Pre-analysis plan for "Jobs and Fertility in Ethiopia"

Andreas Kotsadam[1], Janneke Pieters[2], Pauline Rossi[3], Espen Villanger[4], Céline Zipfel[5], and Tigabu Getahun[6]

[1]Ragnar Frisch Centre for Economic Research
[2]Wageningen University
[3]Ecole Polytechnique-CREST
[4]Chr. Michelsen Institute
[5]Stockholm School of Economics
[6]EconInsight Center for Development Research

August 26, 2024

1 Introduction

We leverage a randomized field experiment in Ethiopia to examine how female wage work opportunities affect the fertility of married women. The experiment, set up in collaboration with 27 large companies, randomly assigned factory worker jobs to equally qualified female applicants. The women who were offered jobs experienced significant increases in income and share of household earnings (Kotsadam and Villanger 2022; Aalen et al. 2023). The baseline survey was collected in 2016 on a sample of 1463 women. Five follow-up surveys were then collected around 6, 12, 18, and 34 months (Waves 2-5) after baseline and 4 years after baseline respectively (Wave 6). We are now planning a seventh survey round in 2024, i.e. 8 years post-baseline.

The long lapse of time since women in the treatment group were randomly offered the jobs will allow us to track the comprehensive impacts of the jobs on women's fertility histories over a period covering a large share of their prime childbearing years. At baseline, women in this sample were on average 25 years old, with two-thirds of them having already given birth at least once.

In the new data collection we will include a detailed fertility module and we will invest heavily in tracking to try to reduce attrition. In particular, we will use previously collected contact information of family members of the interviewee, information from neighbours, and we will complement the in person interviews with phone interviews for the women that are unable to meet us.

This plan will detail our main coding and estimation choices and it is registered before the data collection starts. All deviations from the plan will be highlighted in the paper.

2 Data

Our survey data is rich both in depth and scope. In Table 1 we describe the measurement and definition of some of our main variables.

Table 1: Construction of main variables

Variable	Survey question	Coding
Fertility related variables		
Any child	Have you ever given birth to a living child?	0 if No; 1 if Yes
Number of children	Reported number of live births	
Appropriate age of first birth	What age do you think is a good age for a woman to have her first child?	Reported age
Months to birth	Months between baseline survey and next birth	Calculated months
Wanted number of children	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?*	
Partner wanting a higher number of children	Does your husband/partner want the same number of children that you want, or does he want more or fewer than you want?	1 if more children; 0 otherwise
Other outcome variables		
Social and religious activities	How many hours did you spend on social and religious activities over the last seven days?	Number of hours over the last 7 days
Leisure	How many hours did you spend on social and leisure time (watching TV, reading magazine, playing, exercising, recreation etc.) over the last seven days?	Number of hours over the last 7 days
Employment and income variables		
Any current wage job	During the last six months, how much income did you obtain from factory job employment? How much from other wage employment?	0 if zero income; 1 if positive income
Earnings from wage job last 6 months (in Birr)**	During the last six months, how much income did you obtain from factory job employment? How much from other wage employment?	Total earnings reported
Share of earnings from wage job**	Wives earnings from wage job last 6 months divided by the husbands	Share between 0 and 1
She earns more than him	Wives share of earnings is more than 0.5	0 if No; 1 if Yes
Months factory job	Months since the baseline survey with a factory job	Calculated months
*Baseline variables****		
Any wage job ever	Have you ever had a formal wage job in the past?	0 if No; 1 if Yes
Age	Age	Reported age in years
Religion	Religion	Categories Orthodox Christian, Catholic or Other; Muslim; Protestant
Education	How many years of education have you completed?	Categories Low if $<$10 years; Medium if 10 years; High if $>$10 years of education

Notes: *If she does not have any living children: If you could choose exactly the number of children to have in your whole life, how many would that be?
**We will also have corresponding variables for factory jobs only and earnings and income from any source.
*** In addition we will include baseline values of the fertility outcomes as well as the other outcome variables when they exist.

We show the number of interviews over time by wave in Figure 1. Out of the 1463 randomly assigned women in our baseline sample we managed to interview 1262 in the second wave and 976 in the sixth and final Wave to date. Attrition is not correlated with treatment.

