Essays on household decision-making and women's labor supply in Ethiopia

Sandra Kristine Halvorsen

Acknowledgments

My first and deepest gratitude goes to my main supervisor, Kjetil Bjorvatn, for his extraordinary dedication, support, and guidance. I greatly appreciate our numerous discussions and his attentive feedback and help at all stages of the PhD. I cannot imagine having a better supervisor.

I am also sincerely grateful to my second supervisor, Espen Villanger, for always believing in me and encouraging me to aim higher. Without his leap of faith in me and continued support, I would never have had reached this far. Thank you!

Besides my supervisors, who are also my co-authors, I would like to extend my gratitude to my other co-authors, Andreas Kotsadam and Tigabu Degu Getahun. In particular, Andreas for his encouragement and dedicated involvement in the project. And Tigabu, for his impressive work and devotion holding the project together in Ethiopia.

I have been extremely lucky to be part of FAIR - The Choice Lab, with its unique research environment of high ambitions, excellent research, enthusiasm, and support. I would like to express my appreciation, especially to Bertil Tungodden and Alexander W. Cappelen for creating an inclusive and inspiring research center, and to Erik Sørensen for his interesting and inspiring lectures. I would also like to thank Vincent Somville for his comments in my numerous presentations and sincere interest in my research.

I am also exceptionally fortunate to have the inspiring and wonderful Charlotte Ringdal, Ingrid Hoem Sjursen, and Ingvild Skarpeid as my PhD colleagues and friends. Thank you for sharing your insights and experiences both academically and outside. The environment among the PhD students has been kind and supportive also thanks to Ceren, Charlotte, Erling, Evangelos, Felix, Ole-Andreas, Hallgeir, Ingar, Ingeborg, Mads, Maria, Mirjam, Oda, Ranveig, Sara, Stefan, Timothy, Xu, and Øyvind.

I have additionally been lucky to be part of Chr. Michelsen Institute, with its inclusive, warm, and inspiring research environment, in which Lovise Aalen and Espen Villanger welcomed me. Thanks also to my colleagues at CMI for the fun times at the yearly quiz tournament and beer tastings.

I would also like to thank Øivind Anti Nielsen, who was my master's thesis supervisor. When I came to seek advice about the PhD application, Øivind immediately offered great help and advice. His guidance made a difference.

I am very grateful to all the participants in my studies and the hard work of all enumerators and research assistants making the projects possible. I also thank the Norwegian Research Council, the Centre for Ethics and Economics, the Department of Economics, and Chr. Michelsen Institute for financial support.

I would also like to thank my family, Anne Grethe, Harald, Sissel, Mariann, and Farmor, I am lucky to have all of you backing me and cheering me on.

Finally, I would like to thank Evan. Whose sense of humor and loving support has made this journey all the better.

Contents

Introduction 3

1 Conflict or cooperation? Experimental evidence on intra-household allocations in Ethiopia 7

2 Factory employment and fertility decisions: Experimental evidence from Ethiopia 41

3 Understanding the challenges of high labor turnover in the Ethiopian manufacturing industry 89

Introduction

This thesis consists of three research papers in the field of household economics in a developing country context. All three papers focus on women in Ethiopia: their role and power in the household, labor market participation, and fertility decisions. Each paper employs different research methods. The first paper uses a lab-in-the-field experiment to measure income sharing and cooperation among Ethiopian couples. The second paper is based on a field experiment and investigates the causal impact of industrial employment on fertility, income, and household decision-making power. The third paper combines a panel survey of female factory workers with qualitative interviews to explore the household decision-making processes of women's participation in industrial employment.

The three papers are in the context of Ethiopia, one of the fastest-growing countries in the world, both in terms of economic and population growth. While still being among the poorest in the world, the country has experienced strong and rapid economic growth the past decade, averaging at 9.9% annual GDP growth between 2008-2018 compared to a regional average of 5.4% in the same period.[1] This growth has contributed to significant reductions in poverty and improvements in several human development indicators. For instance, the share of the population living under the national poverty line decreased from 38.7% in 2004 to 23.5% in 2015.[2] Ethiopia is also one of the fastest-growing countries in terms of the population. In 2017, its population exceeded 100 million and it is estimated to double within the next three to four decades.[3] To continue the accelerated economic growth and to absorb part of the fast-growing workforce, the government has a strong focus on industrial policy. In particular, manufacturing sector development is currently creating thousands of much-needed jobs, especially for women and low-educated individuals, and is expected to transform the economy towards the ambitious goal of becoming a middle-income country by 2025. The lessons learned from this thesis may be useful for anyone interested in women's role in the household, their decision-making power, and labor market participation, particularly within the context of the growing manufacturing industry of Ethiopia.

The first paper, titled *Conflict or cooperation? Experimental evidence on intra-household allocations in Ethiopia*, is co-authored with Kjetil Bjorvatn and Tigabu Degu Getahun. It uses a lab-in-the-field experiment to test couples' cooperation and sharing behaviors, and explore possible mechanisms for their choices. The way couples interact in a lab setting may give meaningful insights into real household decision-making processes, which in turn have important implications for individual family members' welfare and the overall household wealth. By understanding these processes, policies may be better designed to improve welfare, for instance by changing the receiver of transfers. The relatively small intra-household lab experimental literature finds that established couples rarely maximize their combined payoffs. For instance, spouses have been found to withhold money or take advantage of asymmetric information to guarantee their own payoffs. Such uncooperative behaviors ap-

[1] World Development Indicators: GDP growth (annual %).

[2] World Development Indicators: Poverty headcount ratio at national poverty lines (% of population).

[3] United Nations, World Population Prospects 2019.

pear to be related to social structures of gender relations, unequal control over household finances, and lack of trust, altruism, or patience.

In the paper, we present additional evidence from a comprehensive lab-in-the-field experiment with couples in Ethiopia. We use public goods and dictator games to measure the couples' willingness to cooperate and their sharing behaviors, combined with spectator games to elicit preferences and social norms. Contrary to the vast majority of the existing literature that focuses on conflicts in the household, we argue for a large degree of cooperation between spouses. We find a general attitude and corresponding behavior that the wife should control a larger share of the household budget as the person managing the everyday household expenses. The latter statement implies that husbands act more generously in the lab than wives in traditional cultures, which aligns well to what we, and other intra-household lab experiments in developing countries, find. Moreover, the difference in sharing rates between men and women only exists in the intra-household games, while it disappears when they play with an anonymous party. Thus, the general attitude to fair allocations and corresponding behaviors is restricted to within households and does not reflect gender differences in general.

The second paper is co-authored with Andreas Kotsadam and Espen Villanger and is titled *Factory Employment and Fertility Decisions: Experimental Evidence from Ethiopia.* Industrialization is expected to have transformational impacts in many ways, not only on the economy as a whole, but also directly on the lives of the poor. For instance, large-scale job creation is expected to improve the welfare of workers and lead to women's empowerment, which can have further impacts on development, such as lower fertility rates and better child outcomes. However, an important challenge is the measurement of the direct effects of employment on the workers, since individuals that choose to participate in industrial employment are most likely different from individuals who do not participate. In the paper, we use a unique opportunity to circumvent the problem of selection bias by employing a randomized controlled trial (RCT). In the field experiment, we randomized job offers among eligible female applicants in the manufacturing industry in Ethiopia. This approach ensures that the women who got the job offer were, on average, similar to the women who did not get the job offer, hence any difference in the two groups over time can be attributed to the job that they got by the random draw.

The paper investigates whether getting a job in the manufacturing industry affects households' decision to have a child, the women's preference for the number of children, and whether the job increased the women's income and decision-making power in the household. It is to the best of our knowledge the first paper in the literture that measures the causal effects of a job on fertility outcomes at the individual level.

We find that receiving a job offer does not affect actual fertility nor preferred lifetime fertility. The women in the group receiving a job offer had higher income and were more likely to be employed in any job the following year. However, we find no impact of job offers on household decision-making power. The results imply that a reduction in fertility is not likely to occur as a direct result of industrialization. On the other hand, the effects of industrialization may operate at a more aggregate level. If that is the case, the results

challenge established theories of direct income effects for reduced fertility.

While the jobs did not affect fertility or decision-making power, they affected labor participation and income. However, we observe very high employee turnover rates, which are a common aspect of manufacturing sectors in developing countries, probably resulting from low wages, long working hours, and unsafe or unhealthy working conditions. The high turnover rates may suggest that manufacturing jobs will not have the expected positive direct impacts on the lives of the poor as policy-makers are hoping, or at least not in the short-run. The third paper, titled *Understanding the challenges of high labor turnover in the Ethiopian manufacturing industry*, combines the panel survey of the female factory workers in the RCT with qualitative interviews to investigate the reasons for entry and exit in the manufacturing industry.

Among 850 newly employed women in our sample, half of them had left the factory job about seven months later than when they were interviewed for the first follow-up survey. The main reason for leaving was that the salary did not match their efforts along with many work related health problems, such as kidney issues and body pains from standing or sitting all day. Another important reason for the high employee turnover was that many workers had unrealistic expectations about the salary and work efforts and they were therefore very disappointed with the factory experience. Furthermore, many also found it challenging to combine their childcare and household responsibilities with the factory job due to the inflexible work schedule and demanding work tasks.

Understanding the welfare impacts of industrialization on workers is important for evaluating the development benefits of manufacturing in poor countries. Moreover, by better understanding the workers' motives and constraints for entering and leaving the industry, practices may be improved to simultaneously decrease costly turnover rates and improve productivity and workers' welfare.

Chapter 1

Conflict or Cooperation? Experimental Evidence on Intra-Household Allocations in Ethiopia

Conflict or Cooperation? Experimental Evidence on Intra-Household Allocations in Ethiopia

Kjetil Bjorvatn, Tigabu Degu Getahun, and Sandra Kristine Halvorsen*

Abstract

We explore intra-household decision-making using lab experiments with more than 200 married couples in Ethiopia. In contrast to much of the literature that highlights conflicting interests between household members, we find striking similarities in household allocation preferences and norms of men and women. In particular, men and women are equally cooperative and altruistic in anonymous games, they place equal emphasis on public household goods relative to private goods, and they have similar risk and time preferences. Men do, however, contribute more than women in intra-household public goods and dictator games. Qualitative evidence suggests that this is due to a commonly held norm that wives should be compensated for their greater responsibility in managing household expenses. A household survey conducted some days after the experiment supports the external validity of our findings.

Keywords: Intra-household allocation, voluntary contribution mechanisms, social norms, preferences, Ethiopia
JEL classification: D13, D31, C71, C91, C92, J16

*The paper is published in *Journal of Behavioral and Experimental Economics*, Vol 85, January 2020. Kjetil Bjorvatn: The Norwegian School of Economics. Tigabu Degu Getahun: Ethiopian Development Research Institute. Sandra Kristine Halvorsen: The Norwegian School of Economics and Christian Michelsen Institute. We would like to thank EDRI (Ethiopian Development Research Institute) and especially Biruk Tekele for their help in the preparation and organization of the experiments; research assistants (in alphabetic order) Addis Kibru Ayele, Meseret Mamo Bazezew, Elshaday Endale, Leti Getachew Hika, Mekdes Regassa, Mengistu Tamirat, Gaddise Tesfaye, and Heidi Frances Thornhill for excellent help in conducting the experiment; Alexander Cappelen, Bereket Kebede, Charlotte Ringdal, Øivind Schøyen, Ingrid Hoem Sjursen, Bertil Tungodden, Espen Villanger, and numerous seminar participants for helpful comments and discussions at various stages of the project. The project was supported by the Centre of Ethics and Economics at the Norwegian School of Economics and The Research Council of Norway, both through the project: "Women in the developmental state: Female employment and empowerment in Ethiopia" project No 240807 and its Centres of Excellence Scheme, FAIR project No 262675.

1 Introduction

Households are often portrayed as arenas of conflict. Spouses have been found to hide money from each other and even be willing to resort to destructive behaviors to improve their relative position (Ashraf, 2009; Castilla and Walker, 2013; Hoel, 2015; Anderson and Baland, 2002; Mani, 2011). Typically, men are seen as more irresponsible and egoistic and less family oriented than women, and, accordingly, initiatives aimed at increasing family welfare generally target women (Eckel and Grossman, 1998, 2008; Duflo and Udry, 2004; Hoddinott and Haddad, 1995; Haddad et al., 1997; Lundberg et al., 1997; Thomas, 1990, 1993; Yoong et al., 2012).

However, more recent studies question this conventional wisdom. Field experiments on cash transfer programs show that the recipients' gender makes almost no difference for family welfare outcomes (Akresh et al., 2016; Benhassine et al., 2015; Haushofer and Shapiro, 2016), and lab experiments show no gender differences in altruism or cooperative behavior (Andreoni and Vesterlund, 2001; Croson and Gneezy, 2009). Given the diverging findings in the literature and the importance of the household for decisions on consumption, education and health, more research into the mechanisms of household allocation decisions is clearly needed, for instance in guiding cash transfer policies to improve household outcomes.

From a theoretical perspective, households should be ideal arenas for cooperation, as spouses benefit from investments in household public goods, such as children, and because repeated interactions allow them to punish deviating behavior (Browning, 2009; Del Boca and Flinn, 2012). The cooperative equilibrium is facilitated by common information, similar preferences, mutual affection, and shared norms of behavior. Absence of these conducive factors can lead to the breaking up of relationships or to non-cooperative behavior between the spouses, resulting in inefficient outcomes such as the undersupply of household public goods (for an overview, see Baland and Ziparo, 2017).

In order to shed further light on household decision-making, we here present evidence from lab experiments with more than 200 married couples in Ethiopia. We use one-shot public goods and dictator games to measure cooperation, where purely egoistic agents would choose not to cooperate, but where altruism and the presence of reciprocity norms could induce people to do so (see, e.g., Camerer, 2003; Ledyard, 1995). To investigate the importance of the social context, we conduct the same cooperative games in both intra- and extra-household contexts (matched with the spouse or an anonymous other participant, respectively). Furthermore, we investigate preferences and social norms through the use of household allocation tasks. We also collect qualitative data from interviews and focus group discussions to learn more about the mechanisms of the decision-making processes.

Our main finding is that there is extensive cooperation between the spouses in our sample, where husbands in fact contribute more than their wives. Our qualitative evidence shows that there is a commonly held norm that husbands, as the main bread winners, should provide for their wives, who typically have greater household responsibilities. This kind of behavior is consistent with the *separate spheres* model of Lundberg and Pollak (1993), where there is a clear division of labor between the spouses and where husbands typically transfer

resources to their wives in order to achieve common goals. Moreover, we find that spouses in our sample have strikingly similar preferences for the allocation of household resources, which is also conducive to cooperation. Finally, a household survey carried out 10 days after the lab supports the external validity of the choices made in the lab, and importantly does not point to any undoing effect, which is a key concern in the experimental literature on household decision-making.

The articles closest to our study are Iversen et al. (2011) and Kebede et al. (2013) who use variations of the public goods game to test household efficiency in Uganda and Ethiopia, respectively. In line with our results, these studies show that wives in many cases contribute less than husbands do to a household common pool. Similarly, Bezu and Holden (2015) compare dictator games between spouses and other villagers in rural Ethiopia. Again, they find that wives are less likely to share the endowment with their spouse, while they find almost no gender difference when sharing with an anonymous villager. In another related study, Ringdal and Sjursen (2017) find that there is no difference between spouses in spending preferences on children.

The main contributions of our study relative to those cited above are (i) that we use third-party allocation tasks to elicit social norms on what is considered fair household allocations; (ii) that we use a follow-up survey to shed light on the external validity of the experiment, and (iii) that we also include both spousal and anonymous versions of the games as well as eliciting risk and time preferences, thus giving us comprehensive evidence on a large set of potential explanations for household cooperation.

The remainder of the paper is organized as follows. We first present a description of the experimental procedures and the sample. In section 3 we present the results, starting with the intra- and extra-household public goods and dictator games, followed by other household allocation tasks to explore mechanisms. We discuss the external validity of the experiment in Section 4, while Section 5 concludes.

2 The experiments

The lab experiments took place in Bishoftu, a semi-urban town in the Oromiya region, south-east of Addis Ababa. We conducted two rounds of experiments, one in March 2016 (Lab 1) and one in March 2017 (Lab 2). Each round stretched over two consecutive days with sessions lasting about 2.5 hours. In total 211 couples (i.e. 422 individuals) participated (130 couples in Lab 1 and 81 couples in Lab 2). Participants were recruited by letter of invitation based on a list of parents obtained from the principals of two elementary schools, one school for each round. The invited couples were required to live together and to be literate (since participants had to read and write in the lab). The letter was framed as an invitation to a workshop on economic choices where participants would earn at least ETB 100 (USD 4.25 at the time) each, and with the opportunity to earn more depending on their choices in the workshop. The average earnings from the lab experiments was ETB 157 (USD 6.70), which corresponds to about two days' income for the median male participant and about three and a half days' income for the median female participant.

The participants' background characteristics are provided in Table 1. The mean age is 39 for the male participants and 32 for the females. On average, the couples have been married 12 years and have 2.3 children. In general, the participants have a relatively high level of education, with 58 percent having completed higher than secondary education. Moreover, most participants, even the women, are employed.[1] The relatively high level of education reflects the facts that the experiments were conducted in a semi-urban context, close to the capital, and that participants were recruited based on literacy and living in a stable relationship. Thus, the couples in our sample are most likely a relatively selected group, and the results do not necessarily generalize to all types of households (more on this in Section 4). Still, with a median income of ETB 2 900 (self-reported), the households in our study are by no means well-off. To perspectivize, given that this income is shared equally between two adults and two children (which is the median number of children in our study), the median participant in our study would be living on ETB 24 per day. The World Bank's International poverty line of USD 1.90 for Ethiopia is about ETB 20 per day per capita in 2017, thus placing our participants just above the poverty line.[2]

When entering the lab, the women were seated on one side and the men on the other side to avoid communication between spouses. The sessions were led by local research assistants and held in the local language, with instructions and examples explained in plenum. The participants were assured that their choices were anonymous, and only observable by the researchers.

The experiments were designed to explore motives for intra-household cooperation and allocation decisions. For this purpose, the participants played public goods games and dictator games to measure cooperativeness and sharing between spouses. In the public goods games the participants were given an endowment and asked to share this between their own private account and a common account, knowing that another participant would make a similar decision. Contributions to the common account were then increased by 50 percent by the experimenters, and shared equally between the two participants. Note that the Nash-equilibrium of this game (without other-regarding preferences) is to contribute zero to the common account. Accordingly, we shall refer to any positive amount contributed to this account as a sign of cooperation, and higher contributions as a sign of more cooperativeness. In the intra-household version of this game, the other participant was the spouse, while in the extra-household version of the game, the other participant was an anonymous other participant in the experiment of the same sex as their spouse. The purpose of the extra-household games was to elicit preferences more generally, that is, outside the sphere of the household.

In the dictator games, the participants were told that they had an endowment and

[1] The Ethiopia 2016 Demographic and Health Survey (DHS) shows that in urban areas, 31.1 percent of the men and 21.1 percent of the women between 15-49 have completed education beyond secondary school, and that the employment rate of men is 80.5 percent, while that of women is 52 percent.

[2] The PPP adjusted poverty line is taken from the "Poverty & Equity Brief" for Ethiopia, October 2018. We have further adjusted the poverty line by inflation rates 10.11 in 2015 and 7.27 in 2016 (IMF). http://povertydata.worldbank.org/poverty/country/ETH

Table 1: Descriptive statistics of participants.

	All	Men	Women
Age	35	39	32
Monthly income (median), ETB	1 400	1 700	900
Education (Highest level completed)			
No education	2%	2%	2%
Primary school (grade 1-8)	21%	14%	27%
Secondary school (grade 9-10)	19%	19%	20%
Higher education	58%	65%	51%
Occupation			
Working on farm	1%	1%	0%
Casual worker	8%	11%	4%
Employee	58%	66%	50%
Self-employed	15%	11%	20%
Childcare/household chores	11%	1%	22%
Other	7%	10%	4%
Couples characteristics			
Number of children	2.3		
Years married	12		
Monthly household income (median), ETB	2 900		

should make a decision on how much to share with another participant in the room who had not received any endowment. Again, in the intra-household version of this game, the partner was the spouse, while in the extra-household version, it was an anonymous other participant.

The theoretical literature, referred to in the Introduction, points to similarity in preferences and shared social norms as important for sustaining a cooperative equilibrium. We measure preferences by presenting the spouses with a (hypothetical) choice between three accounts: husband, wife, and household, while social norms are measured using a third-party design, where the participants made similar (incentivized) choices on behalf of another household.

We also measure risk and time preferences and add these as controls in the regression analysis as they can potentially explain behavior in the public goods games. For instance, placing money in the common account might be considered more risky than keeping it in the private account. Moreover, contributions to the household account may be interpreted as an investment or a savings decision due to the framing and the multiplier effect. We elicit risk and time preferences using simple, non-incentivized survey questions.[3] We also asked about

[3]In particular, participants were categorized as risk averse if they chose a safe option (ETB 100 for sure) over a lottery (ETB 250 or ETB 0 with 50 percent probability). Similarly, time preference (patience) is based

their expectation of the contribution by the other player, which will be used as a measure of reciprocity.

Considerable care was taken to ensure that the games were well understood. In particular, the public goods game was illustrated with a role play where the research assistants acted the roles of husband and wife, allocating money into traditional baskets that are commonly used for storing money in the households. Qualitative evidence collected after the workshop confirms that the participants understood the games well and that they considered the allocation tasks to be closely related to their everyday allocation choices, which is reassuring for the validity of the exercises.

In the first round of sessions (Lab 1) the participants played the public goods game with their spouse and an anonymous other player, and one version of the third-party allocation tasks. The second round (Lab 2) included additional games as well as interviews and group discussions.[4] There was no feedback between the games and in order to ensure anonymity of choices, the earnings were calculated based on two randomly selected games and paid in private envelopes at the end of the session.[5] In this way, their earnings were random, and thus the participants were not able to backtrack any allocation choices made by themselves or their spouse by looking at the earnings.

3 Experimental results

We start by presenting evidence on cooperation and sharing based on public goods and dictator games. We differentiate between *intra-household* and *extra-household* decisions, where the former involve own household members and the latter involve members of another household, which are then either anonymous (someone else in the room) or identified by their role (wife, husband, household expenses/children). We then explore mechanisms, focusing on preferences and norms, measured by household allocation decisions.

on hypothetical questions involving a choice between a smaller amount soon and a larger amount later. The decision to use simple and non-incentivized methods was motivated by the fact that these measures are not at the core of our interest, but rather as background variables which can potentially explain behavior in the lab. Simple methods are useful in lab-in-the-field experiments where the participants may have low numeracy skills, as they are easy to understand, quick to perform, and give some indication of differences in individual preferences (Charness et al., 2013; Dave et al., 2010).

[4]We chose to have the same order of tasks and questions for all participants. This was mainly due to practical reasons, as this was a "pen-and-paper" experiment based on hand-outs that were distributed to the participants sequentially. While the ordering was the same for both men and women, we cannot rule out the possibility of gender specific ordering effects (Croson and Gneezy, 2009).

[5]Each participant received the earnings from one of the public goods or dictator games with spouse or with an anonymous other, in addition to the allocation received from someone else's third-party decision. Furthermore, we set a floor on earnings at ETB 100.

3.1 Public goods and dictator games

Table 2 presents the regression analysis for the public goods games and the dictator games in the intra-household and extra-household cases. Three main insights can be derived. First, there is significantly more cooperation within the households than with participants outside of the household. Second, men contribute more than women in the intra-household games. Third, there is no gender difference in contributions in the extra-household games.

Comparing average contributions in the intra-household setting with the extra-household setting, we find that the difference is 11 percentage points in the public goods game and 17 percentage points in the dictator game, with both differences significant at the one-percent level. Breaking down by gender, we find that women contribute 9 percentage points more in the public goods game when matched with their husband than when matched with an anonymous man, and 8 percentage points more in the dictator game, while for men the corresponding numbers are 15 percentage points in the public goods game and 26 percentage points in the dictator game; all differences significant at the one-percent level.

We also observe in column (1) that wives contribute significantly less to the common account than their husbands do, with a difference of seven percentage points, equivalent to almost 10 percent of the husbands' contributions. In column (2) we control for expectations, background variables, risk and time preferences, and lab fixed effects and observe that the estimated coefficient on the *Female* dummy is somewhat reduced, but still significant. Note that the gender difference in contributions to the common account do not necessarily imply a difference in preference over private and public consumption. As will be clear from the discussion in Section 3.2, husbands contribute more to the common account not only for public consumption purposes but also for transferring resources to the wife.

In line with the standard reciprocity result in such games, public goods contributions correlate positively with expected contributions.[6] We also observe that there is no significant correlation between risk aversion and cooperation in the lab, while patience, perhaps surprisingly, correlates negatively with contributions in the intra-household public good game.[7]

Columns (5)-(6) present the results for the intra-household dictator game. Wives share 16.5 percentage points less with their spouse than husbands do. When adding the control variables, the estimate is slightly lower, but still highly significant.[8]

Interestingly, as shown in columns (3)-(4) for the public goods game and (7)-(8) for the dictator game, there are no statistically significant gender differences in the extra-household games. In fact, the contribution rates by men and women are strikingly similar, and hence

[6] Interestingly, we note that the *Expected contribution* coefficient in the intra-household public goods game is larger compared to the extra-household game, which suggests, perhaps not surprisingly, that the reciprocity norm is stronger in an intra-household context.

[7] There was no difference in risk preferences between men and women in our sample: 71 percent of the men and 66 percent of the women were risk averse (choosing the safe option) (Chi-squared test, p-value: 0.341). In terms of time-preferences, 46 (43) percent of the men (women) were categorized as impatient, 28 (25) percent as patient, 23 (30) percent as hyperbolic, and 3 (1) percent as inconsistent. The difference in categorization between men and women is not statistically significant (Chi-squared test, p-value: 0.400).

[8] In tables A1 and A2 in the appendix we present additional analyses and robustness checks on the intra-household public goods and dictator game.

Table 2: Shares of endowment given in public goods and dictator games.

	Public goods game				Dictator game			
	Intra-household		Extra-household		Intra-household		Extra-household	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	-0.070***	-0.047**	-0.013	-0.003	-0.165***	-0.145***	0.014	0.028
	(0.024)	(0.021)	(0.026)	(0.026)	(0.026)	(0.027)	(0.032)	(0.034)
Expected contribution		0.493***		0.240***				
		(0.055)		(0.052)				
Income		0.031		0.035		-0.007		0.015
		(0.022)		(0.028)		(0.031)		(0.044)
Higher education		0.050**		0.019		0.020		-0.010
		(0.024)		(0.029)		(0.030)		(0.048)
Risk averse		0.004		-0.003		-0.002		-0.024
		(0.024)		(0.028)		(0.030)		(0.035)
Patient		-0.068**		0.004		0.008		0.021
		(0.028)		(0.033)		(0.037)		(0.043)
Observations (individuals)	422	422	422	422	162	162	162	162
Adjusted R-squared	0.02	0.28	-0.00	0.07	0.20	0.20	-0.01	-0.02
Lab dummies	-	Yes	-	Yes	-	-	-	-
Session dummies	-	-	-	-	-	Yes	-	Yes
Additional controls	-	Yes	-	Yes	-	Yes	-	Yes
Mean of contribution by men	0.747	0.747	0.603	0.603	0.644	0.644	0.382	0.382

Notes: OLS regressions of shares of endowment given in public goods and dictator games. Columns (1)-(4) consist of the pooled sample from Lab 1 and Lab 2, columns (5)-(8) include participants from lab 2 as the dictator game was only played here. The income variable takes the value one if the respondent earns more than or equal to the median weekly income of their respective gender (ETB 425 for men and ETB 225 for women); Higher education takes the value one if the respondent has more than secondary schooling; Additional controls include years married, number of children, and endowment treatments in the public goods games which indicate whether the participant was given equal endowment as the partner, or a higher/lower endowment that the partner. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

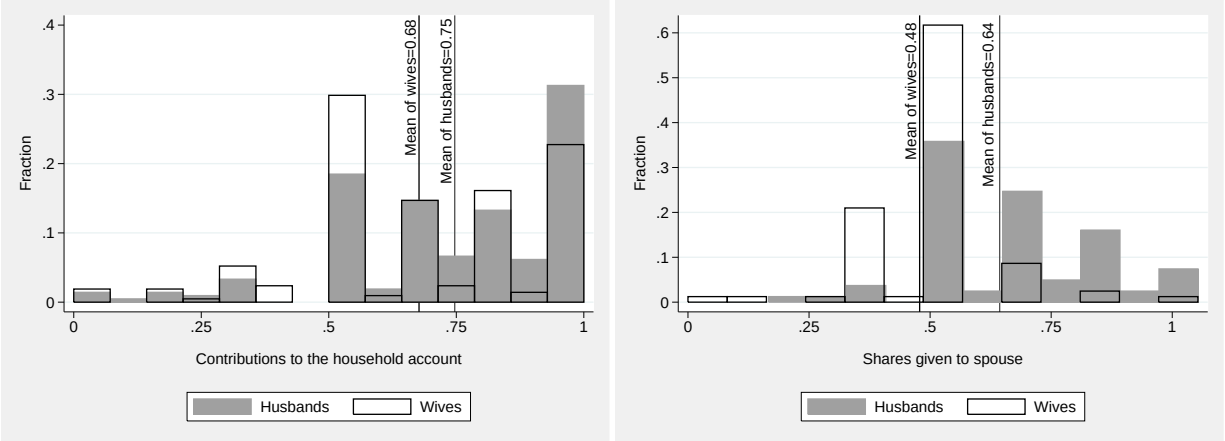
we conclude that there is no gender difference in the general willingness to cooperate or share.

Figure 1 offers a closer look at the decisions in the intra-household games.[9] In the public goods game, as evidenced from the regressions, the husbands on average contribute more of their endowment to the household account than their wives: 75 percent compared to 68 percent (p-value = 0.003). Moreover, we observe that the modal contribution by husbands is 100 percent while the modal contribution by wives is 50 percent.[10] Only 22 couples (10.4

[9]The Appendix Figures A1 and A2 present histograms for all intra- and extra-household games separated by gender.

[10]Chi-squared test of proportions of 100 percent-contributors by gender. Result: p-value = 0.048. Chi-squared

Figure 1: Histograms of contributions in the intra-household games.



(a) Intra-household public goods game

(b) Intra-household dictator game

percent of the sample) achieved the most efficient outcome by both contributing all of their endowment to the public goods account, while at the same time, very few didn't contribute at all (only 7 individuals, that is, 1.7 percent of the sample).

In the dictator game most wives allocate equally (62 percent) or more to themselves (26 percent), while most husbands allocate less to themselves (58 percent) and about a third (36 percent) allocate equal amounts between themselves and their wife. As shown in the regression table, the difference in means is statistically significant (p-value=0.000).

