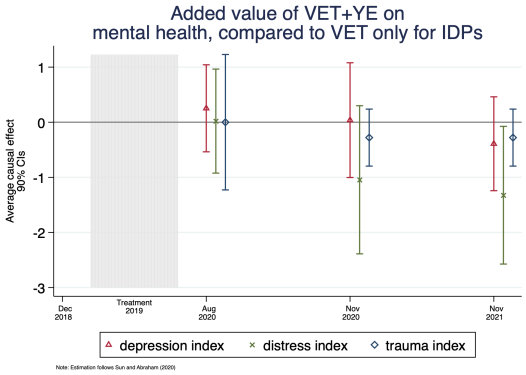
(c) Effect of VET only on mental health



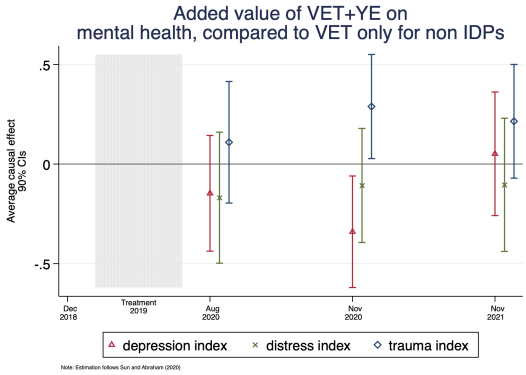
(d) Effect of VET only on mental health



(e) Added value of VET+YE on mental health



(f) Added value of VET+YE on mental health



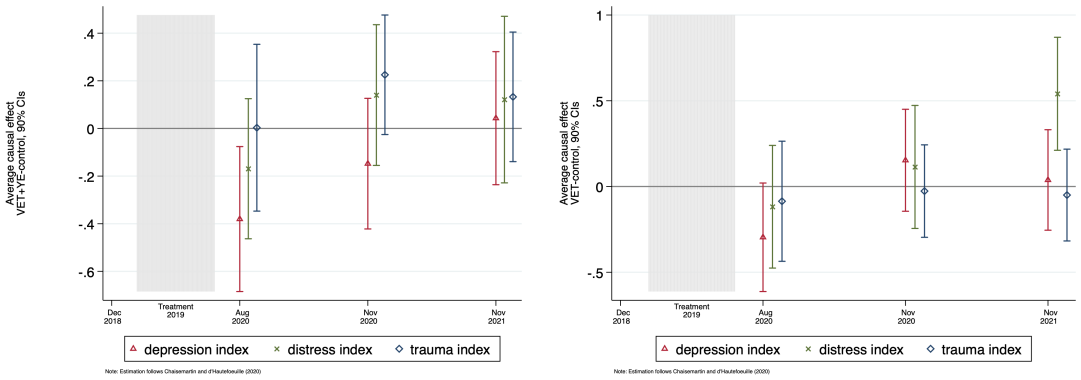
Notes: The graphs illustrate the average causal effect of treatment on mental health over time for women: the trauma index in diamonds, depression index in triangles, and distress in "x"es. Panels a) and b) plot the effect of VET+YE on these three variables compared to the control group over time; panels c) and d) plot the effect of VET only on these three variables compared to the control group over time; and panel e) and f) plot the effect of VET+YE compared to VET only over time. Panels a), c), and e) plot the effect on IDPs; panels b), d), and f) plot the effects on people who are not internally displaced. The capped lines show the 90% confidence intervals.

6.2 Robustness, medium run

Alternative specifications The event study graphs were created following Sun and Abrahams (2020), which hold only the assumptions of parallel trends and no anticipatory behavior (which are standard for differences-in-differences), but also for the more strict assumption of treatment effect homogeneity - or, more precisely, that "each cohort experiences the same path of treatment effects". For this reason, we check the results using the more flexible specification of Chaisemartin and d'Hautefoeuille (2020). Figure 11 runs the decomposition proposed by Chaisemartin and d'Hautefoeuille (2020) for mental health and labor market outcomes. The results do not change significantly.

Figure 11: **Average effect of treatment on the treated on mental health over time**

(a) Effect of VET+YE on mental health (b) Effect of VET only on mental health



Notes: The first two graphs illustrate the average causal effect of treatment on mental health over time: the trauma index in diamonds, depression index in triangles, and distress in "x"es. Panel a) plots the effect of VET+YE on these three variables compared to the control group over time; panel b) plots the effect of VET only on these three variables compared to the control group over time. The capped lines show the 90% confidence intervals.

6.2.1 The labor market value of the effect of treatment on mental health

The effect of VET+YE on mental health can be quantified in terms of earnings when complementing the results above with self-reported data on earnings. In this section, do this in the following way. First, we have estimated the effect of treatment on mental health for August 2020, November 2020, and November 2021 section 5. We can then run simple Ordinary Least Squares regressions of the index of each mental health variable on earnings for each of these time periods to get the correlation between mental health and earnings at each period. This is described in equation 9, where i indicates individuals and t indicates time periods. Multiplying these two figures provides an estimate of the dollar value of the effect of VET+YE on mental health. These are reported in table 10.

$$earnings_{it} = mentalhealth_{it} + \varepsilon_{it} \tag{9}$$

These effects are simply descriptive and do not imply any sort of causal direction. First, this is not a methodology used in other papers and mental health could be promoting earnings but it could be providing other effects. Second, the direction of the effects could go both ways: the effects on income could be driven by better mental health, but better mental health might also be a result of a better income. The purpose of this breakdown is not to disentangle any causal effects but rather just to describe how mental health and earnings relate to each other in this cost-benefit analysis.

Table **??** shows the results of these estimates. We could look at each of these three mental health variables in turn, and then average over all three of them to get an average value of the benefit of VET+YE on mental health. However, the only value that is statistically significant is the effect of VET+YE on depression in August 2020 (see discussion in section 5 for more details). Therefore, it might be safe to say that overall, the value of VET+YE on mental health might be insignificant, except for in times of crisis such as the COVID-19 pandemic. Then, VET+YE is associated with a $3.84 gain in August 2020, compared to participants in the Control group.

Table 10: **Value of effect of training on depression**

	treatment effect	**earnings corr.**	**value of training**
Depression			
Aug 2020	-0.406**	-9.452	$ 3.835
	(0.173)	(5.506)	
Nov 2020	-0.143	-18.810*	$ 2.688
	(0.153)	(7.824)	
Nov 2021	0.016	-21.640*	$ -0.355
	(0.163)	(8.430)	
Distress			
Aug 2020	-0.211	-8.639	$ 1.821
	(0.187)	(5.427)	
Nov 2020	0.154	-15.330	$ -2.364
	(0.181)	(7.924)	
Nov 2021	0.059	-17.750	$ -1.050
	(0.203)	(9.207)	
Trauma			
Aug 2020	-0.095	-1.881	$ 0.178
	(0.153)	(5.686)	
Nov 2020	0.192	-7.201	$ -1.385
	(0.142)	(7.342)	
Nov 2021	0.106	-9.974	$ -1.060
	(0.171)	(7.115)	

Notes: The first column shows the causal effect of treatment on mental health. The second column shows OLS regressions of the index of each mental health variable on earnings for each time periods. The third column shows an estimate of the dollar value of the effect of VET+YE on mental health. Standard errors clustered at the individual level are reported in round brackets. ***, **, and * indicate significance at the 1-, 5-, and 10-% levels, respectively.

7 Conclusion

The results presented in this paper document that vocational education coupled with a youth empowerment program increases the resilience of individuals in the face of extreme shocks such as the COVID-19 pandemic. All participants suffered significantly during the pandemic, but participants who received training suffered less compared to participants in the control group. Results indicate that the treatment programs were particularly helpful for more vulnerable sub-groups of individuals, such as women, displaced individuals, participants with little prior attachment to the labor market, and individuals with pre-existing higher levels of depression. In the medium run, mental health bounces back to baseline values. Treatment no longer seems to affect mental health after the end of the pandemic. There seems to be a strong, acute, cushioning effect of treatment.

Vocational education in and of itself seems to be the force behind the cushion for excess

stress during the crisis, even if it is not a permanent cushion since the lines bounce back to the zero line. What is quite surprising is that, amid this great crisis, there is a greater resilience of the participants in both treatment groups (VET+YE and VET). They are experiencing symptoms similar to others but these tend to be less severe.

Future research is needed to investigate how crucial the youth empowerment component of the program is for the effects we estimate. The limited power of our study does not allow us to fully assess the difference between the treatment for individuals receiving only vocational education and the bundle of vocational education and youth empowerment. Moreover, future research can shed further light on the psychological transformation enchained by program participation, identifying the drivers of higher resilience.

B.1 Supplementary material

This Appendix accompanies the paper "Adding youth empowerment to vocational education provides greater resilience to shocks among vulnerable young adults: evidence from the COVID-19 pandemic in Colombia".

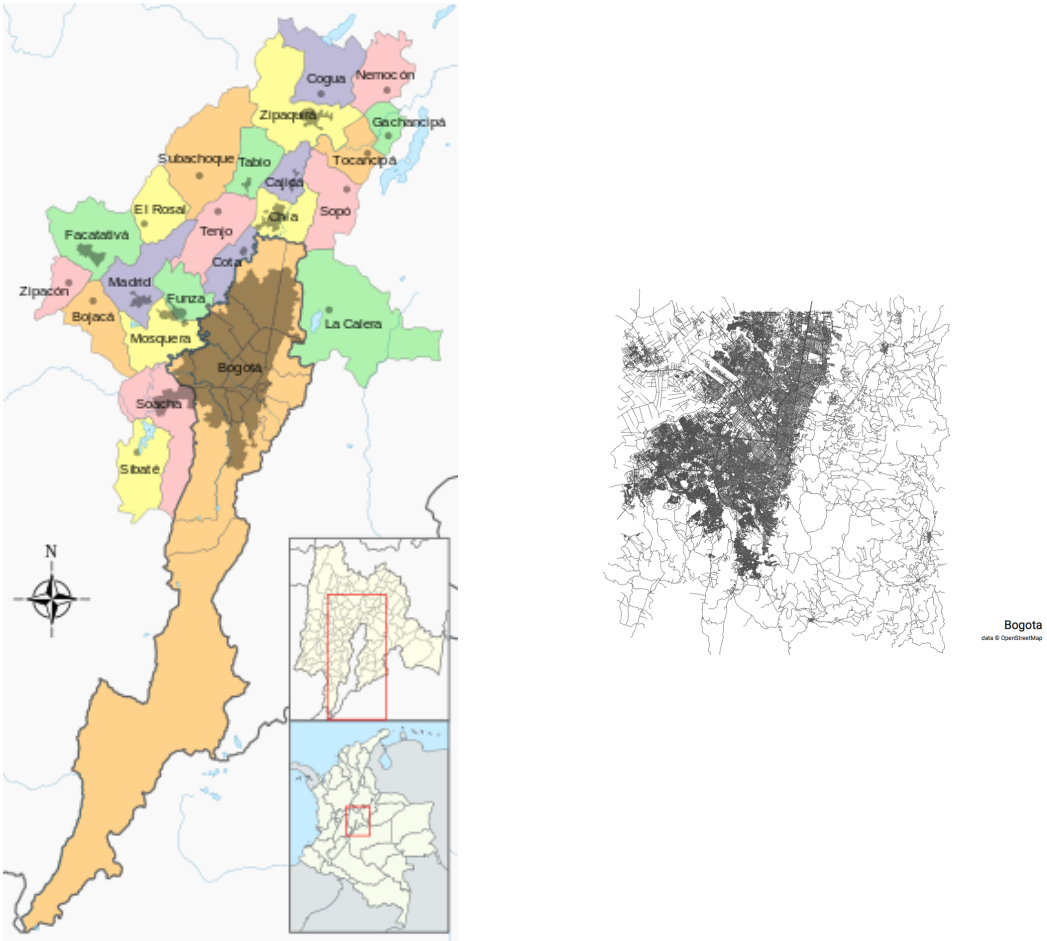
B.1.1 The intervention

B.1.2 The context

B.1.2.1 Bosa And Soacha

The study focuses on two lower-income suburbs of Bogotá: Bosa and Soacha. In the study area, around one-third of the population is eligible for social assistance (e.g. they are ranked SISBEN level 2 or lower), and most people stop studying after high school (DANE, 2019). Furthermore, a large share of the country's internally displaced population is concentrated in the area. At 6 million people, Colombia has one of the highest populations of internally displaced people in the world - a tangible consequence of the fifty-year conflict (DANE, 2019).

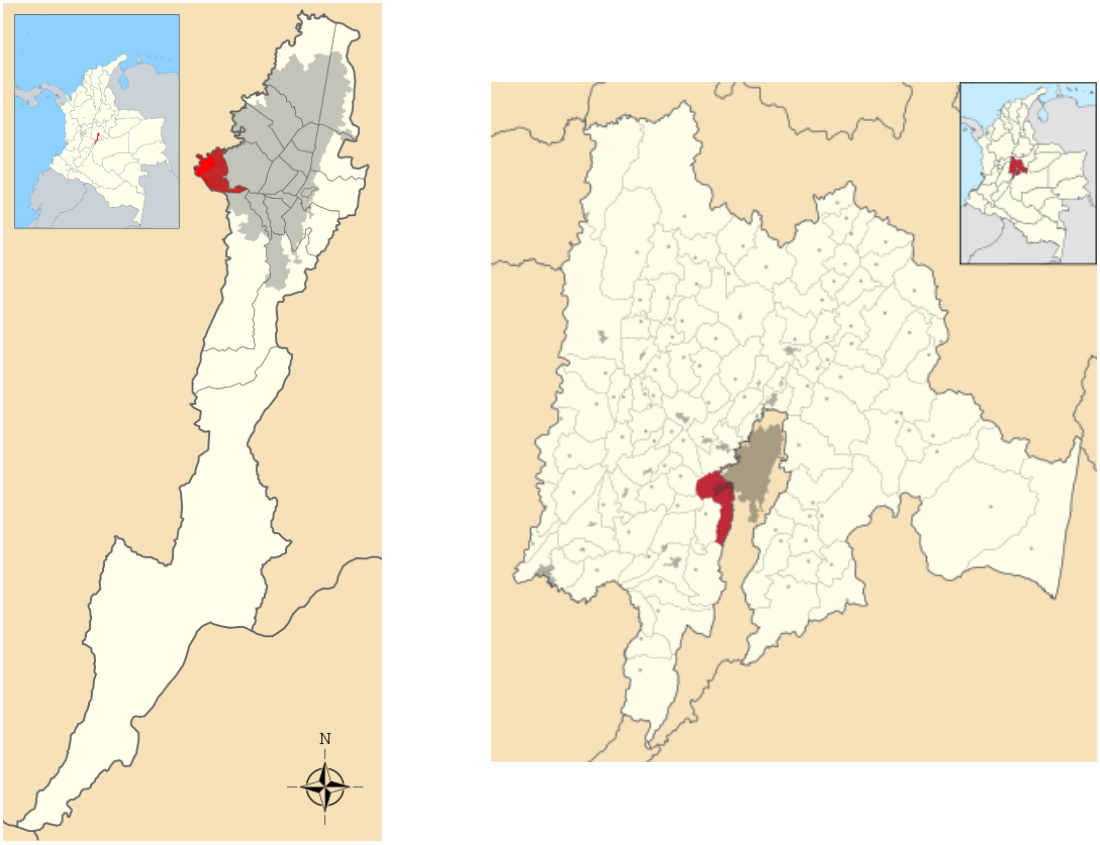
Figure B1: Bogota Municipal Area



Notes: Bogota Municipal Area, administrative map (*Left Figure*) and street map (*Right Figure*).

The theoretical number of participants per group per vocation is listed in table **??**.

Figure B3: Bogota Municipal Area



*Notes:*Location of Bosa in the capital district of Bogota (*Left Figure*), location of Soacha in Cundinamarca (*Right Figure*).

Table B1: Number of people sorted into each group, per vocation and location

Bosa (84)	Group VET + YE (28)	Group VET (28)	Control (28)
Career A (to check) (42)	14	14	14
Career B (to check) (42)	14	14	14
Soacha (216)	Group VET + YE (72)	Group VET (72)	Control 3 (72)
Career A (36)	24	24	24
Career B (36)	24	24	24
Career C (36)	24	24	24

B.1.2.2 COVID-19 in Colombia

The COVID-19 was first confirmed to have reached Colombia on the 6th of March 2020. The government reacted by approving a set of norms aimed at isolating the country from contagions from other countries, protecting older people, school closure, until the national quarantine approved at the end of March 2020.

B.1.2.3 COVID-19 in our study group

The table below shows group averages of how the COVID-19 lockdown in Colombia affected our study participants.

Notes: Timeline of state interventions to control Covid-19 epidemic in Colombia.

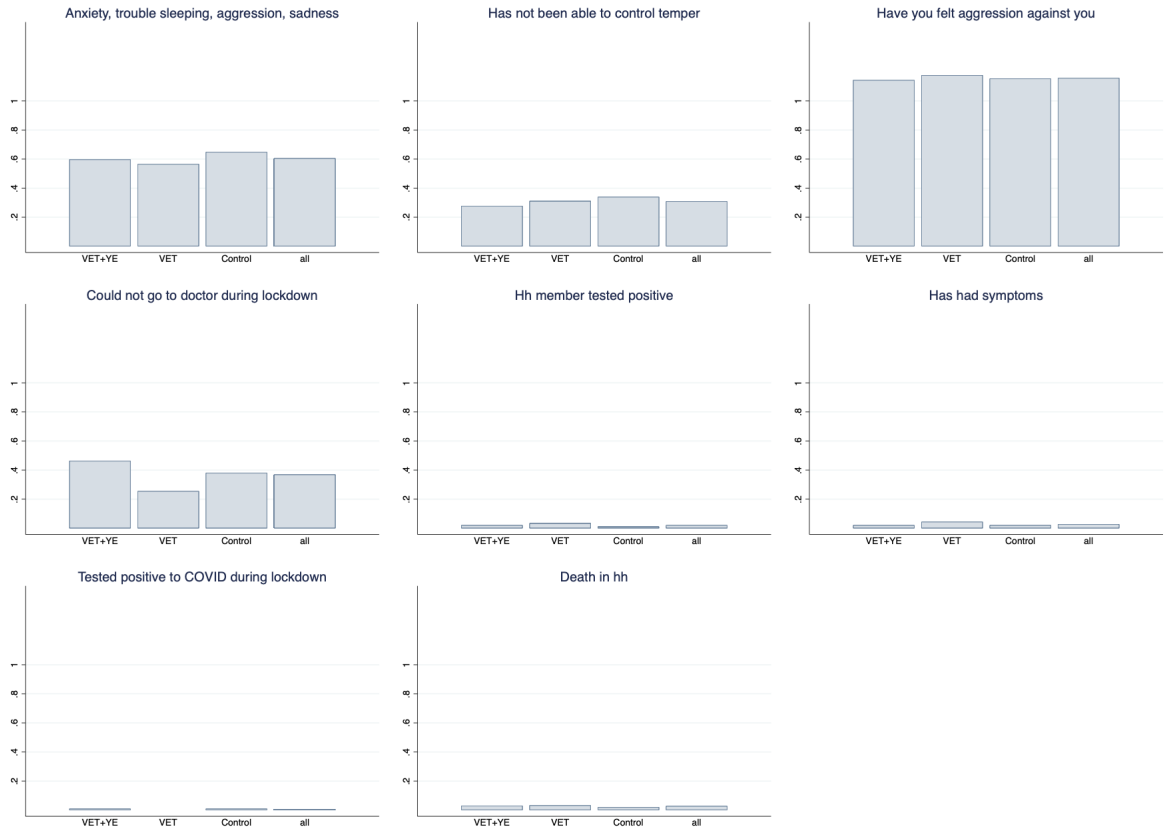
Table B2: How the COVID-19 lockdown has affected participants

	Group VET + YE	Group VET	Control	Total
Is working	37%	39%	41%	39%
Job pays pension	9%	11%	19%	13%
Employment affected due to COVID	60%	67%	63%	63%
Looking for job	78%	68%	68%	71%
Can support family on salary	9%	17%	15%	14%
Received aid	61%	63%	67%	64%
Has moved since pandemic	23%	21%	25%	23%
Who else lives in hh	96%	99%	98%	98%
Partner lives in hh	3%	3%	4%	4%
Hh composition has changed due to COVID	32%	23%	36%	31%
Death in hh	3%	3%	2%	3%
Is vulnerable	4%	9%	7%	7%
Someone in hh vulnerable	34%	33%	23%	30%
Couldn't go to doctor	46%	26%	38%	37%
Has had symptoms	2%	4%	2%	3%
Tested positive	1%	0%	1%	1%
Hh member tested positive	2%	3%	1%	2%
New dependents during lockdown	11%	4%	9%	8%
Have been able to control temper?	28%	31%	34%	31%
Anxiety, trouble sleeping, aggression, sadness	60%	57%	65%	61%
Observations	97	90	97	284

B.1.2.4 Experimental design and data collection

Randomization and baseline data collection was implemented in December 2018/January 2019. The vocational education program and the soft skill program both took place in the second half of 2019. By March 2020, Colombia entered a period of national quarantined, which lasted until August 2020.

Figure B5: Health outcomes



B.1.3 Outcome variables

Depression and distress. To elicit depression and distress we follow Blattman et al. (2017). Each variable is constructed by summing up the answers to each question that applies to it.