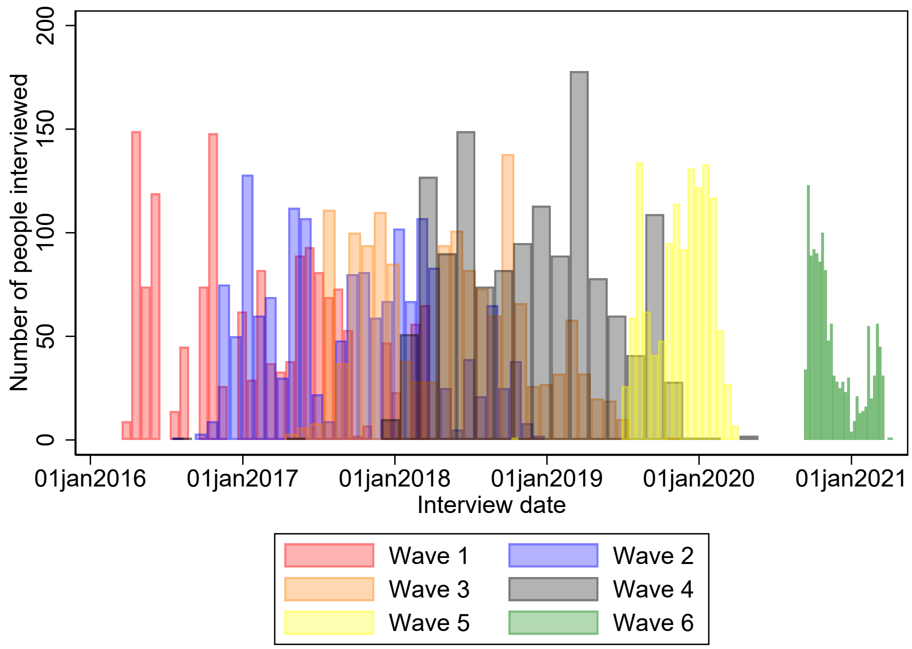

Figure 1: Timing of interviews, by wave

Notes: The bars show the number of people interviewed over different waves and the colours show the different waves.

3 Main empirical specifications and hypotheses

3.1 Cross sectional analyses of the new wave

Our main specification is the following intention to treat (ITT) model:

(1) $Y_{i,t6} = \alpha Y_{i,t0} + \beta Treatment_i + \gamma X_{i,t0} + \delta List_i + \epsilon_{it}$,

where i indexes individuals, $t0$ refers to baseline values, and $t6$ is the sixth follow up (Wave 7). $Treatment_i$ is a dummy variable equal to 1 if the woman was randomized to get the job offer and zero if not. The main coefficient of interest is β, which captures the the total aggregated effect of receiving the randomly assigned job offer on the outcome of interest. We always include list fixed effects (blocking variables) as women are randomized within this unit.

We will present results from estimations without other control variables than the list fixed effects. We will also investigate if we can increase power by adding optimal controls chosen from the total list of controls using a post-double LASSO selection approach (Belloni, Chernozhukov and Hansen, 2014). The LASSO selection approach selects those variables that are correlated with both treatment and the outcomes which may improve precision in the estimates and it also helps to correct for potential imbalances across groups. We use heteroscedasticity robust standard errors in all estimations.

Our first main hypothesis is that job offers affect women's realized fertility. The outcome variable for this hypothesis is *Number of children*. We will also look separately at the extensive margin (i.e. the variable called *Any child*, the probability that women have ever given birth to a living child) and at the intensive margin (i.e. *Number of children*, conditional on having at least one child at baseline). Additionally, we will test whether job offers affect the timing of birth of the next child. We describe how to test this in the next section.

Our second main hypothesis is that job offers affects desired fertility. There are two outcome variables for this hypothesis: *Wanted number of children* and *Partner wanting a higher number of children.*

We will also test for effects of treatment on all other outcomes listed in Table 1 but the results on these outcomes will be viewed as exploratory.

3.2 Duration analysis of fertility timing

To test whether job offers affect the timing of birth of the next child we construct a variable for the *Months to birth*, which is the calculated months between the baseline survey and next birth. The variable will be censored for women who did not have a next birth yet and we use duration models to handle the censoring.