3.2 Preferences and norms in household allocation games

We here document findings from additional household allocation games, in intra-household and extra-household settings. We interpret allocations in intra-household games as revealing *preferences* while decisions in extra-household games inform us about *norms*, since by making decisions for other individuals we exclude self interest in the allocation choices and thus elicit the participants' fairness norms (Konow, 2009). The theoretical literature suggests that both preferences and norms are important for upholding a cooperative equilibrium. When spouses have similar preferences on how to allocate funds across household accounts, the value added of cooperation increases. Similarly, commonly held social norms on cooperation within a household support the cooperative equilibrium by making deviations from it more costly, for instance by triggering social punishment.

In the intra-household allocation game, the participants were given an (hypothetical) endowment of ETB 60 to allocate between the spouses and a household account. As in the public goods game, the amount placed in the household account was multiplied with a factor 1.5 by the experimenters, but in contrast to the standard game, the current version includes

test of proportions of 50 percent-contributors by gender. Result: p-value = 0.006.

a separate fund for the spouse. In addition, the allocation game takes out the strategic component of the public goods game, which complicates the interpretation with regards to preferences. We implemented two extra-household allocation games, both of which were incentivized. The first was a three-accounts game, similar to the one described above, while the second was a two-account allocation task. Here, the participants were presented with a household where the husband earns ETB 60 (20) a day, while the wife earns ETB 20 (60) and then asked how they would like to divide the household income (ETB 80) between the spouses. We report the mean of the two allocation choices by each participant.[11]

Table 3 summarizes the results, with Panel A showing allocations in the intra-household game and Panel B showing allocations in the extra-household games. We can derive three main insights from this table. First, men and women have very similar preferences over household allocations. Second, wives receive a larger share of the household endowment than husbands do. Third, allocations in the extra-household and intra-household decisions are strikingly similar, which is consistent with norms being very influential in determining household behavior.

More specifically, by first comparing the allocation decisions by men and women horizontally we see that there are no gender differences in the allocation decisions: male and female participants agree on how much should be placed in the household account, husband's account, and wife's account. When comparing the amounts placed in the husband's and wife's accounts (vertically) we observe that both men and women allocate less to the husband than to the wife. The results in the extra-household allocation tasks are very similar to those made for their own household. Thus, there seems to be an agreement among the spouses that the wife should have a larger share than the husband, and moreover, that intra-household decisions are largely based on these norms. In an open-ended discussion after the experiment, the most common justification among the men for allocating a larger share of the endowment to the wife is her larger responsibility in the household, while some also argued that the wife was better at handling money.

Surveys conducted immediately after the lab strengthen the impression of widespread agreement between couples (see Appendix B). Here, we find that 89 percent of the men and 84 percent of the women said that in their household, husband and wife have equal say in how to share and spend income, and that relatively few thought that the spouse was hiding money from them (10 percent of the men and 19 percent of the women) or disapproved of the spending decisions of the spouse (21 percent of the men and 23 percent of the women).

Given the widespread cooperation and consensus between spouses, one might wonder why so few place *all* the money in the common account, which is clearly the most efficient choice. In fact, only 4 percent of the participants do so in the intra-household three-account version of the allocation games. Most likely, this is due to mental accounting, where money is not seen as fungible across accounts (Thaler, 1999). Indeed, most participants (61 percent) explained their allocation choice by the need to cover expenses associated with the different 'domains'.

[11] See Table A3 in the appendix for the two-accounts allocation decisions with high and low earnings scenarios separately.

Table 3: Household allocations.

PANEL A. **Preferences: Intra-household allocations**

	Men	Women	Difference
Three accounts:			
Household	0.63 (0.18)	0.62 (0.18)	0.01 [0.03]
Husband	0.16 (0.09)	0.17 (0.08)	-0.01 [0.01]
Wife	0.21 (0.12)	0.21 (0.14)	0.00 [0.02]
Difference in spouses' shares	-0.06 [0.01]***	-0.04 [0.02]***	

PANEL B: **Norms: Extra-household allocations**

	Men	Women	Difference
Two accounts:			
Husband	0.45 (0.17)	0.47 (0.16)	-0.02 [0.02]
Wife	0.55 (0.17)	0.53 (0.16)	0.02 [0.02]
Difference in spouses' shares	-0.1 [0.03]***	-0.06 [0.03]**	
Three accounts:			
Household	0.57 (0.20)	0.59 (0.16)	-0.02 [0.03]
Husband	0.19 (0.11)	0.19 (0.08)	0.00 [0.02]
Wife	0.24 (0.14)	0.22 (0.11)	0.02 [0.02]
Difference in spouses' shares	-0.05 [0.02]***	-0.03 [0.01]***	

Notes: Mean shares allocated to the different accounts. Standard deviations in parenthesis and standard errors in square brackets. *** p<0.01, ** p<0.05, * p<0.1.

4 Discussion

The are two main threats to the external validity of our findings: The undoing problem and generalizability. The undoing problem relates to whether behavior in the lab is driven by the ability to reallocate funds after the lab, while generalizability relates to the question of how representative our findings are for the population as a whole.

4.1 The undoing problem

An important challenge to the external validity of household experiments is the undoing problem: Seemingly cooperative behavior in the lab could simply reflect the behavior of a greedy and dominant spouse (typically the husband) who intends to confiscate the entire household payments ex post (Munro, 2018). To address this issue we carried out a follow-up

survey by telephone on a subsample of the couples (n=94) approximately 10 days after Lab 2.[12]

Great care was taken to ensure independent and truthful answers to the questions in these interviews. First, the respondents were not told beforehand that we would follow up with questions about how they spent their compensation from the lab. Second, wives were contacted first and the husbands immediately afterwards, to avoid the sharing of information between them. Third, we called during daytime, as spouses are less likely to be together at that time (and we kindly asked for privacy if they happened to be together). While these measures should reduce the risk of the spouse influencing the responses, we can of course not rule it out completely.

In the interviews we asked about their experience with the workshop, what they had spent their earnings on, if they knew what their spouse had spent his or her earnings on and if so, on what exactly, and whether they had discussed or disagreed on the use of the earnings. The reported spending can be organized into five different categories: private, spouse, children, savings, and household. Accordingly, each spending category is constructed as a dummy variable equal to one if the individual reported spending in the given category. The category 'spouse' includes both spending on spouse such as a gift, and handing over the earnings. Table 4 presents an overview of how the experimental earning was spent by men and women separately.

There are three main observations from the follow-up survey. First, there is a net transfer from husbands to wives. In particular, 30 percent of the husbands handed over money to their wives, while almost no wife gave money to her husband, with the difference being significant at the one-percent level. Importantly, the transfer from the husband was confirmed by the wife.

Second, women spent a significantly higher share of their earnings on household consumption.[13] More precisely, 75 percent of the women reported to have spent payments from the experiment on household consumption, while 44 percent of the men reported to have done so, with the difference again being highly significant. Third, there is no significant difference between spouses in the propensity to spend money directly on their children. About half of the men and women spent some or all of the experimental money on their children, mostly on snacks, clothes, shoes or school fees. No couple reported any disagreement about the use of the money. In close to all the cases (88 percent) the spouses confirm the expenses made by their partner, which is a strong argument in favour of the reliability of the responses.[14]

[12] One strategy of reducing the undoing problem is to raise the cost of reversibility, for instance by using gendered products or activities rather than money, see for instance Beblo and Beninger (2017), Couprie et al. (2017), Dasgupta and Mani (2015), and Ashraf (2009). However, as Munro (2018) points out, this might not still eradicate the problem. We paid our participants in cash, but in sealed envelopes and individually, in order to minimize the risk of confiscation.

[13] Women's higher spending on household consumption is a well-established finding in the household literature and is thus a common argument for targeting welfare policies to women instead of men (Duflo and Udry, 2004; Haddad et al., 1997; Hoddinott and Haddad, 1995; Lundberg et al., 1997; Thomas, 1990, 1993).

[14] The fact that the spouses have similar allocation responses, may raise concern that this reflects a social desirability bias rather than truthful spending. While it is true that the spending patterns are very similar

Table 4: Use of experimental earnings.

	Men	Women	Difference
Private	0.26 (0.44)	0.21 (0.41)	0.04 [0.09]
Spouse	0.30 (0.46)	0.02 (0.15)	0.28 [0.07]***
Children	0.53 (0.50)	0.45 (0.50)	0.09 [0.10]
Savings	0.34 (0.48)	0.49 (0.51)	0.15 [0.10]
Household	0.44 (0.50)	0.75 (0.44)	0.31 [0.10]***

Notes: Proportion of subjects who spent some or all of their earnings in the given categories. Standard deviations in parenthesis and standard errors in square brackets. *** p<0.01, ** p<0.05, * p<0.1.

Taken together, the household survey clearly supports the validity of our lab experiment. Husbands allocate money to their wives who then spend it on household consumption. This is in line with the various intra-household games in the experiment, where husbands typically contributed more than their wives. Moreover, there is no sign of any undoing effect, in the sense of husbands confiscating or controlling their wives' earnings. Finally, spouses have very similar preferences, as reflected by their similar allocation of earnings from the experiment on their children.

We can explore the undoing problem further by looking at within-sample differences in bargaining strength. If the husbands' contributions to the household funds are driven by their intention to capture the funds ex post, we would expect such contributions to be positively correlated with their relative bargaining strength in the household. In Table 5 we demonstrate that this is not the case. The table shows contributions in the intra-household public goods and dictator games, separately for men and women, using two different proxies for relative bargaining power: relative education and relative income (Browning et al., 2014; Browning and Gørtz, 2012; Schaner, 2017).[15] We observe that there are no significant correlations between any of the bargaining power proxies and public goods or dictator game contributions.

at a general level, in the interviews the spouses were asked open questions about what they spent the earnings from the lab on, both their own earnings and those of the spouse. Hence, the similarity at the general level conceals a lot of variation in the details. Confirmation from the spouse on specific items that were bought strengthens the argument that the answers were truthful, as coordination on the detailed level is much more difficult than on the general level.

[15]The variable *Relative education* is the individual's level of education relative to the spouses' total level of education, where education levels are measured from 1 (no education) to 4 (more than secondary). The variable *Relative income* is the individual's daily income as a share of the spouses' total daily income.

Table 5: Intra-household games and relative bargaining power proxies.

	Public goods game				Dictator game			
	Men		Women		Men		Women	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Relative education	-0.015		0.137		-0.165		-0.086	
	(0.178)		(0.197)		(0.283)		(0.196)	
Relative income		-0.047		0.001		-0.022		-0.025
		(0.070)		(0.066)		(0.072)		(0.052)
Observations	211	211	211	211	81	81	81	81

Notes: OLS regressions of the shares of endowment given in public goods and dictator games by gender. Relative education is the spouse's education level as a share of the spouses' education levels. Income share is the spouse's daily income divided by the sum of both spouses' daily income. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

4.2 Generalizability

Our sample consists of fairly well-educated people in an urban area, and one may be worried that the large degree of intra-household cooperation is driven by an exceptionally strong position of women in our context. Ethiopia is undoubtedly a patriarchal society, ranked 121 out of 189 countries on the 2017 UNDP Gender inequality index, and on large household decisions, the husband is typically in charge. However, on everyday household decisions, which is the relevant level of decisions for our experiment, women *do* have a greater say (Jones et al., 2014; Bourey et al., 2012). This is also evident from questions on decision-making power in the household that we asked the participants in Lab 1, taken from the Demographic and Health Survey. We present the results in Table 6, and compare with the results from the Ethiopian 2016 DHS, broken down by urban and rural. In the table, decision-making power over a specific category means that the wife either decides alone or jointly with her husband.

Two important observations emerge from this table. First, the women in our sample respond very similarly to these questions as the DHS urban sample. Second, the table points to the relative autonomy of women in Ethiopia on every-day household decisions, particularly so in an urban context, but not dramatically different in a rural context. This, together with the fact that our results harmonize well with the findings from other household experiments that we cite in the Introduction, lends support to the external validity of our analysis.

Table 6: Percentage of women with decision-making power in the household.

	Our sample	DHS, Urban	DHS, Rural
Own earnings	89.2	96.2	89.2
Own health care	88.5	91.0	79.6
Major household purchases	85.4	88.1	76.3
Visits to family or relatives	90.8	91.7	82.2

Source: Ethiopia 2016 Demographic and Health Survey.

5 Conclusion

We explore intra-household cooperation of married couples in Ethiopia, contributing to the literature by using a wide range of experimental games and tasks as well as qualitative information and a follow-up survey to evaluate mechanisms and external validity. In contrast to much of the literature that documents conflicts within the household, we find that the spouses in our sample are highly cooperative and strikingly similar in their preferences and norms of behavior. We document such norms using third-party allocation tasks and qualitative evidence from interviews and focus group discussions.[16]

Our evidence points to spouses having clearly defined roles and where the husband makes transfers to his wife in order for her to purchase goods and services for the household running costs, such as for food and for the children, which can be seen as an expression of spousal cooperation in the *separate spheres* model of Lundberg and Pollak (1993).

Our follow-up survey of the participants supports the external validity of the experiment, with spending patterns very much in line with what we observed in the lab. Importantly, the undoing problem does not appear to be a serious concern for our study. Moreover, survey evidence from our sample on bargaining power aligns very well with similar evidence from the national DHS, which points to the representativeness of our sample and hence the generalizability of our results. Hence, while the participants in our study are not among the poorest in the local communities, they clearly do not represent an exclusive elite, and should therefore be relevant target groups for pro-poor policies. While our experiments shed light on the degree of conflict and cooperation in household decision making in a developing country context, further experiments are clearly called for before one can make firm conclusions and derive appropriate policy recommendations.

[16]Note that the observed agreement between spouses in terms of preferences and norms of behavior is not necessarily a sign of gender equality. As argued by Sen (1990), women's perceived preferences and beliefs may be heavily influenced by male dominant gender norms, and not necessarily in line with their preferences and beliefs in a more gender equal setting.

References

Akresh, R., de Walque, D. and Kazianga, H. (2016), 'Evidence from a randomized evaluation of the household welfare impacts of conditional and unconditional cash transfers given to mothers or fathers', *World Bank Policy Research Working Paper No. 7730* .

Anderson, S. and Baland, J.-M. (2002), 'The economics of roscas and intrahousehold resource allocation', *The Quarterly Journal of Economics* **117**(3), 963–995.

Andreoni, J. and Vesterlund, L. (2001), 'Which is the fair sex? Gender differences in altruism', *The Quarterly Journal of Economics* **116**(1), 293–312.

Ashraf, N. (2009), 'Spousal control and intra-household decision making: An experimental study in the Philippines', *The American Economic Review* **99**(4), 1245–1277.

Baland, J.-M. and Ziparo, R. (2017), 'Intra-household bargaining in poor countries', Helsinki: UNU-WIDER.

Beblo, M. and Beninger, D. (2017), 'Do husbands and wives pool their incomes? A couple experiment', *Review of Economics of the Household* **15**(3), 779–805.

Benhassine, N., Devoto, F., Duflo, E., Dupas, P. and Pouliquen, V. (2015), 'Turning a shove into a nudge? A "labeled cash transfer" for education', *American Economic Journal: Economic Policy* **7**(3), 86–125.

Bezu, S. and Holden, S. T. (2015), 'Generosity and sharing among villagers: Do women give more?', *Journal of Behavioral and Experimental Economics* **57**, 103–111.

Bourey, C., Stephenson, R., Bartel, D. and Rubardt, M. (2012), 'Pile sorting innovations: Exploring gender norms, power and equity in sub-Saharan Africa', *Global Public Health* **7**(9), 995–1008.

Browning, M. (2009), 'Love, betrayal and commitment', Unpublished manuscript, University of Oxford, Oxford, UK.

Browning, M., Chiappori, P.-A. and Weiss, Y. (2014), *Economics of the Family*, Cambridge University Press.

Browning, M. and Gørtz, M. (2012), 'Spending time and money within the household', *The Scandinavian Journal of Economics* **114**(3), 681–704.

Camerer, C. F. (2003), *Behavioral Game Theory: Experiments in Strategic Interaction*, Princeton, NJ: Princeton University Press.

Castilla, C. and Walker, T. (2013), 'Is ignorance bliss? The effect of asymmetric information between spouses on intra-household allocations', *The American Economic Review* **103**(3), 263–268.

Central Statistical Agency - CSA/Ethiopia and ICF (2017), 'Ethiopia Demographic and Health Survey 2016', Central Statistical Agency, Addis Ababa, Ethiopia. Addis Ababa, Ethiopia: CSA and ICF.

Charness, G., Gneezy, U. and Imas, A. (2013), 'Experimental methods: Eliciting risk preferences', *Journal of Economic Behavior & Organization* **87**, 43–51.

Couprie, H., Cudeville, E. and Sofer, C. (2017), 'Efficiency versus gender roles and stereotypes: An experiment in domestic production', *Experimental Economics* pp. 1–31.

Croson, R. and Gneezy, U. (2009), 'Gender differences in preferences', *Journal of Economic Literature* **47**(2), 448–474.

Dasgupta, U. and Mani, S. (2015), 'Only mine or all ours: Do stronger entitlements affect altruistic choices in the household', *World Development* **67**, 363–375.

Dave, C., Eckel, C. C., Johnson, C. A. and Rojas, C. (2010), 'Eliciting risk preferences: When is simple better?', *Journal of Risk and Uncertainty* **41**(3), 219–243.

Del Boca, D. and Flinn, C. (2012), 'Endogenous household interaction', *Journal of Econometrics* **166**(1), 49–65.

Duflo, E. and Udry, C. (2004), 'Intrahousehold resource allocation in Cote d'Ivoire: Social norms, separate accounts and consumption choices', NBER Working papers 10498, National Bureau of Economic Research.

Eckel, C. C. and Grossman, P. J. (1998), 'Are women less selfish than men? Evidence from dictator experiments', *The Economic Journal* **108**(448), 726–735.

Eckel, C. C. and Grossman, P. J. (2008), Differences in the economic decisions of men and women: Experimental evidence, *in* 'Handbook of Experimental Economics Results', Vol. 1, Elsevier, pp. 509–519.

Haddad, L., Hoddinott, J. and Alderman, H. (1997), 'Intrahousehold resource allocation in developing countries: Models, methods and policies', *International Food Policy Research Institute* .

Haushofer, J. and Shapiro, J. (2016), 'The short-term impact of unconditional cash transfers to the poor: Experimental evidence from Kenya', *The Quarterly Journal of Economics* **131**(4), 1973–2042.

Hoddinott, J. and Haddad, L. (1995), 'Does female income share influence household expenditures? Evidence from Côte d'Ivoire', *Oxford Bulletin of Economics and Statistics* **57**(1), 77–96.

Hoel, J. B. (2015), 'Heterogeneous households: A within-subject test of asymmetric information between spouses in Kenya', *Journal of Economic Behavior & Organization* **118**, 123–135.

Iversen, V., Jackson, C., Kebede, B., Munro, A. and Verschoor, A. (2011), 'Do spouses realise cooperative gains? Experimental evidence from rural Uganda', *World Development* **39**(4), 569–578.

Jones, N., Tefera, B., Stephenson, J., Gupta, T. and Pereznieto, P. (2014), 'Early marriage and education: The complex role of social norms in shaping Ethiopian adolescent girls' lives', Country Report: Shaping policy for development.

Kebede, B., Tarazona, M., Munro, A. and Verschoor, A. (2013), 'Intra-household efficiency: An experimental study from Ethiopia', *Journal of African Economies* **23**(1), 105–150.

Konow, J. (2009), 'Is fairness in the eye of the beholder? An impartial spectator analysis of justice', *Social Choice and Welfare* **33**(1), 101–127.

Ledyard, J. O. (1995), Public goods: A survey of experimental research], *in* J. Kagel and A. Roth, eds, 'The Handbook of Experimental Economics', Princeton: Princeton University Press.

Lundberg, S. J., Pollak, R. A. and Wales, T. J. (1997), 'Do husbands and wives pool their resources? Evidence from the United Kingdom child benefit', *Journal of Human Resources* **32**(3).

Lundberg, S. and Pollak, R. A. (1993), 'Separate spheres bargaining and the marriage market', *Journal of Political Economy* **101**(6), 988–1010.

Mani, A. (2011), 'Mine, yours or ours? The efficiency of household investment decisions: An experimental approach', CAGE Online Working Paper Series.

Munro, A. (2018), 'Intra-household experiments: A survey', *Journal of Economic Surveys* **32**(1), 134–175.

Ringdal, C. and Sjursen, I. (2017), 'Household bargaining and spending on children: Experimental evidence from Tanzania', WIDER Working Paper 2017/128.

Schaner, S. (2017), 'The cost of convenience? Transaction costs, bargaining power, and savings account use in Kenya', *Journal of Human Resources* **52**(4), 919–945.

Sen, A. (1990), Gender and cooperative conflicts, Chapter 8, *in* I. Tinker, ed., 'Persistent inequalities: Women and World Development', Oxford University Press, New York.

Thaler, R. H. (1999), 'Mental accounting matters', *Journal of Behavioral decision making* **12**(3), 183–206.

Thomas, D. (1990), 'Intra-household resource allocation: An inferential approach', *Journal of Human Resources* pp. 635–664.

Thomas, D. (1993), 'The distribution of income and expenditure within the household', *Annales d'Economie et de Statistique* pp. 109–135.

Yoong, J., Rabinovich, L. and Diepeveen, S. (2012), 'The impact of economic resource transfers to women versus men: A systematic review', London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Appendix A Additional figures and tables

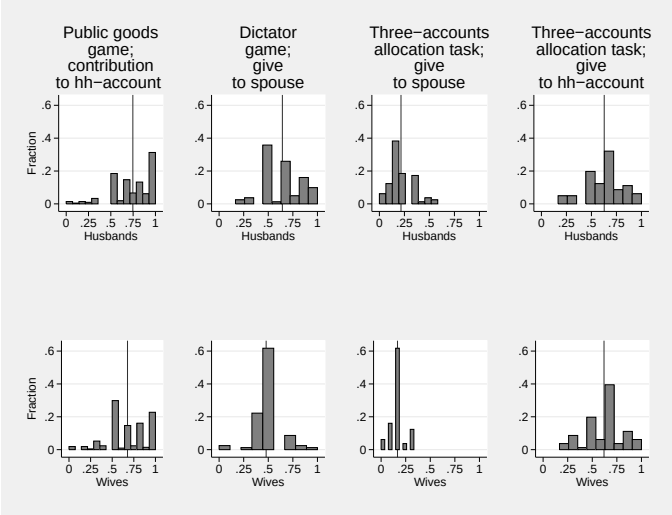


Figure A1: Histograms of contributions in intra-household games, by gender.

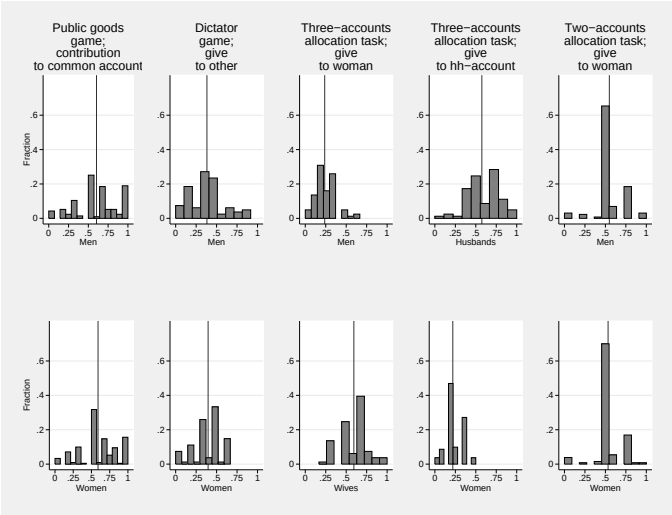


Figure A2: Histograms of contributions in extra-household games, by gender.

Table A1: Robustness to alternative regression specifications: Intra-household public goods and dictator games.

	Tobit Regression		Couple Fixed Effects		Alternative Educ and Income		Include Spouse's Income	
Dependent variable:								
Share of endowment contributed to hh-account (PGG)								
Share of endowment given to spouse (DG)	PGG	DG	PGG	DG	PGG	DG	PGG	DG
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	-0.063**	-0.148***	-0.043**	-0.143***	-0.042*	-0.147***	-0.053**	-0.146***
	(0.028)	(0.027)	(0.022)	(0.034)	(0.023)	(0.029)	(0.022)	(0.030)
Expected contribution	0.694***		0.476***		0.495***		0.489***	
	(0.072)		(0.073)		(0.055)		(0.055)	
Higher education	0.062**	0.024	0.041	0.010	0.030**	-0.019	0.045*	0.020
	(0.031)	(0.030)	(0.039)	(0.062)	(0.013)	(0.017)	(0.024)	(0.032)
Income	0.045	-0.011	-0.009	0.007	0.000	0.000	0.024	-0.008
	(0.031)	(0.032)	(0.033)	(0.058)	(0.000)	(0.000)	(0.023)	(0.034)
Spouse's income							0.025	0.002
							(0.023)	(0.034)
Years married	0.000	-0.000	0.002	-0.003	0.001	-0.001	0.001	-0.000
	(0.002)	(0.001)	(0.005)	(0.009)	(0.002)	(0.001)	(0.002)	(0.001)
Children	0.016	0.006	0.004	-0.004	0.010	0.002	0.010	0.006
	(0.013)	(0.008)	(0.023)	(0.019)	(0.010)	(0.008)	(0.010)	(0.008)
Risk averse	0.013	-0.006	0.036	-0.023	-0.002	-0.009	0.006	-0.002
	(0.031)	(0.030)	(0.031)	(0.052)	(0.024)	(0.030)	(0.024)	(0.031)
Patient	-0.095***	0.005	-0.084*	-0.000	-0.066**	0.008	-0.068**	0.008
	(0.035)	(0.037)	(0.045)	(0.052)	(0.028)	(0.037)	(0.027)	(0.037)
Observations	422	162	422	162	421	161	422	162
Lab dummy	Yes	-	Yes	-	Yes	-	Yes	-
Session dummies	-	Yes	-	Yes	-	Yes	-	Yes
Treatment dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Couple fixed effects	-	-	Yes	Yes	-	-	-	-

Notes: Higher education takes the value one if the respondent has more than secondary school, except for in column (5)-(6) where the variable is a factor variable 1-4 (no education, primary school only, secondary school, or higher than secondary school); The income variable takes the value one if the respondent earns more than or equal to the median weekly income of their respective gender (ETB 425 for men and ETB 225 for women), except for in column (5)-(6) where the variable is continuous in Ethiopian Birr; Children is the number of children reported by the participant. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A2: Public goods games, varying endowments.

	Public goods game			
	Intra-household		Extra-household	
	(1)	(2)	(3)	(4)
Female	-0.030	-0.007	0.025	0.023
	(0.048)	(0.036)	(0.050)	(0.049)
High endowment	0.045	0.033	-0.020	-0.036
	(0.045)	(0.039)	(0.050)	(0.047)
Female*High endowment	-0.123*	-0.101*	-0.098	-0.081
	(0.065)	(0.056)	(0.068)	(0.066)
Expected contribution		0.511***		0.257***
		(0.069)		(0.063)
Income		0.023		0.042
		(0.030)		(0.035)
Higher education		0.059*		0.043
		(0.031)		(0.036)
Risk averse		0.015		-0.025
		(0.034)		(0.037)
Patient		-0.088**		-0.039
		(0.042)		(0.044)
Female + Female*High endowment	-0.153***	-0.109**	-0.072	-0.058
	(0.044)	(0.043)	(0.045)	(0.046)
Observations	260	260	260	260
Adjusted R-squared	0.03	0.32	0.01	0.08
Lab dummies	-	-	-	-
Session dummies	-	Yes	-	Yes
Couple controls	-	Yes	-	Yes
Mean of contribution by low endowed men	0.739	0.739	0.650	0.650
Mean of contribution by high endowed men	0.784	0.784	0.630	0.630

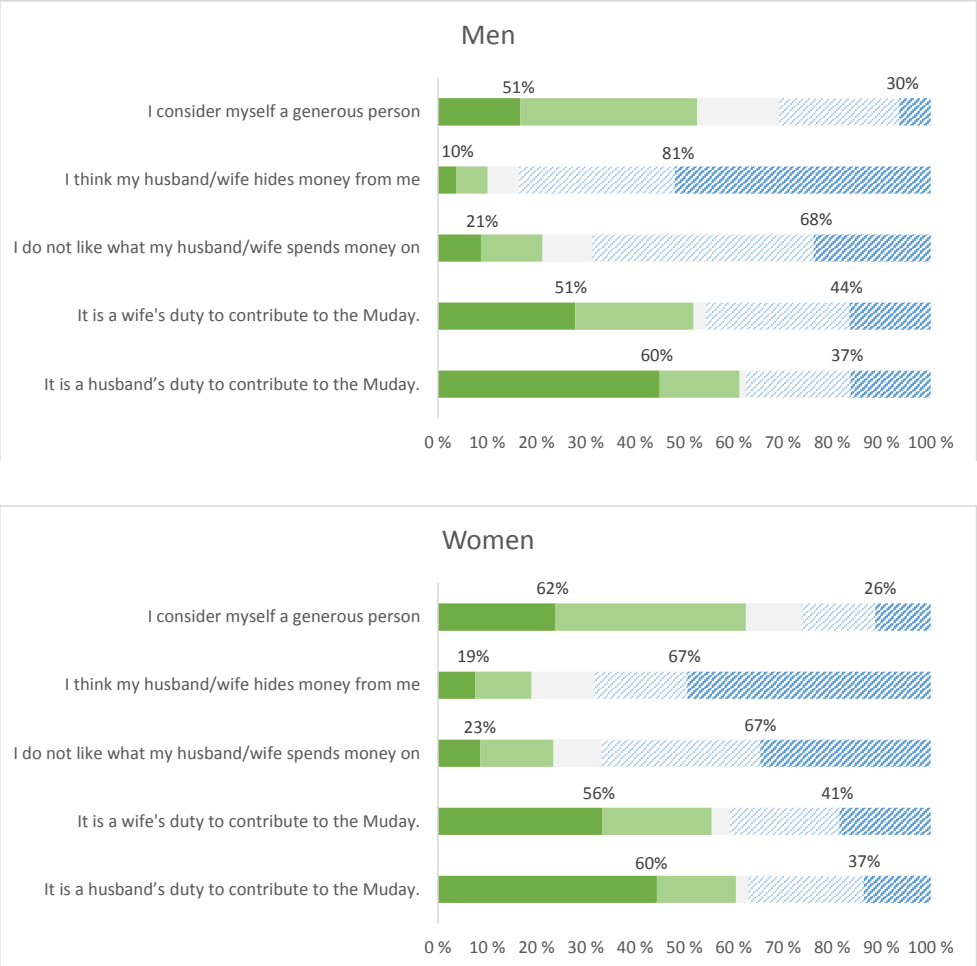
Notes: OLS regressions of shares given in public goods game with unequal endowments (only lab 1 participants). The income variable takes the value one if the respondent earns more than or equal to the median weekly income of their respective gender (ETB 425 for men and ETB 225 for women); Higher education takes the value one if the respondent has more than secondary schooling; Couple controls include years married and number of children reported by the participant. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A3: Third-party two-account allocation tasks.