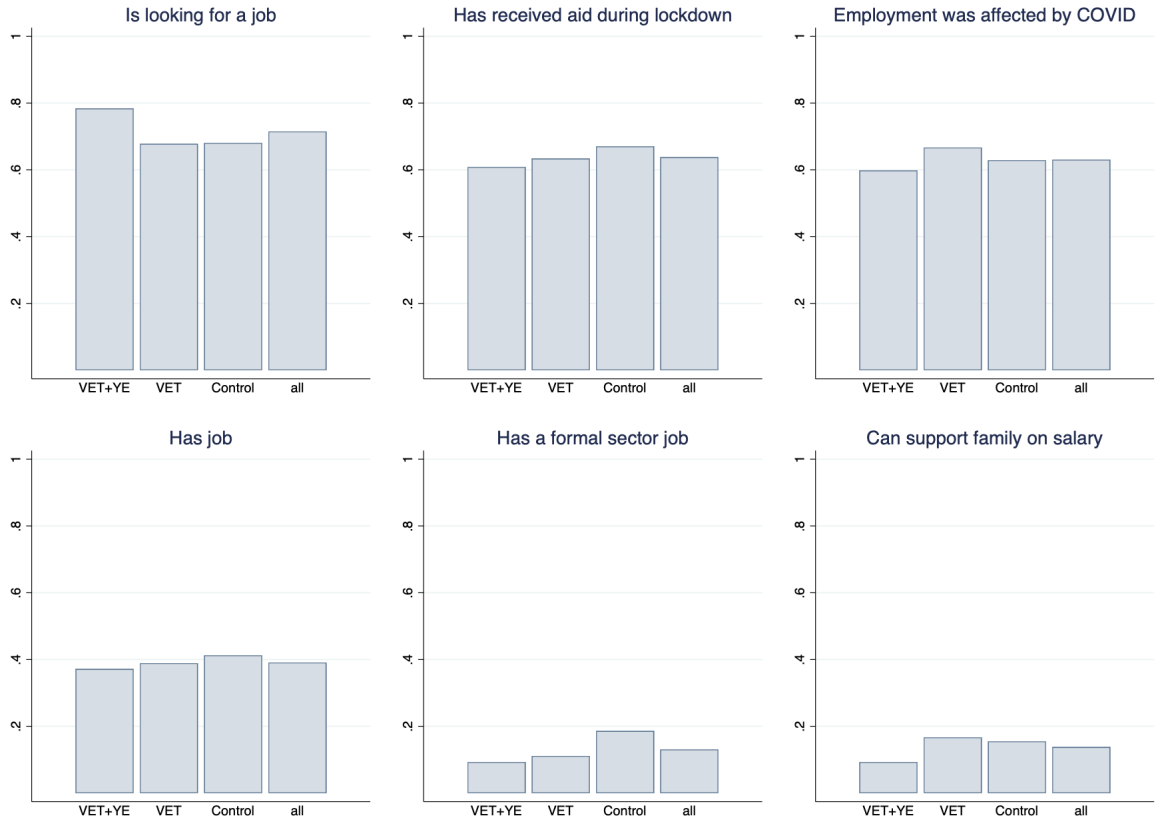
The variable “depression" is built by summing up the answers to the following questions, each of them an index from 0 to 3.

- B63: I feel bad doing things I normally do
- B64: I have a hard time hearing people when thinking about bad things
- B65: I am feeling sad or down hearted
- B66: I feel tired when even not doing anything
- B67: I lose my appetite when feeling bad
- B68: I feel unimportant to everyone

The variable “distress" is built by summing up the answers to the following questions, each of them an index from 0 to 3. We define distress as

- B69: I sit and think bad things
- B70: I have bad dreams
- B71: It seems like bad things are happening again
- B72: I sweat, feel warm, etc when thinking about bad things
- B73: I stay away from certain places
- B74: I talk to myself
- B75: I feel that my heart is spoiled
- B76: I feel frustrated
- B77: I feel like someone’s spirit is hampering me
- B78: My body is dry/tired from worrying
- B79: My heart does not feel satisfied

Figure B6: Labor outcomes



- B80: I feel my heart burning

Post-traumatic stress disorder (PTSD). To explore the levels of post-traumatic stress disorder (PTSD) in our study participants, we administer the PTSD-checklist - civilian version (also known as the PCL-C). It is a standardized, self reported test of 17 questions, to which participants can give answers ranging from 1 'Not at all' to 5 'Extremely'. We use the results of this checklist in two ways in our analysis, according to the PCL-C instructions, which are taken verbatim here:

- The first is the **severity score**, which is a simple sum of each of the 17 answers, and ranges from 17-85. A change of at least 5 points in the severity score is the minimum for the change to be considered relevant, clinically speaking.
- The second is a presumptive diagnosis according to the symptoms laid out in the DSM-IV (the guidebook for clinical diagnoses). We create the **PTSD possible diagnosis** variable according to the guidelines described in the PCL-C. The variable is coded as equal to one if it was marked as 3 or higher[15] to at least one "B" question (questions 1-5), at least three "C" items (questions 6-12), and at least two "D" items (questions 13-17).

Demographics. Because a majority of our outcomes are self-reported, we are complementing survey data with administrative data when we can. In particular, we have collected information on labor market outcomes such as whether participants are in the formal labor market or not (with contributions to the pension system providing an administrative back-up on this information), on socio-economic status (including whether individuals are eligible for social assistance

[15] An answer of "3" is considered to be symptomatic

Figure B7: Household changes outcomes

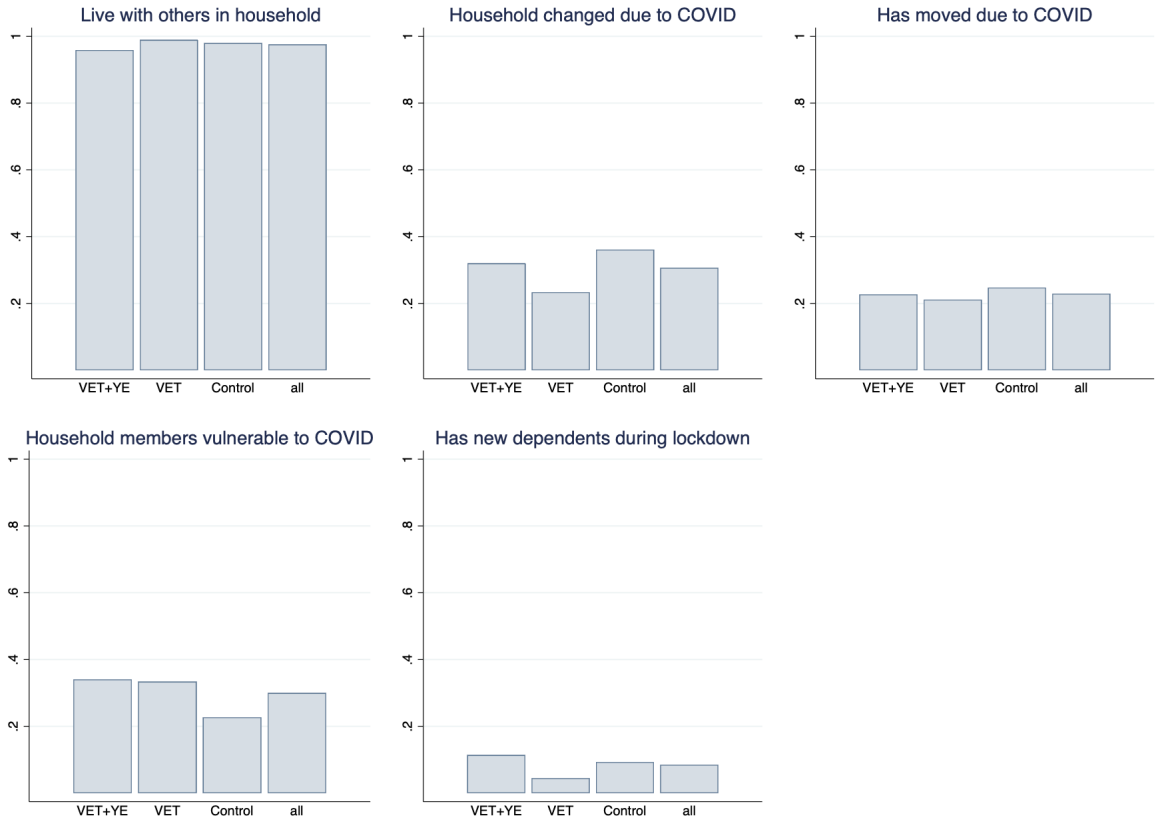
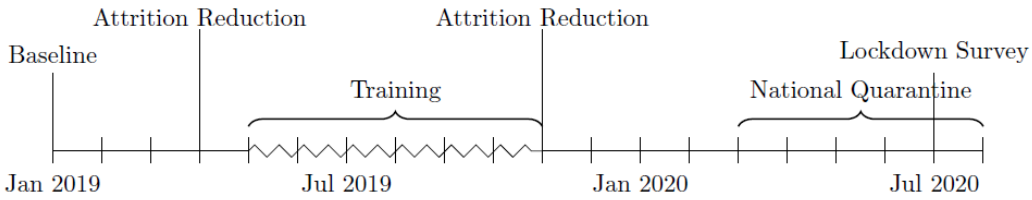


Figure B8: Project timeline



Notes: Figure illustrates the general timeline of the intervention, of the data collection main stages, and of the national quarantine adopted to mitigate the effect of the COVID-19 pandemic.

via the national vulnerability index, SISBEN), on crime and violence (including criminal records and self-reported information on violence), and on education.

We are also interested in the soft skills which will allow us to identify a mechanism that helps vulnerable youth to find decent employment in the formal labor market. To this end, we have collected survey information on preferences, behaviors, and trauma. First, the Global Preferences Survey developed by (Falk et al., 2016), "an experimentally validated survey module to measure six key economic preferences", allows us to measure risk aversion, discounting, trust, altruism, positive and negative reciprocity.

Demographic variables include: gender, age, marital status, number of children, being displaced or not, and socio-economic status (SISBEN, a 1-6 index, given to people by government,

is an indicator of whether or not someone is eligible for government aid).

We also collect administrative labor data, including formal/informal employment (both self reported and verified verifying for formal employment by double-checking that the participant has current records contributing to the retirement pension sysem called PILA, which is the method followed by Attanasio et al. (2017), hourly wages and monthly wages (both formal and informal).

B.1.4 Supplementary results

Table B3: **Mental Health Impacts of COVID-19 Pandemic across Treated with Soft Skills Training and the Control Group**

	Vocational Education versus Vocational Education with Soft Skill Traning					
	Depression		Distress		PTSD Severity	
	Post-COVID19	Δ Post-Pre	Post-COVID19	Δ Post-Pre	Post-COVID19	Δ Post-Pre
VET + YE	-0.411***	-0.333*	-0.283*	-0.213	-0.184	-0.117
	(0.139)	(0.177)	(0.153)	(0.212)	(0.143)	(0.155)
Lagged Dependent Variable	Yes	NA	Yes	NA	Yes	NA
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	183	183	127	127	187	187
R-squared	.202	.0723	.215	.0365	.255	.13

Notes: The sample include participants who received vocational education and soft skill training, and participants that did not receive any training. The dependent variables are the indexes of Depression, Distress and PTSD severity. In even columns the index is measured in July 2020, at the end of the national lockdown. In odd columns, the dependent variable measures the difference between the index measured in July 2020 and January 2019. VET + YE indicates the group of participants who received vocational education and soft skill training, the omitted category is the group that did not receive any training. The Lagged Dependent Variable is the index measured at baseline, on January 2019. Individual controls include age, dummies for gender, displaced individuals, number of children, work experience, earnings and formal employment at baseline. Robust standard errors clustered at the individual level are reported in round brackets. ***, **, and * indicate significance at the 1-, 5-, and 10-% levels, respectively.

Table B4: **Mental Health Impacts of COVID-19 Pandemic across Treated with Vocational Education and the Control Group**

	Vocational Education versus Vocational Education with Soft Skill Traning					
	Depression		Distress		PTSD Severity	
	Post-COVID19	Δ Post-Pre	Post-COVID19	Δ Post-Pre	Post-COVID19	Δ Post-Pre
VET	-0.271*	-0.234	-0.322*	-0.113	-0.129	-0.082
	(0.159)	(0.198)	(0.175)	(0.239)	(0.155)	(0.161)
Lagged Dependent Variable	Yes	NA	Yes	NA	Yes	NA
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	175	175	116	116	181	181
R-squared	.17	.0642	.262	.0832	.288	.0876

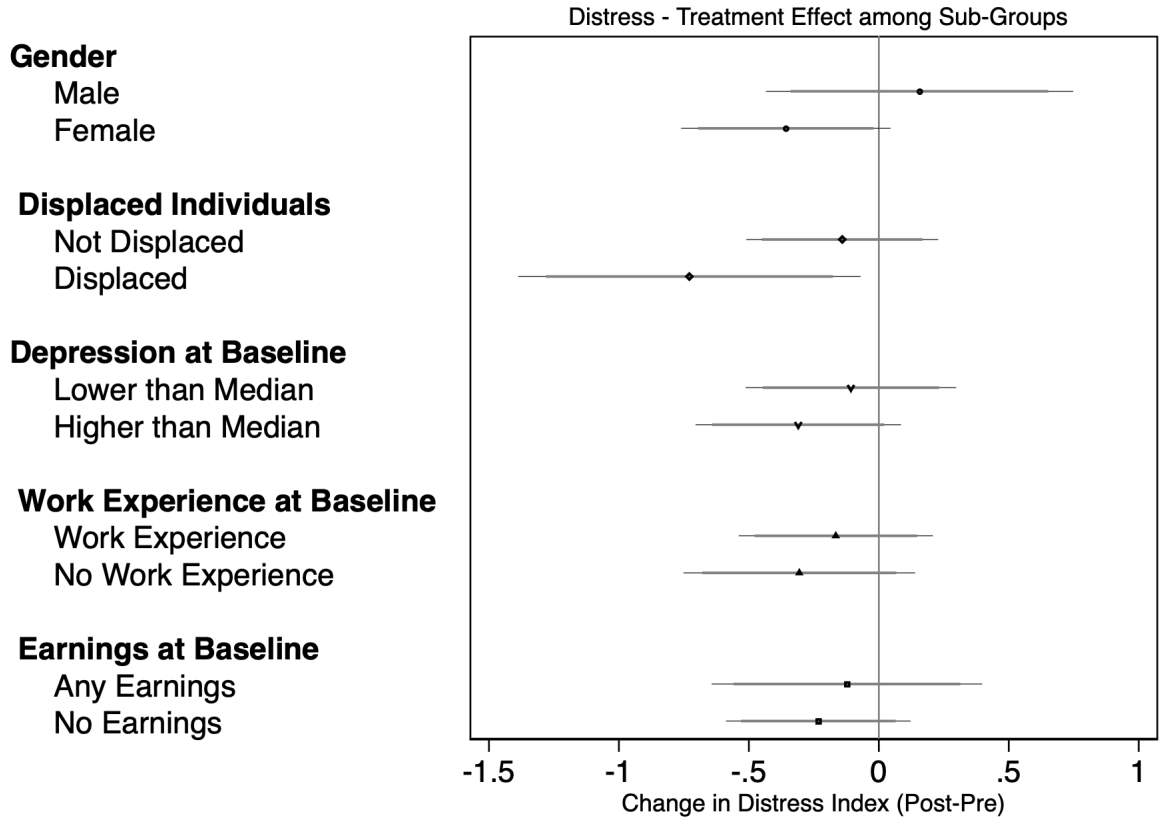
Notes: The sample include participants who received vocational education, and participants that did not receive any training. The dependent variables are the indexes of Depression, Distress and PTSD severity. In even columns the index is measured in July 2020, at the end of the national lockdown. In odd columns, the dependent variable measures the difference between the index measured in July 2020 and January 2019. VET indicates the group of participants who received vocational education, the omitted category is the group that did not receive any training. The Lagged Dependent Variable is the index measured at baseline, on January 2019. Individual controls include age, dummies for gender, displaced individuals, number of children, work experience, earnings and formal employment at baseline. Robust standard errors clustered at the individual level are reported in round brackets. ***, **, and * indicate significance at the 1-, 5-, and 10-% levels, respectively.

Table B5: **Mental Health Impacts of the COVID-19 Pandemic without controls**

	Mental Health Impacts of COVID-19 Pandemic					
	Depression		Distress		PTSD Severity	
	Post-COVID19	Δ Post-Pre	Post-COVID19	Δ Post-Pre	Post-COVID19	Δ Post-Pre
VET + YE	-0.390***	-0.338*	-0.220	-0.132	-0.153	-0.092
	(0.139)	(0.174)	(0.161)	(0.204)	(0.143)	(0.153)
VET	-0.282**	-0.270	-0.302*	-0.030	-0.066	-0.032
	(0.143)	(0.181)	(0.172)	(0.215)	(0.142)	(0.150)
Lagged Dependent Variable	Yes	NA	Yes	NA	Yes	NA
Individual Controls	No	No	No	No	No	No
Observations	269	269	180	180	276	276
R-squared	.145	.0158	.149	.00262	.204	.00148

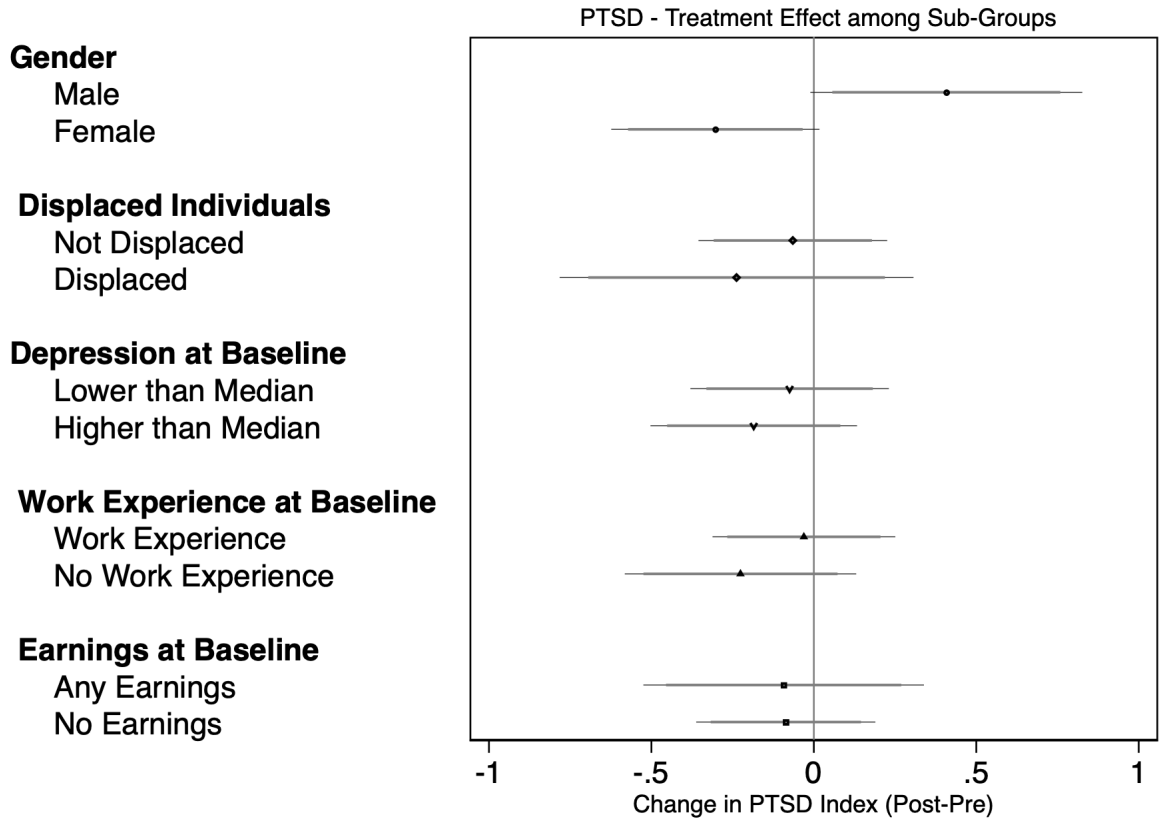
Notes: The dependent variables are the indexes of Depression, Distress and PTSD severity. In even columns the index is measured in July 2020, at the end of the national lockdown. In odd columns, the dependent variable measures the difference between the index measured in July 2020 and January 2019. VET + YE indicates the group of participants who received vocational education and soft skill training, while VET indicates the group of participants that received only vocational education, the omitted category is the control group. The Lagged Dependent Variable is the index measured at baseline, on January 2019. Robust standard errors clustered at the individual level are reported in round brackets. ***, **, and * indicate significance at the 1-, 5-, and 10-% levels, respectively.

Figure B9: **Heterogeneous effect of treatment on distress among sub-groups of participants**



Notes: The figure illustrates the effect of treatment among sub-groups of participants on distress levels during COVID-19. Treatment is defined as both VET + YE and VET, the omitted category is the control group. The dependent variable measures the difference between the distress index measured in August 2020 and January 2019. Regressions account for individual controls, including age, dummies for gender, displaced individuals, number of children, work experience, earnings and formal employment at baseline. The center indicates the coefficient value. Lines indicate 90% (thick lines) and 95% (thin lines) confidence intervals.