The key element of duration models is the hazard function: the instantaneous probability to have another child at date t. In the specification with proportional hazard, the hazard function is modeled as follows:

$$\lambda(t) = \lambda_0(t) \times \exp(\beta Treatment_i + \gamma X_{i,t0} + \delta List_i)$$

Where $\lambda_0(t)$ is the baseline hazard function, common to all individuals, and the other variables are the same as in Equation (1).

Then the survival function – the probability not to have another child at least until t – is as follows:

$$S(t) = \Pr(T > t) = \exp(-\int_0^t \lambda(u)du)$$

We will estimate two duration models:

(1) a non-parametric model: we will compute the Kaplan-Meier estimator of $S(t)$, by treatment arm. Using the Kaplan-Meier estimate of the survival function and its variance, it is possible to test for the equality of distributions between two sub-populations (log-rank tests). We will report the p-value of the test comparing treated and control groups. The survival function will be estimated first using the whole sample of women and then, by parity (i.e. number of births) at baseline.

(2) a semi-parametric Cox model: using a partial likelihood, the method of estimation makes it possible to estimate the coefficient β without imposing a functional form on $\lambda_0(t)$.[1] The main assumption is that the impact of the treatment on the duration is proportional over time. In our main specification, we will pool all parities and include dummies to control for the parity at baseline in $X_{i,t0}$. We will then run separate regressions for women without a child at baseline (parity = 0) and for women with at least one child.

Our prediction is that (i) $\beta = 0$ when $t < 9$: this is a balancing test, the treatment should have no effect on births in the next 9 months after the baseline because not enough time has passed; (ii) $\beta < 0$ when $t \geq 9$: women in the treatment group tend to delay the next birth compared to women in the control group.

4 Planned exploratory analyses

4.1 Exploratory tests of mechanisms and heterogeneity

The net effects of the jobs on women's completed fertility are *a priori* ambiguous. We plan to collect comprehensive survey data to disentangle between at least five key mechanisms that could plausibly contribute to the overall fertility effects of the randomly assigned job offers.

First, according to early fertility choice models, opportunities that differentially raise women's opportunity cost of time could lower their fertility only if the substitution effect offsets the positive income effect of higher income, assuming that children are normal goods (Becker and Lewis 1973; Jones, Schoonbroodt and Tertilt 2011). Here, if the positive income effect from women's increased earnings dominates, we would expect to see more children being born in the treatment group. On the other hand, if the substitution effect

[1] We will use the exact marginal-likelihood method to handle ties among non-censored durations.

dominates, fertility may fall. We expect this substitution effect to be stronger for women who (i) lack access to organized childcare or family support toward childcare and/or (ii) would be working from home or in small-scale self-employment in absence of these formal jobs. The latter is because wage work is arguably less compatible with child-rearing than self-employment (Goldin 1992; Lim 2019; Delecourt and Fitzpatrick 2021).[2] We also expect a more negative impact of the jobs on fertility if the perceived risk of losing one's job due to a pregnancy is higher.

Second, by providing women access to a stable source of income, the jobs may reduce fertility by dampening the old-age security motive for having large numbers of children. The idea that individuals bear children, in particular, as a source of old-age support in absence of formal pension systems, is a well-established hypothesis (Caldwell 1978). Rossi and Godard (2022) provide empirical support for the latter with their recent study of state pensions in Namibia. We plan to test the relevance of this channel for our sample by collecting information on household savings.

Third, these formal job opportunities for women may affect fertility by altering the (perceived) returns to child labor within impacted households. The effect here may be negative or positive, depending on the direction in which the jobs affect, if at all, the relative returns to children's time. For instance, if there are complementarities between women's and children's labor in the type of informal work that women would typically engage in without the factory jobs – i.e their most likely outside option – then the factory jobs may reduce the perceived returns to child labor, which could in turn raise investments in children's schooling. According to the "quantity-quality trade-off" framework (Becker and Lewis 1973), this would be associated with a negative effect on fertility.[3] On the other hand, if child labor within the household, such as performing household chores, gathering fuel and water, or caring for other family members, serves as a replacement for women's time, the time requirements of new jobs for women could potentially *increase* the need for additional children. We would anticipate this channel to be weaker in households cohabiting with extended family members, as they can assist with these additional tasks. Time use questions also allows us to test this hypothesis.