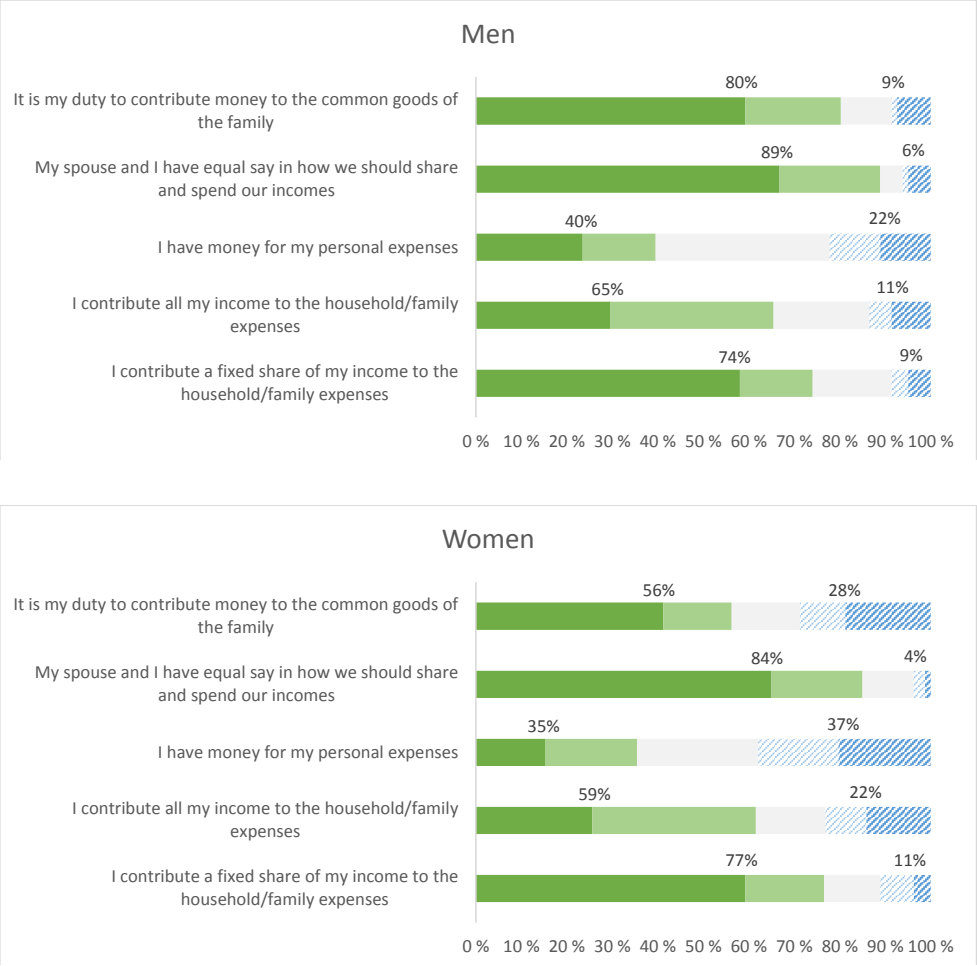
PANEL A: **Pooled (as in Table 3)**	Men	Women	Difference
Two-accounts allocation:			
Husband's share	0.45 (0.17)	0.47 (0.16)	-0.02 [0.02]
Wife's share	0.55 (0.17)	0.53 (0.16)	0.02 [0.02]
Difference in spouses' shares	-0.1 [0.03]***	-0.06 [0.03]**	
PANEL B: **Husband 60 / Wife 20**	Men	Women	Difference
Two-accounts allocation:			
Husband's share	0.45 (0.19)	0.46 (0.17)	-0.01 [0.01]
Wife's share	0.55 (0.19)	0.54 (0.17)	0.01 [0.01]
Difference in spouses' shares	-0.09 [0.03]**	-0.07 [0.03]*	
PANEL C: **Husband 20 / Wife 60**	Men	Women	Difference
Two-accounts allocation:			
Husband's share	0.45 (0.17)	0.47 (0.18)	-0.02 [0.02]
Wife's share	0.55 (0.17)	0.53 (0.18)	0.02 [0.02]
Difference in spouses' shares	-0.1 [0.03]**	-0.06 [0.03]*	

Notes: Mean shares allocated to the different accounts. Standard deviations in parenthesis and standard errors in square brackets. *** p<0.01, ** p<0.05, * p<0.1.

Appendix B Survey questions



Note: Survey questions on intra-household cooperation from Lab 2. The numbers represent the aggregated share of *Strongly agree* and *Agree* (green/full color) and for *Strongly disagree* and *Disagree* (blue/striped) answers.



Note: Survey questions on intra-household cooperation from Lab 2. The numbers represent the aggregated share of *Always* and *Almost always* (green/full color) and for *Rarely* and *Never* (blue/striped) answers.

Appendix C Experimental instructions and handouts

This section provides the instructions given for the lab games. The sequence follows as presented in the article.

Cooperation games (Lab 1 and Lab 2)

The following instructions were given in plenum with use of blackboard, traditional baskets and real money with the help of research assistants role-playing partners in the public goods game.

Public goods games

In this activity, not only your choice, but also the choice of another participant will determine your earnings. You will be paired with another participant. Both of you will receive 60 Birr and you will both allocate the money between yourself and a shared basket. Whatever is placed in the common basket by you and the other person is increased by 50% by the experimenters, and then shared equally between the two of you. This money will be placed in your common-envelope which you will receive at the end. The money you put in the private basket is not increased, but goes directly to your private envelope.
Let us explain with three examples:

Example 1: Let's say that you put your 60 Birr in the common basket, leaving zero for your private basket, and the other person puts 20 Birr in the common basket, leaving 40 Birr in his or her private basket. There will then be $60 + 20 = 80$ Birr to which we will add 50 percent (equal to 40 Birr). The common basket will now contain $80 + 40 = 120$ Birr, which will then be divided equally between you and the other person, so that each of you will get $120/2 = 60$ Birr.

Your earning: 60 birr = 0 Birr in your private basket + 60 Birr from the common basket.

The other person's earning: 100 Birr = 40 Birr in partner's private basket + 60 Birr from the common basket.

Example 2: Let's say you place 0 Birr in the common basket, so that you all your 60 Birr in your private basket, and the other person places 40 Birr in the common basket, leaving 20 for his or her private basket. The common basket will then contain 40 Birr to which we will add 50% (equal to 20 Birr). The common basket will now contain $40 + 20 = 60$ Birr, which will then be divided equally between you and the other person, so that each of you will get $(60/2 =)30$ Birr. Your total earning from this activity would then be:

Your earning: 90 Birr = 60 Birr placed in your private basket + 30 Birr from the group basket.

The other person's earning: 50 Birr = 20 Birr in partner's private basket + 30 Birr from the common basket.

Is this understood? Think about the following example:

Example 3: Let's say that you place 60 Birr in the common basket, leaving zero for your private basket, and the other person places 40 Birr in the common basket, leaving 20 Birr for his or her private basket, how much will you and the other person get?

The common basket will contain _____ (100) Birr to which we will add 50% (equal to _____ (50) Birr). The common basket will now contain _____ (150) Birr, which will then be divided equally between you and the other person, so that each of you will get _____ (75) Birr. Your total earning from this activity would then be:

Your earning: _____ (75) Birr = _____ (0) Birr placed in your private basket + _____ (75) Birr from the common basket.

The other person's earning: _____ (95) = _____ (20) Birr placed in the other person's private basket + _____ (75) Birr from the common basket.

You will do this activity two times, one with a randomly chosen other participant and once with your spouse.

Text on handout
In this activity, you will allocate 60 Birr between your own basket and the common basket. The allocation must sum to 60. Whatever is placed in the common basket by you and the other person is increased by 50% by the experimenters, and then shared equally between the two of you and placed in your private envelopes. The money you put in the private basket is not increased. The choices you make is completely anonymous, that means that no one, not even the experimenter, will know your choice.

1. You and another participant
Of your 60 Birr, how much do you wish to keep for yourself, and how much do you wish to pass to the common basket for yourself and another participant?

If the other person is a *man*:		**If the other person is a *woman*:**	
Keep for self:	_____ Birr	Keep for self:	_____ Birr
Pass to common basket:	_____ Birr	Pass to common basket:	_____ Birr
Total:	60 Birr	Total:	60 Birr

2. You and your spouse

Of your 60 Birr, how much do you wish to keep for yourself, and how much do you wish to pass to the common basket for yourself and your spouse?

Keep for self: _____Birr
Pass to common basket: _____Birr
Total: ___60____Birr

Before we go on to the next activity. We want to ask about your beliefs in the activity we just had. How much do you think the other participant contributed? We will give you an additional 5 Birr if you guess the correct amount your partner contributed to the common basket.

1. Your belief about the other participant
Of the 60 Birr the other participant had, how much do you believe he or she kept for himself or herself, and how much do you believe he or she passed to the common basket?

If the other person was a *man*:	**If the other person was a *woman*:**
Keep for himself: _____Birr	Keep for herself: _____Birr
Pass to common basket: _____Birr	Pass to common basket: _____Birr
Total: ___60____Birr	Total: ___60____Birr

2. Your belief about your spouse
Of the 60 Birr your spouse had, how much do you believe he or she kept for himself or herself, and how much do you believe he or she passed to the common basket?

Keep for himself or herself: _____Birr
Pass to common basket: _____Birr
Total: ___60____Birr

Dictator game

In this activity, you have 60 Birr that you can divide between your own basket and to another participant's basket. You will all do this allocation, but whether you are a giver or receiver will be chosen by a random draw. You will do this allocation choice two times, where the other participant will either be a randomly selected person in this session or your own spouse. You may choose to keep it all for yourself and give nothing to the other participant, or to keep some and pass the remainder to the other participant. Please note that no one else in this room will know your decisions in this, or any other activity, in the session.

Please indicate in the spaces on the sheet how much of the 60 Birr you choose to keep for yourself and how much you choose to pass to the other participant. Once you have made your decisions turn your sheet and wait for the research assistants to collect your sheet.

Text on handout
In this activity you will allocate 60 Birr between yourself and another participant. The allocation must sum to 60.

1. You and another participant
Of your 60 Birr, how much do you wish to put in your personal basket, and how much do you wish to pass to another participant's basket?

Keep for self: _____Birr
Pass to other participant: _____Birr
Total: ___60____Birr

2. You and your spouse
Of your 60 Birr, how much do you wish to put in your personal basket, and how much do you wish to pass to your spouse's basket?

Keep for self: _____Birr
Pass to spouse: _____Birr
Total: ___60____Birr

Allocation tasks (Lab 1 and Lab 2)

Extra-household: Two-accounts allocation (Lab 1)

Now we move on to a new activity. In this activity, you will decide how another couple in this study should divide their incomes between themselves. The other couple, as well as the choice upon which the payment is made, will be randomly chosen by the computer and is anonymous: you will not know who the other couple is and they will not know who you are.

Text on handout
How do you want to divide the money between the husband and wife? Tick off.

A. If the husband has income of 60 Birr and the wife has income of 20 Birr, so in total they have 80 Birr?

B. If the wife has income of 60 Birr and the husband has income of 20 Birr, so in total they have 80 Birr?

Wife	80 Birr	60 Birr	40 Birr	20 Birr	0 Birr
Husband	0 Birr	20 Birr	40 Birr	60 Birr	80 Birr

Extra- and Intra-household: Three-accounts allocation

In this activity you will allocate a sum of money on behalf of another couple in this session. The other couple will be randomly chosen by the computer and is anonymous: you will not know who the other couple is and they will not know who you are.

You have 60 Birr to allocate between the husband, the wife, and their Muday. As in the last activity, whatever you pass to the Muday will be increased with 50% by the experimenters.

Text on handout

In this activity you will allocate 60 Birr on behalf of another couple in this study. You have 60 Birr. How would you like to allocate this between the husband, the wife and their common basket (Muday), which is intended to be spent on household expenses and the children? Note that whatever you pass to the Muday, the experimenters will add 50% of the contribution to the Muday.

1. Allocation on behalf of another couple

You have 60 Birr. How would you like to allocate this between the husband, the wife and their common basket, which is intended to be spent on household expenses and the children? Note that whatever you pass to the Muday, the experimenters will add 50% of the contribution to the Muday.

To the husband: _____Birr
To the wife: _____Birr
To the household account: _____Birr
Total: ___60___Birr

2. Allocation on behalf of your household

If you allocated on behalf of your own household: How would you like to allocate 60 Birr between your spouse, yourself, and your common basket (Muday).

To your spouse: _____Birr
To yourself: _____Birr
To the household account: _____Birr
Total: ___60___Birr

3. Please give a short explanation of your motives in the allocation on behalf of your household.

Preferences (Lab 1 and Lab 2)

Risk aversion and time preference

(No public introductions)

Text on handout

1. Risk
What would you prefer? Tick off your choice:

100 Birr for sure	OR	250 Birr with 50% probability and 0 Birr with 50% probability

1. Time preference
What would you prefer? Tick off your choice:

A. First time choice

100 Birr today	OR	150 Birr one month from now

B. Second time choice

100 Birr one year from now	OR	150 Birr one year and one month from now

Chapter 2

Factory Employment and Fertility Decisions: Experimental Evidence from Ethiopia

Factory Employment and Fertility Decisions: Experimental Evidence from Ethiopia

Sandra Kristine Halvorsen, Andreas Kotsadam and Espen Villanger*

Abstract

We examine the causal relationship between female employment and fertility in a high-fertility, low-employment context by using an experimental approach. The sample consists of 1,140 married women who had applied for jobs in Ethiopia's expanding light manufacturing industry, half of whom were randomly assigned to the treatment group being offered a job and the other half being assigned to the control group. Surveying the participants over a period of fourteen months after the application, we find strong effects of the treatment on income and employment, but no effect on actual fertility or preferred lifetime fertility, investments in children, or household decision-making power.

JEL classification: J13, O14, D12.
Keywords: Female labor force participation, fertility, women's empowerment, household decision-making power, Ethiopia, field experiment.

*Sandra Kristine Halvorsen: The Norwegian School of Economics and Christian Michelsen Institute, e-mail: sandra.halvorsen@cmi.no. Andreas Kotsadam: The Ragnar Frisch Centre for Economic Research, e-mail: andreas.kotsadam@frisch.uio.no. Espen Villanger: Christian Michelsen Institute, e-mail: espen.villanger@cmi.no. We would like to thank EDRI (Ethiopian Development Research Institute) and especially Tigabu Degu Getahun for outstanding field implementation. We also thank Kjetil Bjorvatn, Niclas Buheren, Ingrid Hoem Sjursen, and Maria Polipciuc for valuable comments. The project was supported by The Research Council of Norway through the project: "Women in the developmental state: Female employment and empowerment in Ethiopia" project No 240807. An analysis plan is pre-registered at the AEA RCT registry (https://doi.org/10.1257/rct.2414-1.0.) and deviations from the plan are noted in the text.

1 Introduction

Rapid population growth poses a number of challenges to the economic and social development of many sub-Saharan African countries (Beegle and Christiaensen, 2019; Casterline, 2010; United Nations, 2019). For instance, at 2.7 percent annually, the population growth in Africa considerably raises the need for investments in education and health services and puts increased pressure on urban infrastructure development and labor markets (Beegle and Christiaensen, 2019). Policies aimed at reducing fertility rates may, therefore, have widespread societal benefits in these countries and set them on a path to sustainable development (Dasgupta, 1995; Pritchett, 1994).

Increased female labor force participation is correlated with declining fertility rates across the developing world (Madsen et al., 2018). Hence job creation for women through, for instance, manufacturing sector development, may indirectly affect fertility and thus yield a "double win" for development by reducing fertility rates and promoting economic growth (Ager and Herz, 2019; Franck and Galor, 2015; Heath and Mobarak, 2015). Despite the central role of industrial policy in many developing countries, there is limited evidence on the causal relationship of manufacturing employment on fertility decisions.

We identify the effects of manufacturing jobs in Ethiopia on pregnancies and preferred lifetime fertility using a randomized field experiment.[1] With an oversupply of manufacturing job applicants, we randomly selected eligible applicants to receive a job offer from the factory they applied to, while the other half of our sample of eligible applicants did not receive a job offer and served as a control group. Our data also allows us to explore the labor supply and empowerment channels through which employment may affect fertility choices.

We find that individual level factory employment does not affect the likelihood of pregnancy or childbirth, nor does it affect women's preferred lifetime fertility. When exploring the mediating channels of labor supply and empowerment, we find that being offered a job causes higher income and labor supply, but does not affect investments in children nor household decision-making power. The absence of effects on these variables is a possible explanation for the null result on households' fertility decisions, since these mediators are theoretically and empirically associated with lower fertility, for instance, through the quantity-quality trade-off of children. If so, on the one hand, the results imply that the high hopes for a fertility reduction with increased industrialization should be tempered. On the other hand, it is possible that the effects of industrialization operate at a more aggregate level. If that it is the case, our results challenge the theories of direct income effects for reduced fertility.

By randomizing job offers to eligible applicants we contribute to the literature on the links between female labor force participation and fertility. While a negative correlation between female labor force participation and fertility is well established in the literature, the causal direction within this relationship is complex and inherently difficult to verify empirically (Browning, 1992; Lehrer and Nerlove, 1986). Causal investigations of the relationship

[1] Preference for the number of total children may indicate the long-term effect of the manufacturing jobs on lifetime fertility as people's preferences and social norms are important determinants for their fertility choices (Bongaarts and Watkins, 1996; Casterline, 2010; Pritchett, 1994).

have mainly focused on the effect of children on female labor force participation using various exogenous changes in family size.[2] These studies generally find support for the theoretical predictions of a negative relationship between the number of children a woman has and her labor supply. There are, however, only a few studies investigating the reverse causal direction, namely the effect of labor force participation on fertility. Two prominent examples are Jensen (2012) and Heath and Mobarak (2015) who use exogenous variation in access to jobs. In a field experiment in India, Jensen (2012) randomize recruitment services for young women in the business process outsourcing industry. He finds that the age of marriage and first childbirth is higher in the villages receiving the recruitment service due to higher female labor force participation and more educational attainment. The women also reported higher career aspirations and wanting fewer children. Heath and Mobarak (2015) use variation in proximity to the garment industry as an identification strategy. They find that girls living in villages close to the factories delay marriage and childbirth compared to women living further away, or compared to women in the same village before the factories arrived. These effects occur as a result of younger girls increasing their school attendance and older girls increasing their labor force participation. In contrast to these studies that identify an effect of employment opportunities at the village level, we are the first to identify the causal effect of individual level employment on fertility decisions.

We also contribute to the literature on the impacts of manufacturing employment on medium term earnings and employment. In a similar context to our study, Blattman and Dercon (2018) randomize entry-level applicants in five industrial firms in Ethiopia into one of three treatment arms; an industrial job offer, a self-employment start-up grant, or to a control group. They find no effect of the industrial job offer on wages or employment after one year. In contrast, we find substantial effects after one year. A notable difference between their study and ours is the fact that their study consists of single women living in urban areas, while our study sample is more geographically spread and mainly includes already married women. In an extension of their study, Blattman et al. (2019) report on the outcomes of the trial five years later and find that any differences between the three groups that were present after one year, have disappeared over time.[3]

Our results also speak to the literature on the links between manufacturing employment and household decision-making power. Atkin (2009) exploits the variation in the number of factory jobs available for women in Mexico at the age of 16, the age a woman may first enter the labor market. He finds that women who had their first job in the manufacturing industry have significantly taller children. He further argues that this is not an income effect alone, but that these women had higher bargaining power within their households. In another

[2]Examples include twins (Rosenzweig and Wolpin, 1980*a*; Angrist and Evans, 1998; Bronars and Grogger, 1994; De Jong et al., 2017), the sex composition of the first two children (Angrist and Evans, 1998; Cruces and Galiani, 2007), infertility (Aguero and Marks, 2008), abortion legislation (Bloom et al., 2009; Angrist and Evans, 2000), and access to the contraceptive pill (Bailey, 2006).

[3]Fertility is not a main outcome in their study, yet in Blattman et al. (2019) they report pregnancies and live births outcomes in the Appendix. They find no effect of job offers on fertility after five years, however, since they do not have an effect on employment either, their study does not teach us anything about the relationship between employment and fertility.

study, Majlesi (2016) finds that higher relative labor market opportunities for women than for men, measured by changes in the manufacturing sector composition, increases women's relative decision-making power within households. Again, we contribute by being able to identify the individual level effects of employment.

Ethiopia makes an excellent setting for testing whether employment affects fertility decisions. First, Ethiopia has a high fertility rate. In 2017 the fertility rate was 4.35 children per woman (The World Bank, 2019). Policies aimed at limiting the population growth were first implemented in Ethiopia in 1993 and while the total fertility rate has decreased significantly since then (from 7.1), it still has one of the highest in the world (rank 33 of 200 countries, The World Bank, 2019). Ethiopia is also one of the most populous and fastest-growing countries in the world. Its population exceeded 100 million in 2017 and is estimated to double within the next three to four decades (United Nations, 2019). Such rapid population growth poses a number of challenges for the country including pressure on natural resources, poverty, housing, education, and access to health and employment opportunities (United Nations, 2019). Second, industrial development is core to Ethiopia's structural transformation agenda and the government has outlined a plan to become a leading manufacturing hub in Africa and globally (NPC, 2016). The manufacturing sector has accelerated its growth in the recent years reaching 17.9 percent output growth in 2015-17 (Oqubay, 2018) and is expected to create 1.5 million new jobs for women and youth by 2025 (NPC, 2016). In 2016, when we started the project, only 33% of the women in the Demographic and Health Survey (DHS) reported being currently employed, compared to 88% of the men (DHS, 2016).[4] Given the widespread unemployment among women and the strategic focus on female employment, manufacturing growth has potentially important impacts on the lives of women and their households.

The paper is organized as follows. In Section 2 we discuss the theoretical framework which guides the analysis. In Section 3 we describe the context, experimental design, sample, and estimation strategy. Sections 4 and 5 present the experimental results and sensitivity analyses. The main findings are summarized in the conclusion in Section 6.

2 Theoretical framework

Based on both the theoretical and empirical literature on family economics, women's labor force participation is likely to affect fertility through three channels: The income, substitution, and empowerment (bargaining power) channels (e.g., Blundell et al., 2005; Schultz, 1990). Following the pioneering work of Mincer (1962, 1963), Becker (1965), Becker and Lewis (1973), and Willis (1973) there is now a rich economics literature on household behavior and choices. While the initial theoretical contributions modeled the household as a single decision-making entity, more recent contributions allow for multiple decision makers with

[4]Currently employed is defined in the DHS as: "Respondents who were employed in the 7 days before the survey; includes persons who did not work in the past 7 days but who are regularly employed and were absent from work for leave, illness, vacation, or any other such reason."

different preferences and bargaining power within the same household. Outcomes are then viewed as a result of a bargaining process (Manser and Brown, 1980; McElroy and Horney, 1990; Browning and Chiappori, 1998; Chiappori, 1988, 1992; Lundberg and Pollak, 1993).

Due to women's central role in child rearing, the analysis of household decisions on consumption and labor supply is inseparable from the analysis of fertility choice. Building on the neoclassical framework for labor supply, Mincer (1962) and Becker (1965) argue that a household's allocation of time is not only a choice between labor and leisure time, but it should also include home production. A household is both a consuming and a producing unit, where home production, such as cooking and childcare, contributes to available goods for consumption. A household may thus allocate the time of its members between labor market activities, home production, and leisure, subject to prices and each member's marginal productivity and wage rates. As women's market wage increases, the opportunity cost of leisure and home production rises, which makes it relatively more profitable to work. This effect, in isolation, would induce women to spend less time on childcare and could lead to reduced fertility (the substitution effect). However, the effect of increased income raises the demand for non-market time and could, in isolation, lead to higher demand for leisure and spending time with children (the income effect). This effect could then work in the opposite direction of the substitution effect. Theoretically, the net effect could go in both directions and can only be verified empirically.

Assuming that the substitution effect is higher than the income effect, the net effect of higher paying labor opportunities is an increase in women's labor market participation and hence an increase in households' income, which enables higher household consumption. In economics, children are often considered normal goods, thus an increase in income is likely to increase the amount spent on children. However, considering children as durable goods, increased income leads to higher demand in both the quantity as well as the quality of the children. It is generally assumed that the elasticity for quantity is much smaller than the elasticity for quality (Becker, 1960). Thus, as households are constrained by their budget, an increase in income may lead to a decrease in the demand for children if parents invest more in quality per child (better schooling, health, clothes etc.), instead of increasing the number of children (Becker, 1960; Becker and Lewis, 1973; Becker and Tomes, 1976; Willis, 1973).[5] However, in cases where the initial ideal number of children is high, an increase in income may enable parents to reach their desired number of children, thus increasing the likelihood of pregnancy in the short-term.

Another channel for which increased labor market opportunities and labor participation may affect fertility is by the empowerment of women. In bargaining models of the household, outcomes depend on the relative preferences and power between decision-makers (Browning et al., 2014). Since spouses may have different preferences over fertility, and they need to

[5]The quantity-quality model has been tested in several contexts, although with mixed results. Using twin births or preference for a mixed sibling-sex composition Black et al. (2005) and Angrist et al. (2010) find small or no effects of family size on children's education, thus they disregard the quantity-quality model. In contrast, using data from India and China, Rosenzweig and Wolpin (1980*b*) and (Rosenzweig and Zhang, 2009) do find evidence for the quantity-quality trade-off.

agree in order to realize the outcome (i.e., the desired number of children), incorporating household bargaining in fertility analysis is seen as an important part of research on fertility determinants (Ashraf et al., 2014; Doepke and Tertilt, 2018).[6] Women's participation in the labor market may increase their status and decision-making power in the household through their increased contribution to the household (Getahun and Villanger, 2018; Luke and Munshi, 2011) and increased control over own income (Anderson and Eswaran, 2009).[7] In fact, simply increasing employment opportunities for women in an area may increase a woman's autonomy and household bargaining power regardless of whether she takes employment or not by increasing her outside options (Pollak, 2005; Dharmalingam and Morgan, 1996; Majlesi, 2016). The empirical literature on the relationship between women's empowerment and fertility in developing countries generally finds that female empowerment is associated with smaller family size, longer birth intervals, and lower rates of unintended pregnancies (Doepke and Tertilt, 2018; Upadhyay et al., 2014). Empirical studies have also found that when women control larger shares of the household income due to an exogenous change in the receiver of the income, the household spends more on children's health, nutrition, and education (Duflo, 2003; Duflo and Udry, 2004; Lundberg et al., 1997). Hence, female empowerment may change the relative cost of children in households, which further affects the demand for children through the quantity-quality trade-off.

In sum, the theoretical predictions of the relationship between employment and fertility goes in both directions, and the effects are likely to be context dependent and contingent on the effects of employment on mediating factors. Our data allows us to test whether employment causes higher incomes, labor supply, child expenditures, and decision-making power. We can also explore whether incomes, child expenditures, and decision-making power are correlated with actual and preferred fertility.

3 The background and experimental design

3.1 The manufacturing industry

Ethiopia is in a favorable position to develop a leading global manufacturing industry. It has low labor costs, an abundance of resources, both human and natural, low electricity costs from hydro-power, and high-quality cotton production. Moreover, under the Africa Growth and Opportunity Act and the Cotonou Agreement, Sub-Saharan countries enjoy duty-free and quota-free access to the U.S. and EU markets when exporting light manufacturing goods such as textiles and leather (Dinh et al., 2012). The Ethiopian government is aiming to lift

[6] There are also other channels than different preferences for which empowerment may affect fertility, for instance by increased education which increases the opportunity cost of childbearing. However in this paper we only focus on preferences and household decision-making power as a possible channel for how empowerment can affect fertility.

[7] Female labor participation may also have negative impacts on their relative power. Heath (2014) finds that women who suffer lower agency have higher risks of intimate partner violence after starting work in garment factories in Bangladesh because husbands use violence as an instrument to sustain their relative position.

Ethiopia from today's position as one of the poorest countries in the world, to a middle-income country by 2025 and industrialization and manufacturing play important strategic roles in this development (Growth and Transformation Plan II, NPC, 2016).

The textile and leather sectors are the largest manufacturing exports in Ethiopia.[8] As a policy strategy, these sectors are prioritized by the government due to their labor-intensive nature. The Ethiopian government offers direct support for further development of the industry by accommodating private investors with economic incentives and capacity building.[9] Traditionally, these manufacturing sectors mainly employ women, the argument being that women have a competitive advantage with their 'nimble fingers' and docile behavior (Elson and Pearson, 1981).

One of the main competitive advantages of Ethiopian manufacturing is low labor costs. From the workers' side, however, this may be one of the main drawbacks of the newly established job opportunities the sector offers. In our sample, the mean monthly wage is ETB 964, which amounts to PPP adjusted USD 112 per month, or about 3.7 dollars a day.[10]

3.2 The experimental design

We investigate whether a job offer in the manufacturing sector affects households' fertility decisions and preferences by conducting a randomized field experiment. Twenty-six privately owned factories located in five industrial parks in different regions in Ethiopia participated in our study. The location of the industrial parks is shown in Figure 1.

The majority of the firms (72%) were in the textile and garment industries, while the remaining firms were in plastic, cosmetic, and chemical industries. The factories were all medium or large (>10 employees) and the average number of workers was 1,134 with 79% female workers at the median. The sample includes both international and locally-owned enterprises. Each firm agreed to randomize job-offers to a subset of eligible job applicants consisting of married women (or women living with a partner).

Factories commonly hire in cohorts of applicants at a time. When choosing among applicants, some of the firms have formal requirements such as proof of completed 8th or 10th grade. Some firms also make the applicants perform various tests, such as handling a sewing machine and following patterns. The exact tests are different in each firm, but with the same aim, namely to single out ineligible applicants. At the day of recruitment, married women who passed the formal or informal requirements were directed to our research team and invited to participate in our study. All agreed to participate in the survey and were compensated for their time. The randomization procedure was presented as a lottery which is well understood in the Ethiopian context. Due to the higher number of applicants than jobs available, the lottery was considered a fair way to select workers from the pool of eligible

[8] In 2011, the textile and leather sectors accounted for over 60 percent of the total manufacturing exports (Gebreeyesus, 2013).

[9] The economic incentives include generous credit schemes, exemption from import taxes on investment capital goods and raw materials, and tax holidays on profit for five years.

[10] 31 December 2016: value 8.61. 964/8.61=112.

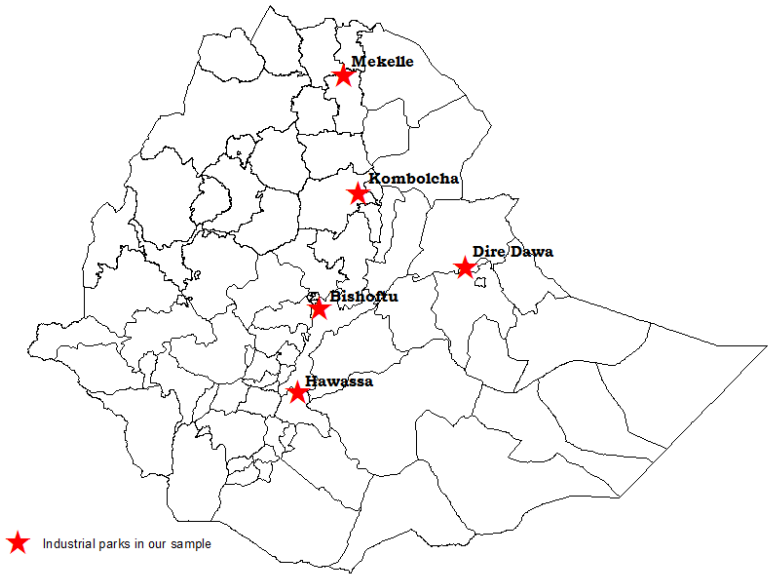


Figure 1: Industrial parks

applicants. Within each hiring block, half were assigned to the treatment group and received a job offer from the firm they had applied to, after the interview had taken place. The other half were assigned to the control group and did not receive any job offer.