Figure B10: **Heterogeneous effect of treatment on trauma among sub-groups of participants**



Notes: The figure illustrates the effect of treatment among sub-groups of participants on PTSD levels during COVID-19. Treatment is defined as both VET + YE and VET, the omitted category is the control group. The dependent variable measures the difference between the PTSD severity index measured in August 2020 and January 2019. Regressions account for individual controls, including age, dummies for gender, displaced individuals, number of children, work experience, earnings and formal employment at baseline. The center indicates the coefficient value. Lines indicate 90% (thick lines) and 95% (thin lines) confidence intervals.

A plague on both your houses! How locust outbreaks affect conflict

Alice Antunes[1], [2]

April 5, 2023

Abstract

Desert locusts have been a source of food insecurity throughout much of the world for millennia. This paper aims to study the effects of locust outbreaks on conflict at the local level in the last twenty years. We use geolocalized data on locust sightings from the FAO LocustHUB, aggregated to the 0.5x0.5 cell level, and match it to cell-level data on conflict. At the country-year level, we use a differences-in-differences specification with cell and country-year fixed effects to explore the yearly effect of locusts on conflict. We find that locusts have a significantly negative correlation with conflict. To explore the monthly effects of locust outbreaks, we run a two-way fixed effects event study. We find that locust outbreaks significantly reduce conflict one month later. Locusts are no longer associated with a change in conflict six months after a locust outbreak. To understand the underlying mechanisms that might be driving this surprising reduction in conflict, we match the locust data to household agricultural data in Ethiopia in 2014. We find that households having experienced a locust outbreak as early as one month before the survey are more likely to report a reduction in crop incomes, an increase in non-agricultural wage incomes, and an increase in food insecurity. The reduction in conflict could arise from three different channels: the locust outbreaks may be causing a reduction in appropriable resources - there are fewer things to steal. It is also possible that there are simply no resources left to compete over. Alternatively, the locust outbreaks might increase collective action and therefore increase social cohesion.

Keywords: agricultural shocks, desert locusts, plagues, food insecurity, conflict.

JEL Codes: Q12, Q18, Q45, D13, Q34, D74

[1] Department of Economics, University of Lausanne.

[2] This is part of a larger project with Ulrich Eberle, Princeton University and Crisis Group, and Dominic Rohner, Department of Economics, University of Lausanne. I would like to thank Elena Esposito for her valuable comments. I would also like to thank Juan Lozano for their invaluable work on matching local and global commodity prices.

1 Introduction

The impact of the environment on human society can be undeniably devastating. Hurricanes, earthquakes, droughts, and pests have always been inextricably linked to human development. In fact, natural disasters are often key players in mythologies around the world. One such story of a natural disaster is recounted in the Book of Exodus in the Christian Bible and the Jewish Torah. In the story, the god of the Israelites punishes the Egyptians for the enslavement of the Jewish people by unleashing ten plagues upon the land. One of these plagues is a plague of locusts, which is recounted as a threat from the god in the following way:

> *Let my people go, so that they may worship me. If you refuse to let them go, I will bring locusts into your country tomorrow. They will cover the face of the ground so that it cannot be seen. They will devour what little you have left after the hail, including every tree that is growing in your fields.* - The Bible, Exodus 10:4-5

While this may be one of the more familiar stories of locusts in modern religion, locust outbreaks are still a cause of human suffering today. The desert locust, or *Schistocerca gregaria*, is a species of grasshopper that is found primarily in the deserts and dry areas of Africa, the Arabian peninsula, and South Asia. Scientists and development agencies warn about the potential consequences of locust outbreaks on food systems (see, for instance, Cressman (2016)). After swarming together, they travel over long distances, eating crops that are also agricultural staples. Locust outbreaks and the direction in which they travel are more or less unpredictable and can therefore be considered to be random income shocks. A key factor of conflict is income shocks, and locust outbreaks affect both incomes and food sources. They may, thus, contribute to social conflict. This study is interested in exploring the effects locusts may have on conflict at the local level.

When solitary, locusts are mostly harmless. When many are born together at the same time in the same place, they swarm and can devastate agricultural systems, contributing to regional food insecurity. Locusts are a type of grasshopper that spawns from buried eggs when soils are particularly moist. When they spawn, they are wingless and brown - these are called *hoppers*. If several hoppers are born at the same time and in the same place, they can turn yellow and move in *bands*. This is when they technically become locusts. When they become *adults*, they grow wings. Several adults together are called *swarms*. These can travel over large distances on wind streams if the speed and temperature of the winds are right (Cressman, 2016).

This study uses the Food and Agriculture Organization's LocustHUB[16], where geolocalized locust sightings have been documented since 1987. Each observation comes with information

[16]The data is available online at https://locust-hub-hqfao.hub.arcgis.com.

on locust appearance (whether it is a hopper, adult, or travels in swarms or bands), date of sighting, and latitude and longitude of the sighting. We aggregate these observations first to the 0.5x0.5 degree cell level. We match these to the Armed Conflict Location & Event Data Project (ACLED) conflict dataset, which we also aggregate to the 0.5x0.5 degree cell level. We define conflict as an event classified as a battle, violence against civilians, a protest, or a riot. We, therefore, have a dataset at the cell-year level that has information on whether a cell has seen a conflict or locust outbreak that year, and how many of either variable were observed. We additionally disaggregate the dataset to the cell-month level to explore the dynamic effects of locust outbreaks.

We use a differences-in-differences approach to explore the relationship between conflict and locust outbreaks at the cell year. We use cell fixed effects to account for any unobserved time-invariant factors. Country-year fixed effects, similarly, account for any unobserved confounders that may vary at the country level that are also time-varying - such as country-wide recessions. By including rainfall and temperature in the specification, we assuage concerns about potential omitted variable bias.

This baseline exercise provides some insights into the relationship between conflict and locust outbreaks. We find that locust sightings are associated with a decreased probability of having conflict in a cell in any given year. More specifically, a locust sighting is associated with a 0.5 percentage point decrease in the probability of having conflict in a cell in any given year. The probability of a conflict occurring in any cell is 6.2%. Following a locust sighting, that becomes 5.7%, which is an 8% decrease in the probability of conflict. A locust sighting is associated with a 0.6 percentage point decrease in the probability of having conflict in a cell in any given year for any cell that has experienced locusts. The probability of conflict at cells that have experienced locust sightings is 7.02%; after a swarm sighting, that becomes 6.42%, which is an 8.5% decrease in the probability of conflict.

Our main concerns are revere causality, measurement error, and omitted variable bias. More peaceful areas could simply have more agriculture for locusts to eat, which would entail reverse causality. Measurement error might arise because the data rely on recording sightings and people might not report locusts in more conflict-prone areas since they might be too busy fighting. Finally, omitted variable bias might arise because locust outbreaks could be more severe following years with more rain.

To begin addressing some of these issues, we exploit monthly variation in locust and conflict outbreaks using a two-way fixed effects model which includes month and cell fixed effects. We relate cell-month variation in locust outbreaks to cell-month variation in conflict in an event study analysis. This allows us to account for time-invariant unobservables (using cell fixed ef-

fects) and time-varying unobservables (using month fixed effects). We also control for rainfall. We do this because the year-level regressions might be smoothing out spikes and dips in conflict that are visible at the monthly level. The month-level analysis gives us some information on the temporal, short-term effects of locust outbreaks since locust outbreaks may change the probability of conflict only in the short run.

The event study model addresses some of the concerns listed above when looking at the pre-trends. Indications of reverse causality and omitted variable bias would appear in the coefficients in the months before a locust sighting. For example, if coefficients in the months before a locust sighting were negative, this would indicate that less fighting might lead to fewer locust sightings. Insignificant coefficients before locust sightings could indicate that concerns about reverse causality, and also omitted variable bias, could be laid to rest.

The event study shows a sudden decrease in the probability of conflict one month after a locust sighting. The effect of a locust sighting on conflict goes back to zero five months after the locust sighting. This seems to indicate a strong but temporary effect of locusts on conflict. More rigorously, one month after a locust sighting, the probability of conflict in a treated cell compared to an untreated cell decreases by 0.42 percentage points, compared to the difference before treatment. The probability of conflict in a given cell in a given month is 2.3% - with a locust sighting, this probability becomes 1.88%, which is an 18.3% decrease in the probability of conflict. While the size of the coefficient is similar to that of the cell-year decomposition explored previously, the change in probability is of a higher magnitude. This could be an indication that the cell-year exercise averages out the effect of the shock. The change in the probability of conflict following treatment, compared to the probability of conflict in untreated cells, goes back to zero after five months following a locust sighting. Additionally, the lack of pre-trends indicates that reverse causality and omitted variable bias may not be important issues.

In the second part of the paper, we explore potential channels and examine more closely how locust sightings affect rural livelihoods. To do so, we aggregate locust sightings to the district level[17]. We match these to survey data from the World Bank's Living Standards Measurement Study - Integrated Surveys on Agriculture (LSMS-ISA). We use the fact that the second wave of the survey in Ethiopia in 2014 matches an upsurge of locust sightings from 2013. This allows us to focus on the second wave of the LSMS-ISA in Ethiopia to get a screenshot of the effect of locusts on the local crop economy. We aggregate crop, livestock, transfer, and other incomes to the district level. Households can also indicate how many months they have spent in food insecurity over the last year.

We can use the fact that the survey data was collected by April 2014, to create a variable

[17]Which is also the third administrative level according to the Database of Global Administrative Areas (GADM).

that counts the number of locust sightings n months before the survey. We do this because we had previously found that locust sightings seem to have the biggest attenuating effect on conflict one to five months after being recorded. In addition, the FAO indicates that the most vulnerable time of a crop's life is during the growing phase, which coincides with the months between October 2013, and April 2014 (FAO, 2020). We can therefore use the panel structure of the LocustHUB data to explore some correlations between the time of locust sightings and rural incomes. We run a simple OLS regression where we correlate incomes from various sources to the number of locust sightings $n \in 1, ..., 12$ months before April 2014 in each district, while controlling for rainfall.

Finally, we match locust sightings to state-level food prices from the World Food Program's food price monitoring system (VAM), which collects year-level time series data on several commodity goods at each state's main market. We run a simple two-way fixed effects regression, controlling for country and year fixed effects, to get an idea of the relationship between state-level prices and locust sightings.

This second part of the paper provides some insights into the relationship between locust sightings and rural livelihoods. We find, first, that locust sightings do not change the prices of consumer crops such as sorghum and millet. This could indicate that at the state level, markets use coping strategies to weather the shock of losing harvests. We do not have the data to know how this is done, but this could be because the main state market is well integrated into a larger system of markets. Enough of the crops lost due to locusts may be replaced from markets that did not have locust outbreaks.

Zooming in, we find that households in districts with locust sightings indicate that they have spent around half a month in food insecurity. We also find that districts with locust sightings are associated with significant losses in crop incomes. In particular, a locust sighting one month previously is associated with a 96.3% decrease in crop incomes. This cross-sectional correlation indicates that there is a strong immediate negative relationship with crop incomes, which becomes less strong one month later and slowly dissipates over time. In addition, a locust sighting one month before the survey is associated with a large increase in non-agricultural wage incomes.

What could be happening? Following a locust sighting, agricultural workers might switch income sources, and might rely on informal insurance networks such as family members in other parts of the country. There is anecdotal evidence that household members depart for urban areas to find non-agricultural wage work following a locust outbreak (Brader et al., 2006). The fact that locust sightings are associated with an increasing reliance on non-agricultural wages seems to indicate that there is also some substitution in sources of income occurring. This

exercise provides some evidence that households seem to rely on social insurance mechanisms and income substitution to offset losses from locust outbreaks.

These findings indicate several channels which may explain the attenuating effect of locust outbreaks on conflict. There is a clear negative income shock associated with locust outbreaks. The negative income shock could indicate that locust plagues leave regions with nothing appropriable - there is less conflict because there is nothing to steal. Locust outbreaks leave regions poorer and hungrier, and with fewer reasons for violence. Similarly, but subtly different, is the idea that locusts destroy any resources that could be competed over - areas for growing crops or grazing livestock. The desert locust is active in the same regions that are known for conflicts between nomadic pastoralists and sedentary agriculturalists. Locusts may eliminate resource competition between these groups of people. There is significant anecdotal evidence that resource competition is a leading source of conflict throughout the Sahel - following droughts, nomadic herders who live off their livestock often encroach into agricultural lands, seeking places for grazing. Examples include the Fulani in Nigeria (ICG, 2017, 2021), the Dinka in Sudan (Eberle et al., 2020), the Tuaregs in Mali[18], and fighting in Laikipia county in Kenya[19]. locust outbreaks may reduce conflict by eliminating the most common source of conflict in the region.

Alternatively, households may cope with locust outbreaks in various ways that increase social cohesion. Households seem to substitute their agricultural incomes with other incomes. They may also rely on social networks - in particular, family in other parts of the country - to weather the income shock from locust outbreaks. A greater sense of social cohesion could be the result. It is hard to disentangle which of these mechanisms might be more important, though all three forces might be acting at the same time.

This paper contributes to several literatures. First, it contributes to the knowledge of the economic effect of locusts. Locusts outbreaks can be very damaging to local economies. The most obvious consequence of locust outbreaks is on food systems. Their food of choice is key crops such as millet and sorghum. A locust swarm of one kilometer squared (which can comprise up to 150 million locusts) can eat the same amount of food per day as 35,000 people and can travel up to 150 kilometers a day. To put this into perspective, a swarm the size of Paris eats the same amount of food in one day as half the population of France (Cressman, 2016). Locust outbreaks are expensive, both directly and indirectly. The 2003-2005 outbreak in the Horn of Africa resulted in crop loss worth $2.5 billion (Brader et al., 2006). Furthermore, the FAO estimates that household expenditures on food increased by three-to-five fold during that same outbreak. During that outbreak, some households had to resort to taking out loans to face the increased burden on their budgets (Brader et al., 2006). The increased pressure on household budgets

[18]https://www.middleeasteye.net/news/sahel-fire-malian-refugee-crisis-world-forgot
[19]https://www.voanews.com/a/kenya-deadly-land-invasions-blamed-on-political-incitement-/3974615.html

following a locust outbreak also has well-documented consequences for health and nutrition. To face the increased pressure on household budgets, households have to reduce expenditures on non-food items such as clothing, travel, and even schooling.

To the best of our knowledge, there are only a few studies that systematically estimate the effect of locust outbreaks on rural households, and most of these studies are on their effect on health. Conte et al. (2021) find that children exposed to a locust outbreak in Mali in 2004 are more stunted than children not affected by the outbreak. Similarly, Marending and Tripodi (2022) find a significant farm profit loss following a locust outbreak in Ethiopia in 2014, which results in stunted growth and body-mass-index scores for children affected by the outbreak. Vreyer et al. (2015) find that the locust outbreak in Mali from 1987-1988 reduced school enrollment of boys living in rural areas affected by the outbreak four years later. Some studies also document the effect of other agricultural pests on rural livelihoods. Banerjee et al. (2007) find that Phylloxera (an insect pest) that attacked French vineyards in the 19th century led to stunting in adults affected by the pest as children. Finally, Baker et al. (2020) find that adults who were children during a boll weevil plague on cotton in the early 20th century American South had increased education compared to people who worked in cotton farms unaffected by the pest.

In political economics, it is well-established that shocks to agricultural production can cause social disruption and changes to social cohesion. While the potential disruption created by locust outbreaks has been a priority of development and aid agencies to help regions affected, not much effort has been put into understanding the impacts of locusts on the economy, and, more particularly, on conflict. To the best of our knowledge, this is the first paper to do so.

Broadly speaking, there are two opposing forces explaining the impact of income shocks on conflict: the opportunity cost effect arises when an increase in prices of labor-intensive goods leads to a decrease in conflict, by increasing wages and thus increasing the potential loss someone may incur if they joined a rebellion, or by encouraging potential rebels to earn money outside of a rebellion. On the other hand, the rapacity effect occurs when prices of (less labor-intensive) goods increase and the gains from appropriation become greater than the gains from wages (thereby encouraging violence). These theories have been developed in the literature for over twenty years, with key works refining subtleties in the two forces. Examples include the importance of feasibility in joining a rebellion (Collier et al., 2009) or the importance of the share of labor in the price shock of a labor-intensive good in generating the opportunity cost effect, and the importance of the share of capital in the price shock of a capital-intensive good in generating the rapacity effect (Bó and Bó, 2011; Dube and Vargas, 2013). Most recently, McGuirk and Burke (2020) use agricultural goods in Africa to explore how the opposing forces of price

shocks on consumption and production on households may affect conflict.

Commodity price markets have been a great setting to explore both mechanisms. Crost and Felter (2020) explore how banana plantations in the Philippines are more likely to cause rebellion, but only if they are exported bananas, and not if they are the locally consumed type of banana, giving a real-life example of the extortion story. Berman and Couttenier (2015), similarly, find that negative income shocks are likely to cause conflict but less so in more remote areas. Yet other papers examine the differential effects on conflict if the commodities are labor-intensive (such as agriculture) or capital-intensive (such as mining) (Bazzi and Blattman, 2014; Bó and Bó, 2011; Dube and Vargas, 2013). Overall, the literature has made a very recent move from cross-country analyses to more disaggregated analyses (Berman and Couttenier, 2015), which is the proposed approach in this application. Berman et al. (2017), for example, look at how world prices of minerals affect local conflict. Camarena (2016) also uses within-country variation in local food prices and refugee influxes to see how local demand shocks can lead to conflict.

Previous literature has focused on how shocks to global food prices affect local conflict. While this is an exogenous measure of income variation, it is unclear to what extent global price shocks affect rural incomes. Countries can use policies such as agricultural subsidies or export restrictions to protect local markets from international price fluctuations (Smith, 2014). To address concerns such as these, studies have used environmental sources of income shocks such as rainfall shocks (Miguel et al., 2004; Miguel and Satyanath, 2011) to proxy local income variation. This study builds on this legacy by using an inherently local price shock to contribute to our understanding of the origins of conflict.

This study also contributes to the larger literature on the economic effects of natural disasters. There is a large body of literature looking at how rainfall, temperature, and climate can affect various economic outcomes, especially economic growth. Dell et al. (2014) find that whereas temperature shocks do not seem to affect growth in rich countries, poorer countries grow less if they experience a hotter year. Hsiang (2010) finds something similar in the Caribbean basin, with an increase in temperature in a given year leading to less economic growth. When it comes to the role of climate in fueling conflict, Burke et al. (2009) find that higher temperatures lead to more conflict, and Miguel et al. (2004) find that less rainfall leads to more conflict. There is likely an exacerbation effect going on: hotter places suffer more if temperatures increase, as dryer places suffer more if rainfall decreases.

In addition, it is well-established that natural disasters can affect social structures. Following natural disasters (such as cyclones) or shocks to temperature or rainfall, some communities can experience more conflict (Hsiang et al., 2011, 2013; Hsiang, 2010; Dell et al., 2014). However, this

is not the only social change that might occur. For example, Ager and Ciccone (2018a) find that communities with greater rainfall variation in the 19th-century United States had larger shares of their populations enrolled in religious organizations. Similarly, Bentzen (2021) finds that districts of the world's neighboring districts hit by earthquakes are more likely to be religious. Finally, different coping strategies to risk can lead to social cohesion. Communities with less access to formal insurance systems tend to create their informal insurance systems in the face of risk, especially rainfall risk (see, for example, Rosenzweig (1988a,b); Rosenzweig and Stark (1989); Townsend (1995)). Our study contributes to the understanding of how social structures change following an unpredictable environmental shock.

The rest of this paper proceeds as follows. Section 2 explains data construction and data sources. Section 3 shows some initial results on the effect of locusts on conflict. Section 4 explores the channels. Section 5 provides a discussion of the methodology used, and section 6 concludes.

2 Data sources

This section explains how each data point was collected, how they were merged, and to which level they are aggregated. We also present some descriptive statistics.

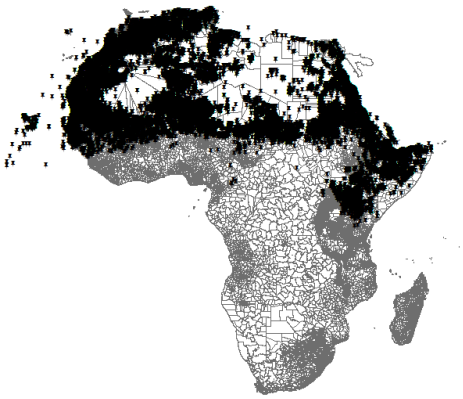
2.1 Locusts

The Food and Agriculture Organization, a branch of the United Nations, is responsible for monitoring locusts throughout the world and providing aid to regions affected. In addition, they monitor the online LocustHUB[20], which has collected information transmitted by all locust-affected countries since 1987. Each locust sighting provided is geolocalized and given information about whether it is a hopper (a young, wingless locust) or adult (a fully matured locust, with wings), and whether it was seen in a band (which is a collection of hoppers) or a swarm (which is a collection of adults). Furthermore, information on when the sighting was logged, and if any pesticides were sprayed in the area, is provided. For swarms/bands and pesticide control entries, an estimate of the hectares affected is provided. We have information on when each sighting was recorded. All locust sightings are mapped in figure 1. Figure 2 shows locust sighting frequency by type. The figure documents the plague of 2003-2005, the plagues in 2013-2014, and the plague of 2019-2020. In the figures, I restrict the sample to locust sightings in Africa. This is because, first, the conflict data that I use is limited to the African continent. Second, the majority of the habitat for the desert locust is in Africa. This removes means that the relationship between locust sightings and conflict found in this paper is limited to Africa.