Randomization of job offers among eligible applicants started in April 2016. The collection of baseline data continued over 22 months until we reached the pre-specified sample size. We conducted two follow-up surveys, the first at seven months and the second at fourteen months after baseline. An illustration of the project's timeline is presented in Appendix A.

3.3 The study participants

In Ethiopia, premarital pregnancies are generally not tolerated, thus we restricted the sampling of applicants to include only married women or women living with a partner as these women make up the relevant group for our research question.[11] We have further restricted the sample to only include women younger than 41 years old.[12] The sample consists of a panel of 1,140 women who had applied and were eligible for a job in a manufacturing firm.

[11] Couples who live together without being married usually have to commit to an informal 'cohabitation contract' which is culturally accepted and considered equivalent to a formal marriage.

[12] The age criterion was chosen to only capture women in their reproductive years. Although women over 40 can also be fertile, menopause is common among women over 40 in Ethiopia. For instance in the 2016 DHS, 11% of women between 35-39 years old report being in menopause, while 71% of women between 40-45 report being in menopause. In our sample, 2% of the women that were interviewed were older than 40, and these are dropped from the analysis. The results do not change whether we include these women or not (not showed).

Measures on income and employment status, consumption, childbearing status, stated ideal number of children, and household decision-making power were collected at baseline and in follow-up surveys at seven and fourteen months after baseline.

Table 1 presents summary statistics on the baseline sample.[13] The women in our study are on average 24.2 years old and about two thirds have finished grade 10 or higher (mean years of schooling 9.3). Most are Christian, either Ethiopian Orthodox or Protestant. Almost all are married and the rest are living with a partner. Two-thirds have ever given birth and about half of them have children living with them; 33% of the households have one child, 22% have two children, and 10% have three or more children under the age of 15. Most of the respondents live in nuclear households, only 11% of the households have other adult household members which in most cases are siblings or an older son or daughter of the husband (counted as an adult if the person is 16 years old or older). Half of the women report they want four children in their lifetime and the mean number of desired children is 3.9.

Only 31% of the respondents had ever had a formal or informal job before the study. Of these, 41% had worked in some unskilled trade, such as domestic workers or daily laborers. 24% had been self-employed and were selling clothes or food on the street or running their own business. Another 16% had previous experience with skilled or semi-skilled professions, such as hairdressing or seamstressing. Only 4% of all the applicants had previous experience with factory work. In the six months prior to baseline, half of the respondents did not have any income. Among the respondents who did have income, their contribution to the household income was on average 27%. At the time of the baseline, only 2% had an unemployed husband and the average monthly income of husbands was about ETB 2,500 (PPP $ 288).

The F-test for joint significance is significant at the 10% level. To account for this imbalance, we include the set of pre-specified individual control variables in all regressions in the main analysis.[14]

3.4 Attrition

We were not able to reach 232 respondents (16%) by the first follow-up and an additional 88 respondents (6%) by the second follow-up (endline). The attrition rate at endline is slightly higher in the treatment group than in the control group: 23% in the treatment group and 21% in the control group, but the attrition rates are not significantly different in the two groups (p-value=0.340).[15] Attrition is however associated with not having children, which might indicate that more mobile individuals move away.[16]

[13] See Appendix Table B1 for balance of the endline sample.

[14] In Appendix E, we perform sensitivity tests by running the main outcome regressions without the control variables and we find that the results are not sensitive to control variables.

[15] The Appendix Table B2 shows the regression analysis of treatment status on attrition controlling for block fixed effects and individual control variables.

[16] In the Appendix Table B3, we have performed a robustness check for attrition on the average effect of the main outcomes, namely the respondents' fertility, desired lifetime fertility, income, and household decision-making power using Lee bounds. Trimming the sample to account for attrition does not change the average

Table 1: Baseline summary means and tests of randomization balance

	Baseline sample			
	Full sample Mean/SE (1)	Control Mean/SE (2)	Treatment Mean/SE (3)	t-test difference (2)-(3)
Age	24.221	24.034	24.406	-0.371
	(0.125)	(0.167)	(0.186)	
Years of schooling	9.327	9.310	9.344	-0.034
	(0.080)	(0.109)	(0.116)	
Muslim	0.142	0.152	0.133	0.020
	(0.009)	(0.013)	(0.013)	
Ethiopian Orthodox	0.629	0.617	0.642	-0.025
	(0.013)	(0.018)	(0.018)	
Protestant	0.225	0.224	0.225	-0.002
	(0.011)	(0.015)	(0.015)	
Married	0.929	0.942	0.917	0.026*
	(0.007)	(0.009)	(0.010)	
Have ever given birth	0.666	0.669	0.663	0.006
	(0.012)	(0.017)	(0.017)	
Number of children	1.153	1.173	1.133	0.041
	(0.032)	(0.046)	(0.044)	
Pregnant at baseline	0.050	0.052	0.048	0.004
	(0.006)	(0.008)	(0.008)	
Number of children living in the household	1.125	1.139	1.112	0.027
	(0.029)	(0.042)	(0.041)	
Number of adults living in the household	2.155	2.163	2.146	0.017
	(0.013)	(0.018)	(0.019)	
Desired number of children	3.906	3.841	3.971	-0.131
	(0.046)	(0.058)	(0.072)	
Decision-making index (0-1)	0.517	0.500	0.534	-0.034*
	(0.009)	(0.013)	(0.012)	
Have ever had a formal job before	0.306	0.294	0.318	-0.024
	(0.012)	(0.017)	(0.017)	
Respondent's income the last six months	2221	2193	2250	-57
	(111)	(169)	(146)	
Husband's income the last six months	15101	15207	14995	212
	(333)	(497)	(445)	
Total recurrent expenses the last months	2521	2505	2537	-32
	(39)	(57)	(53)	
Total infrequent expenses the last six months	1976	2043	1909	134
	(65)	(103)	(78)	
Observations	1460	728	732	
F-test of joint significance (F-stat)				1.488*
F-test, number of observations				1460

The value displayed for t-tests are the differences in the means across the treatment and control groups. The covariate variable *block* is included in all estimation regressions. All missing values in balance variables are treated as zero. $^{***}p < 0.001$,$^{**}p < 0.05$,$^{*}p < 0.1$.

treatment effect conclusions.

3.5 Estimation strategy

The impact of the factory job is evaluated by comparing outcomes across the treatment and control groups in a simple regression framework with baseline control variables. For each individual-level outcome we estimate the following equation for the intention-to-treat (ITT) estimation:

$$Y_{i1} = \beta_0 + \beta_1 T_i + \beta_2 Y_{i0} + \gamma_k \boldsymbol{X}_{ki} + b_l + u_i, \quad (1)$$

Y_{i1} is the outcome variable measured at first follow-up or endline. The dummy-variable T_i indicates whether the respondent i is in the treatment group. In addition to controlling for the baseline value of the outcome variable Y_{i0}, the set of baseline control variables, X_i, includes the respondent's age, religion, education, the number of household members, the total household income the last six months, and a dummy for whether the respondent had wage employment the last six months prior to baseline. The randomization was conducted at each factory hiring round, thus we include the block fixed effects b_l.

We report two main fertility outcomes: actual fertility and preferred lifetime fertility. Actual fertility is constructed based on whether the respondent reports being currently pregnant and whether the household has a new baby at the time of the endline survey.[17] The second main outcome variable for fertility is preferred lifetime fertility. This variable is collected by asking the respondents "If you [could go back to the time you did not have any children and] could choose exactly the number of children to have in your whole life, how many would that be?" where the text in the brackets was included if the respondent already had children. We include this measure for fertility outcome because preferences on the number of children couples wish to have may indicate future fertility rates, for instance, if the preferred number of children decreases, fertility rates are likely to decrease accordingly but with a lag. Hence, an effect of the treatment on desired lifetime fertility may give an indication about the long-term impact on fertility.

In the analysis, we investigate multiple outcomes of being assigned to treatment. Using the free step-down resampling methodology of Westfall and Young (1993), we adjust the p-values in each table for the family-wise error rate (FWER). That is, we adjust the p-values by taking into account the probability of falsely rejecting at least one of the true null hypotheses within a "family" (all outcomes in each table) of hypotheses where the outcomes

[17] In the pre-plan, we specified the main fertility outcome as the number of children at baseline + 1 if answering yes to the question: "Are you pregnant now or have you been pregnant since we last interviewed you?" at either follow-up survey. The number of children at baseline is measured with the survey question: "How many children do you have?" Unexpectedly, the number of children the women report in baseline did not always correspond well with the number of children she reported in follow-up surveys: 2.5% reported fewer children and 1.4% reported two, three or more additional children in the endline survey than in the baseline survey. These may be acceptable measurement errors and can possibly be explained by changes in the respondents' households, for instance due to divorce or taking in younger family members which is common in our context. However, we chose instead to use the alternative measure for fertility, namely whether the respondent reports being currently pregnant or the household has a new baby at the time of either follow-up survey, as the main outcome variable for actual fertility throughout the paper.

are correlated.[18]

The intention-to-treat effect is clean but does not directly give us an estimate of the effects of getting the factory job. To estimate such effects for those that started working as a result of the treatment (the compliers), we run the following IV-model:

$$Z_i = \beta_0 + \beta_1 T_i + \beta_2 Y_{i0} + \gamma_k \boldsymbol{X}_{ki} + b_l + u_i, \tag{2}$$

$$Y_i = \beta_0 + \beta_1 \hat{Z}_i + \beta_2 Y_{i0} + \gamma_k \boldsymbol{X}_{ki} + b_l + v_i, \tag{3}$$

In the model, treatment status is an instrument used to estimate the binary outcome Z_i, which is whether the respondent had "any wage employment since baseline".[19] Under the assumptions of no externalities and monotonicity, the treatment status is a valid instrument as it is exogenous (randomization conducted by the researchers) and correlated with Z_i; being assigned to the treatment group increases the likelihood that the respondent had any wage employment since baseline. The no-externalities assumption requires that a respondent is not affected by the treatment received by another respondent (Stable Unit Treatment Value Assumption). While the monotonicity assumption requires that assignment to treatment changes the likelihood of being employed in a wage job in the same direction for all individuals. Both assumptions seem likely to hold in our setting. First, the respondents in the control group were free to apply and take any other job, both within manufacturing and other sectors. The factories typically operate in industrial clusters that frequently hire new employees. Thus, the employment of the treatment group individuals probably did not limit the opportunity for employment for the control group individuals. Second, being assigned to treatment means that the applicant receives a job offer. Logically, assignment to treatment will increase the likelihood of employment for all individuals in the treatment group. The estimate from the IV regression can then be interpreted as the local average treatment effect (LATE) of the group that went from being unemployed to having wage employment, because of being assigned to the treatment group.[20]

[18] We have calculated the adjusted p-values using the STATA user program **wyoung** (Jones et al., 2019).

[19] The pre-specified instrumented variable is an indicator for having had "any wage employment the last six months" as each survey was scheduled to be carried out with six months intervals. Due to various factors, such as political unrest and difficulties reaching all respondents, some interviews were delayed and thus the average time between each survey was instead 7.2 months. Therefore, we changed the instrumented variable to "any wage employment since baseline" which basically captures what we intended in the pre-plan. Sensitivity tests for the main outcome variables using the pre-specified instrumented variable are included in the tables E1-E4 in the appendix.

[20] More formally, the IV estimate is the ITT effect divided by the difference in the proportion of compliers in the treatment and control group: $\beta_{IV} = \frac{\bar{Y}_1 - \bar{Y}_0}{D_1 - D_0}$.

4 Empirical Results

4.1 Impacts on employment and income

Before exploring the effects of employment on fertility outcomes, we examine whether the randomization actually led to any durable difference in employment and earnings between treated and control women. These analyses were not pre-registered. Of the total baseline sample of 1,460 women, 732 (50%) were offered a job in the factory where they had applied. We do not have perfect compliance as only 63% accepted the job offer and 15% from the control group started working in the sample factories. While we asked the factories not to hire the women from the control group, we could not prevent them from hiring these women at a later point. By the first follow-up survey, 34% of the workers who had accepted the job offer had left the job and by endline, 39% had left the factory job. In Table 2 we present the estimated effects of treatment status on whether the applicants started working in the factory, employment status at the time of the follow-up surveys, total income the last six months, and respondents' income as a share of the household income.

In the first two columns of Panel A we see that women in the treatment group were 46 percentage points more likely to have started working in the factory, and 28 percentage points more likely to be employed in the factory at the time of the first follow-up. While, in the third and fourth columns, we see that any wage employment at the time of the follow-up survey was almost twice as high in the treatment group as in the control group and the average income in the treatment group was 42% higher than income in the control group. Moreover, in column (6) we see that the women's income the past six months contributed on average to 20 percent for the total household income in the control group, while in the treatment group the women's income contributed on average 25 percent for the total household income. Thus, despite imperfect compliance, there is strong employment and income effect from treatment status. However, these estimates also suggest that the factory jobs paid very low wages as their income is 42% higher than those in the control group, but it only translates into an increase of 5 percentage points additional contribution to their total household income. In other words, while the women in the treatment group earned much more compared to the control group, their income was modest relative to their household's income (which in most cases is the husband's income). Moreover, women's outside options for employment seem to be very limited, since the difference in employment status is high (column (3)).[21]

In order to capture the impact on the compliers, we report the LATE estimates (IV) in addition to the ITT estimates (OLS) for income and income shares (columns (5) and (7)). The first stage estimates, using treatment status as an instrument, strongly predicts whether the respondent had any formal wage employment since baseline, which is 37 percentage points higher in the treatment group than in the control group at first follow-up (31 percentage points higher at endline). The IV estimates imply that the effect on income and share of household income is, as expected, much larger for the compliers.

[21] The interested reader is referred to Halvorsen (2020) for a description of the workers' experiences of the job and decisions to leave the factory employment.

Table 2: Impact of the treatment on employment and income

PANEL A: First follow-up (7 months)

	Started working in the factory	Currently employed in the factory	Currently any wage employment	Total income last six months		Share of total household income the last six months	
	OLS	OLS	OLS	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treatment	0.457***	0.282***	0.209***	1,176***		0.050***	
	(0.024)	(0.023)	(0.026)	(220.416)		(0.016)	
Any formal job since baseline					3,156***		0.135***
					(562.655)		(0.040)
Observations	1,228	1,228	1,228	1,228	1,228	1,228	1,228
Control mean	0.149	0.105	0.224	2,798		0.200	
Unadjusted p-value	0.000	0.000	0.000	0.000	0.000	0.001	0.001
Adjusted p-value (West-Young)	0.000	0.000	0.000	0.000	0.000	0.001	0.001
First stage results: Any formal job since baseline							
Treatment					0.372***		0.372***
					(0.026)		(0.026)
p-value from F-test					0.000		0.000

PANEL B: Endline (14 months)

Treatment		0.236***	0.137***	807***		0.023	
		(0.023)	(0.028)	(258.709)		(0.018)	
Any formal job since baseline					2,588***		0.075
					(774.492)		(0.054)
Observations		1,140	1,140	1,140	1,140	1,140	1,140
Control mean		0.108	0.257	3666		0.238	
Unadjusted p-value		0.000	0.000	0.002	0.001	0.186	0.165
Adjusted p-value (West-Young)		0.000	0.000	0.004	0.002	0.182	0.169
First stage results: Any formal job since baseline							
Treatment					0.310***		0.310***
					(0.027)		(0.027)
p-value from F-test					0.000		0.000

All regressions include baseline covariates and block fixed effects. Baseline covariate controls include: age, religion, education level, total household income the last six months, number of household members, and a dummy for whether the respondent had any wage employment the last six months. Robust standard errors in parenthesis. $^{***}p < 0.001$,$^{**}p < 0.05$,$^{*}p < 0.1$.

The results from the endline are presented in Panel B in Table 2. Factory employment status between first follow-up and endline surveys did not change for most respondents

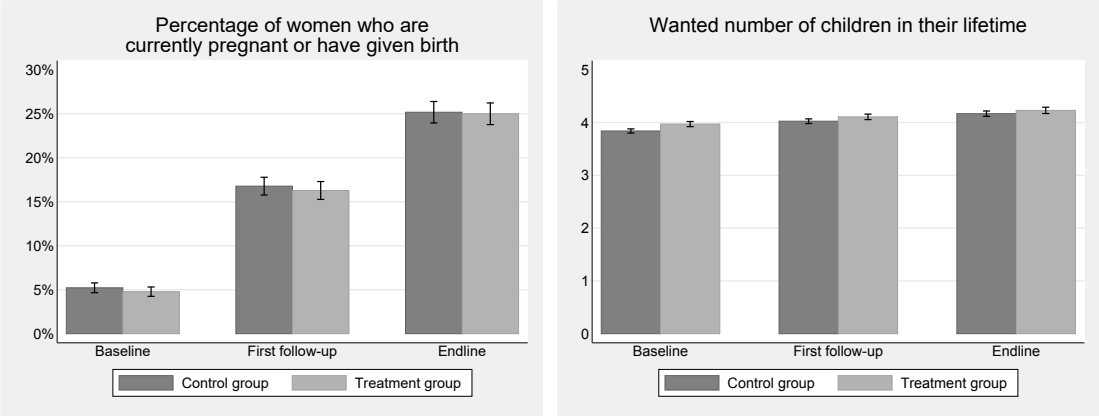
(compare panels A and B in column (2)). In fact, 80% of the factory workers in the first follow-up were still employed in the same factory at endline. However, the income difference between the treatment and control group is somewhat smaller (column (4)), as the employment rate declined in the treatment group and increased in the control group. There is a slight difference in the women's share of household income between the two groups, but no longer statistically significant (column (6)). Although the treatment effects are smaller at endline than at the first follow-up, most of the measures are still financially and statistically significant. In the remaining analyses of the paper, we focus on the outcomes of the endline survey.

It is appropriate to compare our first stage results on income and employment with those of Blattman and Dercon (2018). Contrary to us, they do not find any income difference between the ones who were offered a job in an industrial firm and a control group at about one year after baseline. There are two potential reasons for this. First, we measure income as total income (including both earned and unearned income) the past six months, while Blattman and Dercon measure income as the average weekly earning the past two weeks. As we do find a declining difference in income between first follow-up and the endline survey, a shorter income window may have yielded different results. However, we do not have data to do this comparison. If instead, we compare the shares of women who are currently unemployed, Blattman and Dercon report that 34% of the control group and 33% of the treatment group had been unemployed the past two weeks prior to endline (Blattman and Dercon, 2018, p. 20). In our sample, we find that 74% of the women in the control group were currently unemployed at endline and 58% of the women in the treatment group were currently unemployed, hence significantly higher shares than what Blattman and Dercon find in their study. This leads us to the second reason why our results may differ from each other: the sample. Blattman and Deron's sample includes mostly young, unmarried women, living close to the capital, Addis Ababa. While being on average only two years older, the women in our sample are all married or living with a partner. Most already have children and are thus less mobile. They also live in areas more distant from the capital and potentially have fewer alternative labor opportunities, which might explain the large difference in unemployment between the two samples. Thus, while many of the workers in Blattman and Dercon's study found alternative jobs, making the difference in the groups smaller, the workers in our study struggled with finding alternative income opportunities or preferred to stay at home taking care of their household and children instead.

4.2 Fertility outcomes

At endline, 8% of the respondents reported being currently pregnant and 17% had given birth in the period between baseline and endline. In Figure 2, we show the main outcome variables for fertility at each survey round by treatment. Panel (a) shows the percentage of women who reported being currently pregnant or had given birth at any time between baseline and endline. We see that 5% were already pregnant when applying for the factory job. Actual fertility in the control and treatment groups follow each other very closely over

the study period, thus indicating that there is no treatment effect. Both groups increase the percentage of women who are pregnant or have given birth by about 10% for each survey round. In panel (b), we show the average number of children the women report wanting to have in their lifetime. While the average is about four children, the number increases slightly over time, but with no difference across treatment and control groups.



(a) **(b)**

Figure 2: Actual fertility and desired fertility by survey rounds

More formally, Table 3 reports the OLS and IV regressions for the main fertility outcomes. Evidently, there are no differences in fertility between treatment and control women when controlling for covariates for total fertility (are currently pregnant or have have given birth since baseline) nor for the pre-specified fertility measure (columns (1)-(4)). Conducting an equivalence test for the ITT results with two one-sided t-tests (TOST), we can reject that the effects are more negative than -0.037 and more positive than 0.043. These effects correspond to around 15 percent of the mean in the control group so the study is powered to reject moderately sized effects. The IV estimates capture the impact on the compliers and confirm the ITT results, but with larger point estimates.

In columns (5)-(6), we report the estimated treatment effect on desired lifetime fertility. As Figure 2 suggested, the treatment effect is very close to zero and not significant in either estimation method. For the ITT estimation we can here reject that the effects are more negative than -0.057 and more positive than 0.024. As the mean number of desired fertility in the control group is over four children, we conclude that our study is powered to reject small effects for this outcome.

4.2.1 Heterogeneity analysis

The ITT estimates on actual and desired fertility show no effects of factory employment after fourteen months. It might be that some subgroups of women are more affected by the treatment than others. In the appendix, Table C1 presents heterogeneity analyses on women less than 25 years old (57% of the sample), who have less than 10 years of schooling (26% of the sample), who were Orthodox Christian (61% of the sample), who had no children living

Table 3: Impact of the job offer on fertility

	Fertility measure as in pre-plan[1]		Currently pregnant or have given birth after baseline		Desired lifetime fertility	
	OLS (1)	IV (2)	OLS (3)	IV (4)	OLS (5)	IV (6)
Treatment	-0.054		0.003		0.089	
	(0.039)		(0.024)		(0.089)	
Any formal wage employment since baseline		-0.175		0.009		0.282
		(0.122)		(0.076)		(0.277)
Observations	1,140	1,140	1,140	1,140	1,138	1,138
Control mean	1.365		0.252		4.168	
Unadjusted p-value	0.164	0.150	0.904	0.901	0.317	0.309
Adjusted p-value (West-Young)	0.489	0.461	0.907	0.898	0.641	0.636
First stage results: Any formal job since baseline						
Treatment		0.310***		0.310***		0.314***
		(0.027)		(0.027)		(0.027)
p-value from F-test		0.000		0.000		0.000

[1] The number of children at baseline + 1 if answering yes to the question: "Are you pregnant now or have you been pregnant since we last interviewed you?", at either follow-up survey.
All regressions include baseline covariates and block fixed effects. Baseline covariate controls include: whether the respondent was pregnant at baseline (in columns (1)-(4)), age, religion, education level, total household income the last six months, number of household members, and a dummy for whether the respondent had wage employment the last six months. Robust standard errors in parenthesis. $^{***}p < 0.001$,$^{**}p < 0.05$,$^{*}p < 0.1$.

in the household at baseline (31% of the sample), whose household had earned less than the median income of all households in our sample in the six month prior to baseline (50% of the sample), and who had wage employment in the six months prior to baseline (19% of the sample) separately. However we see no treatment effects even when dividing the sample into these subgroups.

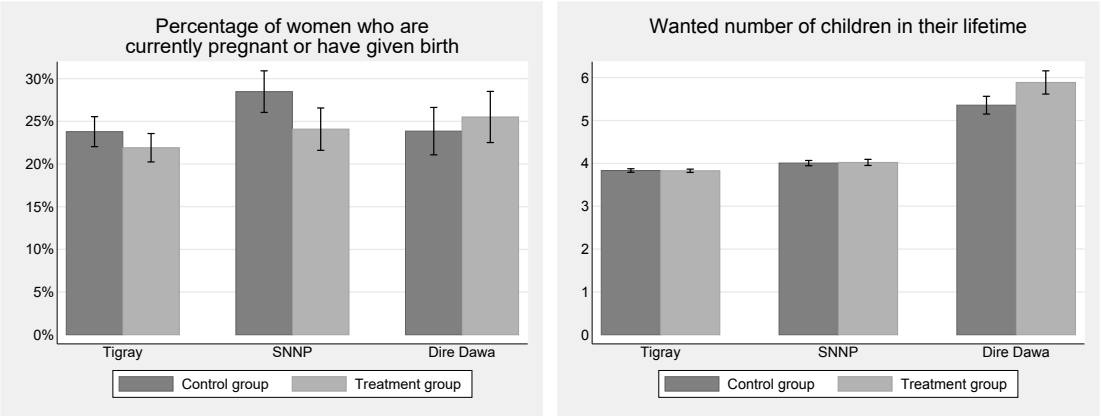
The impact of employment on fertility may also vary by region. In Figure 3, we show the outcomes of actual and desired fertility in the control and treatment groups in the three largest regions.[22] The Figure shows that while outcomes differ some by region, there is no significant difference between the treatment and control groups in any region.

4.3 Other outcomes

4.3.1 Treatment effects on child expenditures

The cost of children is likely to influence the demand for children. For instance, in the quantity-quality trade-off model, higher investments in children (quality) is associated with

[22] We do not include the regions Oromia and Amhara in the figure due to few observations (n=20 and n=65). We do however, show the full regression results by region in Appendix Tables C2.



(a) **(b)**

Figure 3: Actual and desired fertility by regions

lower number of children (quantity). In this section, we investigate whether the assignment to treatment affects child expenditures on clothes, schooling, and health in the six months before the endline survey.

We find no difference in expenditures on children by treatment as shown in Table 4. In columns (5)-(6), we do see that the households in the treatment group spent more per child on health expenditures than households in the control group. This seems to be be driven by worse health among the children in the treatment group. By endline, the treatment group children had 86% (3.5 percentage points, p-value<0.028) more instances of illness or injuries. The effect is nonetheless small in absolute terms and there are no differences in the other expenditure categories, nor in the aggregated measure.

Table 4: OLS regressions on the impact of the job offer on child expenditures

	Children's total expenditure		Children's clothes and shoes expenditure		Children's health expenditure		Children's school expenditure	
	OLS (1)	IV (2)	OLS (3)	IV (4)	OLS (5)	IV (6)	OLS (7)	IV (8)
Treatment	6.470		-21.606		34.006*		1.005	
	(90.158)		(46.151)		(17.456)		(65.990)	
Any formal job since baseline		20.199		-67.605		106.178**		3.141
		(272.114)		(139.679)		(53.664)		(199.362)
Observations	845	845	845	845	845	845	845	845
Control mean	1082		524.4		38.35		519.8	
Unadjusted p-value	0.943	0.941	0.640	0.628	0.052	0.048	0.988	0.987
Adjusted p-value (West-Young)	0.996	0.996	0.887	0.887	0.175	0.173	0.996	0.996

All regressions include baseline covariates and block fixed effects. Baseline covariate controls include: age, religion, education level, total household income the last six months, number of household members, and a dummy for whether the respondent had wage employment the last six months. Robust standard errors in parenthesis. $^{***}p < 0.001, ^{**}p < 0.05, ^{*}p < 0.1$.

Table 5: The impact of the job offer on empowerment proxies

	Decision-making index (higher number higher dm power)		Gender attitude index (1-4; higher index is more favorable gender equal attitudes)		Women who work for a salary outside the home are more respected in the local community (1/0)		In my village, it is generally preferred that married women should work in hh (1/0)		Have you and your husband discussed the use and methods of contracep-tives since we last interviewed you (1/0)		Are you permitted by your husband to visit a health clinic alone (1/0)	
	OLS (1)	IV (2)	OLS (3)	IV (4)	OLS (5)	IV (6)	OLS (7)	IV (8)	OLS (9)	IV (10)	OLS (11)	IV (12)
Treatment	-0.020		0.029		0.032*		0.002		0.011		-0.017	
	(0.018)		(0.022)		(0.019)		(0.014)		(0.030)		(0.022)	
Any formal job since baseline		-0.064		0.094		0.105*		0.007		0.036		-0.055
		(0.056)		(0.069)		(0.063)		(0.046)		(0.093)		(0.069)
Observations	1,140	1,140	1,140	1,140	1,086	1,086	1,075	1,075	1,087	1,087	1,140	1,140
Control mean	0.643		0.616		0.858		0.0567		0.498		0.212	
Unadjusted p-value	0.263	0.255	0.184	0.174	0.098	0.093	0.881	0.878	0.708	0.701	0.440	0.429
Adjusted p-value (West-Young)	0.670	0.670	0.631	0.625	0.446	0.455	0.916	0.915	0.916	0.915	0.825	0.821

All regressions include baseline covariates and block fixed effects. Baseline covariate controls include: age, religion, education level, total household income the last six months, number of household members, and a dummy for whether the respondent had wage employment the last six months. Robust standard errors in parenthesis. $^{***}p < 0.001$,$^{**}p < 0.05$,$^{*}p < 0.1$.

4.3.2 Treatment effects on household decision-making power and mindset

Empowerment refers to the ability to make choices (Kabeer, 1999). In our analysis, we use the involvement in household decisions as a proxy for empowerment. More specifically, from a list of 15 different household decisions, we coded each decision as 1 if the respondent is the main decision-maker or has "a lot of input" to the decision. The 15 decisions are summarized in one index ranging from 0 to 1.

In Table 5 we report the treatment effect of the factory employment on the household decision-making index (columns (1)-(2)).[23] Alternative proxies for empowerment are also included as changes in attitudes to gender equality, feeling respected by the society, or communication with the main decision-maker may empower women and simultaneously impact fertility decisions. We find, however, no effect of the factory job on any of these outcomes, except for a small effect on feeling respected in the local community (columns (5)-(6)). As the treatment did not affect household decision-making power much, nor various attitudes and beliefs, it is unlikely that fertility outcomes are affected by a factory job offer through the empowerment channel.

Our finding of no effect of employment on household decision-making power or any other empowerment measure contrasts with much of the existing literature. While there are very few papers that identify the causal effect of employment on empowerment, previous studies find increased household decision-making power as a result of increased outside employment options (Majlesi, 2016) and increased visible contribution to the household income (Atkin,

[23] In the appendix, Table D1 shows the treatment estimates for each decision separately.

2009; Anderson and Eswaran, 2009). One reason for our finding may be that the index score of decision-making power for the women in our sample is quite high, averaging at 0.65. That is, on average, the women report being the main decision-maker or having a lot of input in 9.75 of the 15 household decisions listed in our survey. In fact, 26% report having high decision-making power in all 15 decisions and 45% have high decision-making power in at least 12 out of 15 decisions. However, the results do not change if we only include women with low baseline decision-making power (not displayed in the table). Another reason may be the short employment spells in our sample. For instance, 34% of those who were offered the job did not even start and 49% of those who did start left within three months. Such short employment spells are unlikely to change household dynamics significantly.