[20]The data is available online at https://locust-hub-hqfao.hub.arcgis.com.

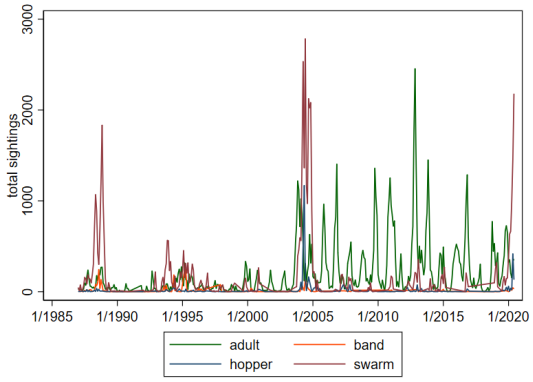
The relationship between desert locusts and conflict in the Middle East and South Asia may very well be the same as it is in this analysis, but it may also differ.

Figure 1: Locust sightings 1987-2020



Notes: The figure illustrates, in a map, all sightings of desert locusts between 1987 and 2020 collected by the FAO and registered in the FAO LocustHub database in Africa. Black pins indicate instances of locust sightings, regardless of type and number of sightings.

Figure 2: Locust sightings in time, by type



Notes: The figure illustrates, over time, all locust sightings, disaggregated by type of sighting in terms of the lifecycle of the locust, limited to the African continent.

We aggregate these sightings in two different ways. First, we match these locusts to the PRIO grid, which is a grid developed by the Peace Research Institute Oslo that divides the world into 0.5x0.5 cells, which are equivalent to 55km by 55km squares at the equator. We thus have a dataset at the month-cell level between 1987 and 2020. We construct two variables with this information: first, a locust dummy that is equal to one if there was a locust sighting in cell i during month m; second, we build a cell-month locust intensity variable that is simply the sum of locust sightings in a cell i during month m.

Matching locusts to the PRIO grid gives two benefits: we can easily match other geo-

localized variables to the same grid, and we can also benefit from the variables collected by PRIO, such as average temperature and rainfall in the cell each year. We make use of this to control for weather in our regressions, given that locusts are born under moist conditions, especially following years with floods or cyclones. Since the conflict dataset (described below) only covers Africa, we restrict our sample to Africa only. In addition, since the conflict dataset only covers the period from 1997 on, we drop any locust observations before 1997.

Second, we aggregate these locusts to the district level (which is equivalent to the third administrative level of GADM, or the Database of Global Administrative Areas). We thus have a dataset at the month-district level between 1987 and 2020. Similarly to the cell-month level, we construct a dummy that is equal to one if a district d has experienced a locust sighting during month m; and a locust intensity variable that is the sum of locusts sighted in district d during month m. We also drop observations outside of Africa and before 1997.

For any given cell, the probability of seeing a locust in any given month is 0.05%. There are 17,226 cell-months in our sample that have seen locusts. For any given cell that has seen at least one locust sighting ever, the probability of a locust sighting in any given month is 1.9%. The probability of seeing a locust in any given year in any given cell is 3.5%; there are and 9,092 cell-years that have seen locusts. The probability of seeing a locust in any given year in cells that have seen at least one locust ever is 11.9%[21]. Table 1 provides additional information on locust events in our sample.

2.2 Conflict

We make use of the dataset on conflict created by the Armed Conflict Location & Event Data Project (ACLED) - which, in their own words, is a "disaggregated data collection, analysis, and crisis mapping project. ACLED collects information on the dates, actors, locations, fatalities, and types of all reported political violence and protest events around the world.[22]" Each observation is geo-localized and coded as a battle, violence against civilians, riot, or protest. Like for the locust sightings, we match ACLED to the PRIO grid and aggregate conflicts to the 0.5x0.5 cell level. We create a variable that is equal to one if if least one event is classified as battle, violence against civilians, riot, or protest in a cell-month-year or in a district-month-year. We also create a variable that is equal to the number of conflict events in a given cell-month or cell-year. The number of cells in Africa is 10,744.

For our entire sample, we have 3,094,272 cells, 12 times for each of the 10,744 cells in the continent for each month from 1997 to 2020. There is a probability of 2.3% that there will be conflict in any cell in any month; for cells that have seen at least one locust sighting ever, that

[21] That is, $9,092/75,840 * 100$

[22] https://acleddata.com/about-acled/

probability goes up to 2.6%. For any given year, the probability of a conflict event occurring in any cell is 6.2%; for cells that have seen at least one locust sighting ever, that probability goes up to 7.02%. Table 1 provides some information on conflict and locust events in our sample.

Table 1: **Descriptive statistics of conflict and locusts at different levels of observation**

	Observations	Mean	Std. Dev.	Min	Max
At month level					
Locust dummy	3,094,272	0.006	0.074	0	1
Conflict dummy	3,094,272	0.023	0.150	0	1
Sum of locust sightings	3,094,272	0.037	1.233	0	489
Sum of conflict events	3,094,272	0.058	0.751	0	205
At year level					
Locust dummy	257,856	0.035	0.184	0	1
Conflict dummy	257,856	0.102	0.303	0	1
Sum of locust sightings	257,856	0.444	6.568	0	644
Sum of conflict events	257,856	0.693	7.244	0	863

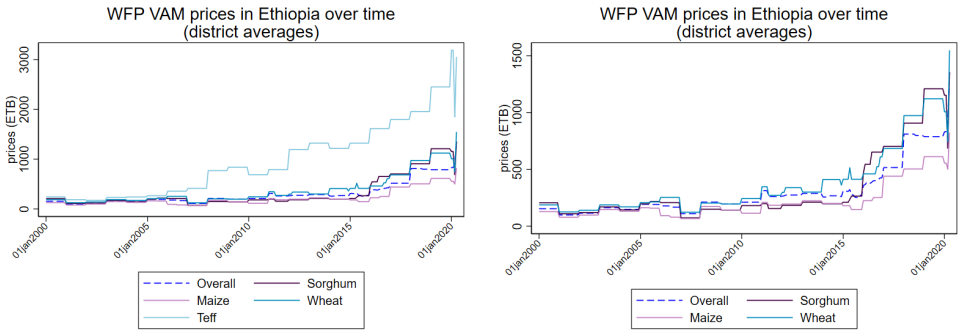
Notes: The table gives the descriptive statistics of conflict and locust events at cell-month level and at cell-year level.

2.3 Food prices

We collect food prices from two sources: first, the World Food Program's VAM, which has collected food prices sold at main markets at the state level (this is equivalent to the first administrative level) from 2000 to 2020. The data collected vary for each market, and the panel is highly unbalanced - that is, there are gaps when prices for some goods are collected. However, this is the most comprehensive dataset for local food prices and allows us a first glance at the correlation between locust sightings and local shocks. In particular, we have information on 2,174 markets in 796 regions on the African continent. We have also matched these to the global food prices tracked by the Food and Agriculture Organization's Food Price Monitoring and Analysis Tool (FPMA).

Figure 3 illustrates the time series we obtain for prices of certain grains in Ethiopia. These prices are not adjusted for inflation but show the real price of goods at the markets. Not all grains are accounted for since the beginning of the time series (e.g. rice), so we only include some grains that we have full time series for that are also known to be locust favorites. The figure shows that there has been a gradual increase in the price of grains, with a jump in the price of teff in 2005. There are some spikes in the prices of the main grains, and removing teff from the sample we can see these spikes and dips more clearly in the right hand side panel. In particular, there was a big dip in grain prices in 2007, and big spikes in 2011, 2016, 2018, 2019, and 2020.

Figure 3: **Grain prices, Ethiopia, 2000-2020**



Notes: The table shows a time series of prices for some grains in Ethiopia since 2000. The left hand side shows staple grains, including teff. The right hand panel removes teff to explore the fluctuations in price more closely.

2.4 Agriculture and income

To further explore how locusts may affect individuals, we use the Living Standards Measurement Study - Integrated Surveys on Agriculture (LSMS-ISA)[23]. The LSMS-ISA is a survey projected funded by the Bill and Melinda Gates foundation and implemented by the World Bank's LSMS team that provides data on household welfare and income-generating activities at the agricultural level. Eight countries in Sub-Saharan Africa are partners in the project and provide data in waves: Burkina Faso, Ethiopia, Malawi, Mali, Niger, Nigeria, Tanzania, and Uganda. Figure 4 shows which countries are in the LSMS-ISA. These figures plot conflict frequency (the solid lines), locust sighting severity (the dots), and maps when each wave of the LSMS-ISA is undertaken. Unfortunately, several countries do not have a wave following locust sightings; and several countries in the LSMS-ISA have never had locust sightings reported. For this reason, we focus on the second wave of the LSMS-ISA in Ethiopia.

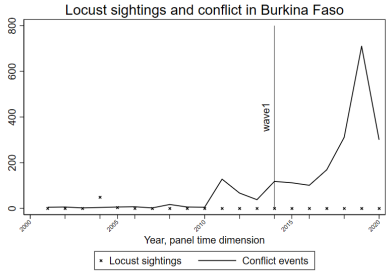
We chose to use only Ethiopia's Socioeconomic Survey for 2013-2014 (ESS2), which is the second round of the LSMS-ISA survey. The first reason for this is because the other countries in the LSMS-ISA do not have locust sightings recorded for their respective waves (see figure 4), so it does not make sense to include the other countries in the sample. Mali and Niger could be candidates for further inspection to see if the relationships we find in ESS2 hold. Second, the locust surges experienced by Ethiopia only coincide with the second wave of the survey - figure 5 makes this clear. Locust sightings are disaggregated by type, and the timing for the three waves of the survey are shown in gray. The locust upsurge in 2013 is a good setting to explore the relationship between the agricultural economy and locust sightings.

The timing for the second wave of the Ethiopia Socio-Economic Survey (which is the sec-

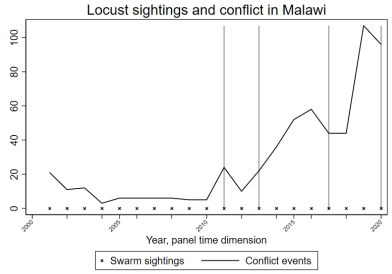
[23]https://www.worldbank.org/en/programs/lsms/initiatives/lsms-ISA

Figure 4: **Locust sightings and conflict events in countries surveyed by LSMS-ISA**

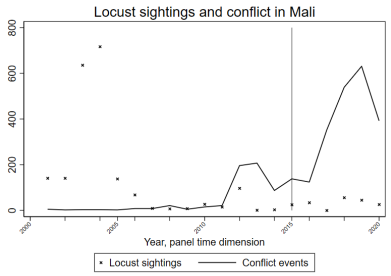
(a) Burkina Faso



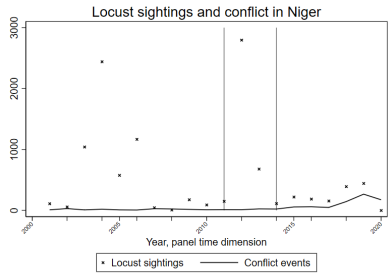
(b) Malawi



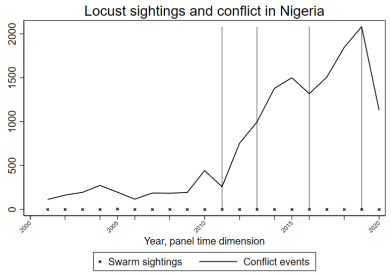
(c) Mali



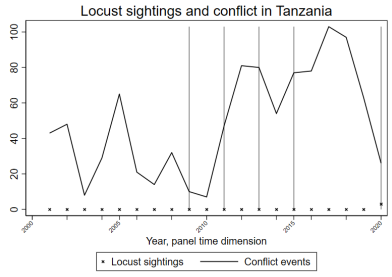
(d) Niger



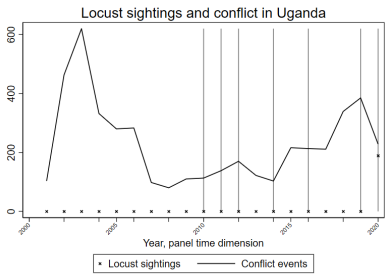
(e) Nigeria



(f) Tanzania



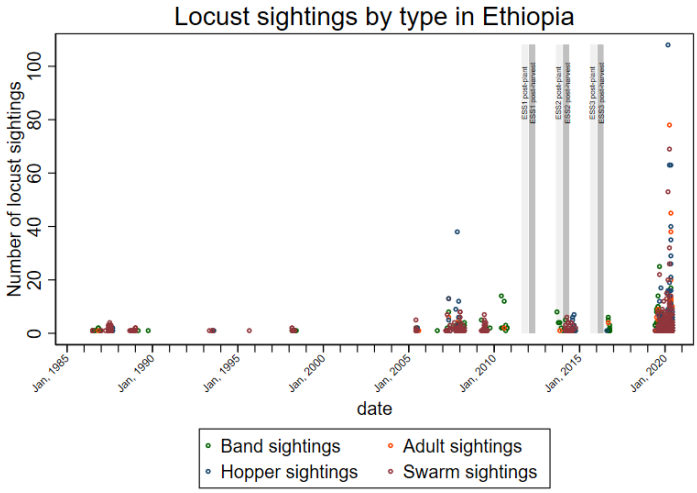
(g) Uganda



Notes: The graphs illustrate the frequency of conflict events and frequency of locust sightings in each country surveyed by the LSMS-ISA, each year. Black lines are conflict events, dots are locust sightings, and vertical grey lines are the dates for each wave of the survey in each country.

ond wave of the LSMS-ISA, and from now on we will refer to the second wave of the survey as ESS2) was the following: enumerators conducted a post-planting and post crop-cuting survey in October-November 2013, and the post-harvest questionnaire was administered between March 2014 and April 2014. For this reason, we only include in our sample locust sightings that

Figure 5: Locust sightings by type in Ethiopia



Notes: The figure illustrates sightings of locusts for each month in Ethiopia, decomposed by type of locust. The vertical axis shows how many locust sightings were reported in that month in Ethiopia. The horizontal axis shows months from 1985 to 2020. The bars show when Ethiopia collected its data for the Ethiopia Socio-Economic Survey (ESS), where ESS1 is the first wave, ESS2 is the second wave and ESS3 is the third wave. "Post-plant & cut" stands for the post-planting and post cutting questionnaires, and "post-harvest" stands for the post-harvest questionnaires.

were registered before May 2014.

We are interested in incomes from different sources for each household surveyed. In particular, we have information on income from crops, livestock, transfers, and "other" sources. Table 2 shows descriptive statistics of these variables in Ethiopia in 2014, in Ethiopian Birr[24]. Both crop and livestock incomes are higher than transfer and other incomes, but they can also be negative - these are more risky sources of income than transfers and "all others".

2.5 Other data

Finally, we also include information on rainfall and temperature from PRIO. We use the average cell-year values for rainfall and temperature provided by PRIO. These are measured at the 0.5x0.5 degree, and provide yearly values of both variables.

3 Initial results

We have some hypotheses for how locusts may affect conflict. The fact that we use to kick-start our thinking is that locusts, by construction, destroy harvests and reduce incomes. They are, in our assumptions, a negative income shock.

[24]The exchange rate of Ethiopian Birr to US dollars in 2014 was, on average, 0.0509 Birr to USD

Table 2: **Descriptive statistics of incomes from different sources in Ethiopia 2014, Ethiopian Birr**

	mean	sd	min	max
non-ag wage income	7260	60474	0	2945576
self-employment income	36758	2183792	-7123200	155902304
ag. wages income	501	7286	0	437400
crop income	3876	14578	-228396	584232
livestock income	3691	40729	-31438	2654187
transfers income	2720	15020	0	600000
total income	54806	2184835	-7096940	155902304
non-farm income	46738	2184546	-7107000	155902304
farm income	8068	44029	-221896	2656033

Notes: The table gives the descriptive statistics of incomes from different sources aggregated to the district (administrative 3) level, for households surveyed in the second wave of the LSMS-ISA in Ethiopia, in 2014.

Hypothesis 1: Locusts contribute to food riots. Locusts reduce incomes and increase food prices, which exacerbates food insecurity. Grievances increase, and the opportunity cost to join protests decreases. There would be a negative correlation between income and conflict. This would be the "food riots" story, where rising food prices lead to social unrest. One example of an infamous food riot is the Arab Spring, where it is possible that rising food prices led to the protests that kickstarted the revolution (though evidence on this is mixed (Smith, 2014)[25]).

Hypothesis 2: Locusts devour any incentive for violence. Our second hypothesis is that locusts reduce conflict by destroying anything to appropriate. In this scenario, locust plagues decrease harvests (and therefore incomes) and decrease conflict. We would see a positive correlation between locust events, incomes, and conflict (that is to say, these all would move in the same direction). A related, but subtly different mechanism, is the following: because locusts also affect the fodder of livestock, the usual story of conflict between agriculturalists and pastoralists is less likely to occur: there are no resources to compete over.

Hypothesis 3: Locusts increase social cohesion via informal insurance networks. It is possible that following a locust outbreak, which is a covariate environmental shock, social cohesion increases because households rely more on each other and on informal insurance networks. There could be a sense of community and trust that is formed in the wake of the natural disaster, and neighbors that would usually fight amongst each other may help each other out in times of need. Alternatively, social cohesion could increase because households send out members of their family to other parts of the community to work off the farm or in urban centers -

[25] See, for instance, https://www.foreignaffairs.com/articles/middle-east/2011-03-23/let-them-eat-bread

households become more reliant on other sources of income besides their crops and get greater exposure to other people and other livelihoods. As more households substitute crop incomes for non-agricultural wage incomes or familial transfers, there could be a lessening of tensions via a dismantling of social barriers and via a greater reliance on informal insurance networks.

3.1 Initial results

At the most basic level, we can estimate the effect of locust plagues on conflict and economic development at the cell and year level using a differences-in-differences model comparing treated and untreated cells (where treatment = locust sighting) before and after a sighting. This gives an initial understanding of the effect locusts have on conflict. To do so, at the cell-year level, for swarms only, we estimate the following regression, where results are reported in table 3:

$$Y_{ict} = locust_{ict} + \gamma_t + \gamma_i + \varepsilon_{ict}$$

Where Y is either a conflict dummy that is equal to one if there has been a conflict event in a cell-year; *locust* is a dummy equal to one if there has been a locust sighting in a cell-year; γ_t and γ_i are year and cell fixed effects, respectively. Furthermore, i stands for cell, c for country and t for year. We begin by running this regression; second, we add weather controls from PRIO; third, we replace time fixed effects by country-year fixed effects. We run this regression on all cells in the PRIO database for Africa, and also by restricting to only cells that have experienced a locust attack ever to get an idea of an average treatment effect. We end up with six different regressions.

This regression includes cell fixed effects which take into account any confounding variables that are time-invariant at the cell level. Time fixed effects account for any confounding variables that are time-varying at the yearly level. Any yearly shocks that affect all of the African continent, such as El Niño weather events, would be controlled for. Finally, we look at two different samples: first, we look at the entire sample in the PRIO dataset for the African continent. Since this might average out the effect of locust events, we also include a regression that includes only cells that are in the LocustHUB dataset. This allows us to proxy for the effect of locust outbreaks in the desert locust's habitat. This could be an oversimplification, yet it makes little sense to include places (such as South Africa) that are not likely to ever experience a locust sighting in the specification.

Table 3 shows that locust sightings are associated with a decreased probability of having conflict in a cell in any given year. Looking at column 5, a locust sighting is associated with a 0.5 percentage point decrease in the probability of having conflict in a cell in any given year. The probability of a conflict occurring in any cell is 6.2%. Following a locust sighting, that becomes

Table 3: **The effect of swarms on conflict**

	Conflict		Conflict		Conflict	
	All	Restricted	All	Restricted	All	Restricted
Locust dummy	-0.015***	-0.018***	-0.011***	-0.016***	-0.005	-0.006*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Year FE	Yes	Yes	Yes	Yes	NA	NA
Cell FE	Yes	Yes	Yes	Yes	Yes	Yes
Climate controls	NA	NA	Yes	Yes	Yes	Yes
Country-year FE	NA	NA	NA	NA	Yes	Yes
Observations	257,856	75,840	178,640	53,411	178,626	53,377
R-squared	.406	.433	.398	.438	.438	.478

Notes: The table regresses the equation $Y_{ict} = locust_{ict} + \gamma_t + \gamma_i + \varepsilon_{ict}$. The first and second columns show the effect of locusts on conflict with cell and year fixed effects; the third and fourth columns show the same regression with climate controls added; and the fifth and sixth columns show the effect with cell fixed effects, country-year trends, and climate controls. Even columns show regressions restricted to cells that have experienced locust events in the past; odd columns show the effect on all cells. The unit of observation is the cell-year. Standard errors are clustered at the cell level.