4.4 Channels

In this section, we explore the labor supply and empowerment channels, which potentially affect the decision-making process for fertility choices in the households of our sample. These analyses were not pre-registered.[24]

4.4.1 The labor supply channel

In Section 4.1, we showed that factory employment had a causal positive impact on the women's labor supply and income. The theoretical prediction of the relationship between income and the number of children, however, is not clear. One could argue that when income rises parents will trade off quantity for quality and thus reduce their demand for a high number of children. However, when there is a high desired fertility rate, we may expect a positive relationship between income and fertility, especially among poor households like in our sample, as income may be a constraint for households to reach their desired number of children.

We begin investigating the labor supply channel by exploring the correlations between household income and the number of children, the lifetime number of desired children and the total expenditures per child on clothes, schooling, and health in the past six months, and the women's decision-making power measured at baseline. Table 6 reports the estimates. As the identity of the earner may matter for the relationships, in the even-numbered columns we report correlations with the respondents' and other household members' income (in most cases this only includes the husband) separately. First, we see that household income is positively correlated with having more children (significant at 10%-level) when controlling for the couple's age and education (Column (1)) and the wife's education is negatively correlated with the number of children. In the second column, we see that in addition to the wife's education, her income is negatively correlated with the number of children she has.

[24] In the pre-analysis plan we wrote that if there is a treatment effect on either of the channels, we would include them (separately) in the main specification with the outcome variables fertility and preferred lifetime fertility. This would, however, lead to a mediation bias (post treatment bias) so we chose to not do this even though income is affected.

Table 6: Correlations of income at baseline

	Number of children		Desired lifetime fertility		Expenditures per child (log)		Decision-making power index	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Household's total income last 12 months (log)	0.052*		-0.020		0.615***		-0.022**	
	(0.029)		(0.052)		(0.192)		(0.011)	
Respondent's total income last 12 months (log)		-0.026***		-0.005		0.016		-0.002
		(0.006)		(0.010)		(0.016)		(0.002)
Other hh-members' total income last 12 months (log)		0.036**		-0.003		0.320***		-0.014**
		(0.017)		(0.034)		(0.070)		(0.006)
Respondent's age	0.133***	0.135***	0.052***	0.052***	0.008	0.009	0.008***	0.008***
	(0.010)	(0.010)	(0.019)	(0.019)	(0.018)	(0.018)	(0.003)	(0.003)
Respondent's years of schooling	-0.060***	-0.065***	-0.159***	-0.159***	0.104***	0.103***	0.004	0.004
	(0.012)	(0.012)	(0.026)	(0.026)	(0.024)	(0.024)	(0.003)	(0.003)
Husband's age	0.023***	0.023***	-0.004	-0.004	0.016	0.020	0.002	0.002
	(0.006)	(0.006)	(0.011)	(0.011)	(0.014)	(0.015)	(0.002)	(0.002)
Husband's years of schooling	-0.011	-0.010	-0.035**	-0.036**	0.034*	0.048***	0.002	0.002
	(0.008)	(0.008)	(0.014)	(0.014)	(0.019)	(0.018)	(0.003)	(0.003)
Number of children living in the household					-0.155**	-0.155**		
					(0.072)	(0.075)		
Observations	1,460	1,460	1,459	1,459	950	950	1,460	1,460

OLS regressions. Robust standard errors in parenthesis. $^{***}p < 0.001$,$^{**}p < 0.05$,$^{*}p < 0.1$.

A wife's employment in the industry will both increase the household income, which is correlated with a higher number of children, and increase the wife's income, which is correlated with fewer children. One possible interpretation of these correlations is that the positive income effect of higher household income on fertility is counterbalanced by the negative substitution effect of the woman's labor supply (measured by income).

In the third and fourth columns, we observe that the number of desired children is negatively correlated with household income, but not significantly so. The number of desired children is negatively correlated with both the wife's and husband's education. Hence, we do not believe factory employment will directly affect fertility in the long term. However, if manufacturing growth leads to increases in women's (and men's) educational attainment, which is a reasonable assumption based on findings in the literature (e.g., Heath and Mobarak, 2015), the industry may indirectly affect fertility rates in the long-run, as the correlation between education and desired children is negative. The negative relationship between schooling and fertility is also supported by other empirical studies, for instance Baird et al. (2010), Duflo et al. (2015), and Osili and Long (2008).

When the household income increases, the demand for higher quality in children may increase. In columns (5)-(6), we use expenditures per child as a measure for quality or investments in the children. The measure is the log of total expenditures spent the last six months before baseline on children's clothing and footwear, health, and schooling divided by the number of children living in the household. We see from the table that higher-income households do spend more on each child, but this is primarily driven by the husband's income and not the wife's income.

In the last two columns, we see that the women's household decision-making power is negatively correlated with husbands' income, and positively correlated with the respondents' age. Quite surprisingly neither the respondents' income nor education are correlated with their decision-making power. However, this corresponds well with the null result of the treatment effect presented in Section 4.3.2.

4.4.2 The empowerment channel

While we did not find any treatment effect of the factory job offer on empowerment, nor a correlation between income and decision-making power at baseline, it may still be interesting to explore the baseline correlation between decision-making power and fertility. Empowerment is likely to affect fertility if there is a discrepancy in the preferences in the number of and spacing between children within couples. In our sample, at baseline, 22% of the women reported that their husbands wanted more children than they wanted, 8% said their husbands wanted fewer children, 59% wanted the same number, and 11% said they did not know.

In Table 7, we show that, for the full sample (Columns (1)-(3)), the wife's decision-making power is not correlated with the number of children the couple has, the number of children she wishes to have, nor the total expenditures per child, when controlling for age and education of the wife and the husband. However, if we only include couples where there is a discrepancy between the number of desired children (Columns (4)-(6)), there is a significant negative correlation between the wife's decision-making power and the number of children, if the wife prefers fewer children than her husband (coefficient (a) in Column (4)). If on the other hand, the wife reports wanting more children than her husband, the relationship is also negative, but not statistically significant (coefficient (a+b) in Column (4)). The negative relationship between the woman's decision-making power and the number of children the couple has in both discrepancy cases, may indicate that when spouses decide together (wife has high decision-making power), rather than the husband deciding alone, the couple will have the number of children that corresponds to the lowest number of children preferred by either spouse. In Column (5) we see a weak positive correlation between the woman's decision-making power and desired lifetime fertility.

Table 7: Correlations of decision-making power index at baseline

	Full sample			Only couples with discrepancy in preference for lifetime fertility		
	Number of children	Desired lifetime fertility	Expenditures per child	Number of children	Desired lifetime fertility	Expenditures per child
	(1)	(2)	(3)	(4)	(5)	(6)
Decision-making index (0-1) (a)	-0.107	0.153	0.120	-0.341**	0.490*	-0.081
	(0.068)	(0.133)	(0.178)	(0.146)	(0.268)	(0.391)
Wife wants more children than husband				0.124	1.456***	0.580
				(0.226)	(0.395)	(0.384)
Interaction (b)				0.173	-0.038	-0.033
				(0.338)	(0.696)	(0.616)
Age	0.134***	0.050***	-0.002	0.136***	0.041	0.006
	(0.010)	(0.019)	(0.019)	(0.017)	(0.030)	(0.030)
Husband's age	0.024***	-0.005	0.022	0.023**	-0.017	0.006
	(0.006)	(0.011)	(0.015)	(0.011)	(0.018)	(0.022)
Years of schooling	-0.059***	-0.159***	0.118***	-0.067***	-0.220***	0.118***
	(0.012)	(0.026)	(0.024)	(0.018)	(0.039)	(0.036)
Husband's years of schooling	-0.009	-0.036**	0.053***	-0.003	0.011	0.030
	(0.008)	(0.014)	(0.018)	(0.014)	(0.026)	(0.031)
(a+b)				-0.167	0.452	-0.114
				(0.314)	(0.640)	(0.463)
Observations	1,460	1,459	950	439	439	307

OLS regressions. The decision-making index is constructed by 15 different household decisions in which each decision is coded as 1 if the respondent is the main decision-maker or has "a lot of input" to the decision. The 15 decisions are then averaged in one index ranging from 0 to 1. The interaction variable (b) is the decision-making index multiplied by an indicator of whether the respondent prefers more children in her lifetime than her husband. In Columns (4)-(5) the coefficients for the variable Decision-making index (0-1) (a) should be interpreted as the relationship between the dependent variable and wife's decision-making power in the cases where the wife prefers fewer children than her husband. The opposite case, when the wife prefers more children than her husband is shown by the sum of the coefficients on the decision-making power index and the interaction variable (a+b). Robust standard errors in parenthesis. $^{***}p < 0.001,^{**}p < 0.05,^{*}p < 0.1$.

5 Sensitivity analyses

In Appendix Tables E1-E4, we report sensitivity analyses of the main outcomes; respondents' income the last six months, fertility, desired lifetime fertility, and household decision-making power, by running both OLS and IV specifications without control variables, changing the sample size to different criteria, removing the region with the weakest first stage, and using the pre-specified IV model. Each table has two panels: panel A presents the results at the first follow-up and panel B presents the endline results. The first two columns in each table reproduce the estimates on the main outcome variables from the tables throughout the paper for comparison. In columns (3)-(4) we run the regressions without individual control variables. In columns (5)-(6) we use the sample of individuals available by the endline (n=1,140) to estimate the treatment effect at first follow-up. Furthermore, in one of the areas, some firms had not followed the randomization procedure as instructed. This was discovered by the research team when the first set of data became available and the field supervisor and firms where contacted. The 373 observations that were not randomized have been dropped from the sample of this paper but included in the sensitivity tests in columns (7)-(8). The first stage of the IV model, the effect of being assigned to the treatment group on whether the respondent had any wage employment in the project period, is weakest in Oromia and Tigray regions. While Oromia had a very small sample (n=20), in Tigray the treatment was only associated with a 14 percentage points more wage employment compared to the control group (see Appendix Table C2). In columns (9)-(10) in the sensitivity analyses tables, we report the treatment effects on the main outcomes after dropping the two regions, Oromia and Tigray. Last, in column (11) we use the pre-specified IV model which uses the treatment status as an instrument for having had any formal wage employment the last six months, instead of 'since baseline'.

In general, we see little change in the estimates. While in some cases the estimates are larger, in other cases they are smaller, and the significance stays the same in all sensitivity checks. One exception is the analyses when we drop the two regions with the weakest first stage (columns (9)-(10)). In these regressions, the ITT and LATE coefficients are much larger than when we use the full sample, which might suggest that there could be impacts on factory employment in these regions, however the estimates are not significant, possibly due to the lower sample size. We do however find a negative significant impact of the factory employment on decision-making power at endline. This might suggest that there is a negative effect on decision-making power as a result of the employment opportunity, which we could not capture that in the larger dataset. While this result is the opposite of what we would expect, the finding might reflect that when the respondent spends less time at home, she is less involved in household-decisions due to time constraints or fatigue. This is just a speculative explanation which needs further investigation to get a better understanding of the mechanisms at work.

6 Conclusion

In this paper we use a randomized field experiment to study the impact of working in the manufacturing industry on fertility. This is the first paper that measures the causal effects of a job on fertility outcomes at the individual level. In the experiment, entry-level applicants in the manufacturing industry in Ethiopia were randomized to either receive a job offer or to be in the control group. The main outcome variables are childbearing and desired lifetime fertility fourteen months after baseline.

We find strong effects on income, with the treatment group earning 42% more than the control group after seven months and 22% more after fourteen months. Despite this, we are able to reject relatively small effects on fertility and especially on fertility preferences for lifetime number of children.

Based on previous empirical and theoretical work, jobs are expected to affect fertility decisions through three channels; the substitution channel, the income channel, and the empowerment channel. Our data allows us to explore the channels. We find that there is a positive effect of being offered the job on labor supply. Furthermore, we find a negative correlation between women's income and the number of children they have. This suggest that we could expect a negative effect of the treatment on fertility (substitution effect). However, the household income is positively correlated with more children and the desired number of children in their lifetime is generally very high (four children on average). Since few of the respondents have reached their desired number of children, the income effect may drive fertility in the opposite direction. The two effects may thus cancel each other out. When investigating proxies for women's empowerment, we find no support for a relationship between factory jobs or income on decision-making power or other factors to empowerment. Nor is there any relationship between women's decision-making power and the number of children they have, their fertility preferences, or expenditures spent per child (after controlling for age and education).

Our results clearly show that individual level employment does not necessarily affect fertility. In order to draw broader conclusions about the effects of industrial policy, a few caveats about external validity are necessary. First of all, we study the effects for already partnered women. It is possible that employment has different effects on initially single women. Secondly, it is also possible that employment affects fertility at a higher level of aggregation (via substitution) and over a longer period. We know from previous field experiments that aggregate level employment in India affected education and fertility of young women (Jensen, 2012) and that education reduces fertility of young women (Duflo et al., 2015). Thirdly, it may be that the effects of employment on fertility are different in Ethiopia than in other places. After all, the fertility transition in sub-Saharan Africa displays a lower correlation with economic development than in other regions of the world (Bongaarts, 2017; Shapiro and Hinde, 2017). We hope that future studies will use individual level randomization in other settings so that we will learn to what extent the effects we find are generalizable.

References

Ager, P. and Herz, B. (2019), 'Structural change and the fertility transition', *Review of Economics and Statistics* **0**, 1–45.

Aguero, J. M. and Marks, M. S. (2008), 'Motherhood and female labor force participation: Evidence from infertility shocks', *American Economic Review* **98**(2), 500–504.

Anderson, S. and Eswaran, M. (2009), 'What determines female autonomy? Evidence from Bangladesh', *Journal of Development Economics* **90**(2), 179–191.

Angrist, J. D. and Evans, W. N. (1998), 'Children and their parents' labor supply: Evidence from exogenous variation in family size', *American Economic Review* **88**(3), 450–477.

Angrist, J. D. and Evans, W. N. (2000), Schooling and labor market consequences of the 1970 state abortion reforms, *in* 'Research in Labor Economics (Research in Labor Economics, Vol. 18)', Emerald Group Publishing Limited, pp. 75–113.

Angrist, J. D., Lavy, V. and Schlosser, A. (2010), 'Multiple experiments for the causal link between the quantity and quality of children', *Journal of Labor Economics* **28**(4), 773–824.

Ashraf, N., Field, E. and Lee, J. (2014), 'Household bargaining and excess fertility: An experimental study in Zambia', *American Economic Review* **104**(7), 2210–37.

Atkin, D. (2009), 'Working for the future: Female factory work and child health in Mexico', Unpublished Manuscript, Yale University.

Bailey, M. J. (2006), 'More power to the pill: The impact of contraceptive freedom on women's life cycle labor supply', *Quarterly Journal of Economics* **121**(1), 289–320.

Baird, S., Chirwa, E., McIntosh, C. and Özler, B. (2010), 'The short-term impacts of a schooling conditional cash transfer program on the sexual behavior of young women', *Health Economics* **19**(S1), 55–68.

Becker, G. S. (1960), An economic analysis of fertility, *in* 'Demographic and Economic Change in Developed Countries', Columbia University Press, pp. 209–240.

Becker, G. S. (1965), 'A theory of the allocation of time', *The Economic Journal* **75**(299), 493–517.

Becker, G. S. and Lewis, H. G. (1973), 'On the interaction between the quantity and quality of children', *Journal of Political Economy* **81**(2, Part 2), 279–288.

Becker, G. S. and Tomes, N. (1976), 'Child endowments and the quantity and quality of children', *Journal of Political Economy* **84**(4, Part 2), 143–162.

Beegle, K. and Christiaensen, L. (2019), *Accelerating Poverty Reduction in Africa*, Washington DC: World Bank.

Black, S. E., Devereux, P. J. and Salvanes, K. G. (2005), 'The more the merrier? The effect of family size and birth order on children's education', *Quarterly Journal of Economics* **120**(2), 669–700.

Blattman, C. and Dercon, S. (2018), 'The impacts of industrial and entrepreneurial work on income and health: Experimental evidence from Ethiopia', *American Economic Journal: Applied Economics* **10**(3), 1–38.

Blattman, C., Dercon, S. and Franklin, S. (2019), 'Impacts of industrial and entrepreneurial jobs on youth: 5-year experimental evidence on factory job offers and cash grants in Ethiopia', NBER Working Paper 25788. National Bureau of Economic Research, Cambridge, MA.

Bloom, D. E., Canning, D., Fink, G. and Finlay, J. E. (2009), 'Fertility, female labor force participation, and the demographic dividend', *Journal of Economic Growth* **14**(2), 79–101.

Blundell, R., Chiappori, P. A. and Meghir, C. (2005), 'Collective labor supply with children', *Journal of Political Economy* **113**(6), 1277–1306.

Bongaarts, J. (2017), 'Africa's unique fertility transition', *Population and Development Review* **43**(S1), 39–58.

Bongaarts, J. and Watkins, S. C. (1996), 'Social interactions and contemporary fertility transitions', *Population and Development Review* **22**(4), 639–683.

Bronars, S. G. and Grogger, J. (1994), 'The economic consequences of unwed motherhood: Using twin births as a natural experiment', *American Economic Review* **84**(5), 1141–1156.

Browning, M. (1992), 'Children and household economic behavior', *Journal of Economic Literature* **30**(3), 1434–1475.

Browning, M. and Chiappori, P.-A. (1998), 'Efficient intra-household allocations: A general characterization and empirical tests', *Econometrica* **66**(6), 1241–1278.

Browning, M., Chiappori, P. A. and Weiss, Y. (2014), *Economics of the Family*, Cambridge University Press, New York.

Casterline, J. B. (2010), 'Determinants and consequences of high fertility: A synopsis of the evidence - portfolio review', Washington DC: World Bank.

Central Statistical Agency/CSA/Ethiopia and ICF (2016), 'Ethiopia Demographic and Health Survey 2016 (DHS 2016)', Addis Ababa, Ethiopia, and Rockville, Maryland, USA: CSA and ICF.

Chiappori, P. A. (1988), 'Rational household labor supply', *Econometrica* **56**(1), 63–90.

Chiappori, P. A. (1992), 'Collective labor supply and welfare', *Journal of Political Economy* **100**(3), 437–467.

Cruces, G. and Galiani, S. (2007), 'Fertility and female labor supply in Latin America: New causal evidence', *Labour Economics* **14**(3), 565–573.

Dasgupta, P. (1995), 'The population problem: Theory and evidence', *Journal of Economic Literature* **33**(4), 1879–1902.

De Jong, E., Smits, J. and Longwe, A. (2017), 'Estimating the causal effect of fertility on women's employment in Africa using twins', *World Development* **90**, 360–368.

Dharmalingam, A. and Morgan, S. P. (1996), 'Women's work, autonomy, and birth control: Evidence from two South Indian villages', *Population Studies* **50**(2), 187–201.

Dinh, H. T., Palmade, V., Chandra, V. and Cossar, F. (2012), *Light Manufacturing in Africa: T]argeted Policies to Enhance Private Investment and Create Jobs*, Washington, DC: World Bank.

Doepke, M. and Tertilt, M. (2018), 'Women's empowerment, the gender gap in desired fertility, and fertility outcomes in developing countries', *AEA Papers and Proceedings* **108**, 358–62.

Duflo, E. (2003), 'Grandmothers and granddaughters: Old-age pensions and intrahousehold allocation in South Africa', *World Bank Economic Review* **17**(1), 1–25.

Duflo, E., Dupas, P. and Kremer, M. (2015), 'Education, HIV, and early fertility: Experimental evidence from Kenya', *American Economic Review* **105**(9), 2757–97.

Duflo, E. and Udry, C. (2004), 'Intrahousehold resource allocation in Cote d'Ivoire: Social norms, separate accounts and consumption choices', NBER Working Paper 10498. National Bureau of Economic Research, Cambridge, MA.

Elson, D. and Pearson, R. (1981), "Nimble fingers make cheap workers': An analysis of women's employment in third world export manufacturing', *Feminist Review* **7**(1), 87–107.

Franck, R. and Galor, O. (2015), 'Industrialization and the fertility decline', Working paper. Brown University.

Gebreeyesus, M. (2013), 'Industrial policy and development in Ethiopia: Evolution and present experimentation', WIDER Working Paper 2013/125. Helsinki: UNU-WIDER.

Getahun, T. D. and Villanger, E. (2018), 'Labour-intensive jobs for women and development: Intra-household welfare effects and its transmission channels', *The Journal of Development Studies* **54**(7), 1232–1252.

Halvorsen, S. (2020), 'Understanding the challenges of high labor turnover in the Ethiopian manufacturing industry', Unpublished Manuscript. Norwegian School of Economics.

Heath, R. (2014), 'Women's access to labor market opportunities, control of household resources, and domestic violence: Evidence from Bangladesh', *World Development* **57**, 32–46.

Heath, R. and Mobarak, A. M. (2015), 'Manufacturing growth and the lives of Bangladeshi women', *Journal of Development Economics* **115**, 1–15.

Jensen, R. (2012), 'Do labor market opportunities affect young women's work and family decisions? Experimental evidence from India', *The Quarterly Journal of Economics* **127**(2), 753–792.

Jones, D., Molitor, D. and Reif, J. (2019), 'What do workplace wellness programs do? evidence from the illinois workplace wellness study', *Quarterly Journal of Economics* **134**(4), 1747–1791.

Kabeer, N. (1999), 'Resources, agency, achievements: Reflections on the measurement of women's empowerment', *Development and Change* **30**(3), 435–464.

Lee, D. S. (2009), 'Training, wages, and sample selection: Estimating sharp bounds on treatment effects', *The Review of Economic Studies* **76**(3), 1071–1102.

Lehrer, E. and Nerlove, M. (1986), 'Female labor force behavior and fertility in the United States', *Annual Review of Sociology* **12**(1), 181–204.

Luke, N. and Munshi, K. (2011), 'Women as agents of change: Female income and mobility in India', *Journal of Development Economics* **94**(1), 1–17.

Lundberg, S. J., Pollak, R. A. and Wales, T. J. (1997), 'Do husbands and wives pool their resources? Evidence from the United Kingdom child benefit.', *Journal of Human Resources* **32**(3).

Lundberg, S. and Pollak, R. A. (1993), 'Separate spheres bargaining and the marriage market', *Journal of Political Economy* **101**(6), 988–1010.

Madsen, J. B., Moslehi, S. and Wang, C. (2018), 'What has driven the great fertility decline in developing countries since 1960?', *The Journal of Development Studies* **54**(4), 738–757.

Majlesi, K. (2016), 'Labor market opportunities and women's decision making power within households', *Journal of Development Economics* **119**, 34–47.

Manser, M. and Brown, M. (1980), 'Marriage and household decision-making: A bargaining analysis', *International Economic Review* **21**(1), 31–44.

McElroy, M. B. and Horney, M. J. (1990), 'Nash-bargained household decisions: Reply', *International Economic Review* **33**(1), 237–242.

Mincer, J. (1962), Labor force participation of married women: A study of labor supply, *in* 'Aspects of Labor Economics', Princeton University Press, pp. 63–105.

Mincer, J. (1963), Market prices, opportunity costs, and income effects, *in* C. F. Christ, ed., 'Measurement in Economics: Studies in Mathematical Economics and Econometrics in Memory of Yehuda Grunfeld', Stanford University Press, pp. 67–82.

National Planning Commission (2016), 'Growth and Transformation Plan II (GTP II), (2015/16-2019/20)', Vol. I, (main text), National Planning Commission of the Federal Democratic Republic of Ethiopia, Addis Ababa.

Oqubay, A. (2018), The structure and performance of the Ethiopian manufacturing sector, *in* F. Cheru, C. Cramer and A. Oqubay, eds, 'The Oxford Handbook of the Ethiopian Economy', 299, Oxford University Press.

Osili, U. O. and Long, B. T. (2008), 'Does female schooling reduce fertility? Evidence from Nigeria', *Journal of Development Economics* **87**(1), 57–75.

Pollak, R. A. (2005), 'Bargaining power in marriage: Earnings, wage rates and household production', NBER Working Paper 11239. National Bureau of Economic Research, Cambridge, MA.

Pritchett, L. H. (1994), 'Desired fertility and the impact of population policies', *Population and Development Review* **20**(1), 1–55.

Rosenzweig, M. R. and Wolpin, K. I. (1980*a*), 'Life-cycle labor supply and fertility: Causal inferences from household models', *Journal of Political Economy* **88**(2), 328–348.

Rosenzweig, M. R. and Wolpin, K. I. (1980*b*), 'Testing the quantity-quality fertility model: The use of twins as a natural experiment', *Econometrica: Journal of the Econometric Society* **48**(1), 227–240.

Rosenzweig, M. R. and Zhang, J. (2009), 'Do population control policies induce more human capital investment? Twins, birth weight and China's "one-child" policy', *The Review of Economic Studies* **76**(3), 1149–1174.

Schultz, T. P. (1990), 'Testing the neoclassical model of family labor supply and fertility', *Journal of Human Resources* pp. 599–634.

Shapiro, D. and Hinde, A. (2017), 'On the pace of fertility decline in sub-Saharan Africa', *Demographic Research* **40**, 1327–1338.

The World Bank (2019), 'Fertility rate, total (births per woman)'. Data retrieved from World Development Indicators, 23.10.2019, `https://data.worldbank.org/indicator/SP.DYN.TFRT.IN`.

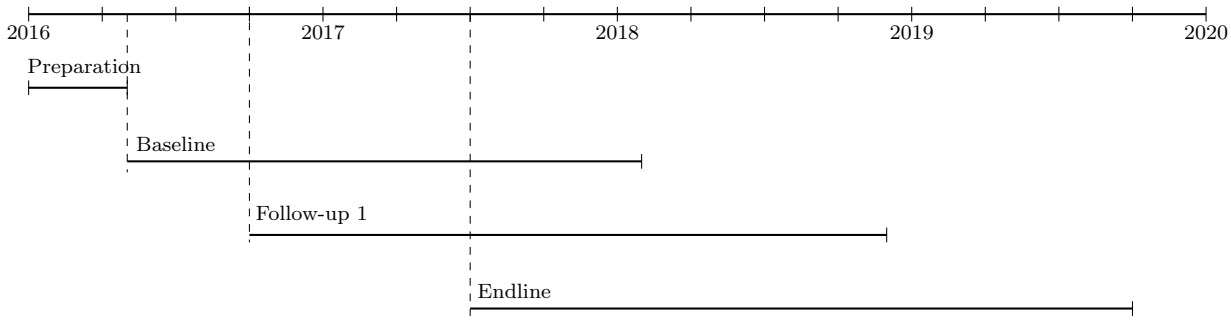
United Nations (2019), *World Population Prospects 2019: Highlights*, United Nations, Department of Economic and Social Affairs, Population Division. ST/ESA/SER.A/423.

Upadhyay, U. D., Gipson, J. D., Withers, M., Lewis, S., Ciaraldi, E. J., Fraser, A., Huchko, M. J. and Prata, N. (2014), 'Women's empowerment and fertility: A review of the literature', *Social Science & Medicine* **115**, 111–120.

Westfall, P. H. and Young, S. S. (1993), *Resampling-Based Multiple Testing: Examples and Methods for p-value Adjustment*, Vol. 279, John Wiley & Sons.

Willis, R. J. (1973), 'A new approach to the economic theory of fertility behavior', *Journal of Political Economy* **81**(2, Part 2), S14–S64.

A Timeline


2016
2017
2018
2019
2020
Preparation
Baseline
Follow-up 1
Endline

B Test for balance of endline sample and attrition

Table B1: Baseline summary means and tests of randomization balance of endline sample

	Endline sample		
	Control Mean/SE (1)	Treatment Mean/SE (2)	t-test difference (1)-(2)
Age	24.436	24.362	0.074
	(0.194)	(0.206)	
Years of schooling	9.252	9.528	-0.277*
	(0.126)	(0.124)	
Muslim	0.160	0.121	0.039**
	(0.015)	(0.014)	
Ethiopian orthodox	0.587	0.635	-0.048*
	(0.021)	(0.020)	
Protestant	0.245	0.245	0.000
	(0.018)	(0.018)	
Married	0.951	0.926	0.026*
	(0.009)	(0.011)	
Have ever given birth	0.729	0.679	0.050*
	(0.019)	(0.020)	
Number of children	1.297	1.149	0.148**
	(0.053)	(0.050)	
Pregnant at baseline	0.056	0.060	-0.005
	(0.010)	(0.010)	
Number of children living in the household	1.248	1.138	0.110*
	(0.047)	(0.046)	
Number of adults living in the household	2.184	2.129	0.055*
	(0.021)	(0.020)	
Desired number of children	3.868	3.902	-0.035
	(0.069)	(0.078)	
Decision-making index (0-1)	0.511	0.554	-0.043**
	(0.015)	(0.014)	
Have ever had a formal job before	0.281	0.316	-0.034
	(0.019)	(0.020)	
Respondent's income the last six months	2137.226	1985.258	151.968
	(187.335)	(136.996)	
Husband's income the last six months	15193.516	14808.039	385.477
	(539.194)	(509.297)	
Total recurrent expenses the last months	2524.599	2506.440	18.159
	(61.020)	(60.879)	
Total infrequent expenses the last six months	2037.516	1889.450	148.065
	(114.068)	(87.664)	
N	576	564	
			1.606*
			1 140

The value displayed for t-tests are the differences in the means across the treatment and control groups. The covariate variable *block* is included in all estimation regressions. All missing values in balance variables are treated as zero. $^{***}p < 0.001$,$^{**}p < 0.05$,$^{*}p < 0.1$.

Table B2: Correlates with attrition

	First follow-up		Endline	
	(1)	(2)	(3)	(4)
Treatment	-0.004	-0.009	-0.019	-0.023
	(0.018)	(0.018)	(0.021)	(0.021)
Age		-0.003		-0.007*
		(0.003)		(0.003)
Years of schooling		0.003		0.000
		(0.005)		(0.005)
Muslim		0.158		0.185
		(0.163)		(0.187)
Ethiopian orthodox		0.169		0.209
		(0.161)		(0.185)
Protestant		0.138		0.142
		(0.157)		(0.181)
Married		-0.058		-0.096*
		(0.037)		(0.043)
Have ever given birth		-0.102***		-0.086**
		(0.029)		(0.033)
Number of children		-0.030		-0.025
		(0.019)		(0.022)
Pregnant at baseline		-0.131**		-0.110*
		(0.041)		(0.048)
Number of children living in the household		0.023		0.007
		(0.017)		(0.020)
Number of adults living in the household		0.011		0.017
		(0.023)		(0.027)
Desired number of children		0.011		0.011
		(0.006)		(0.007)
Decision-making index (0-1)		-0.050		-0.085*
		(0.030)		(0.034)
Have ever had a formal job before		-0.008		-0.008
		(0.022)		(0.026)
Respondent's income the last six months		0.000		0.000*
		(0.000)		(0.000)
Husband's income the last six months		0.000		0.000
		(0.000)		(0.000)
Total recurrent expenses the last months		-0.000		-0.000
		(0.000)		(0.000)
Total infrequent expenses the last six months		0.000		0.000
		(0.000)		(0.000)
Observations	1460	1460	1460	1460
p-value from F-test for joint significance	0.001	0.000	0.000	0.000
Dependent variable mean	0.159	0.159	0.219	0.219

OLS regressions of an indicator for attrition at first follow-up and endline. Block fixed effects are included in all regressions (not displayed). Standard errors in parentheses. $^{***}p < 0.001$,$^{**}p < 0.05$,$^{*}p < 0.1$.