5.7%, which is an 8% decrease in the probability of conflict. A locust sighting is associated with a 0.6 percentage point decrease in the probability of having conflict in a cell in any given year for cells in the locustHUB dataset. The probability of conflict at cells that have experienced locust sightings is 7.02%; after a swarm sighting, that becomes 6.42%, which is an 8.5% decrease in the probability of conflict.

Our main concerns are reverse causality, measurement error, and omitted variable bias. More peaceful areas could simply have more agriculture for locusts to eat, which would entail reverse causality. Measurement error might arise because the data rely on recording sightings and people might not report locusts in more conflict-prone areas since they might be too busy fighting. Finally, omitted variable bias might arise because locust outbreaks could be more severe following years with more rain.

3.2 Event study

We can further decompose locust sightings and conflict to the monthly level. To do so, we use a two-way fixed effects event study and estimate the effect of locust sightings on time trends. We run the event study regression by hand. In addition, this approach begins addressing some of the endogeneity issues. We relate cell-month variation in locust outbreaks to cell-month variation in conflict in an event study analysis. This allows us to account for time-invariant unobservables (using cell fixed effects) and time-varying unobservables (using month fixed effects).

We also control for rainfall. We do this because the year-level regressions might be smoothing out spikes and dips in conflict that are visible at the month level. The month-level analysis gives us some information on the temporal, short-term effects of locust outbreaks since it is possible that locust outbreaks change the probability of conflict only in the short run. For instance, farmers might substitute crops that are less attractive to locusts; farmers might diversify their income in a year where there are predictions for upsurges; or the FAO might increase their deployment of pesticides to prevent locust outbreaks from becoming problematic.

We run the following regression:

$$Y_{i,t} = \gamma_i + \gamma_t + \sum_{l \neq -1} \beta_l 1\{F_i = t - l\} + \varepsilon_{i,t}$$

Where $Y_{i,t}$ is a dummy that is equal to one if there is conflict in a particular cell i at month t, γ_i are cell fixed effects, γ_t are month fixed effects, l is the month relative to the current month t and F_i is equal to one if there is a locust sighting in a cell i at month t (which is our treatment variable). Because of the importance of soil moisture for locust birth, we show these regressions with and without precipitation controlled for. We also show these regressions for all cells in ACLED, and ever-treated cells only. Each coefficient should be read as: the change in probability of conflict occurring in a cell l months after a locust sighting, compared to the probability of conflict occurring in a cell without a locust sighting - compared to the difference between the two before the locust sighting.

Figure 6 shows the regression results in graphical form. These are collected in table form in table C1 in the appendix. The figures in panel 6b and panel 6d document a sudden decrease in the probability of conflict up to four months after a locust sighting. This sudden drop stabilizes six months after the locust sighting. This seems to indicate a strong but temporary effect of locusts on conflict. The first thing that should be noted is that the results are remarkably similar for treated cells and all cells. Coefficients are similar (though not identical), but standard errors change. The second thing that should be noted is that without rainfall, the parallel pre-trends assumption does not necessarily hold, since it seems five months before treatment there may be a decreased probability of rainfall in treated cells compared to untreated cells. There is some omitted variable bias that is captured by controlling for rainfall.

The results show that one month after a locust sighting, the probability of conflict in a treated cell compared to an untreated cell decreases by 0.42 percentage points, compared to the difference before treatment. Recall that the probability of conflict in a given cell in a given month is 2.3% - with a locust sighting, this probability becomes 1.88%, which is an 18.3% decrease in the probability of conflict. While the size of the coefficient is similar to that of the cell-year decom-

position explored previously, the change in probability is of a higher magnitude. Two months after a locust sighting, there is no change in the probability of conflict; but three months after a locust sighting, the probability of conflict once again decreases by half a percentage point. The change in the probability of conflict following treatment, compared to the probability of conflict in untreated cells, goes back to zero after six months following a locust sighting.

The event study model addresses some of the concerns listed above when looking at the pre-trends. Indications of reverse causality and omitted variable bias would appear in the coefficients in the months before a locust sighting. For example, if coefficients in the months before a locust sighting were negative, this would indicate that less fighting might lead to fewer locust sightings. The absence of statistically significant pre-trends is a positive sign for excluding the threat of reverse causality. Insignificant coefficients before locust sightings could indicate that concerns about reverse causality, and also omitted variable bias, could be laid to rest. Regarding measurement error, cell fixed effects can partially account for the concern that, for example, international observers would be in areas that see locusts but not conflict. While these concerns are by no means entirely assuaged by this approach, it is reassuring that there are no pre-trends in the results.

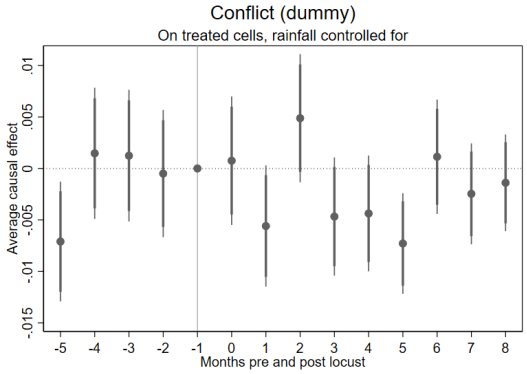
What could be the causing these results? This seems to point in the direction that locust outbreaks do not change the opportunity cost of violence. Rather, the decrease in the probability of conflict in treated areas compared to untreated areas indicates that either there are no goods left to appropriate or there are no resources to compete over (this was our second hypothesis). Alternatively, households may cope with the covariate shock by turning to community-oriented coping strategies, which then increase social cohesion and reduce the probability of conflict. In the subsequent sections, we attempt to follow several threads that may help explain the mechanisms that might be driving the results. First, we present a case study showing the relationship between incomes and locust outbreaks during the locust upsurge in 2013/2014 in Ethiopia. Then, we document some anecdotes which explain common narratives of conflict in the region. We finally present some shortcomings of our approach and present a solution for future research.

3.3 Robustness of results

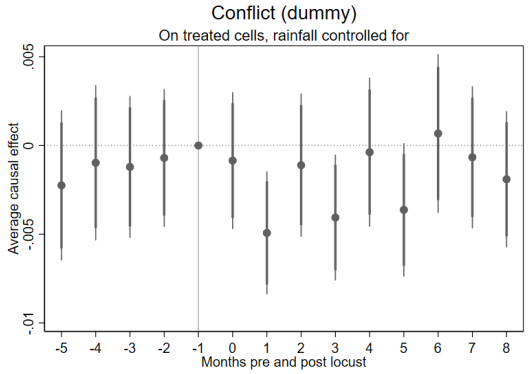
Recently, several studies have found that two-way fixed effects models with staggered treatment and multiple time periods - such as the one above - are unbiased for an average treatment effect if two conditions hold: that of parallel trends before treatment, but also that of constant treatment effect between groups and over time (De Chaisemartin and d'Haultfoeuille, 2022). This second condition is additional to standard differences-in-differences and is not always

Figure 6: **Effect of locust sightings on conflict**

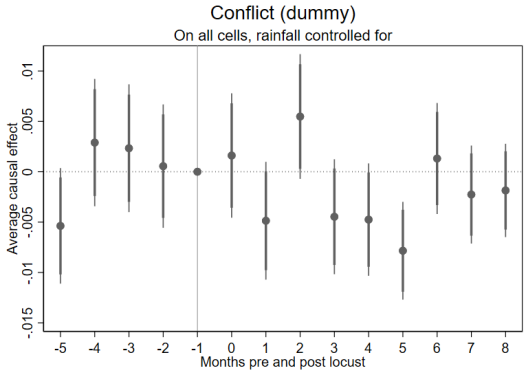
(a) Results, no controls, treated cells



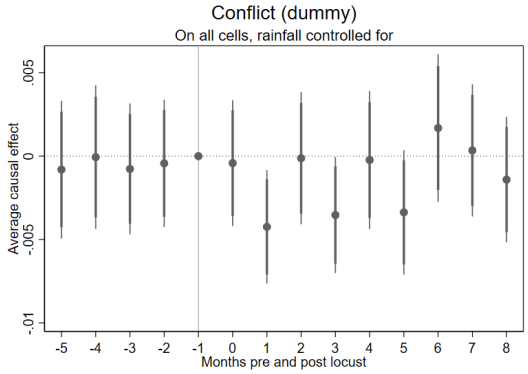
(b) Results, treated cells



(c) Results, no controls, all cells



(d) Results, all cells



Notes: The figure illustrates an event study of the probability of conflict following a locust sighting, by months. The first row shows results for treated cells; the second shows results for all cells. The first column shows results without controls; the second shows results with rainfall controlled for. Standard errors are robust. Dots are coefficients; lines show confidence intervals. Thin lines show confidence intervals at the 5% significance level, and thicker lines show confidence intervals at 10% significance.

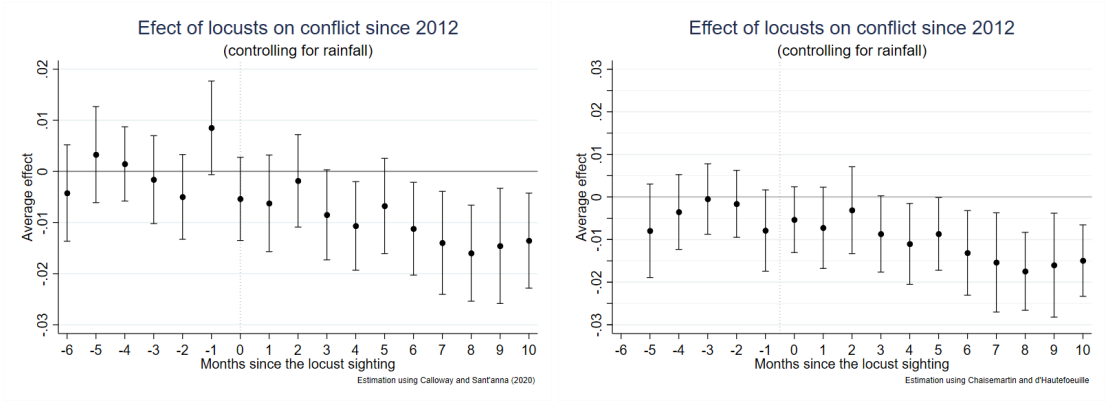
likely to hold. To make sure the results presented above are robust to such errors, we estimate a TWFE model following de Chaisemartin and D'Haultfœuille (2020) and Callaway and Sant'Anna (2021).

The results are presented below. Because the estimation proposed by Callaway and Sant'Anna (2021) does not handle a dataset such as the one we have well, we have estimated the effect of locust outbreaks in figure 7 from 2012 to 2020. We have run the same sample on the de Chaisemartin and D'Haultfœuille (2020) command to compare results. It needs to be noted that these event studies use varying base periods, which is the default in the `csdid` and `did_multiplegt` commands, meaning that the base period is the immediately preceding period. It is for this reason that there is no 0 coefficient at the (-1) time period like in the event studies in figure 6. Instead, we have a dotted line passing between the period before treatment (-1) and the period of treatment (0). This does not affect the interpretation of coefficients. Regardless, we have run the same regression using de Chaisemartin and D'Haultfœuille (2020) with a universal base pe-

riod to explore whether there is any difference between the results, which does not seem to be the case. This can be found in the appendix, in figure C1.

Reassuringly, the results are similar to the ones we had found previously, with a slight difference in terms of the timing of the effect. While previously, we had found an effect as early as one month after a locust sighting, now the negative effect is pushed back to four months after the locust sighting. The effect is also much smaller. Four months after a locust sighting, a cell with a locust sighting has an almost 1 percentage point lower probability of conflict compared to a cell that has not had a locust sighting. This effect remains constant up to ten months after the sighting. One could say, for instance, that ten months after a locust sighting, the probability of conflict in a cell with that locust sighting is 0.15 percentage points smaller than the probability of conflict in a cell that has not seen a locust sighting, compared to the difference before the locust sighting. In addition, both specifications reaffirm the parallel trends assumption is valid.

Figure 7: **Effect of locusts on conflict since 2012**



Notes: The dependent variable is a dummy for conflict. The figures illustrate an event study of the probability of conflict following a locust sighting, by months. The left-hand side shows estimations using Callaway and Sant'anna (2020). The right-hand side shows estimations using Chaisemartin and d'Hautefoeuille (2020). Standard errors are robust. Dots are coefficients; capped lines show confidence intervals at 5% significance.

4 Channels: Agriculture and the crop economy

In this section, we explore the effect of locusts on agriculture. We begin by exploring the effect of locust plagues on the prices of agricultural goods. We then make use of the LSMS-ISA and zoom into the case of Ethiopia in 2014 to explore the effects of locust events on incomes.

4.1 Prices

What could be the effect of locusts on local prices? We have been assuming until now that locusts reduce harvests and, thus, push up the prices of local goods. However, we would like

to see if that is actually the case. Using state-level data collected in a time series by the World Food Program, we can match locust events to local food prices. We have information on sales in 750 districts for 96 goods between 2000 and 2020. Descriptive statistics of these can be found in the appendix, section C.1.2.

Table 4 estimates the following regression to explore the relationship between specific crop prices and swarm sightings at the district-year level:

$$log(P_{sct} + 1) = locust_{sct} + \gamma_{ct} + \varepsilon_{ict}$$

Where $locust$ = 1 if a locust sighting is reported, the independent variable is $log(prices + 1)$, s is district (admin level 1), c is country, t is year, and γ_{ct} represents country-year fixed effects. Country-year fixed effects capture any linear trends that are at the country level and that varies from year to year that might be driving the results. We run the same regression on a greater number of commodities which are tracked by the World Food Program in figure 8. A figure of the effect of a locust sighting on all commodities we have price information for can be found in the appendix, table C2.

Table 4: **The effect of locusts on food prices**

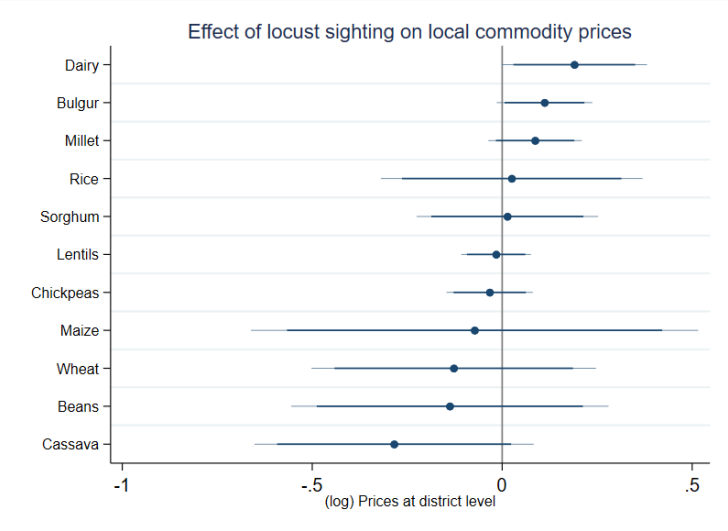
	Average	Millet	Sorghum	Maize	Wheat	Cassava	Bulgur	Rice	Chickpeas	Lentils	Maize	Beans	Wheat
Locust sighting	0.411*	0.087	0.014	-0.073	-0.127	-0.284	0.112	0.025	-0.033	-0.016	-0.073	-0.138	-0.127
	(0.205)	(0.062)	(0.121)	(0.299)	(0.190)	(0.186)	(0.063)	(0.175)	(0.057)	(0.046)	(0.299)	(0.212)	(0.190)
Controls	No	No	No	No	No	No	No	No	No	No	No	No	No
Country-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,579	960	1,347	3,439	2,752	1,234	154	5,420	361	666	3,439	2,849	2,752
R-squared	.918	.837	.891	.968	.919	.942	.998	.944	.966	.953	.968	.969	.919

Notes: The table explores the equation $logP_{sct} = locust_{sct} + \gamma_{ct} + \varepsilon_{ict}$, at the state-year level. Country-year fixed effects are included. Standard errors are clustered at the state level.

Table 4 shows that, at the district-year level, locust plagues increase prices overall, as well as some crops that are known to be preferred by locusts - such as millet and sorghum, though these increases are insignificant. Table 4 indicates that a locust sighting is associated with a 41.1% increase in average food prices, driven mainly by the increase in prices of teff, though the small sample size indicates caution when looking at this. However, recalling figure 3, there has been a steady increase in the price of teff throughout the last decade - the trend would be partially captured by the country-year fixed effects included in the regression.

Figure 8 shows that most commodity prices are unaffected by a locust outbreak, except dairy, bulgur, and millet. The figure indicates that a locust outbreak is associated with an around 10% increase in the price for dairy, bulgur, and millet (statistically significant at the 10% level only for dairy and bulgur). This shows that locust outbreaks do affect local staple

Figure 8: **The effect of locusts on food prices, expanded**



Notes: The table explores the equation $log(P_{sct} + 1) = locust_{sct} + \gamma_{ct} + \varepsilon_{ict}$, at the state-year level. Country-year fixed effects are included. Standard errors are clustered at the state level. The lines are confidence intervals at the 5% signficance level for light blue and 10% significance level for dark blue.

commodities by making them more expensive, and possibly rarer. The fact that the rest of the commodities are unaffected by a locust outbreak could be explained by connected markets. It is possible that for important commodities, a state's main market can make up for any missing quantities by bringing in commodities from other, unaffected regions. Therefore, the difference between affected and unaffected regions following a locust outbreak remains null. One curious effect, which can be found in the appendix, in figure C2, is that locust outbreaks are associated with a decrease in the price of some oil nuts such as groundnuts and sesame. While this result should be taken with a grain of salt, it could indicate that oil nuts are used as coping mechanisms. That is to say, households may dump them in markets to have any income at all for the season - the act of dumping these oil nuts in the market pushes their price down but gives households some safety net. There is anecdotal evidence that households usually use livestock in such a coping strategy (FAO, 2020), and there is some quantitative evidence backing this up (Helgeson et al., 2013). Perhaps oil nuts serve a similar purpose in this scenario.

4.2 Case study: Ethiopia 2014 – Mechanism

In the following section, we explore the relationship between locust sightings and agricultural incomes in a case study of Ethiopia in 2014. As has already been discussed, we chose Ethiopia for data availability reasons: the second wave of the LSMS-ISA survey conducted in Ethiopia, called ESS2, coincides with the beginning of an upsurge of locusts that was estimated to be particularly damaging in 2014. Furthermore, Ethiopia has been historically plagued by locusts,

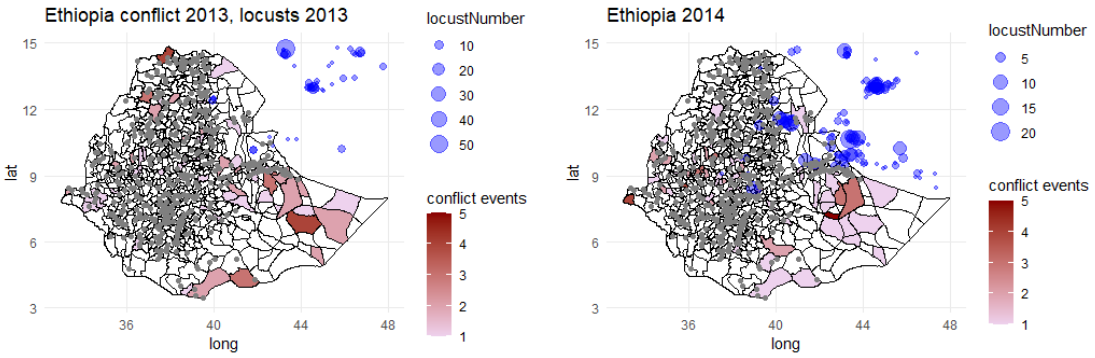
has a long history of conflict, and has a large share of agricultural homes, which makes it an appropriate setting to understand how locusts may affect the local rural economy.

4.2.1 The setting

Figure 9 maps conflict events in red shading, locust sightings in blue bubbles (where bigger bubbles represent more locusts sighted in that location), and households surveyed in Ethiopia in 2014 in gray dots. The map for 2013 (panel 9a) shows locust sightings and conflict events at the district level for 2013. The map for 2014 (panel 9b) shows locust sightings and conflict events for 2014. We can see that locust sightings and surveyed households intersect quite significantly. It is clear from the evolution of blue circles between 2013 and 2014 that there was an upsurge in locust sightings during these years. The north and east of Ethiopia, which is heavily surveyed in the LSMS-ISA, had the most sightings of locusts. In addition, conflict seems to have been important in the east of the country.

Figure 9: **A snapshot of Ethiopia in 2013 and 2014**

(a) Ethiopia in 2013 (b) Ethiopia in 2014



Notes: The figure maps conflict events in red shading, locust sightings in blue bubbles (where bigger bubbles represent more locusts sighted in that location), and households surveyed in Ethiopia in 2014 in gray dots. The map for 2013 (panel a) shows locust sightings and conflict events at the district level for 2013. The map for 2014 (panel b) shows locust sightings and conflict events for 2014.

Ethiopia has two main grain growing seasons, which coincide with the rainy seasons, called *belg* and *meher*[26]. The shorter season is *belg*, and runs from February, when crops are sown, to July, when farmers begin harvesting crops. The main growing season, which is also the most economically important one, is *meher*, which coincides with the ESS survey waves. *Meher* runs from May, when farmers begin sowing crops, and ends in February, at the end of the harvesting

[26]For more information, see: https://essp.ifpri.info/files/2011/02/ESSP2_WP16_Crop-Production-in-Ethiopia-Regional-Patterns-and-Trends.pdf and https://www.africa.upenn.edu/Hornet/welo1804.html

season (harvesting begins around November and ends in February). Any crop harvested between September and February can be considered a *meher* crop. Correspondingly, the first ESS survey begins around October, when all farmers have finished sowing and growing their *meher* crops, and the second ESS survey takes place in March-April when all farmers have finished harvesting their *meher* crops.

4.2.2 Results

We can begin by exploring the relationship between household incomes, their sources, and locust sightings. We estimate the regression below, plotted in table 10.

$$I(hj)_{d,2014} = \sum_{t}^{april} locust_{d,t} + \gamma_s + \varepsilon_d$$

Where $I(hj)$ is income from $j \in \{$crop, livestock, transfers, non-agricultural wage$\}$ source, d is district, γ_s are state fixed effects, and $\sum_t^{april} L_{d,t}$ is the sum of locust sightings $t \in 1, ..., 7, 12, 18$ months before April, 2014. We cannot control for district fixed effects, even if we run this regression at the household level, because all the variation in locust sightings is at the district level. For this reason, this should be read as a correlation.

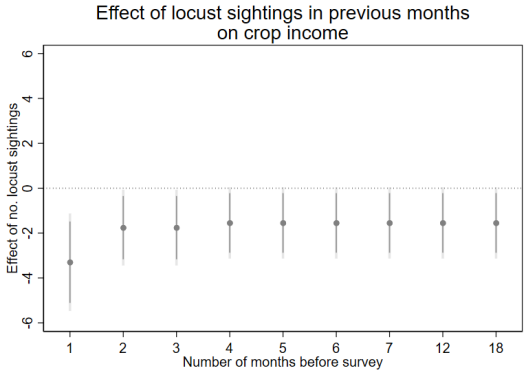
We run the regression on the income of all households surveyed by the second wave of the LSMS-ISA in Ethiopia. Locust sightings are at the district level (called woreda, or admin 3). Income variables were collected during the third and last visit of enumerators to households between March and April 2014. They were disaggregated by crop income, livestock income, income from transfers coming to the household, and non-agricultural wages. We then built a variable that is equal to the number of locust sightings $t \in 1, ..., 7, 12, 18$ months before April 2014 (when the households would have been surveyed at the latest), no matter what kind of locust sighting it is. We aggregated this to the district level.

We use the natural logarithm of income for ease of interpretation of the coefficient. Other specifications (e.g. using the natural logarithm of locust sightings) do not give different results (see appendix, table C3). Thus, each of our coefficients below should be read as: A change in locust sightings by one unit (Δ *locust*=1) is associated with a $(exp^{(\beta_1)} - 1) \times 100\%$ change in income. We cluster standard errors at the state (admin 2) level since we have 67 states in our sample.

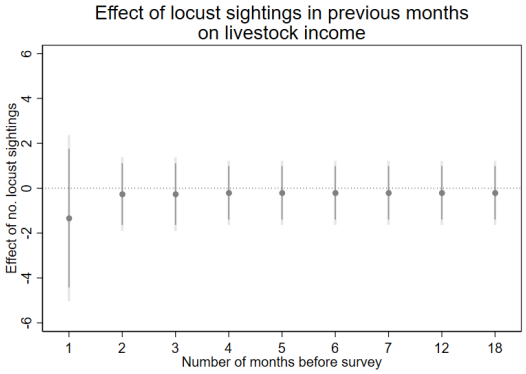
Table 10a shows that locust sightings affect crop incomes right away significantly and negatively. In particular, a locust sighting one month previously is associated with a 96.3% decrease in crop incomes. This is not a dynamic effect. Locust sightings 2 months or more previously have the same effect - this cross-sectional correlation indicates that there is a strong immediate negative association of locust sightings with crop incomes, which is less strong one month later

Figure 10: **Relationship between incomes and locust sightings**

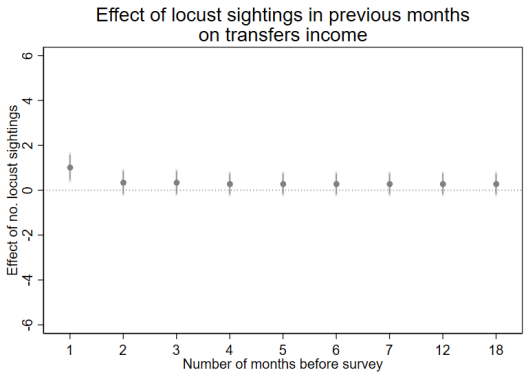
(a) Crop incomes



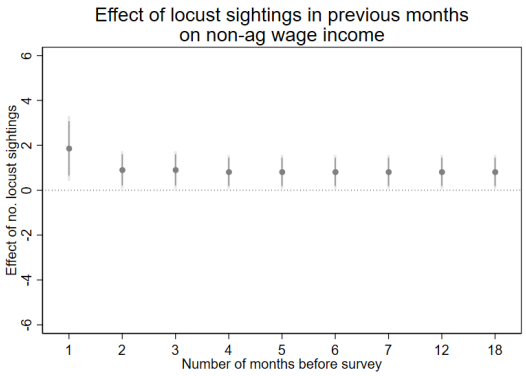
(b) Livestock incomes



(c) Transfers incomes



(d) Non-agricultural wage incomes



Notes: The figure shows the effect of locust sightings $t \in 1, ..., 7, 12, 18, ever$ on crop incomes (panel a), livestock incomes (panel b), transfers incomes (panel c) and all other incomes (panel d) of households surveyed March-April 2014. Dots show the coefficients, light gray lines show 95% confidence intervals, and dark gray lines show 90% confidence intervals. Standard errors are clustered at the state (admin 2) level.

and slowly dissipates over time.

Livestock incomes (panel 10b) do not seem to react to locust sightings, which is surprising - one would think that farmers would sell their livestock quickly following a locust event in order to meet basic income needs to feed their remaining livestock, thus pushing the value of livestock down (Brader et al., 2006; FAO, 2020)[27]. In fact, there is extensive evidence that livestock is used as insurance against weather shocks (Helgeson et al., 2013). It is possible that in a panel setting one would see such a negative effect on livestock income.

The response on transfers (panel 10c) and non-agricultural wage incomes (panel 10d) is intuitive - following a locust sighting, agricultural workers might rely on greater income from informal insurance networks such as family members in other parts of the country. There is anecdotal evidence that household members depart for urban areas to find non-agricultural wage work following a locust plague (Brader et al., 2006). The fact that locust sightings are as-

[27] see https://www.alnap.org/system/files/content/resource/files/main/796.pdf, page 8

sociated with an increasing reliance on non-agricultural wages seems to indicate that there is, to some extent, some substitution in sources of income occurring. The coefficient indicates that an increase of one locust sighting one month before the survey is associated with a 543.6% increase in non-agricultural wage incomes. Households may rely on social insurance mechanisms, state insurance mechanisms, or even loans (as anecdotal evidence suggests, according to Brader et al. (2006)), to offset losses from locust sightings, as well as substitution towards wage incomes.

Does this translate to food insecurity at the household level? To answer this last question, we can run the following regression at the household level to get an idea of the correlation between food insecurity and locust sightings:

$$Y_h = locust_h + \gamma_s + \varepsilon_h$$

Where Y_h is the number of months a household has been food insecure, locust is either the dummy variable that is equal to one if there has been a locust sighting in a district, or the number of locust sightings in a district, and γ_s are state fixed effects. Standard errors are robust. Running this regression indicates a locust sighting is associated with an increase in 0.63 months of food insecurity, and an increase of locust sightings by one is associated with an increase of 0.51 months of food insecurity (see table 5).

Table 5: **Months of food insecurity associated with locust sightings**

	Locust dummy	Locust sightings number
Months food insecurity	0.631*	0.519***
	(0.258)	(0.128)
N	5256	5256

Notes: The table shows the regression results for $Y_h = locust_h + \gamma_s + \varepsilon_h$. The dependent variable is months of food insecurity at the household level. The independent variable is a dummy equal to one in column 1 if there has been a locust sighting and the number of locust sightings that month in column 2, at the district level. Robust standard errors in parentheses, $*p < 0.05, **p < 0.01, ***p < 0.001$

5 Implications and discussion

The results above paint a consistent picture of how locusts affect people. To summarize, we found that a locust sighting is associated with a 0.5 percentage point decrease in the probability of having conflict in a cell in any given year, pushing the probability of a conflict occurring in any cell in any given year from 6.2% down to 5.7%. The event study showed that one month after a locust sighting, the probability of conflict in an affected cell compared to an unaffected cell decreased by 0.42 percentage points, compared to the difference before a locust outbreak.

This pushed the probability of conflict in any given cell in any given month from 2.3% to 1.8%, which is an 18.3% decrease in the probability of conflict.

We had three hypotheses for channels that could explain our results. The *appropriation* channel was that locusts reduce conflict by destroying anything to appropriate. In this scenario, locust plagues decrease harvests (and therefore incomes) and decrease conflict. Similarly, but subtly different, is the *resource competition* channel, which was the idea that locusts destroy any resources that could be competed over - areas for growing crops or grazing livestock. The *social cohesion* channel was that locust outbreaks increase social cohesion because households rely on informal insurance mechanisms such as family or community members who might be able to hire people from households that had crop losses. We would like to discuss the implication of our results for either of these stories, and for ease of reading, we will split them into two: a social cohesion channel, and an appropriation/resource competition channel.

The social cohesion channel: Our case study of Ethiopia in 2014, and the exercise correlating locust outbreaks and commodity prices, provide some evidence for the social cohesion hypothesis. We found that locust sightings do not change the prices of consumer goods at the state level. Markets may use coping strategies to weather the shock of losing harvests. Furthermore, in Ethiopia in 2014, we found that households in districts with locust sightings indicated that they have spent around half a month in food insecurity. We also found that districts with locust sightings are associated with significant losses in crop incomes and a large increase in non-agricultural wage incomes and transfer incomes.

This indicates several things. First, at the state level, markets hedge for commodity price volatility. We see this because commodity prices are not strongly affected by locust outbreaks. The commodities that see an increase in prices see a barely statistically significant increase in prices. This could be explained if markets were connected and could replace missing commodities from an affected region with commodities from an unaffected region, pushing price differences between unaffected and affected areas to zero. There is evidence in other parts of the world that it is possible for prices between regional markets to converge - for example, cell phone communication allows fishermen in Kerala to sell their fish in markets that are not saturated (Jensen, 2007). There is no reason to think no such similar thing occurs following locust outbreaks, especially if aid organizations such as the FAO get involved and aim to decrease barriers between markets. It seems unlikely that states that rely on food from other regions would get involved in conflict with other regions which might have necessary food.

The second thing to note is the shift to non-agricultural wage income. This indicates that following a locust outbreak, farmers may seek employment from other members of the community who do not rely on crop income. The increased interaction between members of a commu-

nity that had different economic activities could be creating a greater source of community in a region. The substitution in income sources could be creating social cohesion. The third thing to note is that the increase in transfer income supports anecdotal evidence that family members will often leave a locust-affected region to seek employment in urban areas (Brader et al., 2006). They are likely to send transfers to the family members that stayed behind. In addition, households may rely on family members that already live in other parts of the country. These informal insurance mechanisms may reduce the likelihood of conflict because people simply rely on one another more.

The evidence points in the direction that a locust outbreak is associated with a greater reliance on networks that reduces the probability of conflict by increasing, at least temporarily, social cohesion. This falls in line with literature such as Ager and Ciccone (2018a), where a community with greater environmental risk is more likely to be more tight-knit (in the case of Ager and Ciccone (2018a), greater environmental risk is associated with more people in religious communities). There is also anecdotal evidence that natural disaster relief efforts create a sense of community. Following hurricanes, people will often come together to help each other out and reconstruct their communities together.

The appropriation and resource competition channel: This hypothesis goes back to the economic origins of conflict. There are two aspects to this hypothesis. First, it could be that following a locust outbreak there are no goods left to appropriate. That is, any crops that might be attractive to thieves would have been eaten by locusts. This would be in line with literature such as Crost and Felter (2020), where banana plantations in the Philippines are more likely to cause rebellion, but only if they are exported bananas, and not if they are the locally consumed type of banana. The mechanism here would be that in the absence of valuable exportable goods, there would not be any conflict. We do not find evidence for this narrative per se, but it does not mean that it is not one factor driving our results. Further venues of research would disaggregate by crop income, based on which crop is grown in the plot of land. This would provide further information on whether the appropriation is an incentive.

More convincing, perhaps, is the narrative that locusts destroy the resource that leads to competition and therefore to conflict - land. The desert locust is active in the same regions that are known for conflicts between nomadic pastoralists and sedentary agriculturalists. Locusts may eliminate resource competition between these groups of people. There is significant anecdotal evidence that resource competition is a leading source of conflict throughout the Sahel - following droughts, nomadic herders who live off their livestock often encroach into agricultural lands, seeking places for grazing. Examples include the Fulani in Nigeria(ICG, 2017, 2021),

the Dinka in Sudan (Eberle et al., 2020), the Tuaregs in Mali[28], and fighting in Laikipia county in Kenya[29]. This kind of conflict is extensively documented by Eberle et al. (2020).

The common thread in these stories is resource competition: valuable fertile land is coveted by both pastoralists and agriculturalists. If there are no resources to encroach upon, there would simply not be any fighting. Take the example of fighting in Laikipia in Kenya since September 2022. Following a drought in Olmoran, herders were faced with a lack of grazing areas for their livestock. To find ways to provide for their families, herders moved to the less affected regions in Laikipia, which are areas settled by farmers who control water sources and use them for their crops. This led to clashes between both groups. Each group accused the other of antagonistic behavior: cattle raiding by agriculturalists and so-called illegal grazing by herders. The fighting has led to at least 95 deaths since September[30]. Similar stories occur all over the Sahel, though locust outbreaks are not drivers. The lack of a resource to compete over might reduce tensions. Locust outbreaks may reduce conflict by eliminating the most common source of conflict in the region.

Both mechanisms have their own merits. There is evidence for both the appropriation mechanism and the social cohesion mechanism. Disentangling between them is not possible without additional analyses, but it could be an area for further research. In fact, it is also possible that both are occurring simultaneously. For instance, the situation could be so dire in these conflict-ridden regions that even locust outbreaks cannot make conflict worse. The policy implication here is that conflict and poverty both need to be addressed in a systemic manner so that communities hard-hit by locust outbreaks can focus on their recovery. Conflict and natural disasters make everyone worse off. The main takeaway of this study should not be that locust outbreaks reduce conflict, but that locust outbreaks at least do not contribute to more conflict. Rather, locust outbreaks destroy livelihoods and contribute to famine, and policymakers can focus their limited resources on those issues instead.

Endogeneity: Our results should be taken with a grain of salt. There are some sources of endogeneity in the study which have not been addressed. First, locust reportings could be happening in places that have more economic development and less conflict - more peaceful areas could simply have more agriculture for locusts to eat. This would indicate some reverse causality. Furthermore, there may be reporting bias in how the dataset is constructed. The data rely on inputting sightings, which would entail either having access to technology that can record the sightings (a cell phone, computer, or landline), or access to a government worker

[28]https://www.middleeasteye.net/news/sahel-fire-malian-refugee-crisis-world-forgot
[29]https://www.voanews.com/a/kenya-deadly-land-invasions-blamed-on-political-incitement-/3974615.html
[30]https://www.crisisgroup.org/africa/horn-africa/kenya/drought-and-conflict-laikipia-kenya

Table 6: **The effect of locust swarms on conflict**

	Conflict
Swarm dummy	-.0196**
	(.0061)
Cell FE	✓
Country-year FE	✓
Observations	257760

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

who might be able to record the sightings. If a sighting is recorded, it entails that there is an enumerator present. There is a saying that goes, "if a tree falls in the forest and there is no one there to hear it, did it make a sound?". It is possible that there are locusts flying past regions that have conflict but people do not record them. The way this could be addressed is by looking at how the most extreme locust events affect conflict. These extreme events are more likely to be reported no matter what else is going on at the moment. If these also lead to similar results, it could indicate that measurement error is not hugely consequential in our analysis.

For example, if we limit our sample to only swarm sightings, which are much less likely to be overlooked - since these are large gatherings of flying locusts -, we can already see that the yearly differences-in-differences estimate in table 6 is very similar to the one we had found earlier. A swarm sighting is associated with a 1.9 percentage point decrease in the probability of conflict in any given cell in any given year. However, a more methodical examination is required to establish the robustness of our results in the face of potential measurement error.

We would like to address the potential bias in our estimation by using a projected locust observation as an instrumental variable in future research. To do so, we could rely on a model created by the National Oceanic and Atmospheric Association (NOAA) which simulates locust movements called HYSPLIT. The model uses climate inputs to project where a locust may go three days into the future. Figure C4 in the appendix shows an example of an output given by the model based on an actual locust observation. Locusts travel far distances over a short period of time. We are therefore able to input each locust observation into HYSPLIT and receive a landing latitude and longitude three days later. We can, in addition, simulate the landing location, three days later, of these simulated locusts. We can iterate these over several times and create a hypothetical locust observation n months into the future. This would allow us to build a new, hypothetical, world of locust observations one month in advance. The instrument would be observed at the cell-month level. For a given cell i in a given month m, the instrument would be the number of locust observations n months into the future. Therefore, the instrument would be labeled $z_{i,m+n}$.

The rationale for the instrument is that locusts move downwind and these models can simulate, but not predict entirely, where locusts may land. The relevance of the instrument would be that hypothetical locusts would be strongly correlated to their real observations. The exogeneity of the instrument would be that locusts n months in advance do not affect conflict today. The exclusion restriction is, thus, that locusts n months in advance only affect conflict today via their actual location today[31].

6 Conclusion

We were interested in exploring the effects of desert locust outbreaks on conflict over the last two decades in Africa. We found that a locust sighting is associated with a 0.5 percentage point decrease in the probability of having conflict in a cell in any given year, pushing the probability of a conflict occurring in any cell in any given year from 6.2% down to 5.7%. The event study showed that one month after a locust sighting, the probability of conflict in an affected cell compared to an unaffected cell decreased by 0.42 percentage points, compared to the difference before a locust outbreak. This pushed the probability of conflict in any given cell in any given month from 2.3% to 1.8%, which is an 18.3% decrease in the probability of conflict.

Exploring the effects of the locust outbreak in Ethiopia in 2014, we found that locust sightings do not change the prices of consumer goods at the state level. Markets may use coping strategies to weather the shock of losing harvests. Furthermore, in Ethiopia in 2014, we found that households in districts with locust sightings indicated that they have spent around half a month in food insecurity. We also found that districts with locust sightings are associated with significant losses in crop incomes and a large increase in non-agricultural wage incomes and transfer incomes.

We explored some mechanisms that could explain our results. The appropriation mechanism was that locusts reduce conflict by destroying anything to appropriate. In this scenario, locust plagues decrease harvests (and therefore incomes) and decrease conflict. The resource competition mechanism was the idea that locusts destroy any resources that could be competed over - areas for growing crops or grazing livestock. In fact, the desert locust is active in the same regions that are known for conflicts between nomadic pastoralists and sedentary

[31]The science behind using HYSPLIT as an instrument relies on the idea that locusts only travel in specific kinds of winds. They move, for example, northwards from West Africa across the Sahara and the Mediterranean Sea during the few days in which there are low-pressure points (hot wind going towards cold areas) because the winds are warm. Locust populations move downwind on warm winds (FAO and WMO, 2016). They usually take off about 2 to 3 hours after sunrise. In sunny conditions, they can take off in temperatures of at least 15 to 17 degrees Celsius. Under cloudy conditions, take-off occurs when temperatures reach 23 to 26 degrees Celsius. Under cooler conditions, take-off can be delayed to some 4 to 6 hours after sunrise. Locusts generally will not take off if winds are greater than 6 to 7 m/s (11.7 to 13.6 knots). The hotter the wind, the greater the distance traveled per day (FAO and WMO, 2016). Turbulent winds reduce the area density of swarms (FAO and WMO, 2016). Downwind movement eventually brings locusts into zones of wind convergence, where they accumulate (FAO and WMO, 2016).

agriculturalists. Locusts may eliminate resource competition between these groups of people. The social cohesion mechanism explored the idea that locust outbreaks increase social cohesion because households rely on informal insurance mechanisms such as family or community members who might be able to hire people from households that had crop losses. There was some evidence in our analysis that locust outbreaks are associated with income substitution and reliance on networks.

Further research would address some shortcomings in our study, in particular by instrumenting for locust outbreaks by using simulated locust outbreaks based on a model of wind direction and speed. In addition, further research could address each of the mechanisms listed above to quantify to what extent each of these may be playing a role in the results.

Our study contributes to the growing literature on the social and economic effects of natural disasters, and in particular of locust outbreaks. A better understanding of these can help prevent loss of life and conflict. We know locust outbreaks are associated with suffering, worsening poverty, and famines. A lot of the regions affected by locust outbreaks are also conflict-ridden. The fact that a locust outbreak reduces the probability of conflict in such regions could indicate that the situation could be so dire that even locust outbreaks cannot make conflict worse. The main takeaway of this study should not be that locust outbreaks reduce conflict, but that locust outbreaks at least do not contribute to more conflict. Rather, locust outbreaks destroy livelihoods and contribute to famine, and policymakers can focus their limited resources on those issues instead.