Lee bounds

As a robustness check to the average treatment effects of the main outcome variables, we apply Lee bounds (Lee, 2009) for upper and lower distribution attrition scenarios. The estimates are presented in the Table B3. The Lee bounds is a non-parametric trimming technique that rests on the assumption of monotonicity. This means that treatment status makes attrition either more or less likely. In our setting, this means that we assume that there are some individuals who attrited if they were assigned to the control group, but that no individuals attrited because of being assigned to treatment. This seems plausible in our setting and the bounds can be interpreted as the estimates for "never-attriters". The Lee bounds technique trims, either from below or above, the group that has less attrition such that the groups have equal shares of missing observations. The lower bounds occur only if the individuals with the highest values of the outcome variable attrit (highest income, bearing child, highest number of desired lifetime fertility, or having most household decision-making power), while the upper bound occurs if only individuals with the lowest values attrit. The table shows the point estimates for the main outcome variables do no change significantly when trimming the sample to account for attrition.

Table B3: Bounds for treatment effects at endline

	Average treatment effect (1)	Lower bounds (2)	Upper bounds (3)	Trimming portion
Fertility	-0.002	-0.009	0.018	0.026
	(0.026)	(0.027)	(0.034)	
Desired lifetime fertility	0.063	-0.014	0.289	0.028
	(0.116)	(0.132)	(0.201)	
Respondent's income the last six months	1009***	910***	1396**	0.026
	(262)	(285)	(378)	
Household decision-making index	0.016	-0.001	0.026	0.026
	(0.020)	(0.027)	(0.023)	

The table reports bounds computed using the Lee bounds (2009). The lower and upper bound, respectively, correspond to extreme assumptions about the missing information that is consistent with the observed data. Means and standard errors are shown in column (1). Estimated bounds and standard errors are shown in columns (2)-(3). The estimates are without covariates. $^{***}p < 0.001$,$^{**}p < 0.05$,$^{*}p < 0.1$.

C Additional analyses on fertility

C.1 Heterogeneity in treatment effects on fertility

C.2 Treatment effects on fertility by region

Table C1: Heterogeneiety in treatment by baseline age, education, religion, childless, husband's income, and work experience.

	var=Age<25		var=Low educ. level		var=Orthodox		var=No child at baseline		var=Household's income<median		var= Had wage employment the last six months before baseline	
Dependent variable:	AF	DF	AF	DF	AF	DF	AF	DF	AF	DF	AF	DF
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treatment (a)	0.057	0.263	-0.022	0.005	-0.005	0.312	0.017	0.162	-0.018	0.129	0.002	0.075
	(0.036)	(0.167)	(0.028)	(0.072)	(0.044)	(0.203)	(0.030)	(0.113)	(0.036)	(0.120)	(0.026)	(0.099)
var	0.074*	0.228	-0.005	-0.024	-0.144***	-0.467**	0.039	0.044	-0.070**	0.004	0.008	-0.053
	(0.039)	(0.178)	(0.051)	(0.181)	(0.047)	(0.200)	(0.041)	(0.164)	(0.033)	(0.124)	(0.051)	(0.175)
Treatment*var (b)	-0.094**	-0.299	0.099*	0.323	0.017	-0.369*	-0.041	-0.264	0.041	-0.084	0.005	0.073
	(0.047)	(0.184)	(0.055)	(0.282)	(0.052)	(0.219)	(0.051)	(0.186)	(0.047)	(0.174)	(0.065)	(0.247)
(a+b)	-0.037	-0.036	0.077	0.329	0.012	-0.056	-0.024	-0.102	0.023	0.046	0.007	0.148
	(0.032)	(0.087)	(0.047)	(0.273)	(0.028)	(0.080)	(0.041)	(0.145)	(0.032)	(0.128)	(0.059)	(0.222)
Observations	1,140	1,138	1,140	1,138	1,140	1,138	1,140	1,138	1,140	1,138	1,140	1,138
Control mean	0.252	4.168	0.252	4.168	0.252	4.168	0.252	4.168	0.252	4.168	0.252	4.168

OLS regressions. Outcome variables: AF = Actual fertility (currently pregnant or have given birth after baseline measured at endline). DF = Desired lifetime fertility at endline. The coefficient of (a) is the treatment effect on the women who are older than 25, have medium or higher education, are not Orthodox Christian, have children in their household, have higher than median total household income, or who had wage employment in the six months before baseline. While the coefficient of (a+b) is the treatment effect on the opposite groups, that is, the women who are younger than 25, who have low education and so on. All regressions include baseline covariates and block fixed effects. Baseline covariate controls include: age, religion, education level, total household income the last six months, number of household members, and a dummy for whether the respondent had wage employment the last six months. The covariate with the corresponding heterogeneity variable is dropped as a control variable, e.g., in columns (1) and (2) the covariate Age is not included. Robust standard errors in parenthesis. $^{***}p < 0.001$,$^{**}p < 0.05$,$^{*}p < 0.1$.

Table C2: The impact of the job offer on fertility by region

	Tigray		Amhara		Oromia		SNNP		Dire Dawa	
	AF	DF	AF	DF	AF	DF	AF	DF	AF	DF
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
PANEL A: OLS regressions										
Treatment	-0.009	-0.040	0.084	1.124	-0.179	-0.213	-0.046	0.063	0.110**	0.524
	(0.032)	(0.073)	(0.152)	(0.695)	(0.480)	(0.784)	(0.055)	(0.130)	(0.054)	(0.400)
Observations	552	551	66	65	20	20	295	295	207	207
Control mean	0.252	4.168	0.252	4.168	0.252	4.168	0.252	4.168	0.252	4.168
PANEL B: IV regressions										
Any formal job since baseline	-0.067	-0.271	0.125	1.620*	-0.686	26.397	-0.073	0.101	0.368*	1.714
	(0.218)	(0.489)	(0.210)	(0.866)	(1.090)	(326.699)	(0.083)	(0.199)	(0.200)	(1.382)
First stage results: Any formal job since baseline										
Treatment	0.139***	0.148***	0.668**	0.694**	0.261	-0.008	0.627***	0.625***	0.300***	0.306***
	(0.038)	(0.038)	(0.305)	(0.297)	.	.	(0.055)	(0.056)	(0.080)	(0.080)
p-value from F-test	0.000	0.000	0.000	0.000	.	.	0.000	0.000	0.000	0.000

Outcome variables: AF = Actual fertility (currently pregnant or have given birth after baseline measured at endline). DF = Desired lifetime fertility at endline. All regressions include baseline covariates and block fixed effects. Baseline covariate controls include: age, religion, education level, total household income the last six months, number of household members, and a dummy for whether the respondent had wage employment the last six months. Robust standard errors in parenthesis. $^{***}p < 0.001$,$^{**}p < 0.05$,$^{*}p < 0.1$.

D Treatment effects on household decision-making power

Table D1: The impact of the job offer on household decision-making power

	Decision-making index (higher number higher dm power)	Whether to send or not send children to school	What to do if a child falls sick	What to do if the respondent falls sick	Whether to have children or to have more children	Which family planning methods to use	Whether or not you should earn money outside the house	Whether you can visit your family or relatives	The use of the wife's income	The use of the husband's income	Purchase of small daily food purchases	Purchase of bulk or expensive food items	Large purchases of items like furniture, cattle, TV, or other assets	Purchase of children's clothing and shoes	Whether to open a bank account or borrow money	Whether to start a new business
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
PANEL A: OLS regressions																
Treatment	-0.020	-0.066**	-0.041	-0.019	-0.050*	-0.024	-0.067***	-0.008	-0.020	-0.023	-0.008	0.024	-0.008	-0.014	0.016	0.002
	(0.018)	(0.031)	(0.031)	(0.025)	(0.026)	(0.022)	(0.025)	(0.025)	(0.025)	(0.026)	(0.020)	(0.026)	(0.028)	(0.034)	(0.027)	(0.027)
Observations	1,140	792	792	1,138	1,118	1,127	1,140	1,140	925	1,054	1,140	1,137	1,127	792	1,135	1,133
Control mean	0.643	0.657	0.670	0.729	0.688	0.809	0.741	0.710	0.781	0.294	0.851	0.679	0.484	0.526	0.486	0.483
Unadjusted p-value	0.263	0.035	0.187	0.448	0.053	0.260	0.008	0.739	0.429	0.385	0.679	0.369	0.780	0.690	0.562	0.942
Adjusted p-value (West-Young)	0.901	0.277	0.808	0.966	0.369	0.901	0.080	0.994	0.966	0.966	0.994	0.966	0.994	0.994	0.979	0.994
PANEL B: IV regressions																
Any formal job since baseline	-0.064	-0.196**	-0.123	-0.062	-0.161**	-0.078	-0.218***	-0.027	-0.058	-0.071	-0.027	0.077	-0.026	-0.041	0.050	0.006
	(0.056)	(0.092)	(0.091)	(0.080)	(0.082)	(0.068)	(0.083)	(0.079)	(0.072)	(0.080)	(0.063)	(0.083)	(0.090)	(0.099)	(0.084)	(0.084)
Observations	1,140	792	792	1,138	1,118	1,127	1,140	1,140	925	1,054	1,140	1,137	1,127	792	1,135	1,133
Unadjusted p-value	0.255	0.034	0.179	0.439	0.049	0.251	0.009	0.732	0.417	0.373	0.672	0.355	0.774	0.680	0.552	0.940
Adjusted p-value (West-Young)	0.900	0.277	0.811	0.967	0.363	0.900	0.088	0.993	0.967	0.967	0.993	0.967	0.993	0.993	0.981	0.993

The decision-making index is constructed by 15 different household decisions in which each decision is coded as 1 if the respondent is the main decision-maker or has "a lot of input" to the decision. The 15 decisions are then averaged in one index ranging from 0 to 1. All regressions include baseline covariates and block fixed effects. Baseline covariate controls include: age, religion, education level, total household income the last six months, number of household members, and a dummy for whether the respondent had wage employment the last six months. Robust standard errors in parenthesis. $^{***}p < 0.001$,$^{**}p < 0.05$,$^{*}p < 0.1$.

E Sensitivity analysis

E.1 Income

E.2 Fertility

E.3 Desired Lifetime Fertility

E.4 Household Decision-Making Power

Table E1: Sensitivity analysis of treatment effects on income by alternative sample size and variable specification

	Income as reported in Table 2		Without control variables		Using the no-attrition sample		Do not drop observations that were not randomized		Drop observations from regions with weakest first stage		Pre-specified IV-model
	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
PANEL A: First follow-up (7 months)											
Treatment	1,176***		1,251***		1,293***		1,569***		1,762***		
	(220.416)		(227.733)		(228.604)		(235.341)		(341.429)		
Any formal job since baseline		3,156***		3,009***		3,534***		4,211***		2,957***	
		(562.655)		(529.388)		(595.442)		(613.891)		(540.034)	
Any wage employment the last six months											4,263***
											(716.397)
Observations	1,228	1,228	1,228	1,228	1,140	1,140	1,562	1,562	591	591	1,228
Control mean	2798		2798		2760		2765		2774		
PANEL B: Endline (14 months)											
Treatment	807***		1,037***				919***		1,965***		
	(258.709)		(257.312)				(288.339)		(374.713)		
Any formal job since baseline		2,588***		2,832***				2,933***		3,812***	
		(774.492)		(664.028)				(873.762)		(686.846)	
Any wage employment the last six months											4,036***
											(1,114.589)
Observations	1,140	1,140	1,140	1,140			1,453	1,453	568	568	1,140
Control mean	3666		3666				3851		3010		

The pre-specified IV-model uses the randomization as instrument for having had any formal wage employment the last six months, instead of 'since baseline' which is used in the rest of the paper. All regressions include baseline covariates and block fixed effects, except in Columns (3)-(4). Baseline covariate controls include: age, religion, education level, total household income the last six months, number of household members, and a dummy for whether the respondent had wage employment the last six months. Standard errors in parentheses. $^{***}p < 0.001$,$^{**}p < 0.05$,$^{*}p < 0.1$.

Table E2: Sensitivity analysis of treatment effects on fertility by alternative sample size and variable specification

	Income as reported in Table 3		Without control variables		Using the no-attrition sample		Do not drop observations that were not randomized		Drop observations from regions with weakest first stage		Pre-specified IV-model
	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
PANEL A: First follow-up (7 months)											
Treatment	-0.008		-0.004		-0.014		-0.006		0.020		
	(0.019)		(0.018)		(0.020)		(0.016)		(0.030)		
Any formal job since baseline		-0.023		-0.010		-0.038		-0.017		0.033	
		(0.049)		(0.043)		(0.053)		(0.043)		(0.049)	
Any formal job the last six months											-0.031
											(0.066)
Observations	1,228	1,228	1,228	1,228	1,140	1,140	1,562	1,562	591	591	1,228
Control mean	0.168		0.168		0.174		0.181		0.163		
PANEL B: Endline (14 months)											
Treatment	0.003		-0.005				0.017		0.027		
	(0.024)		(0.023)				(0.021)		(0.037)		
Any formal job since baseline		0.009		-0.015				0.056		0.052	
		(0.076)		(0.064)				(0.068)		(0.070)	
Any formal job the last six months											0.015
											(0.118)
Observations	1,140	1,140	1,140	1,140			1,453	1,453	568	568	1,140
Control mean	0.252		0.252				0.256		0.267		

The pre-specified IV-model uses the randomization as instrument for having had any formal wage employment the last six months, instead of 'since baseline' which is used in the rest of the paper. All regressions include baseline covariates and block fixed effects, except in Columns (3)-(4). Baseline covariate controls include: age, religion, education level, total household income the last six months, number of household members, and a dummy for whether the respondent had wage employment the last six months. Standard errors in parentheses. $^{***}p < 0.001$,$^{**}p < 0.05$,$^{*}p < 0.1$.

Table E3: Sensitivity analysis of treatment effects on desired lifetime fertility by alternative sample size and variable specification

	Income as reported in Table 3		Without control variables		Using the no-attrition sample		Do not drop observations that were not randomized		Drop observations from regions with weakest first stage		Pre-specified IV-model
	OLS (1)	IV (2)	OLS (3)	IV (4)	OLS (5)	IV (6)	OLS (7)	IV (8)	OLS (9)	IV (10)	IV (11)
PANEL A: First follow-up (7 months)											
Treatment	0.047		0.039		0.059		0.028		0.152		
	(0.080)		(0.080)		(0.087)		(0.070)		(0.170)		
Any formal job since baseline		0.125		0.093		0.160		0.075		0.254	
		(0.210)		(0.191)		(0.230)		(0.182)		(0.278)	
Any formal job the last six months											0.169
											(0.284)
Observations	1,225	1,225	1,225	1,225	1,137	1,137	1,556	1,556	589	589	1,225
Control mean	4.026		4.026		4.047		4.013		4.300		
PANEL B: Endline (14 months)											
Treatment	0.089		0.032				0.091		0.259		
	(0.089)		(0.089)				(0.075)		(0.180)		
Any formal job since baseline		0.282		0.087				0.287		0.501	
		(0.277)		(0.243)				(0.233)		(0.342)	
Any formal job the last six months											0.442
											(0.437)
Observations	1,138	1,138	1,138	1,138			1,448	1,448	567	567	1,138
Control mean	4.168		4.168				4.138		4.497		

The pre-specified IV-model uses the randomization as instrument for having had any formal wage employment the last six months, instead of 'since baseline' which is used in the rest of the paper. All regressions include baseline covariates and block fixed effects, except in Columns (3)-(4). Baseline covariate controls include: age, religion, education level, total household income the last six months, number of household members, and a dummy for whether the respondent had wage employment the last six months. Standard errors in parentheses. $^{***}p < 0.001, ^{**}p < 0.05, ^{*}p < 0.1$.

Table E4: Sensitivity analysis of treatment effects on household decision-making power by alternative sample size and variable specification

	Income as reported in Table D1		Without control variables		Using the no-attrition sample		Do not drop observations that were not randomized		Drop observations from regions with weakest first stage		Pre-specified IV-model
	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
PANEL A: First follow-up (7 months)											
Treatment	0.003		0.010		0.008		0.012		-0.012		
	(0.017)		(0.017)		(0.018)		(0.015)		(0.029)		
Any formal job since baseline		0.008		0.024		0.021		0.031		-0.020	
		(0.044)		(0.042)		(0.047)		(0.039)		(0.046)	
Any formal job the last six months											0.011
											(0.060)
Observations	1,228	1,228	1,228	1,228	1,140	1,140	1,562	1,562	591	591	1,228
Control mean	0.616		0.616		0.615		0.590		0.568		
PANEL B: Endline (14 months)											
Treatment	-0.020		0.002				-0.009		-0.061**		
	(0.018)		(0.019)				(0.016)		(0.030)		
Any formal job since baseline		-0.064		0.006				-0.028		-0.118**	
		(0.056)		(0.051)				(0.050)		(0.058)	
Any formal job the last six months											-0.099
											(0.088)
Observations	1,140	1,140	1,140	1,140			1,453	1,453	568	568	1,140
Control mean	0.643		0.643				0.612		0.598		

The pre-specified IV-model uses the randomization as instrument for having had any formal wage employment the last six months, instead of 'since baseline' which is used in the rest of the paper. All regressions include baseline covariates and block fixed effects, except in Columns (3)-(4). Baseline covariate controls include: age, religion, education level, total household income the last six months, number of household members, and a dummy for whether the respondent had wage employment the last six months. Standard errors in parentheses. $^{***}p < 0.001$,$^{**}p < 0.05$,$^{*}p < 0.1$.

Chapter 3

Understanding the Challenges of High Labor Turnover in the Ethiopian Manufacturing Industry

Understanding the challenges of high labor turnover in the Ethiopian manufacturing industry

Sandra Kristine Halvorsen*

Abstract

Manufacturing industry expansion is a central part of Ethiopia's growth and transformation agenda due to its potential for accelerated economic development and large-scale job creation, in particular for women. However, the industry is experiencing extremely high labor turnover rates, which is hampering the prospects of a successful industrialization of the country. Using a combination of survey data and qualitative interviews, the study highlights three main reasons for the high turnover; unrealistic expectations about wages and work efforts, poor working conditions, and difficulties combining domestic responsibilities with factory employment. The analysis suggests that the manufacturing industry has important welfare impacts for many poor households due to low labor participation of women, however, the low wages and poor working conditions lead to high turnover as many workers do not find the efforts worth the compensation. Firm managers should consider measures to handle the turnover problem as continued high turnover will likely limit the capacity of firms to participate in global markets and hamper economic development from industrialization.

Keywords: Labor turnover, female labor supply, household decision-making, manufacturing industry development

*Norwegian School of Economics and Chr. Michelsen Institute (sandra.halvorsen@nhh.no).

I would like to thank Kjetil Bjorvatn for invaluable advice and support. I also thank Gaddise Tesfaye for excellent research assistance and the Ethiopian Development Research Institute (EDRI) for outstanding data collection and collaboration. The paper has benefited from helpful comments by Charlotte Ringdal and Espen Villanger. The project would not have been possible without the financial support provided by the Centre of Ethics and Economics at the Norwegian School of Economics and The Research Council of Norway through the project: "Women in the developmental state: Female employment and empowerment in Ethiopia" project No 240807.

1 Introduction

Labor intensive manufacturing can be a powerful source of employment opportunities for women in developing countries, which may further lead to positive societal change. For instance, greater economic opportunities for women have been found to contribute to poverty reduction, promote women's empowerment, and increase investments in children, especially in girls' health and education.[1] Accordingly, the inclusion of women in the labor force is an important policy goal in most countries (World Bank, 2011, 2012).

While job creation in manufacturing has the potential of leading to positive societal change, such jobs are also often characterized by very high labor turnover.[2] For instance, recent studies on the rapidly expanding manufacturing sector in Ethiopia, document labor turnover rates of around 80-100% annually, mainly driven by low wages and poor working conditions (Barrett and Baumann-Pauly, 2019; Blattman and Dercon, 2018; Hailu et al., 2018; Hardy and Hauge, 2019; Yost and Shields, 2017). High labor turnover rates are costly both for the firms and for the workers. In particular, turnover curtails productivity, in addition to frequent investments in finding and training new staff.[3] At the same time, workers incur the costs of searching and switching jobs, while wages are often lower during the training period.[4] It is therefore important to get a comprehensive understanding of the challenges faced by the workers in order to understand why the turnover rates are so high and what can be done to counter them.

Ethiopia is an interesting case in this context, as it is currently experiencing rapid manufacturing growth and with a declared ambition of becoming the largest manufacturing hub in Africa (NPC, 2016). The Ethiopian industrial development agenda focuses on labor-intensive, export-led manufacturing in textiles, garments, leather, food and beverages, and pharmaceuticals. These sectors are intended to absorb part of the rapidly growing workforce and aim at employing young individuals where the goal is to employ women in 60% of the floor worker positions and 30% in managing positions (NPC, 2016). Through the establishment of large industrial parks and generous incentive schemes, total FDI inflow to the manufacturing sector in Ethiopia has increased more than sixfold in a nine year period: from 570.2 Million USD in 2007/8 to 3,712 Million USD in 2016/17 with an FDI share in

[1] See, e.g., Anderson and Eswaran (2009); Dharmalingam and Philip Morgan (1996); Dollar and Gatti (1999); Duflo (2012); Heath and Mobarak (2015); Jensen (2012); Kabeer (2002); Luke and Munshi (2011); Rosenzweig and Schultz (1982); Thomas (1993); Qian (2008).

[2] High labor turnover has been a characteristic in manufacturing since the earliest factories in the industrial revolution, and is a major concern in manufacturing sectors all over the world also today (see, Moon et al., 2018, and references therein).

[3] While the cost varies across firms, IFC (2013) provides several examples of high costs related to labor turnover and estimates, for instance, that for Nalt Enterprise, a Vietnamese textile company, a 10 percent reduction in labor turnover would translate into an 8.5 percent saving of the total annual wage bill when assuming that it takes up to three months for a new textile worker to reach full productivity.

[4] In Ethiopia, search costs poses significant constraints for unemployed individuals in both time and monetary terms as there are few jobs available online and many need to rely on social networks or travel to the city centre to find vacancies and to register and make formal applications (Abebe et al., 2018; Franklin, 2017; Meyer, 2018).

manufacturing of 89% (EIC, 2018). In the same period, manufacturing output grew by 17.9% (Oqubay, 2018). However, high labor turnover rates shed some doubt on both the positive welfare impact of industrialization, as they are likely to be the result of worker dissatisfaction, and on the economic feasibility of the industrialization strategy for economic growth, as high turnover rates reduce profitability of manufacturing by driving up the costs of production. In a recent study, Hardy and Hauge (2019) report that employee turnover is the most common and costly labor issue, and it is considered the main reason for poor firm performance and a barrier to competing in global markets, according to managers in the textile and leather sectors in Ethiopia. Additionally, Blattman and Dercon (2018) report that 77% of newly employed workers in their study left within a year and many left the industrial sector altogether, rather than switching firms, indicating that the problem with labor turnover is not in general specific to firms, but for the sector as a whole.

The aim of the present paper is to shed further light on labor turnover in the Ethiopian manufacturing sector and the household decision-making process regarding the women's labor supply. It adds to the existing literature methodologically by combining data from a panel survey and qualitative interviews of both workers and their spouses across five different regions. The few existing studies, including reports, which describe the workers' experience of the Ethiopian manufacturing industry mainly rely on qualitative data from a few factories only (Barrett and Baumann-Pauly, 2019; Hardy and Hauge, 2019; Yost and Shields, 2017).[5] The main benefit of combining the two types of data is their complementarity: the quantitative data of the full sample shows correlates for a large sample, while the qualitative interviews shed additional light on the mechanisms at work. Using a mixed methods approach thus allows me to explore the issue of labor turnover by mapping the most common reasons and at the same time learning the respondents' own stories associated with the categorical measures for a better understanding and complete analysis.

The paper also contributes to the literature by investigating the household dimension of the decision-making process regarding women's labor supply in the manufacturing industry. While the majority of factory workers in Ethiopia are young, unmarried women who have migrated to the cities for work, such as those studied by Blattman and Dercon (2018), my analysis focuses instead on a less-studied group, namely the married women. Today, these women have a lower likelihood of being recruited due to their status of relationship or motherhood.[6] Since marriage and childbearing is nearly universal in Ethiopia, even among young women, a large pool of potential workers are thus held back from participating.[7] However, these women may be very valuable to the companies. For instance, married women may be less likely to migrate for other labor opportunities and they may have longer time-perspectives regarding their employment than single women (Djamba et al., 2006).

[5]Two exceptions are Blattman and Dercon (2018) and Hailu et al. (2018) who combine survey data with qualitative interviews.

[6]In many firms the applicants are given a score based on their experience and qualifications. Being single and childless give higher scores than being married or having children.

[7]According to the most recent Demographic and Health Survey (DHS) data from 2016, 63.9% of all women between 15-49 were currently married (65.2% currently in a union). Ethiopian women also marry early, the median age of marriage in 2016 was 17.1 years old, and by the age of 24, 69% have married.

The two studies closest to the current paper are Blattman and Dercon (2018) and Hailu et al. (2018). Both papers describe the working conditions in the Ethiopian manufacturing industry and discuss possible reasons for the high labor turnover. The current paper adds to these studies by focusing on the importance of the household as a decision-making unit, and in particular the role of domestic obligations and husbands' preferences and beliefs for the women's job decision, which potentially play important roles in women's labor market decisions. Collective models of household behavior typically consider the spouses' market wages and relative power to be important for the members' labor supply (Browning et al., 2011). A difference in market wages leads spouses to specialize in their comparative advantages, since this will maximize the household output (Becker, 1985, 1991). Women generally earn less in the market and have a biological and cultural comparative advantage for childcare, hence most women thus specialize in household production and most men specialize in market activities. When women's market wage increases, their time spent on home production and leisure becomes relatively more expensive and more women will enter the labor market. Since a shift in the time allocated to home production and market activities of one household member affects all household members, the decision to enter the labor market is dependent on all household members' preferences and relative power. Due to the added complexity of household labor supply decisions, compared to single decision-makers, it is important to include the husbands' voice and opinions into the analysis of women's labor supply.

The main sample in this study consists of 850 married women who were offered jobs in the Ethiopian manufacturing sector (full sample). The women were surveyed on the day of recruitment (baseline) and again about seven months later (endline). To explore their labor market decisions in further detail, and from a household perspective, a sub-sample of 24 women and 11 of their husbands were individually interviewed using semi-structured interviews (focus sample).

Only half of the women who were offered jobs were working in the factory at the endline survey. The study highlights three main reasons for the turnover; unrealistic expectations about wages and work efforts, unhealthy working conditions, and intra-household decisions on time allocations. First, before starting the factory job, many applicants had over-optimistic expectations regarding the wage level and work efforts, especially in areas with newly established manufacturing industry. Unrealistic expectations then led to short employment intervals as many quickly realized that the jobs were not what they had imagined. By comparing the stayers and the leavers, I find that women coming from poorer households and who had higher earnings potential within the factory were more likely to remain in the job.

Second, unhealthy working conditions, long and inflexible working hours, few opportunities for advancement, and disrespectful treatment by supervisors were also important reasons for leaving the job. Combined with inadequate compensation, many chose to leave the job even without having an alternative income. A somewhat contradictory finding is that despite the low wages and poor working conditions, most still described themselves as happy with the factory job (83% of the stayers and 44% of the leavers). A possible explanation is that the women operate in labor markets with few employment opportunities; for most

the alternative to factory employment is to stay at home or participate in erratic work in the informal sector. In contrast to their alternative options, the factories, in general, offer a number of benefits such as a longer-term contracts, steady income, pension, maternity leave, compensation for sick days, and annual leave which likely helped, but were not sufficient to retain most workers.

Third, the qualitative assessment indicated that, in general, the husbands were supportive of the wives' employment decision, but that they would prefer that their wives were at home or had a less demanding job. The factories' inflexible work schedules and heavy workload make it problematic to combine the job with household responsibilities such as cooking and childcare, especially since it is difficult for many to find reliable childcare options.

In sum, the manufacturing jobs seem to be good for many workers, in that they offer higher income and stable employment in a market with high unemployment. At the same time, poor working conditions and low wages make the jobs too hard and not worth the effort for others, and instead, many of the women choose to work at home or find other jobs. To retain workers, the factories could benefit from targeting medium educated women (completed 10th grade) coming from poorer households. The hiring managers also need to be more explicit about the wage levels and expected work effort so that the applicants can better consider whether the job will fit their preferences and if they should accept the job offer. Furthermore, due to high incidents of work-related health issues, improving working conditions would likely lower turnover. Finally, providing childcare options for the workers may be an effective way to retain female workers, as for many, home-based childcare is their main alternative to the factory employment. While factory managers may be opposed to higher labor costs, whether from higher wages or improved working conditions, the cost of not handling the high labor turnover rates may be even larger and a trade-off between higher labor costs and further expansion of the industry may exist.

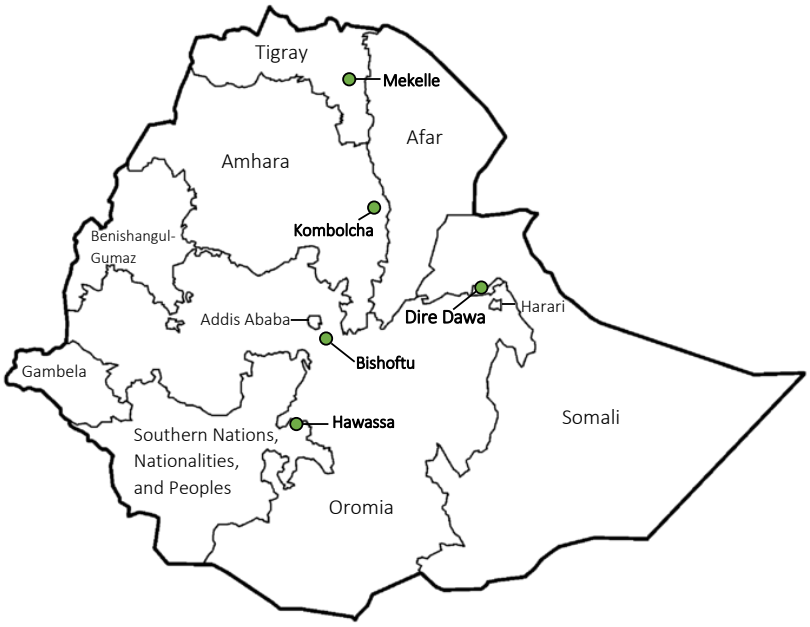
The remainder of the paper is organized as follows. Section 2 introduces the study setting and data collection, as well as the participants' background and their motives for applying for jobs in the manufacturing industry. Section 3 presents the findings and analysis, by first describing the workers' expectations and experiences of the job, and then continues to describe the reasons for labor turnover. Section 4 summarizes the implications of the findings and concludes.