C.1 Supplementary material

This Appendix accompanies the paper "A plague on both your houses! How locust outbreaks affect conflict".

C.1.1 Results, expanded

The following table shows the results of the following regression:

$$Y_{i,t} = \gamma_i + \gamma_t + \sum_{l \neq -1} \beta_l 1\{F_i = t - l\} + \varepsilon_{i,t}$$

Where $Y_{i,t}$ is a dummy that is equal to one if there is conflict in a particular cell i at month t, γ_i are cell fixed effects, γ_t are month fixed effects, l is the month relative to the current month t and F_i is equal to one if there is a locust sighting in a cell i at month t (which is our treatment variable). Because of the importance of soil moisture for locust birth, we show these regressions with and without precipitation controlled for. We also show these regressions for all cells in ACLED, and for ever-treated cells only. Each coefficient should be read as: the change in probability of conflict occurring in a cell l months after a locust sighting, compared to the probability of conflict occurring a cell without a locust sighting - compared to the difference between the two before the locust sighting.

Table C1: **Regression results, dependent variable: conflict**

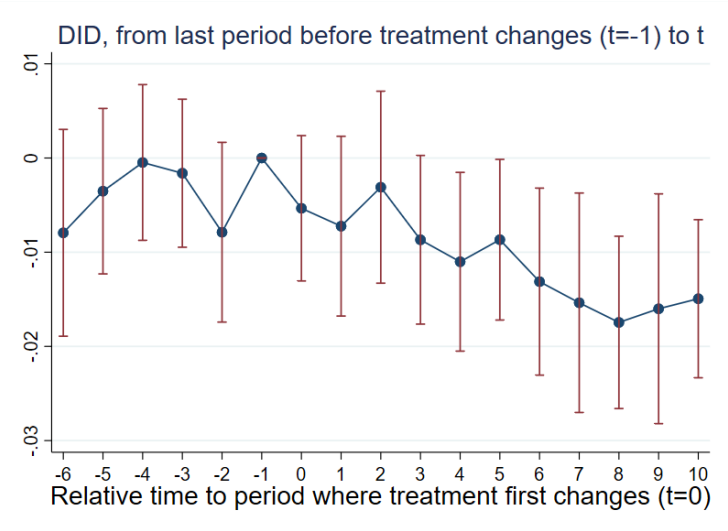
	(1)	(2)	(1)	(2)
1 month later	-0.006	-0.00493**	-0.005	-0.00424*
	(0.003	0.002	0.003	0.002
2 months later	0.005	-0.001	0.005	0.000
	0.003	0.002	0.003	0.002
3 months later	-0.005	-0.00406*	-0.004	-0.00353*
	0.003	0.002	0.003	0.002
4 months later	-0.004	0.000	-0.005	0.000
	0.003	0.002	0.003	0.002
5 months later	-0.00729**	-0.004	-0.00784**	-0.003
	0.002	0.002	0.002	0.002
6 months later	0.001	0.001	0.001	0.002
	0.003	0.002	0.003	0.002
7 months later	-0.002	-0.001	-0.002	0.000
	0.003	0.002	0.002	0.002
8 months later	-0.001	-0.002	-0.002	-0.001
	0.002	0.002	0.002	0.002
0 months	0.001	-0.001	0.002	0.000
	0.003	0.002	0.003	0.002
2 months previously	0.000	-0.001	0.001	0.000
	0.003	-0.002	-0.003	-0.002
3 months previously	0.001	-0.001	0.002	-0.001
	0.003	0.002	-0.003	-0.002
4 months previously	0.001	-0.001	0.003	0.000
	0.003	0.002	0.003	0.002
5 months previously	-0.00710*	-0.002	-0.005	-0.001
	0.003	0.002	0.003	0.002
Rainfall	0.0000263**		0.0000231***	
	0.00		0.00	
N	869000	619164	2954600	2089360

Standard errors in alternating rows

*$p < 0.05$, *$p < 0.01$, *p*< 0.001

Notes: The table gives the values of the regression shown in figure 6

Figure C1: **Effect of locusts on conflict since 2012 using Chaisemartin and d'Hautefoeuille (2020) universal base period**

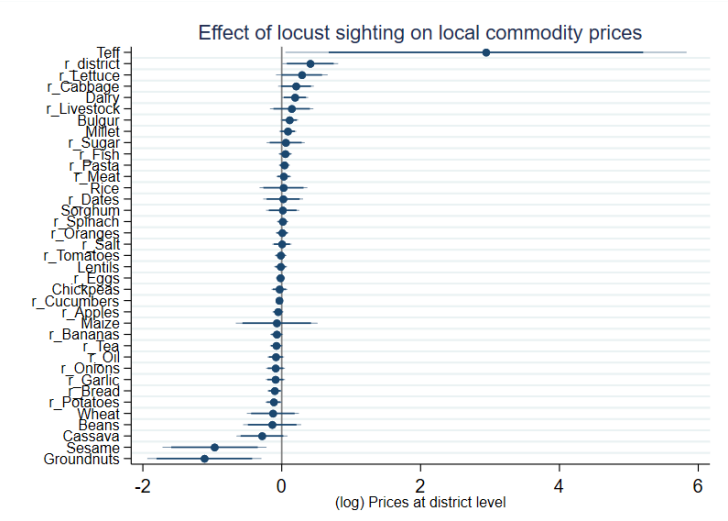


Notes: The dependent variable is a dummy for conflict. The figures illustrate an event study of the probability of conflict following a locust sighting, by months. Standard errors are robust. Dots are coefficients; capped lines show confidence intervals at 5% significance.

C.1.2 Prices and locusts

Figure C2 shows the effect of locusts on more food prices.

Figure C2: **The effect of locusts on food prices, expanded**



Notes: The table explores the equation $logP_{sct} = locust_{sct} + \gamma_{ct} + \varepsilon_{ict}$, at the state-year level. Country-year fixed effects are included. Standard errors are clustered at the state level. The lines are confidence intervals at the 5% signficance level for light blue and 10% significance leve for dark blue.

The table C3 shows how locust sightings at the district level t months before the ESS2 affect different sources of income at the household level. Both locust numbers and incomes are log transformed and the coefficient should thus be read as: a 1% increase in number of locust sightings t months before the survey is associated with a β% change in income.

Table C2: **Descriptive statistics of prices of all goods in WFP dataset**

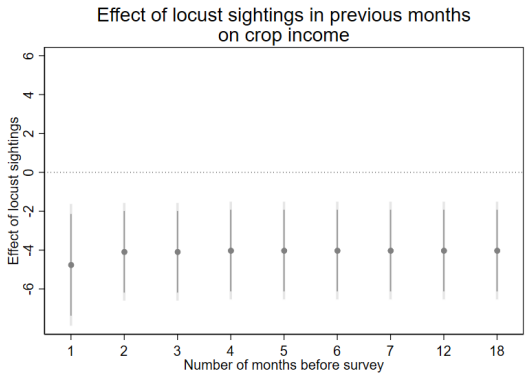
	n	Mean	sd	Min	Max
p_Apples	327	4.418145	2.16427	0.5128236	8.131824
p_Avocados	132	5.905297	0.5802842	4.97097	7.601402
p_Bananas	322	5.022785	2.163726	0.5255172	7.299347
p_Beans	3041	5.753051	2.645254	0.3138509	12.40434
p_Beetroots	133	2.724224	1.58949	1.350063	5.774086
p_Bitterball	30	5.969256	0.4446642	4.847136	6.59868
p_Blackberry	18	7.773374	0.1267054	7.4868	7.967904
p_Bread	1336	3.0511	2.060093	0.14842	8.131824
p_Broccoli	18	7.840347	0.2565605	7.449498	8.188967
p_Buckwheat	164	3.904919	1.381053	2.166049	6.789028
p_Bulgur	164	5.856142	4.171709	0.538363	12.72738
p_Cabbage	714	3.2456	1.593916	0.1849223	7.644441
p_Carrots	666	3.675611	1.679522	0.1590103	7.567518
p_Cashew	12	5.03006	0.995307	3.771323	6.428105
p_Cassava	1366	5.640108	1.770155	0.2426847	10.53489
p_Cauliflo∼r	51	4.29803	2.559918	1.114141	7.863651
p_Causeni	1	2.014903	.	2.014903	2.014903
p_Charcoal	72	7.500761	1.259997	5.527841	8.839567
p_Chicken	11	7.73593	0.0834864	7.615287	7.876549
p_Chickpeas	388	4.764681	2.657033	0.6061358	8.629628
p_Chili	112	6.78808	3.689422	1.340979	10.93134
p_Chisinau	1	2.008447	.	2.008447	2.008447
p_Cocoa	89	5.338338	1.414157	3.628543	8.63297
p_Coconut	24	5.755747	0.2335564	5.393324	6.121835
p_Coffee	75	4.397138	2.701754	1.394673	9.667829
p_Cornstarch	21	5.749527	0.233278	5.344724	6.127493
p_Cotton	4	5.727144	0.1260217	5.583496	5.860786
p_Couscous	86	3.966572	2.254801	1.343735	6.775366
p_Cowpeas	330	5.171191	2.563127	0.2363888	10.42492
p_Cucumbers	209	4.427542	2.551416	0.3701833	7.540444
p_Dairy	1230	4.744622	2.483065	0.168831	10.90643
p_Dates	100	7.202075	0.488613	5.929589	8.370821
p_Eggplants	101	4.894583	2.014591	1.070605	9.903538
p_Eggs	1147	3.909106	2.319491	0.2701607	12.27817
p_Fat	101	3.590652	0.2552131	2.977823	4.286737
p_Feed	10	8.82692	0.251795	8.43282	9.242824
p_Fish	1068	6.96491	2.694639	0.6417223	10.60496
p_Fonio	48	7.056882	1.614093	3.625496	9.633489
p_Gari	82	8.277564	1.812947	1.342865	10.248
p_Garlic	396	6.09757	2.709388	1.071612	9.964159
p_Groundnuts	827	5.463341	2.452427	0.2070142	10.72353
p_Guava	27	5.838375	0.9015623	4.834845	7.545742
p_Leafyveg	12	5.47249	0.282446	4.960499	5.851363
p_Lemons	23	5.744008	0.3464663	4.900752	6.50279
p_Lentils	749	5.026291	1.914539	0.5873421	11.49337
p_Lettuce	17	7.04034	0.1834134	6.782755	7.367919
p_Linseed	2	8.103772	0.013284	8.094378	8.113165
p_Livestock	309	10.61121	1.369295	6.534515	15.0293
p_Maize	3751	4.794173	2.730145	0.1211551	11.90159
p_Mangoes	43	6.942619	0.8543683	5.841794	8.227108
p_Maracuja	40	6.531755	0.3464174	5.794427	7.030465
p_Meat	2002	6.756425	2.631051	0.9776548	15.54947
p_Millet	1039	5.146416	2.233918	0.1947441	12.18076
p_Nigerseed	5	7.769956	0.5547591	6.781058	8.055475
p_Noodles	99	2.247879	0.4407166	1.480468	3.196494
p_Oil	2955	5.666963	2.241845	0.7167032	11.22526
p_Okra	24	5.680836	0.3474513	5.126729	6.331946
p_Onions	1215	4.007875	2.163515	0.233078	11.19954
p_Oranges	155	3.182398	2.622898	0.1768597	7.244942
p_Papaya	60	5.988133	0.8468668	4.612021	7.840969
p_Parsley	57	4.174727	0.4249506	3.093107	5.17754
p_Pasta	520	4.16979	2.528325	0.2623643	10.35824
p_Peanut	76	6.610895	0.7358284	5.288477	7.908383
p_Peas	355	4.540521	2.187581	1.219383	10.33947
p_Peppers	135	5.773549	2.641911	0.5772392	9.734398
p_Pigeonpeas	40	8.210125	1.073195	5.347417	9.093919
p_Plantains	369	4.56235	2.168923	0.2164155	8.068573
p_Potatoes	1788	3.927763	2.201907	0.1185235	11.6742
p_Poultry	5	5.065672	0.0872883	4.939561	5.160893
p_Prawn	15	10.70965	0.3860761	10.11797	11.346
p_Pulses	44	6.769358	1.367151	0.7935806	7.539849
p_Pumpkin	51	5.805965	0.687118	4.787492	7.108829
p_Quinoa	30	6.54288	0.464122	5.476463	7.554668
p_Radish	14	5.17538	0.1203273	4.976734	5.406424
p_Rice	5811	5.539475	2.75776	0.3435593	14.56
p_Salt	1003	3.951629	2.275246	0.103098	8.537192
p_Semolina	48	4.755728	1.05106	1.722771	6.692084
p_Sesame	140	5.981968	1.834784	1.807606	10.87964
p_Shrimps	8	8.609143	1.240166	6.111467	9.542159
p_Snail	4	7.16918	0.2764601	6.814817	7.414282
p_Sorghum	1549	5.4978	2.087319	0.189242	12.2019
p_Soy	113	6.592689	1.916933	0.5074867	10.68481
p_Spinach	141	3.998742	2.702428	0.2613381	8.216358
p_Squashes	32	2.390906	3.040054	0.187842	6.94751
p_Sugar	2175	4.698501	2.360208	0.2847695	10.77155
p_Taro	19	5.247773	0.417329	4.792203	6.16804
p_Tea	508	4.99044	2.032619	1.460842	9.263383
p_Teff	18	6.906704	1.773039	3.650658	8.12148
p_Tomatoes	1068	4.603935	2.057217	0.2723146	9.79586
p_Tortilla	53	1.283421	0.8882028	0.3602643	2.677362
p_Walnuts	31	5.040175	0.1771994	4.672119	5.254017
p_Water	55	1.337334	1.579447	0.1906203	6.862105
p_Wheat	3045	4.321038	2.175854	0.1816269	12.37765
p_Yam	753	5.206619	1.821036	0.4205362	11.10297
p_Yeast	13	6.881091	1.560567	3.54818	8.810634
p_Zucchini	40	4.121303	1.933091	1.452741	5.992173

C.1.3 HYSPLIT instrument

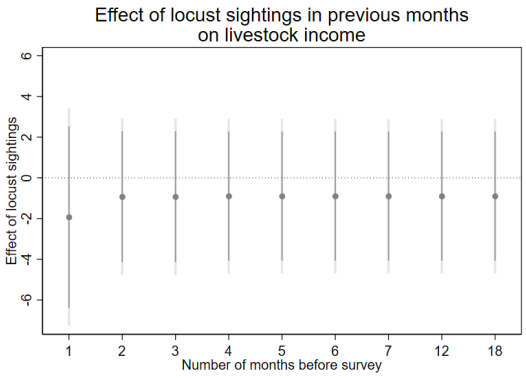
We would like to address the potential bias in our estimation by using a projected locust observation as an instrumental variable in future research. To do so, we could rely on a model created by the National Oceanic and Atmospheric Association (NOAA) which simulates locust movements called HYSPLIT. The model uses climate inputs to project where a locust may go three days into the future. Figure C4 shows an example of an output given by the model based

Figure C3: **Relationship between incomes and locust sightings (locust number log transformed)**

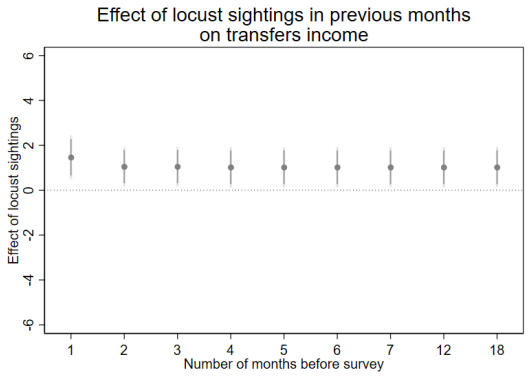
(a) Crop incomes



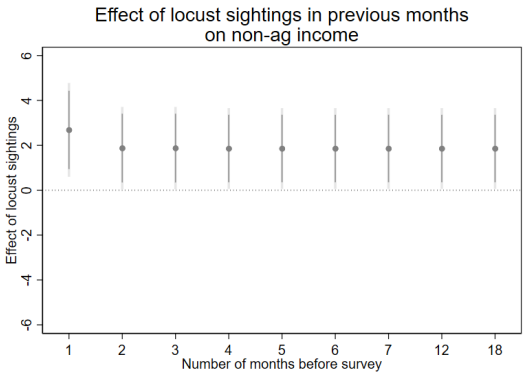
(b) Livestock incomes



(c) Transfers incomes



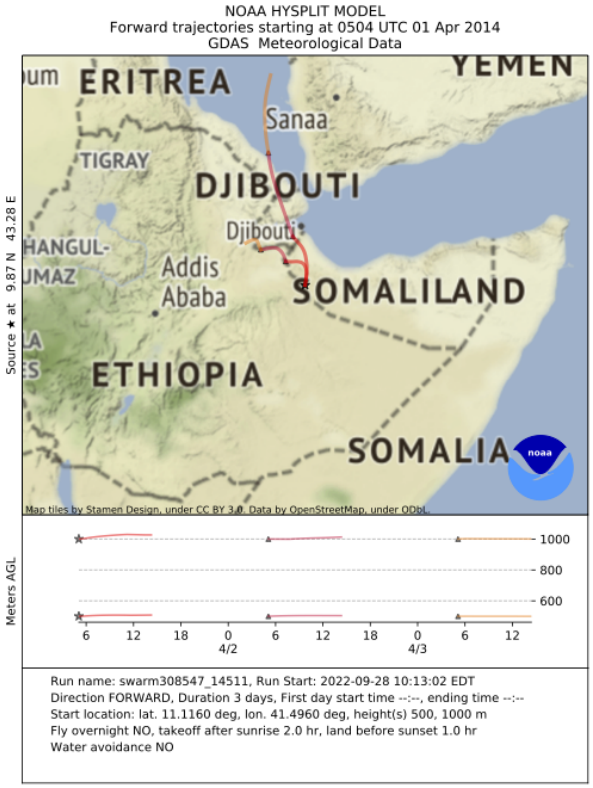
(d) Non-agricultural wage incomes



Notes: The figure shows the effect of log (locust sightings + 1) $t \in 1, ..., 7, 12, 18$ on crop incomes (panel a), livestock incomes (panel b), transfers incomes (panel c) and all other incomes (panel d) of households surveyed March-April 2014. Dots show the coefficients, light gray lines show 95% confidence intervals, dark gray lines show 90% confidence intervals. Standard errors are clustered at the state (admin 2) level.

on an actual locust observation. Locusts clearly travel far distances over a short period of time. We are therefore able to input each locust observation into HYSPLIT, and receive a landing latitude and longitude three days later. We can, in addition, simulate the landing location, three days later, of these simulated locusts. We can iterate these over several times and create a hypothetical locust observation n months into the future. This would allow us to build a new, hypothetical, world of locust observations one month in advance. The instrument would be observed at the cell-month level. For a given cell i in a given month m, the instrument would be number of locust observations n months into the future. Therefore, the instrument would be labeled $z_{i,m+n}$.

Figure C4: Illustration of output from HYSPLIT



Notes: The figure shows output from the HYSPLIT model from a locust swarm and adult groups from the third of March, 2013 in Cairo, Egypt. This shows, at the 0.5x0.5 cell level, how a these groups might moved over the course of several days given wind direction, temperature, and pressure. Source: FAO and WMO, 2016.

References

Abbott, A. (2021). Covid's mental-health toll: how scientists are tracking a surge in depression. *Nature 590*(7845), 194–195.

Adoho, F., S. Chakravarty, D. T. Korkoyah, M. Lundberg, and A. Tasneem (2014). *The impact of an adolescent girls employment program: The EPAG project in Liberia*. The World Bank.

Ager, P. and A. Ciccone (2018a). Agricultural risk and the spread of religious communities. *Journal of the European Economic Association 16*(4), 1021–1068.

Ager, P. and A. Ciccone (2018b). Agricultural risk and the spread of religious communities. *Journal of the European Economic Association 16*(4), 1021–1068.

Alon, T., S. Coskun, M. Doepke, D. Koll, and M. Tertilt (2021, April). From mancession to shecession: Women's employment in regular and pandemic recessions. Working Paper 28632, National Bureau of Economic Research.

Alon, T. M., M. Doepke, J. Olmstead-Rumsey, and M. Tertilt (2020). The impact of covid-19 on gender equality. Technical report, National Bureau of economic research.

Anderson, W., C. Taylor, S. McDermid, E. Ilboudo-Nébié, R. Seager, W. Schlenker, F. Cottier, A. de Sherbinin, D. Mendeloff, and K. Markey (2021). Violent conflict exacerbated drought-related food insecurity between 2009 and 2019 in sub-saharan africa. *Nature Food 2*(8), 603–615.

Attanasio, O., A. Guarín, C. Medina, and C. Meghir (2017). Vocational training for disadvantaged youth in colombia: A long-term follow-up. *American Economic Journal: Applied Economics 9*(2), 131–43.

Attanasio, O., A. Kugler, and C. Meghir (2011). Subsidizing vocational training for disadvantaged youth in colombia: Evidence from a randomized trial. *American Economic Journal: Applied Economics 3*(3), 188–220.