2 Background

2.1 Setting and data collection

The findings and analyses in this paper are based on a panel survey of manufacturing workers in Ethiopia and qualitative interviews with a selection of current or previous workers and their husbands from the same survey. The survey was carried out with workers from twenty-seven manufacturing firms in five different regions between March 2016 and December 2018 (an illustration of the timeline is included in Appendix A). The sample is mainly drawn from Mekelle Industrial Park (n=417), Hawassa Industrial Park (n=282), and Dire Dawa (n=102),

Figure 1: Location of samples



while a few respondents are from Kombolcha Industrial Park (n=15) and Bishoftu (n=34). Mekelle Industrial park, located in Tigray in the North part of Ethiopia and Kombolcha Industrial Park, located in Amhara, were both inaugurated in July 2017 and are expected to create more than 20,000 jobs each, with special emphasis on the export garment industry. Hawassa Industrial Park was opened just one month earlier in the region of Southern Nations, Nationalities and Peoples (SNNP). The park is expected to employ 60,000 workers at full capacity and is already hosting important global companies, such as PVH, the world's second-largest apparel company, and the Swedish clothing giant H&M. In Dire Dawa, there is so far no industrial park although the area has a long tradition of manufacturing due to its direct rail-route to the Djibouti port. A 4,000 acre industrial park is contracted to be built in the near future. The last area includes three factories located around Bishoftu and outside Addis Ababa.

The firms in this study were mainly in the textile and garment industries with both foreign and domestic ownership. The main sample consists of 850 women who had applied for, and were offered, positions in one of the factories. The respondents were surveyed at the time of employment (baseline) and the endline survey was carried out about seven months later. The survey was collected as part of a larger study on manufacturing workers in Ethiopia and includes a comprehensive set of information about the workers' welfare, living arrangements, and domestic relations.[8] The sample includes only married women or women

[8]In the larger study, the sample of eligible entry-level applicants in the Ethiopian manufacturing industry was randomized into receiving a job offer or serving in a control group. The main findings from the larger

living with a partner at the time they applied for the job.

In addition to the survey, individual qualitative interviews were conducted in September and October 2019 with 24 women who were working or had been working in one of the factories and 11 of their husbands. These will be referred to as the focus sample. The number of respondents was based on budget restrictions and anticipated saturation. Following Guest et al. (2006), the Appendix Figure B1 documents the saturation of code creation. That is, for each successive set of five interviews (either of couples or the women alone), I document the number of thematic codes (similar answers) that evolved during the analysis. The figure shows that after 15 interviews only one more thematic code was added to the analysis, thus including more interviews would likely have had limited added value. Respondents were sampled from two different areas; Hawassa and Dire Dawa. In Hawassa, manufacturing is relatively new and concentrated in an industrial park with mainly foreign owners, while Dire Dawa has a longer tradition in manufacturing and the firms are not clustered as the firms are in Hawassa. To ensure variation in the workers' experiences, the informants were from seven different factories and with varying duration of employment. The interviews were semi-structured and focused on the household decision-making processes of the women's entry and exit in the industry and their experiences of the work, especially in regard to their domestic roles as wives and mothers.[9] The interviews were carried out by myself together with a research assistant for translation purposes. Each interview lasted between an hour and an hour and a half and were carried out in the respondents' own home or in a nearby café. While we tried to reach the husbands of all the respondents, we were only able to interview half of the husbands as, in many cases, they were working late or working out of town. The interview notes were transcribed on the same day and analyzed in an inductive manner (data-driven) using thematic analysis (Guest et al., 2011).

2.2 Study participants

Table 1 summarizes some key characteristics of the respondents. Column (1) describes the full sample which is used for the quantitative analysis, while columns (2) and (3) describe the sub-sample of Dire Dawa and Hawassa and compares the respondents from the survey sample (column (2)) with the qualitative interviews (column (3)) to see whether the qualitative interviews are representative of the larger survey sample. Significance of difference in means between the survey and interview samples are indicated in the last column.

In the full sample, the women's ages ranges from 18 to 60 years old, but 91% were aged thirty or younger and on average the respondents were 25 years old. Almost 80% of the full sample had completed at least 10th grade, reflecting that most factories required completion of 10th grade. Only 2% of the women were household heads and by design, almost all of the women in the full sample were married, and the remaining were living with a partner. Most of the respondents lived in nuclear households with their spouse and their children. The last

project are reported in other articles: Halvorsen et al. (2020) on fertility outcomes and Kotsadam and Villanger (2020) on intimate partner violence.

[9]The interview guide is provided in Appendix D.

Table 1: Summary statistics of the sample at baseline

	Full sample	Focus sample		
	All	Survey	Qual. Interviews	Sig. diff.
	(1)	(2)	(3)	(2)-(3)
Panel A. Respondents				
Age	24.69	26.77	29.71	*
Low education (<10 years completed)	0.21	0.30	0.63	***
Medium education (10 years completed)	0.53	0.41	0.21	**
High education (>10 years completed)	0.26	0.29	0.17	
Muslim	0.07	0.09	0.21	*
Ethiopian Orthodox	0.65	0.29	0.38	
Protestant	0.28	0.60	0.38	**
Panel B. Households				
Respondent is the household head	0.02	0.02	0.00	
Married	0.94	0.95	0.96	
Number of adults living in the household	2.15	2.29	2.42	
Number of children living in the household	1.10	1.46	1.46	
Panel C. Employment status at endline				
Working in the factory	0.51	0.51	0.50	
Self employed	0.05	0.10	0.00	
Other employment	0.07	0.11	0.08	
No job	0.38	0.29	0.42	
Observations	850	360	24	

Significance of difference in means indicated by asterisks: *** significant at the .01 level; ** significant at the .05 level; * significant at the .10 level.

panel of the table describes the respondents' employment status at the endline survey. In the full sample, half of the respondents were not working in any factory and most of those who had left factory work had not taken other employment or self-employment.

The qualitative interviews were conducted to gain more insights into the workers' decisions to leave or stay in the factory. Since the selection of respondents to the qualitative interviews was based on quota sampling, instead of purposeful sampling, it is relevant to test the representativeness of the qualitative sample to the survey sample. By comparing the two sub-samples of the survey and qualitative interviews in Dire Dawa and Hawassa along some main characteristics, we see that the qualitative interviewed informants are similar to the survey respondents on most household observables and their employment status at endline (comparing columns (2) and (3)). The qualitative interviewed informants are however on average about three years older, less educated, and of different religion than the surveyed respondents in their areas (see significance of difference in the last column). This difference is likely because 14 of the 24 interviews took place in Dire Dawa where the respondents (both survey and interviewed sub-samples) were on average older, less education, Muslim,

and from poorer households, than respondents from Hawassa or the other regions. If we instead split the sub-sample by region, the survey and qualitative samples are very similar on all characteristics (not shown because of small sample sizes).

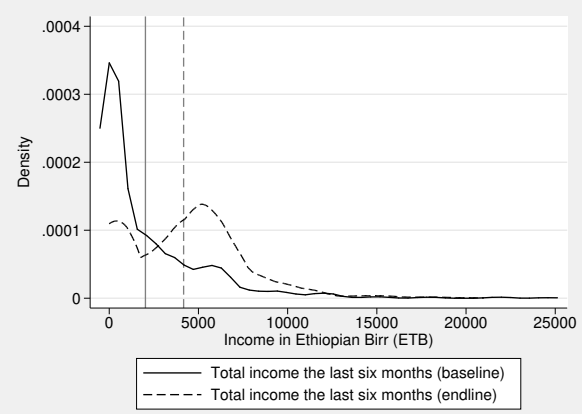
2.3 The women's economic background and motive for applying for the factory job

Before entering the factory job, less than one-third of the women had ever had a formal or informal job (29%) and only 6% had experience with other factory work (see Table 2). The most common previous jobs were in unskilled trades (33%), such as daily worker in construction or domestic worker, or in skilled and semi-skilled trades (24%), such as factory work or hairdresser. On average their income the past six months before baseline was 2,108 Ethiopian Birr (ETB), however, 46% of all the applicants in the sample did not have any income the previous six months, and by excluding the ones who did not have any income, the average income the past six months before baseline almost doubles to ETB 3,870. The husbands' average income the past six months at baseline was ETB 15,130, about seven times more than that of their wives. Among those respondents who did have a job prior to the factory, the average (basic) monthly salary was ETB 747, which is about 25% lower than the average (basic) salary in the factories of ETB 964 measured at the endline survey (p-value<0.000). While the respondents who were interviewed in the focus sample had on

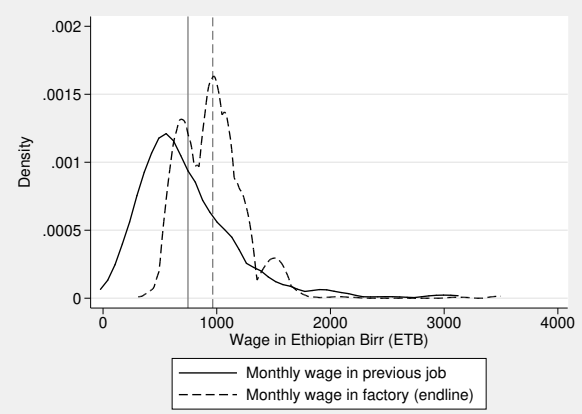
Table 2: Economic background of the sample at baseline

	Main sample	Focus sample		
	All	Survey	Qual. Interviews	Sig. diff.
	(1)	(2)	(3)	(2)-(3)
Ever had a job before	0.29	0.30	0.33	
Agriculture and fishing	0.03	0.03	-	
Retail and commercial (Self employed)	0.20	0.27	0.25	
Retail and commercial (Employed)	0.12	0.12	0.13	
Unskilled trade	0.33	0.28	0.25	
Skilled and semi-skilled trades	0.24	0.22	0.38	
Professionals	0.08	0.7	-	
Student	0.01	0.01	-	
Ever worked in a factory	0.06	0.06	0.13	
Respondent's income the last six months	2,108	2,501	3,722	
Husband's income the last six months	15,130	13,956	14,755	
Monthly salary in previous job	747	718	704	
Monthly salary in factory (endline)	964	861	910	
Observations	850	360	24	

All income and salary measures are in Ethiopian Birr (ETB). Significance of difference in means indicated by asterisks: *** significant at the .01 level; ** significant at the .05 level; * significant at the .1 level.



(a) Respondents' total income in the six months prior to baseline and endline



(b) Pre-factory and factory jobs' monthly wage

Figure 2: Distributions of incomes before and after entering factory employment

average higher income than the larger survey, the measures are not significantly different.

Figure 2 displays the distribution of incomes before and after entering factory employment. Panel 2(a) shows the distribution of the respondents' total income in the six months prior to baseline and endline for the full sample. It is evident that the number of respondents who had no, or very low, income before entering the factory employment is much lower by endline; at baseline, 46% had no income in the six months prior to the survey, while 20% had no income in the six months prior to the endline survey. On average, the total income of the respondents increased by 115%, from ETB 2,108 to ETB 4,539 (the mean total incomes are illustrated with vertical lines). While much of this increase is coming from women who entered the labor market for the first time, women who had some income before entering the factory employment also experienced higher incomes. Panel 2(b) shows the distribution of the monthly wage in the jobs they had prior to the factory employment, and the distribution of monthly wage at the factory reported by the respondents at endline. We see that the factory jobs, in general, were better paid than their previous jobs, with the factory wage distribution spiking at around ETB 700 and ETB 1000.

To provide further insights into how their labor marked decisions were made, I started the qualitative interviews by asking them to tell me about their lives; their previous occupation history, their household situation, and key events in their lives leading them up to the present time. The answers that emerged under each theme are summarized in Table 3. Need for money, either to supplement their husbands' income or to improve their lives, was the most common (21/24) reason for their decision to apply for the factory job. Some also said they applied because they wanted to have something to do or they were curious about the factory jobs, telling us that staying at home makes them feel isolated and bored (7/24). There were few alternative opportunities for employment for the women in the sample and apart from factory employment, erratic work in the informal sector was the immediate job

Table 3: Motivation for applying for factory jobs (focus sample)

Themes (Questions)	Answers (codes)	Count
Motivation for applying	We needed money	21/24
	The only job available	3/24
	To save money to start my own business	4/24
	I was curious/I wanted to work	7/24
Alternative to factory job	Start my own business	10/24
	Stay at home	6/24
	Other employment (maid, daily worker in construction, waitress, cleaner, nurse)	12/24
The role of the husband in the decision to apply for the factory job	Supportive	8/24
	No role	15/24
Husband's opinion about factory jobs	Wants her to work in the factory	4/11
	Does not want her to work in the factory	5/11
	Has mixed opinion about her working in the factory	2/11

The respondents' answers could fit into several codes, thus the sum of answers exceeds the number of respondents. For instance if a respondent answers "I applied for the job because we needed more income and because I don't have education, factory job is the only available job for me," the respondent's answer would then be coded as both "We needed money" and "The only job available."

alternative for many. Such jobs included being a maid or doing daily construction work. However these jobs were described as very demanding and unstable (W3, W4, W12, and H2).[10] For instance, we were told that maids are not respected and are expected to do all kinds of work at all times of the day (W12 and H2). The preferred alternative for many (10/24) was to run their own business, either selling vegetables on the street or having a small coffee stand as that would give them the freedom to be independent and get a quite decent income. When asked why they did not do that instead, they said they were lacking the initial capital needed to start and therefore chose the factory employment. Additionally, some considered running their own business as more risky and unstable than factory work (W11 and W22). Only four out of the twenty-four said they started the job to save up money for their prospective business. Most of those who stated another type of employment as their alternative, said that they had chosen factory employment because they believed it would give them a higher salary (10/12), more security (7/12), or because they could not get employed in the other occupation or were offered the factory job first (5/12). Finally, six out of the twenty-four said that if there were no factory jobs, they would just stay at home.

For most of the women we interviewed, the husband had no role in the decision to apply

[10]The abbreviation refers to qualitative interviews with wives in couples no. 3, 4, and 12 and the husband in couple no. 2.

for the job. Several women (15/24) told us something like this: "My husband had no role in the decision to apply for the job, I just decided myself and told him about it later." Others (8/24) had a more active and supportive husband in the decision. They typically told us that the two of them had discussed that she should find a job to help with the expenses, but that she could decide for herself which type of job to take. Many of the husbands (7/11) expressed concerns saying that they knew factory work would be very hard and that they did not want their wife to work there, but that they really needed the income and the decision was hers.

3 Analysis

3.1 Experiences of the job

3.1.1 The full sample

From the survey we find that both those who had previous work experience and those who did not, reported higher salary as the most frequent argument for why factory employment was preferable to their previous job or their current situation at baseline. Future prospects, stability, and good working environment were also important factors they expected the factory would offer at baseline. While optimistic expectations about the factory job were high for most, many expressed disappointment as the salary and working conditions did not match what they expected before they started. This was especially true among younger and the ones who were inexperienced with formal jobs. Half of the respondents had learned about the factory job from a social contact and only 15% had learned about the job from a social contact already working in the factory, which may suggest that only a few had a good understanding of what the job would be like.

Table 4 compares their expectations at baseline with the realized measures of salary and job satisfaction at endline. First we see that the mean monthly salary of ETB 964 is almost 30% lower than the baseline expectation of ETB 1,348, the difference is statistically significant (p-value<0.000).[11] We further see that while 40% had expected to earn more than ETB 1,500, only 7% ended up earning more than this. Their wage expectations also varied across the different areas (not shown in the table). In places where the factories had existed for a longer time (Oromia and Dire Dawa), the average wage expectations were significantly lower than in areas where industrial parks were newly established (Hawassa, Mekelle, and Kombolcha). In particular, in Oromia and Dire Dawa, the average expected monthly wage was ETB 1,011, while in the three areas with new industrial parks, the average expected wage was ETB 1,398 (p-value<0.000). The perhaps more realistic expectations in Oromia and Dire Dawa are probably due to better information; in these regions 51% had learned about the job from a social contact already working in the factory, while in the other regions

[11]The wage is the basic wage reported by the respondent. If we add the monetary value of other benefits such as bonuses, overtime, lunch, and transportation, the mean of the reported monthly wage is ETB 1,083, with a standard deviation of 372.

Table 4: The women's expectations (full sample)

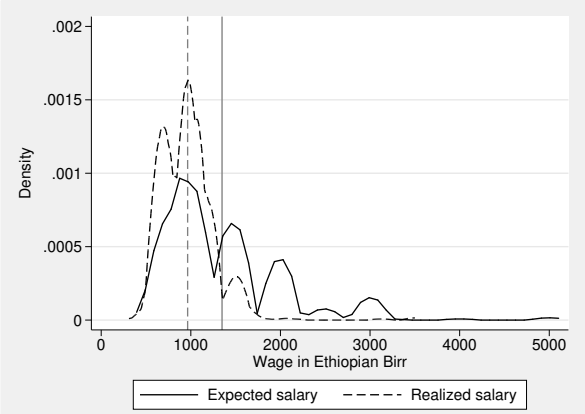
	Baseline expectations		Endline realizations		
	Mean (1)	Percent (2)	Mean (3)	Percent (4)	Sig. diff. (1)-(3)
What is your expected basic salary in this factory? (Ethiopian Birr)	1,348 (668)		964 (311)		***
310 - 499[1]		-		0.5%	
500 - 999		28%		49%	
1,000 - 1,499		32%		43.5%	
1,500 - 1,990		19%		6%	
2,000 - 5,000		21%		1%	
Rank of factory job from worst to best given your qualifications. (1=Worst to 10=Best possible job)	8.54 (1.90)		8.54 (1.92)		
10		51%		48.5%	
5-9		46%		48.5%	
1-4		3%		3%	
Taking all things together, do you think you will be happy in this job? (1=Not at all happy to 4=Very happy)	2.7 (0.56)		1.9 (0.99)		***
Very happy		75%		33%	
Quite happy		21%		36%	
Not very happy		4%		19%	
Not at all happy		0%		22%	
How does your husband feel about you working in this factory? (1=Very unhappy to 5=Very happy)	3.38 (0.93)		2.4 (1.29)		***
Very happy		60%		21%	
Happy		25%		33%	
Fair		10%		18%	
Unhappy		3%		17%	
Very unhappy		2%		11%	
Number of respondents	850		663		

[1] Wage ranges created ex-post. The number of realized observations is less than the full sample of 850 because some (179) decided not to start the job and there are a few missing values on some of the variables. Standard deviations in parenthesis. Significance of difference in means indicated by asterisks: *** significant at the .01 level; ** significant at the .05 level; * significant at the .1 level.

only 9% had learned about the job from such sources and instead, most (55%) learned about the job from an official source. To get a better sense of the mismatched wage expectations, Figure 3 displays the density plots of expected and realized monthly wage at the factory. The

applicants' expectation about the wage from the factory ranged from ETB 500 per month to ETB 5,000 with the mean at 1,348 (vertical solid line), while the realized wages ranged from ETB 310 per month to ETB 3500 with the mean at ETB 964 (vertical dashed line).

Figure 3: Density plots of expected and realized monthly wage



Before starting the job, the factory employment appears to have been the preferable option at baseline for the women receiving a job offer. On a scale from 1 to 10 where 1 being the worst possible job for their qualification and 10 being the best, on average the women placed the factory job at 8.54, and 51% ranged it at 10 (fewer than 3% ranged it at less than 5). When asked how happy they thought they will be in the factory job, 96% answered either happy or very happy. Similarly, 95% stated that their husband was very happy, happy or OK (fair) with her employment at the factory. Their most referred to second best alternative to the factory job was to work in a commercial business, running their own business, selling clothes or food on the street, or having no job at all.

When we interviewed the women again at the endline, the rank of whether the factory job was their best or worst option had not changed notably. However, their positive expectations turned out to have been too optimistic for most. While 75% had expected to be very happy in the factory job, only 33% reported being very happy with the job at the endline. The ratio of unhappy rose from 4% (expected) to 31% (realized). We see similar trends for how the respondents perceived their husbands' opinions.

3.1.2 Qualitative interviews

The mismatch between expectations and realizations was also discussed in the qualitative interviews. Those who had learned about the job from family or friends working inside the factory said they generally knew about the salary levels, but not the specifics about the job task as that varied according to which production stage they would be allocated to work in (W1, W5, and W7). Women who instead learned about the job from official channels such as their Kebele (neighborhood) office, were not informed about the salary or the type of job

they were getting into: "All I knew was that it was a factory," many told us (W4, W15, W19, W21, W22, and W23).

In Hawassa, which had a fairly recently established industrial park (opened in 2017), there were many rumors among the applicants, which inflated the expectations about salary and working conditions (W18, W22, and W23). For instance, one factory sent the first cohort of workers to Indonesia for training. In the qualitative interviews, some of the respondents stated that this gave the workers and other applicants the impression that the job contracts would be luxurious and generous (W17 and H15). Furthermore, with the opening of the industrial park, the jobs were promoted on the television as new, modern, and life-changing opportunities for young and especially female Ethiopians. The people were led to think the tasks would be easy and that the salaries would be much higher than they were in reality (W24 and H15). The disappointment was large for many (13/24), but when asked why they stayed and not left immediately, some answered that they were expecting salary increases in time, but in most cases, this did not happen or it would not realize until after one year (W10, W14, W19, and H15).

When asked to describe the work environment, the women were almost unanimous, "The work in the factory is very tiresome. I have to stand/walk/sit all day. The machines are very loud and it is too hot there." Table 5 summarizes the answers (codes) that emerged under each theme concerning the work environment and their decision about working in the factory. Cutters and assistants have to stand or walk all day and lift and carry heavy loads. These women struggle with leg swellings and back pain. Sewing machine operators sit all day and have to work fast to reach strict daily quotas. Some firms offer group-based incentives for certain production goals, but the workers we interviewed said the goals were almost impossible and that they seldom manage to reach the goals (W15, W20, and W23). The machine operators experience high levels of stress and many (10/24) complained about back and shoulder pain, hurting eyes and headaches. Workers are not allowed to drink water while working and many suffer different health issues from lack of water and food (W16, W19, W20, and W24). Shifts in low and high temperatures, dust particles, and smoke were mentioned as reasons for respiratory issues, allergies, rashes, eye infections, and even miscarriages (W11, W14, and W19).

Most (18/24) worked six days a week. In general, the lunch break was 30 minutes, but there were instances where the break was shorter. In Hawassa, the lunch was generally provided, but not in Dire Dawa. Although overtime happened only in certain periods or not at all, some factories had shift work, which was considered problematic for many. For instance, several (7/24) told us that it was dangerous for them to walk outside at night and that many female workers resigned from the job because of this. We were also told by some about robbers that would know the workers' pay-day and who would wait in the dark for them to get off the factory transportation bus and take their money (W21, W22, and H22).

Only a few women (5/24) told us the supervisors or leaders treated them nicely, while most (18/24) told us that the supervisors often shouted, insulted, and sometimes used physical force on the workers to make them work faster or to punish them if they made a mistake. The supervisors often shouted things like, "Why are you like this, don't you want to work

Table 5: Work environment and job satisfaction (focus sample)

Themes (Questions)	Answers (codes)	Count
Work environment	Have to stand/walk/sit all day	10/24
	Loud noise	17/24
	Too hot	15/24
Treatment by supervisors	Good	5/24
	Bad, yell/shout/insult/threaten	18/24
	Bad, use cuts in salary to discipline	3/24
Opportunity for advancement	Possible	15/24
	Not possible	8/24
	Have experienced advancement	7/24
All in all, was/is it a good decision to work in the factory?	Likes the job	14/24
	More economic freedom	9/24
	More economic security	2/24
	More social/good for the mind	4/24
	Does not like the job	7/24
Husband's opinion about the impact of her job on the family	Good because of the extra income	9/11
	Bad for the relationship	4/11
	Bad for the children	4/8
	Bad because she does not have time for household chores	3/11
Husband's opinion about how the factory job compares to her other alternatives	The factory has better working conditions	2/11
	The factory has better salary	1/11
	The factory is better for socialization	1/11
	The factory is worse than her alternative	5/11

The respondents' answers could fit into several codes, thus the sum of answers exceeds the number of respondents.

here? I can easily replace you. You need to work better or I will fire you" (W17, W20, W22, and W24). The respondents told us that any attempts to talk or fight back would result in the worker losing their job (W2 and W6). Cuts in the salary were also frequently used as a disciplinary method for arriving late, taking a break or making mistakes (W2, W6, W7, W13, W14, and W24). Yet, the supervisors' behavior was excused by some respondents: "They behave like that because it is their job and they hope to get promoted if they do it well," two women told us (W20 and W24).

There were also few opportunities for advancement in their positions. In most factories there was only one advancement opportunity: from transporting fabric to the cotton machines, to working with the machines (this could happen through experience), or from being a machine operator (sewing) to being promoted to a trainer or line supervisor (this could happen through experience conditional on having completed 12 years of schooling). Out of the twenty-four women we interviewed, six had once advanced in position.

On the positive side, the factories in general offered stable working hours and long term contracts with benefits such as pension, maternity leave, compensation for sick days, and annual leave, as opposed to most of the alternative income activities relevant for the women in this sample. When asked the question, "All in all, was/is it a good decision to work in the factory?" most (14/24) answered that they were happy with the job as they needed the income or because they at least got out of the house. The ones who were not happy said that the jobs were too hard, they were treated badly, and the salary was too low. They had had enough of factory work and said they would not go back even if the conditions changed for the better.

While some of the husbands we interviewed told us that he helped with the household chores or childcare when their wife was at work (4/11), for most women, the factory work simply added to their domestic responsibilities. The husbands were asked about the impact of their wives' factory employment on their relationship and the children (if relevant). While most (9/11) said that it was good that she worked there, as she contributed with much needed extra income, a few (4/11) also showed discontent over her lack of time to socialize with him or to make coffee and meals. In such instances, the husbands would buy food or coffee from outside, which increased their expenses. Some (4/8) were also concerned about the children because other caretakers, whether family, neighbors or hired help, did not take well-enough care of the children. However, most (9/11) of the husbands still wanted their wives to work, either because they needed money or because they thought it was good for them to get out of the house and talk with other people. Although, factory work was not considered as the best alternative (6/11).

3.1.3 Summary

The workers' experience of the job is considered in light of their expectations before starting the job. Many applicants had too optimistic expectations regarding salary level and work efforts. In fact, 62% expected more than their realized salary. The mismatch between expected and realized salary can be explained by inexperience, lack of information, and rumors. Less than one third had ever had a job before and only six percent had ever worked in a factory. The manufacturing industry at scale is a new type of occupation opportunity in most parts of Ethiopia and in the qualitative interviews many respondents said that they did not know what to expect. In the qualitative interviews it also became clear that official sources for job vacancies, such as unemployment offices, newspapers or posters did not give sufficient (if any) information about the salary and work tasks. Hence, many applicants showed up to the job without knowing what they were getting into. The qualitative interviews also revealed that rumors based on media attention had inflated expectations among people (in Hawassa) about wages and working conditions. The women in the focus sample describe heavy workloads under stressful and unhealthy conditions, with rigid work schedules and few opportunities for advancement. Despite these descriptions most still considered themselves as happy with the factory job (83% of the stayers and 44% of the leavers). Based on the qualitative interviews, this seemingly contradictory finding is most likely explained

by the fact that their alternative to factory employment is in many cases worse. The most common alternative to the factory is to stay at home without any income or erratic work without security or with worse working conditions. As one husband said, "The job at the factory is very tiresome and in addition she has to work at home. But her income is very important for us. If it was not that important for us I wouldn't have let her work there" (H9).

Most entered factory employment because they needed money. Although Ethiopia is a patriarchal society, most of the women in the qualitative interviews stated that they had made the decision to enter factory employment themselves without discussing it first with their husband. Long and inflexible work schedules and heavy workload made some of the husbands negative towards the factories, claiming that, besides the extra income, the factory employment was not good for the family as it took too much time away from childcare, their relationship, and cooking. While couples without, or with older, children did not find the work schedule problematic. In general the couples in the focus sample said that they often discuss income needs and possible jobs, but that the wives themselves are free to choose the type of job they would enter or leave.

3.2 Turnover

3.2.1 The full sample

By the time of the endline survey, seven months after baseline, half of the respondents who were offered the job did not work there. While 21% of those who were offered the job did not even start, among those who did start the job, the average time of working in the factory was three months. In this section, I explore correlates between the workers' baseline characteristics and their decision to stay in the job. Table 6 reports the result from a logistic regression where the dependent variable is an indicator for being employed in the factory at endline. In column (1), the regression includes all respondents, while column (2), only includes the respondents who actually started the job. The questions which make up the group variables are described in Appendix C.

The first set of covariates in Table 6 focuses on the women's background variables and z-scores for physical health, cognitive abilities and gender equality attitudes. Women with 10 years of schooling were more likely to stay in the job than women with fewer years of schooling. This may be because women with higher education were better paid due to the type of position or type of factory they were employed in. This point is illustrated in Figure 4. The figure shows first that the wage level is correlated with years of schooling, and second, that the stayers had higher salaries than the leavers in all education levels. Thus the women with lower earnings potential in the factory, either across or within their education levels, were more likely to leave (or not even start). In the table, we also see a weak correlation between women with higher income in the period prior to the factory employment and their retention, this could be because they had more work experience and thus knew better what kind of job they were applying for. However this variable is not significant when we drop the ones who never started (column (2)), nor is it robust to alternative specifications and

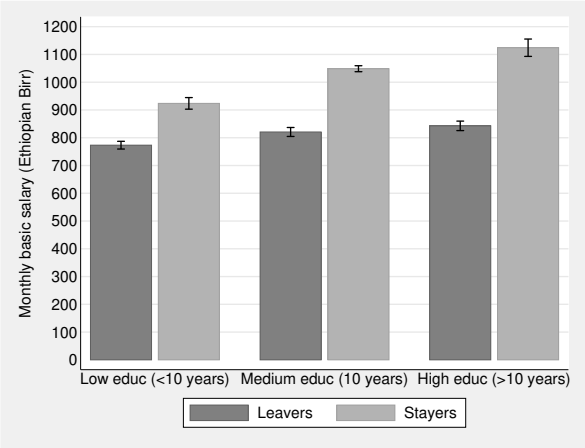
Table 6: Logit regression for choosing to stay in the factory employment

	Stayers	
	All who were offered a job (1)	All who started the job (2)
Panel A. Individual background variables		
Age	-0.009	-0.007
	(0.019)	(0.020)
Medium education (10 years completed)	0.881***	0.815***
	(0.217)	(0.270)
High education (>10 years completed)	0.368	0.089
	(0.238)	(0.287)
Respondent's total income last 6 months (log)	0.035*	0.026
	(0.020)	(0.023)
Physical health (4 questions), z-score	-0.091	0.042
	(0.071)	(0.079)
Cognitive ability (5 questions), z-score	-0.090	-0.010
	(0.078)	(0.090)
Gender attitude (11 questions), z-score	0.021	-0.050
	(0.080)	(0.095)
Panel B. Household characteristics		
Respondent is household head	-1.086**	-1.377**
	(0.496)	(0.551)
Number of children living in the household (under the age of 10)	-0.181**	-0.186*
	(0.089)	(0.104)
Number of household member older than the age of 10	-0.167	0-.378**
	(0.135)	(0.154)
Husband's total income last 6 months (log)	-0.117**	-0.065
	(0.052)	(0.055)
Wife's household decision-making power (15 questions), z-score	-0.144*	-0.025
	(0.080)	(0.091)
Panel C. Expectations		
Difference in expected and realized monthly wage in factory (log)[1]	-0.176**	-0.171*
	(0.069)	(0.089)
Learned about the job from a social contact inside the factory	-0.050	-0.349
	(0.228)	(0.252)
Factory job is the worst-best job for my qualifications (1-10)	0.004	-0.028
	(0.045)	(0.083)
Taking all things together, will you be happy with this job? z-score	-0.087	-0.434
	(0.335)	(0.371)
Is your husband/partner happy about you working in the factory? z-score	0.043	-0.106
	(0.049)	(0.108)
Observations	841	671
Average of dependent variable	0.5	0.63

[1] The difference in expected and realized monthly wage is constructed by the difference in the logs of the stated wages. In cases where the respondent did not start the factory job (n=179), I do not have a measure for the realized wage, the missing value is then replace by the mean log wage reported by the other workers in the factory which the respondent had applied to. In a few cases (n=9), no applicant in the sample started the job at the factory, thus these nine observations are dropped from the regression.
The *Stayers* are the respondents who were still employed in the factory at endline. Time between surveys is added as a control variable. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

should be interpreted with caution.

Figure 4: Monthly basic salary in factory by education level for stayers and leavers



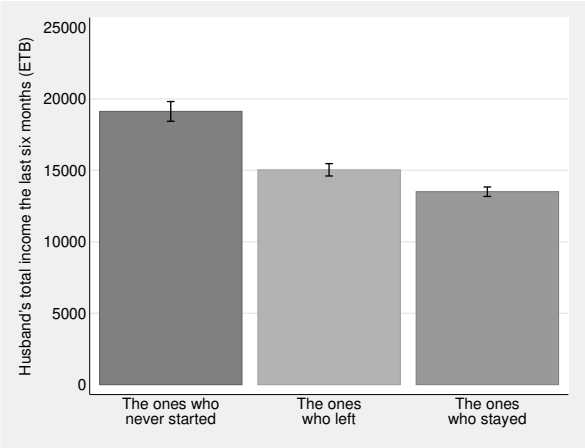
The second set of covariates focuses on the respondents' household characteristics. The regression shows that women who were household heads or had more children, were less likely to stay in the job, possibly due to their larger responsibilities for childcare and running the household than women in smaller households. Among those who started the job (column (2)), the ones who stayed in the job had more household members over the age of 10, which follows the argument that many women who stayed in the job had less domestic responsibilities. Furthermore, women with husbands who earned better were less likely to be working in the factory at endline. This variable is only significant when including the ones how never started, which suggests that women from richer households were more likely to turn down the offer after learning about the wage level, while women from poorer households were the ones who accepted the job offer. Figure 5 illustrates the statement and shows that the husband's of women who never started (first bar) earned on average 27% more than the husband's of the ones who started but who later left (second bar), and 42% more than the husband's of the stayers (third bar). Last, we also see a weak negative correlation between household decision-making power and retention. This correlation, however, is not robust to alternative specifications and I therefore do not explore this variable further.

Panel C of the regression explores applicants' expectations and whether the ones who are better informed remain in the factory job for longer. The first variable is an indication of whether their wage expectations were realistic or not, by taking the difference in the (log) expected wage with the (log) realized wage measured at endline. The regression shows that higher unrealistic wage expectations decreases the likelihood of retaining the job.

In sum, the regressions show that women with higher earnings potential within the factory, who came from poorer households and with fewer (young) children, and who had realistic wage expectations were the ones most likely to remain in the factory job for longer.

The correlates can be complemented with the respondents' own views on reasons for why they left the factory, which are listed in Panel (a) in Figure 6. The most common reason

Figure 5: Husbands' income by wives' factory employment decisions



was the low salary (36%).

Due to the work, there was also a high occurrence of health-related issues among the factory workers (21%), and it is the second most reported reason for leaving the job (16%). The health issues experienced by workers are reported in Panel (b) in Figure 6 and the most common issues were kidney problems[12] (49.5%), leg and back pains, head aches and hunger (28.5%), and respiratory issues (9.5%).[13]

Long and inconvenient working hours (15.5%), such as night shifts or weekends, and an unattractive working environment were also frequently stated reasons for resigning (11%).

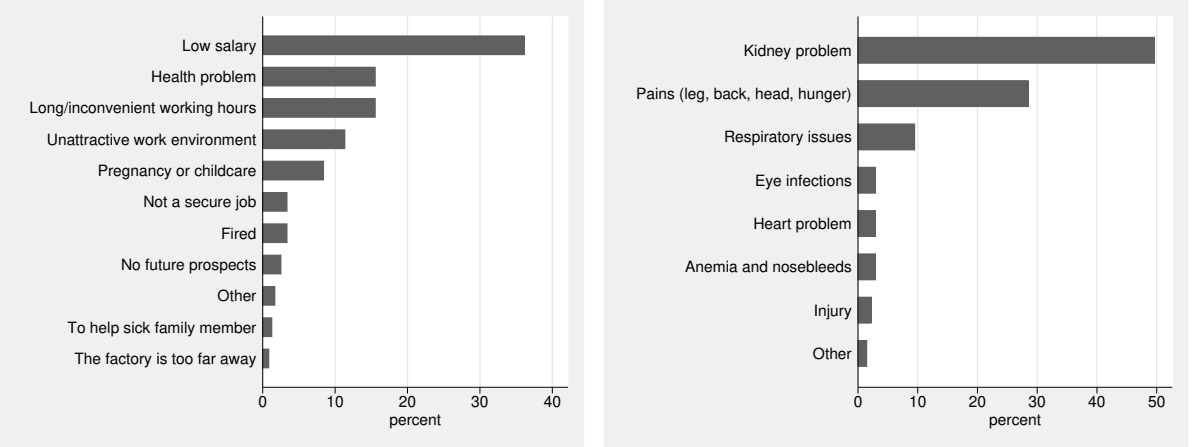
Some also left the job because of pregnancy or the challenges of combining the factory employment with childcare (8%). Among the women in the sample who were still employed by the endline survey, their children in most cases were under the care of other family members (67.5%), a neighbor (11%), or alone (15%) while they were at work. 10.6% had a maid or other paid caretaker and the monthly cost for a maid or paid caretaker was on average ETB 300 or 25% of their own salary from the factory. Thus, without the help of relatives or neighbors, there were no or few opportunities for childcare services as these would be too expensive.

No one mentioned another job offer as a reason for leaving the factory. In fact, 61% of the women who decided not to start or who resigned from the factory job did not enter into any alternative formal or informal job, while 12% did get wage employment, 11% had some daily labor employment or training, and 16% earned income through self-employment.

[12]The frequent occurrence of kidney problems is due to inadequate access to bathroom facilities, dehydration, and poor hygiene. This has also been noted in other studies, for instance, Hailu et al. (2018) report in their study, that 33% of workers in Ethiopian manufacturing firms report a lack of adequate bathroom facilities. For instance, bathroom breaks would be time-restricted and closely monitored, often by a male supervisor, or the factory had too few working toilets.

[13]High incidents of health issues have been described in other studies as well, for instance, Blattman and Dercon (2018) report that for every month of industrial work, workers were 1.1 percentage points more likely to report disability of any sorts.

Figure 6: Reason for leaving and health issues experienced

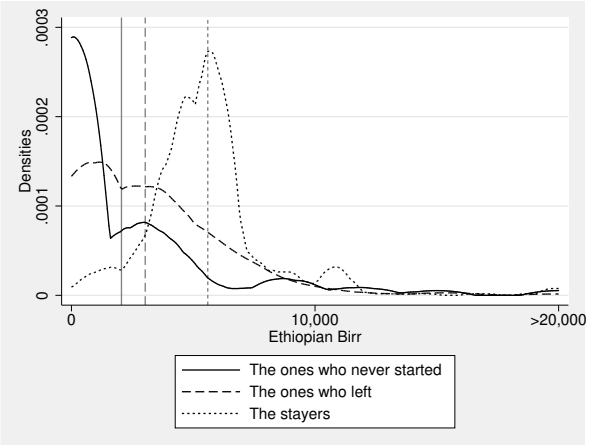


(a) Reason for leaving (n=238)

(b) Health issues experienced (n=137)

The average income from these sources was ETB 4,636 (st.dev = 5,282) over a six-months period. While in the same period, the stayers earned on average ETB 5,661 (st.dev = 3,957) from working in the factories. Figure 7 shows the density plots of the total income from wage- and self-employment in the last six months, measured at endline, by three groups: the ones who were offered the job, but never started, the ones who started in the factory but who left, and the ones who were still employed in the factory by the endline survey. While we see that there are a few outliers, most had either no (or close to zero) income in the six months prior to the endline survey, or they earned somewhere between ETB 3,000-7,000 (or on average ETB 500-1,166 per month) from different income-generating activities.

Figure 7: Distributions of total incomes from wage- and self-employment in the six months prior to the endline survey by factory employment decision



3.2.2 Qualitative interviews

In the qualitative interviews, we asked both the women and their husbands about the role of the husband in the decision to leave the factory. In many cases (6/11), the husbands expressed concern that the factory job was too hard and that his wife often came home from work exhausted. However, in no case did the husband demand that his wife leave the job. In the cases in which he suggested that she should leave the job, the wife was either happy to leave or answered that she would continue because they needed the money, and then she continued. About half of the husbands (5/11) had mixed feelings regarding the factory employment; on the one hand, it was not a good job and he would prefer to have her at home taking care of the children or running her own business, on the other hand, the extra income was important for them. While we only conducted 24 interviews, in all cases the wife expressed that she decided herself which job to take and whether she should leave the job or not. In some couples (4/11), the husband said that the decision was hers completely, they did not even discuss it, while other couples decided together, but it was clear that the wife had the final say in decisions regarding her labor options.

3.2.3 Summary

In section 3.1, we discussed the workers' mismatched expectations and experiences of the job. These descriptions are closely related to the workers' decisions to leave the job, which are presented in this section. Half of the women who were offered jobs in the factories left within seven months. In general, the low salary combined with unhealthy work conditions made the jobs "not worth it" for many of the respondents. Instead they preferred to use their time to take care of the household and family or look for other jobs. Few respondents started any other wage employment or self-employment after leaving the factory. Based on the qualitative interviews, this is likely not because they do not want another job, but that there are very few options. Those who had higher earnings potential in the factory or who came from poorer households were more likely to remain in the job, since the job had higher welfare impact in these households.

Through interviews with the women and their husbands, the women seem to have strong decision-making power over their own labor supply. The spouses support and help each other as much as possible to make enough money for daily expenses and to bring up their children. Hence, while combining factory employment with childcare and other household responsibilities is sometimes challenging, the couples try to make the optimal decisions together as a team. And while the husbands' were generally more sceptical about the factories due to the hard working conditions, it was the women themselves deciding how long they would remain in employment.

4 Implications and conclusion

The paper describes the household labor supply decisions and experiences of female workers in the manufacturing industry of Ethiopia using a combination of quantitative and qualitative data. The manufacturing industry has a large potential for welfare increasing impacts among poor households, since it has the opportunity to increase women's labor market participation by large numbers. For instance, in the this study, only 29% had ever had a job before entering into factory employment, and in the focus sample, six out of the twenty four women (25%) said that if there were no factory jobs available, they would just stay at home. The description of whether manufacturing jobs are "good" for the workers is somewhat mixed; on the one hand, most of the current and former workers in this study describe themselves as happy or very happy with the factory job (83% of the stayers and 44% of the leavers). In general, the factories also offer stable working hours and long term contracts with benefits such as pension, maternity leave, compensation for sick days, and annual leave, as opposed to most of the alternative income activities relevant for the women in this sample. On the other hand, complaints about low wages, hard and unhealthy working conditions, few opportunities for advancement and inflexibly were common and, as a result, employee turnover is high, which may pose major challenges for the profitability and further expansion of the industry.

Most applicants were very optimistic about the job opportunity before starting, but few had real information about the type of work, the working conditions or the wage level, especially in areas with new industrialization. The work experience was therefore a large disappointment for many. As a result, turnover rates can probably be lowered if the factories signal better the true conditions to applicants, thus avoiding hiring workers with too high expectations who will leave after a short time. For instance, by disclosing up front or in the job announcements the wage structure and possibilities for advancement for entry level workers or introducing expectation management for newly employed workers.

It is evident that unhealthy working conditions is an important reason for the high turnover rates. To handle some of these issues, regulations concerning healthy working conditions should be considered. For instance, due to the high frequency of kidney diseases, adequate bathroom facilities and decent bathroom breaks may help. While manual work is hard, overload and long-term disabilities may be prevented, for instance, by varying the workers' tasks or introducing short breaks. Moreover, improvements in leadership style and work facilities could be implemented to reduce stress and improve health among the workers.

Finally, the women in our sample seem to have strong decision-making power regarding their labor supply decisions. This observation is important because of the patriarchal context of Ethiopia where men are traditionally the breadwinners. While the husbands interviewed for this study expressed that they preferred that their wives could stay at home to take care of the household and the children or to have a less demanding job, they also said that the extra income was important for the household and that they appreciated their wives' efforts. They further said that the type of job the wives would take was up to them, as long as they were healthy. While the jobs offered significant improvements to their income, many women still opted to leave and live without the income. According to respondents we

interviewed, the factories' inflexible work schedules and heavy workload on the one hand, and women's central role in childcare and the few alternatives for childcare services on the other, factory employment turned out to be difficult to combine with the upbringing of children. As marriage and childbearing are common among young women in Ethiopia, the industry may consider accommodating the needs of mothers better in order to attract and retain more female workers.

The findings in this paper are based on qualitative interviews with former and present workers and a larger survey of manufacturing employees. A word of caution is in place as the analysis is of a descriptive nature and does not offer causal interpretations of the welfare impacts of the jobs. Descriptive studies may nevertheless be helpful to better understand the situation and the labor choices coming from the workers' own testimonies in the particular context. Future research in this area should aim to causally measure the effects of possible interventions for reducing turnover, for instance the provision of information to applicants or childcare options. The findings in this paper are specific to the Ethiopian labor market around the industrial parks, and may not be generalizable to other countries. However the industrial development experiences in Ethiopia have been similar to other countries', especially other SSAs (Lawrence, 2005), hence lessons may be learned and also used as a reference in other similar contexts.

References

Abebe, G. T., Caria, S., Fafchamps, M., Falco, P., Franklin, S. and Quinn, S. (2018), 'Anonymity or distance? Job search and labour market exclusion in a growing African city', CEPR Discussion Paper No. DP13136.

Anderson, S. and Eswaran, M. (2009), 'What determines female autonomy? Evidence from Bangladesh', *Journal of Development Economics* **90**(2), 179–191.

Barrett, P. M. and Baumann-Pauly, D. (2019), 'Made in Ethiopia: Challenges in the garment industry's new frontier', Center for Business and Human Rights, NYU-Stern.

Becker, G. S. (1985), 'Human capital, effort, and the sexual division of labor', *Journal of Labor Economics* **3**(1, Part 2), 33–58.

Becker, G. S. (1991), *A Treatise on the Family*, Harvard University Press.

Blattman, C. and Dercon, S. (2018), 'The impacts of industrial and entrepreneurial work on income and health: Experimental evidence from Ethiopia', *American Economic Journal: Applied Economics* **10**(3), 1–38.

Browning, M., Chiappori, P. A. and Weiss, Y. (2011), *Family Economics*, Cambridge: Cambridge University Press.

Central Statistical Agency/CSA/Ethiopia and ICF (2016), 'Ethiopia Demographic and Health Survey 2016 (DHS 2016)', Addis Ababa, Ethiopia, and Rockville, Maryland, USA: CSA and ICF.

Dharmalingam, A. and Philip Morgan, S. (1996), 'Women's work, autonomy, and birth control: Evidence from two South Indian villages', *Population Studies* **50**(2), 187–201.

Djamba, Y. K., Goldstein, S. and Goldstein, A. (2006), 'Gender differences in occupational mobility in Ethiopia: The effects of migration and economic and political change', *Genus* **62**(2), 65–88.

Dollar, D. and Gatti, R. (1999), 'Gender inequality, income, and growth: Are good times good for women?', Policy research report on gender and development working paper series; no. 1. Washington, D.C. : The World Bank.

Duflo, E. (2012), 'Women empowerment and economic development', *Journal of Economic Literature* **50**(4), 1051–79.

EIC, Ethiopian Investment Commission (2018), 'The 2017 Ethiopian investment report', Series1/18, Addis Ababa.

Franklin, S. (2017), 'Location, search costs and youth unemployment: Experimental evidence from transport subsidies', *The Economic Journal* **128**(614), 2353–2379.

Guest, G., Bunce, A. and Johnson, L. (2006), 'How many interviews are enough? An experiment with data saturation and variability', *Field Methods* **18**(1), 59–82.

Guest, G., MacQueen, K. M. and Namey, E. E. (2011), *Applied Thematic Analysis*, Sage Publications.

Hailu, Y., Tilaye, A., Worku, I., Birhanu, M. and Daniel, S. (2018), 'A Study on Women in Manufacturing in Ethiopia: Opportunities, challenges & strategic interventions', Ministry of Trade and Industry and UNDP Ethiopia.

Halvorsen, S. K., Kotsadam, A. and Villanger, E. (2020), 'Factory employment and fertility decisions: Experimental evidence from Ethiopia', *Unpublished manuscript* .

Hardy, V. and Hauge, J. (2019), 'Labour challenges in Ethiopia's textile and leather industries: No voice, no loyalty, no exit?', *African Affairs* **118**(473), 1–25.

Heath, R. and Mobarak, A. M. (2015), 'Manufacturing growth and the lives of Bangladeshi women', *Journal of Development Economics* **115**, 1–15.

IFC, International Finance Corporation (2013), 'Investing in women's employment: Good for business, good for development', Washington, DC.

Jensen, R. (2012), 'Do labor market opportunities affect young women's work and family decisions? Experimental evidence from India', *The Quarterly Journal of Economics* **127**(2), 753–792.

Kabeer, N. (2002), *The Power to Choose: Bangladeshi Women and Labor Market Decisions in London and Dhaka*, Verso.

Kotsadam, A. and Villanger, E. (2020), 'Jobs and intimate partner violence - Evidence from a field experiment in Ethiopia', *Unpublished manuscript* .

Lawrence, P. (2005), 'Explaining sub-Saharan Africa's manufacturing performance', *Development and Change* **36**(6), 1121–1141.

Luke, N. and Munshi, K. (2011), 'Women as agents of change: Female income and mobility in India', *Journal of Development Economics* **94**(1), 1–17.

Meyer, C. J. (2018), 'In search of a better life: Self-control in the Ethiopian labor market', Job Market Paper, European University Institute.

Moon, K., Bergemann, P., Brown, D., Chen, A., Chu, J., Eisen, E., Fischer, G., Loyalka, P. K., Rho, S. and Cohen, J. (2018), 'Manufacturing productivity with worker turnover', SSRN Working paper No.3248075.

National Planning Commission (2016), 'Growth and Transformation Plan II (GTP II), (2015/16-2019/20)', Vol. I, (main text), National Planning Commission of the Federal Democratic Republic of Ethiopia, Addis Ababa.

Oqubay, A. (2018), The structure and performance of the Ethiopian manufacturing sector, *in* F. Cheru, C. Cramer and A. Oqubay, eds, 'The Oxford Handbook of the Ethiopian Economy', Oxford University Press.

Qian, N. (2008), 'Missing women and the price of tea in China: The effect of sex-specific earnings on sex imbalance', *The Quarterly Journal of Economics* **123**(3), 1251–1285.

Rosenzweig, M. R. and Schultz, T. P. (1982), 'Market opportunities, genetic endowments, and intrafamily resource distribution: Child survival in rural India', *The American Economic Review* **72**(4), 803–815.

Thomas, D. (1993), 'The distribution of income and expenditure within the household', *Annales d'Economie et de Statistique* **29**, 109–135.

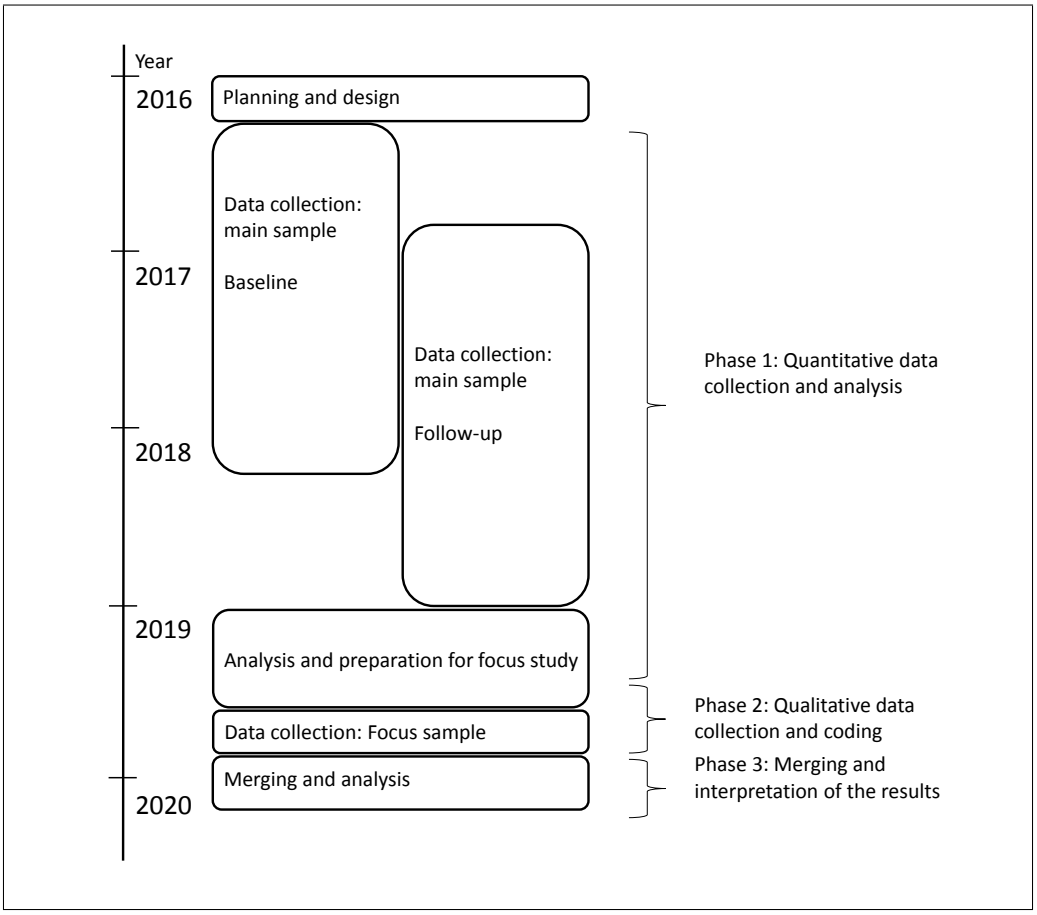
World Bank (2011), 'World Development Report 2012: Gender Equality and Development', World Development Report. Washington, DC: World Bank Group.

World Bank (2012), 'World Development Report 2013: Jobs', Washington, DC: World Bank Group.

Yost, M. and Shields, L. (2017), 'Ethiopia's emerging apparel industry: Options for better business and women's empowerment in a frontier market', BSR, Paris.

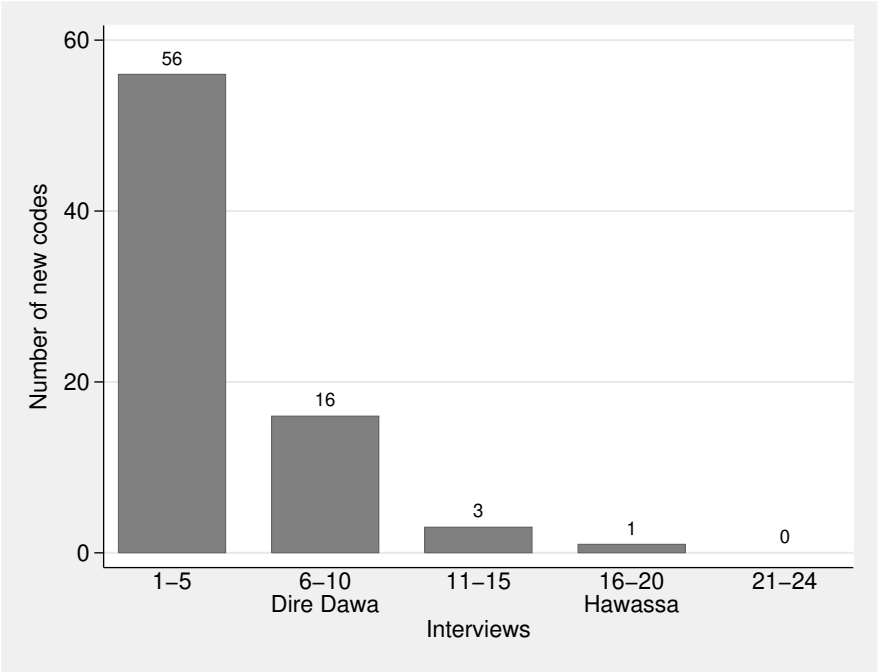
A Timeline of the project

Figure A1: Timeline of the project phases



B Saturation

Figure B1: Number of new codes over the course of data analysis



C Attitude and ability survey questions

(Main sample only)

Gender attitude

I will read some statements about men and women. Please say whether you strongly disagree, disagree, agree or strongly agree with these statements.

- It is better to send a son to school than it is to send a daughter.
- It is okay for women to work outside of the home.
- It is okay for women to earn more money than men.
- Women have a right to decide what to do with the money they earn.
- It is okay for women to travel or to leave the house for several nights to do business.
- Men should be responsible to help with childcare when his wife is busy with business or factory job.
- Men should be responsible to help with domestic duties when his wife is busy with business or factory job.
- The important decisions of the family should be made by the men of the family only.
- A wife should tolerate being beaten by her husband/partner to keep the family together.
- Woman should seek help if she encounters sexual harassment.
- Woman should seek legal recourse if she encounters sexual harassment.

Household decision-making index

Who in your household usually has the final say about the following decisions? (Coded as 1 if the respondent is the main decision-maker or has a lot of input into the decision, coded 0 otherwise).

- Whether to send or not send children to school.
- What to do if a child falls sick.
- What to do if the respondent falls sick.
- Whether to have children or to have more children.
- Which family planning methods to use.
- Whether or not you should earn money outside the house.
- Whether you can visit your family or relatives.
- The use of the wife's earned income.
- The use of the man's /husband's earned income.
- Purchase of small daily food purchases.
- Purchase of bulk or expensive food items.
- Large purchases of items like furniture, cattle, TV, or other assets.
- Purchase of children's clothing and shoes.
- Whether to open bank account or borrow money.
- Whether to start a new business.

Physical health

Responses are on a scale of one to four, with one indicating "unable", two is "very difficult", three is "Slightly difficult," and four is "easy."

- Are you able to walk for 2 kilometers? (pick a Landmark that is 2km distance from the interview location)
- Are you able to carry a 20-liter container of water for 20 meters?
- Are you able to carry out your usual daily activities by yourself?
- Are you be able to stand at a workbench or assembly line for 6 to 8 hours?

Cognitive ability

- If you buy goods for 600 Birr and then sell them for 650 Birr, what is your profit?
- If you earn 100 Birr a day for 30 days what is your total monthly income?
- Selam bought a 1000 birr dress at a 20% discount. How much did she pay for the dress?
- What is 5 times 10, divided by 2?
- Suppose you have 5000 birr in a savings account. The account earns 4 percent interest per year. How much would you have in the account at the end of the year?"

D Interview guides for the qualitative data collection (Semi-structured interviews)

Women

About the decision to enter factory work and the expectations

- Maybe we could start with you telling us about how you got here, I mean how did you grow up, when did you marry your husband, what is your family- and household situation?
- Can you tell me more about your occupation history the past five years until now, in a chronological order? For instance what have you been working with, how long did you work there, what were your main tasks, what did you earn per month and how did you like it.
- Do you remember when you applied for the factory job? Can you tell me a little bit about that?
- Can you tell me more specifically about the events and circumstances in your life that led you to apply for this job?
- What was the role of your husband or others in the decision to apply for the job?
- Can you tell me more about how the decision was made?
- Before starting in the factory job, back when you had never worked in a factory before, what did you know about the work that you were going to do in this factory? Where did you have this knowledge from?
- If there were no jobs available in these factories, what would you have done at that time?
- In what way, did you think before you started the job, that the factory job would be better or worse than your alternatives at that time?
- How long did you plan to stay in the job?
- What were your expectations about the wage and the work that you were going to do in the factory? What were these beliefs based on?
- How did these beliefs match with what you learned about the job on the recruitment day or your experiences working in the job? (Remember to follow-up all the beliefs from last questions)

The experience from factory work and/or other jobs in the past years

- Tell me about how a normal work day for you looks like working in the factory, from you get up in the morning to you go to bed in the night.
- How is the payment structure at the factory?
- How are you treated by managers? Is there any physical or verbal abuse?
- Have you had any opportunity to advance in position? Have you seen other workers in a similar position as you advance?
- Have you had any health issues coming from the job? What caused that?

- What happens if you have/had to stay home from work if you are sick?
- What happens to women who work in the factory and become pregnant?
- Why did you choose to quit the factory job?
- Would you have been willing to continue in the job if the factory management improved things for you? What would these things be that they could improve so that you had continued in the job?
- What was the role of your husband or others in the decision to quit the job?
- Can you tell me how the decision was made?
- What did you do instead after you quit?
- How do you think this occupation compares to the factory job?
- How does/did your work at the factory affect your family and your family's welfare compared to how it was before you started working there?
- What are your aspirations for the future, for yourself and for your children? How realistic do you think that is?
- All in all, are you happy to be working at the factory? Why do think that?
- Is there anything more you would like to tell me or add to this conversation?

Men

About the decision to enter factory work and the expectations

- Maybe you could start with you telling me about yourself, I mean how did you grow up, when did you marry your wife, and what is your family- and household situation now. What is your occupation?
- A few years ago your wife applied for a job in a factory. Do you remember that? Can you tell me what you remember from that?
- Can you tell me more specifically about the events and circumstances in your and your wife's life that led her to apply for this job?
- What were your thoughts and feelings regarding this decision to seek employment in the factory?
- Can you tell me more about how this decision was made?
- Think back before she started the job in the factory. What did you know about the job? Where did you have this knowledge from?
- What were your expectations? Probe salary, tasks,working conditions. What were these beliefs based on?

The experience from factory work and/or other jobs in the past years

- In what ways do you think your wife working in the factory affected you?
- How did it affect your relationship? Let him answer first, then probe, was she often tired, did she have time to spend with you? Did you like her working there?
- In what ways do you think your wife working in the factory affected the children?
- How did her working there affected the amount and division of household work?

- How was her salary normally spent? How important do you think it is for your household that she has regular income?
- What would you tell me if asked what kind of job alternatives does your wife have?
- How do you think working in the factory is different to her other occupations the last five years?
- Can you tell me about the decision to quit the factory job?
- How do you and your wife normally discuss your and her different job decisions?
- All in all, are you happy that your wife is working at the factory? Why do think that?
- Is there anything more you would like to tell me or add to this conversation?