Baker, R. B., J. Blanchette, and K. Eriksson (2020). Long-run impacts of agricultural shocks on educational attainment: Evidence from the boll weevil. *The Journal of Economic History 80*(1), 136–174.

Bandiera, O., R. Burgess, M. Goldstein, N. Buehren, S. Gulesci, I. Rasul, and M. Sulaiman (2014). Women's empowerment in action: evidence from a randomized control trial in africa.

Banerjee, A., E. Duflo, G. Postel-Vinay, and T. Watts (2007, feb). Long run health impacts of income shocks: Wine and phylloxera in 19th century france. Technical report.

Bank, W. (2018, November). *Migration from Venezuela to Colombia*. World Bank, Colombia.

Bazzi, S., R. A. Blair, C. Blattman, O. Dube, M. Gudgeon, and R. Peck (2022). The promise and pitfalls of conflict prediction: evidence from colombia and indonesia. *Review of Economics and Statistics 104*(4), 764–779.

Bazzi, S. and C. Blattman (2014). Economic shocks and conflict: Evidence from commodity prices. *American Economic Journal: Macroeconomics 6*(4), 1–38.

Becerra, O., M. Eslava, M. Fernandez, M. Isaacs, and D. Perez-Reyna (2020, November). *Empleo femenino durante la crisis del COVID-19*. Universidad de los Andes.

Becerra, O., M. Eslava, M. Fernández, M. Isaacs, D. Perez-Reyna, et al. (2020). Empleo femenino durante la crisis del covid-19.

Becerra, O., M. R. C. Garcia, N. Romero, and C. Pecha (2021). Mercado laboral en la crisis del covid-19. Technical report, IPA.

Beerli, A., P. Ruh, D. Kistler, and A. Rom (2017). Review on the impact of vocational skills development (vsd) in developing countries. Commissioned by the Swiss Agency forDevelopment and Cooperation (DEZA).

Bentzen, J. (2020). In crisis, we pray: Religiosity and the covid-19 pandemic.

Bentzen, J. S. (2021). In crisis, we pray: Religiosity and the covid-19 pandemic. *Journal of Economic Behavior & Organization 192*, 541–583.

Berman, N. and M. Couttenier (2015, oct). External shocks, internal shots: The geography of civil conflicts. *Review of Economics and Statistics 97*(4), 758–776.

Berman, N., M. Couttenier, D. Rohner, and M. Thoenig (2017, jun). This mine is mine! how minerals fuel conflicts in africa. *American Economic Review 107*(6), 1564–1610.

Blattman, C. and J. Annan (2015). Can employment reduce lawlessness and rebellion? a field experiment with high-risk men in a fragile state. Technical report, National Bureau of Economic Research.

Blattman, C. and J. Annan (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.

Blattman, C., J. C. Jamison, and M. Sheridan (2017). Reducing crime and violence: Experimental evidence from cognitive behavioral therapy in liberia. *American Economic Review 107*(4), 1165–1206.

Blattman, C. and L. Ralston (2015). Generating employment in poor and fragile states: Evidence from labor market and entrepreneurship programs. *Available at SSRN 2622220*.

Blewitt, C., M. Fuller-Tyszkiewicz, A. Nolan, H. Bergmeier, D. Vicary, T. Huang, P. McCabe, T. McKay, and H. Skouteris (2018). Social and emotional learning associated with universal curriculum-based interventions in early childhood education and care centers: A systematic review and meta-analysis. *JAMA network open 1*(8), e185727–e185727.

Bó, E. D. and P. D. Bó (2011, August). Workers, Warriors, And Criminals: Social Conflict In General Equilibrium. *Journal of the European Economic Association 9*(4), 646–677.

Bornacelly, I. (2013, jan). Educación técnica y tecnológica para la reducción de la desigualdad salarial y la pobreza. *Revista Desarrollo y Sociedad* (71), 83–121.

Bowles, S. (2012). Warriors, levelers, and the role of conflict in human social evolution. *Science 336*(6083), 876–879.

Bowles, S. and J.-K. Choi (2013). Coevolution of farming and private property during the early holocene. *Proceedings of the National Academy of Sciences 110*(22), 8830–8835.

Boyd, R., H. Gintis, and S. Bowles (2010). Coordinated punishment of defectors sustains cooperation and can proliferate when rare. *Science 328*(5978), 617–620.

Brader, L., H. Djibo, F. Faye, S. Ghaout, M. Lazar, P. Luzietoso, and M. O. Babah (2006). Towards a more effective response to desert locusts and their impacts on food security, livelihoods and poverty. Technical report, FAO.

Brodeur, A., A. E. Clark, S. Fleche, and N. Powdthavee (2021). Covid-19, lockdowns and well-being: Evidence from google trends. *Journal of public economics 193*, 104346.

Brückner, M. and A. Ciccone (2010). International commodity prices, growth and the outbreak of civil war in sub-saharan africa. *The Economic Journal 120*(544), 519–534.

Brunnschweiler, C. N. and E. H. Bulte (2008). Linking natural resources to slow growth and more conflict.

Burke, M., F. González, P. Baylis, S. Heft-Neal, C. Baysan, S. Basu, and S. Hsiang (2018). Higher temperatures increase suicide rates in the united states and mexico. *Nature climate change 8*(8), 723–729.

Burke, M., S. M. Hsiang, and E. Miguel (2015). Global non-linear effect of temperature on economic production. *Nature 527*(7577), 235–239.

Burke, M. B., E. Miguel, S. Satyanath, J. A. Dykema, and D. B. Lobell (2009). Warming increases the risk of civil war in africa. *Proceedings of the national Academy of sciences 106*(49), 20670–20674.

Callaway, B. and P. H. Sant'Anna (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics 225*(2), 200–230.

Camarena, K. R. (2016). Refugee return and food insecurity: Explaining increases in violence in ongoing civil wars.

Campo-Arias, A. and C. T. De Mendieta (2021). Social determinants of mental health and the covid-19 pandemic in low-income and middle-income countries. *The Lancet. Global Health 9*(8), e1029.

Campos, F., M. Frese, M. Goldstein, L. Iacovone, H. C. Johnson, D. McKenzie, and M. Mensmann (2018, may). Is personal initiative training a substitute or complement to the existing human capital of women? results from a randomized trial in togo. *AEA Papers and Proceedings 108*, 256–261.

Card, D., P. Ibarrarán, F. Regalia, D. Rosas-Shady, and Y. Soares (2011). The labor market impacts of youth training in the dominican republic. *Journal of Labor Economics 29*(2), 267–300.

Cassar, A., P. Grosjean, and S. Whitt (2013, jun). Legacies of violence: trust and market development. *Journal of Economic Growth 18*(3), 285–318.

Cho, Y., D. Kalomba, A. M. Mobarak, and V. Orozco (2013). *Gender differences in the effects of vocational training: Constraints on women and drop-out behavior*. The World Bank.

Cilliers, J., O. Dube, and B. Siddiqi (2016). Reconciling after civil conflict increases social capital but decreases individual well-being. *Science 352*(6287), 787–794.

Collier, P., A. Hoeffler, and D. Rohner (2009). Beyond greed and grievance: Feasibility and civil war. *Oxford Economic Papers 61*(1), 1–27.

Conte, B., A. Tapsola, and L. Piemontese (2021). The power of markets: Impact of desert locust invasions on child health. *SSRN Electronic Journal*.

Couttenier, M. and R. Soubeyran (2014). Drought and civil war in sub-saharan africa. *The Economic Journal 124*(575), 201–244.

Cressman, K. (2016). Chapter 4.2 - desert locust. In J. F. Shroder and R. Sivanpillai (Eds.), *Biological and Environmental Hazards, Risks, and Disasters*, pp. 87–105. Boston: Academic Press.

Cressman, K., A. V. der Elstraeten, and C. Pedrick (2016). elocust3: an innovative tool for crop pest control. good practice fact sheet. Technical report, FAO.

Crisi, A. and F. Dentale (2016). The wartegg drawing completion test: inter-rater agreement and criterion validity of three new scoring categories. *International Journal of Psychology and Psychological Therapy 16*(1), 83–90.

Crost, B. and J. H. Felter (2020). Extractive resource policy and civil conflict: Evidence from mining reform in the philippines. *Journal of Development Economics 144*, 102443.

DANE (2018). Censo nacional de población y vivienda.

DANE (2019). Gran encuesta integrada de hogares (geih). dirección de metodología y producción estadística (dimpe).

de Chaisemartin, C. and X. D'Haultfœuille (2020, sep). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review 110*(9), 2964–2996.

De Chaisemartin, C. and X. d'Haultfoeuille (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review 110*(9), 2964–96.

De Chaisemartin, C. and X. d'Haultfoeuille (2022). Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey. Technical report, National Bureau of Economic Research.

de Chaisemartin, C. and X. D'Haultfoeuille (2022, jan). Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey. Technical report.

Deaton, A., R. I. Miller, et al. (1995). International commodity prices, macroeconomic performance, and politics in sub-saharan africa. *Princeton Studies in International Finance 79*.

Dell, M., B. F. Jones, and B. A. Olken (2014). What do we learn from the weather? the new climate-economy literature. *Journal of Economic Literature 52*(3), 740–98.

Depetris-Chauvin, E. and R. Durante (2017). One team, one nation: Football, ethnic identity and conflict in africa.

Dube, O. and J. F. Vargas (2013). Commodity price shocks and civil conflict: Evidence from colombia. *The review of economic studies 80*(4), 1384–1421.

Duckworth, A. L. and P. D. Quinn (2009, Feb). Development and validation of the short grit scale (grit–s). *Journal of Personality Assessment 91*(2), 166–174.

Durlak, J. A., R. P. Weissberg, A. B. Dymnicki, R. D. Taylor, and K. B. Schellinger (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child development 82*(1), 405–432.

Eberle, U. J., D. Rohner, and M. Thoenig (2020). Heat and hate: Climate security and farmer-herder conflicts in africa. *Available at SSRN 3753942*.

Edelson, M. G., R. Polania, C. C. Ruff, E. Fehr, and T. A. Hare (2018). Computational and neurobiological foundations of leadership decisions. *Science 361*(6401), eaat0036.

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

Falk, A., A. Becker, T. J. Dohmen, D. Huffman, and U. Sunde (2016). The preference survey module: A validated instrument for measuring risk, time, and social preferences.

FAO (2020). East africa regional desert locust impact monitoring, round 1.

Garcia, M. R. C., L. Debenedetti, A. Farfan, C. Pecha, and C. Rodriguez (2021). *Seguridad Alimentaria y Proteccion Social en Colombia durante el COVID-19*. IPA.

Giuntella, O., K. Hyde, S. Saccardo, and S. Sadoff (2021). Lifestyle and mental health disruptions during covid-19. *Proceedings of the National Academy of Sciences 118*(9).

Griffin, J. M., S. Kruger, and G. Maturana (2019). Personal infidelity and professional conduct in 4 settings. *Proceedings of the National Academy of Sciences 116*(33), 16268–16273.

Hamadani, J. D., M. I. Hasan, A. J. Baldi, S. J. Hossain, S. Shiraji, M. S. A. Bhuiyan, S. F. Mehrin, J. Fisher, F. Tofail, S. M. U. Tipu, et al. (2020). Immediate impact of stay-at-home orders to control covid-19 transmission on socioeconomic conditions, food insecurity, mental health, and intimate partner violence in bangladeshi women and their families: an interrupted time series. *The Lancet Global Health 8*(11), e1380–e1389.

Harari, M. and E. L. Ferrara (2018). Conflict, climate, and cells: a disaggregated analysis. *Review of Economics and Statistics 100*(4), 594–608.

Harris, I., T. J. Osborn, P. Jones, and D. Lister (2020, apr). Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data 7*(1).

Haushofer, J. and E. Fehr (2014). On the psychology of poverty. *Science 344*(6186), 862–867.

Helgeson, J. F., S. Dietz, and S. Hochrainer-Stigler (2013). Vulnerability to weather disasters: the choice of coping strategies in rural uganda. *Ecology and Society 18*(2).

Holmes, E. A., R. C. O'Connor, V. H. Perry, I. Tracey, S. Wessely, L. Arseneault, C. Ballard, H. Christensen, R. C. Silver, I. Everall, et al. (2020). Multidisciplinary research priorities for the covid-19 pandemic: a call for action for mental health science. *The Lancet Psychiatry 7*(6), 547–560.

Hsiang, S., R. Kopp, A. Jina, J. Rising, M. Delgado, S. Mohan, D. Rasmussen, R. Muir-Wood, P. Wilson, M. Oppenheimer, et al. (2017). Estimating economic damage from climate change in the united states. *Science 356*(6345), 1362–1369.

Hsiang, S. M. (2010). Temperatures and cyclones strongly associated with economic production in the caribbean and central america. *Proceedings of the National Academy of sciences 107*(35), 15367–15372.

Hsiang, S. M., M. Burke, and E. Miguel (2013). Quantifying the influence of climate on human conflict. *Science 341*(6151), 1235367.

Hsiang, S. M. and K. C. Meng (2014). Reconciling disagreement over climate–conflict results in africa. *Proceedings of the National Academy of Sciences 111*(6), 2100–2103.

Hsiang, S. M., K. C. Meng, and M. A. Cane (2011). Civil conflicts are associated with the global climate. *Nature 476*(7361), 438–441.

Ibarrarán, P. and D. Rosas Shady (2009). Evaluating the impact of job training programmes in latin america: evidence from idb funded operations. *Journal of Development Effectiveness 1*(2), 195–216.

ICG (2017). Herders against farmers: Nigeria's expanding deadly conflict. Technical report, International Crisis Group.

ICG (2021). Ending nigeria's herder-farmer crisis: The livestock reform plan. Africa Report 302, International Crisis Group.

ILO. Sisben: A unified vulnerability assessment and identification system for social assistance.

Inter-American Development Bank (2019). *El futuro ya está aquí: Habilidades transversales de América Latina y el Caribe en el siglo XXI*. Inter-American Development Bank.

Jackson, M. O. and S. Nei (2015). Networks of military alliances, wars, and international trade. *Proceedings of the National Academy of Sciences 112*(50), 15277–15284.

Jensen, R. (2007). The digital provide: Information (technology), market performance, and welfare in the south indian fisheries sector. *The Quarterly Journal of Economics 122*(3), 879–924.

Kola, L., B. A. Kohrt, C. Hanlon, J. A. Naslund, S. Sikander, M. Balaji, C. Benjet, E. Y. L. Cheung, J. Eaton, P. Gonsalves, et al. (2021). Covid-19 mental health impact and responses in low-income and middle-income countries: reimagining global mental health. *The Lancet Psychiatry*.

Larsen, A. E. and F. Noack (2017). Identifying the landscape drivers of agricultural insecticide use leveraging evidence from 100,000 fields. *Proceedings of the National Academy of Sciences 114*(21), 5473–5478.

Mamun, M. A., N. Sakib, D. Gozal, A. I. Bhuiyan, S. Hossain, M. Bodrud-Doza, F. Al Mamun, I. Hosen, M. B. Safiq, A. H. Abdullah, et al. (2021). The covid-19 pandemic and serious psychological consequences in bangladesh: a population-based nationwide study. *Journal of affective disorders 279*, 462–472.

Marending, M. and S. Tripodi (2022, January). Gone with the Wind: The Welfare Effect of Desert Locust Outbreaks. Working Papers 1-2022, Copenhagen Business School, Department of Economics.

McGuirk, E. and M. Burke (2020, oct). The economic origins of conflict in africa. *Journal of Political Economy 128*(10), 3940–3997.

McKenzie, D. (2012). Beyond baseline and follow-up: The case for more t in experiments. *Journal of development Economics 99*(2), 210–221.

Miguel, E. and S. Satyanath (2011). Re-examining economic shocks and civil conflict. *American Economic Journal: Applied Economics 3*(4), 228–32.

Miguel, E., S. Satyanath, and E. Sergenti (2004). Economic shocks and civil conflict: An instrumental variables approach. *Journal of political Economy 112*(4), 725–753.

Moreno, C., T. Wykes, S. Galderisi, M. Nordentoft, N. Crossley, N. Jones, M. Cannon, C. U. Correll, L. Byrne, S. Carr, et al. (2020). How mental health care should change as a consequence of the covid-19 pandemic. *The Lancet Psychiatry*.

Moya, A. (2018). Violence, psychological trauma, and risk attitudes: Evidence from victims of violence in colombia. *Journal of Development Economics 131*, 15–27.

Moya, A. and M. R. Carter (2019). Violence and the formation of hopelessness: Evidence from internally displaced persons in colombia. *World Development 113*, 100–115.

Moya, A., P. Serneels, A. Desrosiers, V. Reyes, M. J. Torres, and A. Lieberman (2021). The covid-19 pandemic and maternal mental health in a fragile and conflict-affected setting in tumaco, colombia: a cohort study. *The Lancet Global Health 9*(8), e1068–e1076.

Novella, R., A. Repetto, C. Robino, and G. Rucci (2018). *Millennials en América Latina y el Caribe:¿ trabajar o estudiar?* Inter-American Development Bank.

OECD. *Tackling Vulnerability in the Informal Economy*. OECD.

Pandey, M., B. Suwal, P. Kayastha, G. Suwal, and D. Khanal (2021). Desert locust invasion in nepal and possible management strategies: A review. *Journal of Agriculture and Food Research 5*, 100166.

Parra-Saavedra, M. and J. Miranda (2021). Maternal mental health is being affected by poverty and covid-19. *The Lancet Global Health 9*(8), e1031–e1032.

Pierce, M., H. Hope, T. Ford, S. Hatch, M. Hotopf, A. John, E. Kontopantelis, R. Webb, S. Wessely, S. McManus, et al. (2020). Mental health before and during the covid-19 pandemic: a longitudinal probability sample survey of the uk population. *The Lancet Psychiatry 7*(10), 883–892.

Rosenzweig, M. R. (1988a). Risk, implicit contracts and the family in rural areas of low-income countries. *The Economic Journal 98*(393), 1148–1170.

Rosenzweig, M. R. (1988b). Risk, private information, and the family. *The American Economic Review 78*(2), 245–250.

Rosenzweig, M. R. and O. Stark (1989). Consumption smoothing, migration, and marriage: Evidence from rural india. *Journal of political Economy 97*(4), 905–926.

Salari, N., A. Hosseinian-Far, R. Jalali, A. Vaisi-Raygani, S. Rasoulpoor, M. Mohammadi, S. Rasoulpoor, and B. Khaledi-Paveh (2020). Prevalence of stress, anxiety, depression among the general population during the covid-19 pandemic: a systematic review and meta-analysis. *Globalization and health 16*(1), 1–11.

Serafini, G., B. Parmigiani, A. Amerio, A. Aguglia, L. Sher, and M. Amore (2020). The psychological impact of covid-19 on the mental health in the general population. *QJM: An International Journal of Medicine 113*(8), 531–537.

Silverio-Murillo, A., J. R. Balmori de la Miyar, and L. Hoehn-Velasco (2020). Families under confinement: Covid-19, domestic violence, and alcohol consumption. *Domestic Violence, and Alcohol Consumption (September 7, 2020)*.

Sinding Bentzen, J. (2019, 05). Acts of God? Religiosity and Natural Disasters Across Subnational World Districts*. *The Economic Journal 129*(622), 2295–2321.

Smith, T. G. (2014, oct). Feeding unrest. *Journal of Peace Research 51*(6), 679–695.

Soilevuo Grønnerød, J. and C. Grønnerød (2012). The wartegg zeichen test: A literature overview and a meta-analysis of reliability and validity. *Psychological Assessment 24*(2), 476.

Sorrenti, G., D. Ribeaud, and M. Eisner (2020). The causal impact of socio-emotional skills training on educational success.

Souza, C. V. R. d., R. Primi, and F. K. Miguel (2007). Validade do teste wartegg: correlação com 16pf, bpr-5 e desempenho profissional. *Avaliação psicológica 6*(1), 39–49.

Sun, L. and S. Abraham (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics 225*(2), 175–199. Themed Issue: Treatment Effect 1.

Taylor, R. D., E. Oberle, J. A. Durlak, and R. P. Weissberg (2017). Promoting positive youth development through school-based social and emotional learning interventions: A meta-analysis of follow-up effects. *Child development 88*(4), 1156–1171.

Thunström, L. and S. Noy (2019). The value of thoughts and prayers. *Proceedings of the National Academy of Sciences 116*(40), 19797–19798.

Townsend, R. M. (1995). Consumption insurance: An evaluation of risk-bearing systems in low-income economies. *Journal of Economic perspectives 9*(3), 83–102.

UNDP (2014). Ormet: Población desplazada y mercado de trabajo. Technical report, UNDP, https://www.co.undp.org/content/colombia/es/home/presscenter/articles/2014/02/13/ormet-poblaci-n-desplazada-y-mercado-de-trabajo.html.

Vreyer, P. D., N. Guilbert, and S. Mesple-Somps (2015). Impact of Natural Disasters on Education Outcomes: Evidence from the 1987–89 Locust Plague in Mali. *Journal of African Economies 24*(1), 57–100.

Zwane, A. P., J. Zinman, E. Van Dusen, W. Pariente, C. Null, E. Miguel, M. Kremer, D. S. Karlan, R. Hornbeck, X. Giné, et al. (2011). Being surveyed can change later behavior and related parameter estimates. *Proceedings of the National Academy of Sciences 108*(5), 1821–1826.