Fourth, we anticipate that the job offers may affect fertility if women's employment and earnings impact intra-household power dynamics surrounding fertility choices. Recent evidence documenting the existence of significant spousal differences in ideal family size (Doepke and Tertilt 2018; Doepke and Kindermann 2019; Ashraf, Field and Lee 2014; Ashraf et al. 2022) underscores the potential for household bargaining to influence fertility outcomes. The level of influence women have in making decisions within their households may thus be a significant factor in shaping choices related to fertility (Rasul 2008). One test of the salience of this channel will be to explore heterogeneity in the fertility impacts of the jobs by the woman's share of household earnings. Another will be heterogeneity by decision-making power measured with DHS type questions. We will use information collected in earlier survey rounds on the optimal family size of women and their male partners to test whether any reductions in fertility are more concentrated or larger among couples where the spousal gap in fertility preferences was initially larger. Data on male partners' controlling behavior will also allow us to estimate impacts of the treatment on gender attitudes, which would help shed further light on the salience of intra-household power dynamics as a possible channel of impact. We will also collect information on women's marital status, including tracking separation rates from their baseline male partners, which, aside from their direct implications for fertility outcomes, may also allow us to identify whether the jobs enhanced women's economic independence. This could have led them to adjust their fertility ideals and intentions downward and to change their aspirations for family formation more broadly.

Fifth, we hypothesize that jobs could affect fertility through migration. Giving a job to young women may reduce their propensity to migrate, and therefore increase their likelihood to stay with the current partner and start a family. We will collect data on migration and relationship history to understand what happened to women in the control group who were not offered a job. An important question will be to understand, for the compliers (those who did not migrate because they got a formal job), whether the migration moves would

[2]Note that sub-Saharan Africa's comparatively slow fertility decline is characterized by both a significantly higher desired fertility among poorer women in SSA compared to other low- and middle-income regions, and lower rates of female wage employment, which negatively correlates with fertility intentions across SSA (Zipfel 2023).

[3]Empirical evidence of the quantity-quality trade-off includes Bleakley and Lange (2009); Aaronson, Lange and Mazumder (2014) and Ager, Herz and Brueckner (2020), all in the context of 19th or early 20th century US.

have been favorable or defavorable. We will look at several outcomes: probability to move, probability to move to pursue a work opportunity, probability to move for education etc. We will test whether the movers have a better job and a different types of partners than the non-movers, and whether the movers' advantage is the same in treated and control groups.

More specifically, we will estimate the impact of the treatment on the following potential mediators: (1) women's earnings, (2) women's share in household earnings, (3) women's leisure time, (4) aspirations for children (5) expected returns to schooling, (6) savings, (7) divorce, (8) women's participation in decision making, (9) men's controlling behavior, (10) contraceptive use, (11) migration. We will also estimate the impact of the treatment on the following potential consequences for children: (1) education, (2) child labor, (3) time use, (4) child mortality.

We will test for heterogeneous treatment effects first by the following theoretically motivated tests: (1) extended vs nuclear families; (2) other members' participation in household chores; (3) perceived cost of pregnancy at Factory X; (4) spousal disagreement in terms of ideal number of children, ideal timing and appropriate age at first birth.

We will also test if there is detectable heterogeneity in the treatment effect when using the omnibus test proposed by Chernozhukov et al. (2018) in an honest random forest framework (Athey, Tibshirani and Wager, 2019; Wager and Athey, 2018). If such heterogeneity is detectable we will characterize the most influential variables and the differences in background characteristics of the least and most affected women.

4.2 Exploratory analyses of effects of working when instrumented by random job offers

As not all women offered a job started working and as some women not offered a job at this time were able to find another job, treatment does not perfectly predict job status. To measure the effects of having a job we will also estimate Instrumental Variables (IV) models. It should be noted that the exclusion restriction need not hold for variables such as earnings and income shares as it is likely that getting a job affects a person's identity in addition to the effects it has on income. We therefore pre-specify that the intention to treat specification is the main specification. The IV models should rather be seen as explorative tests of mechanisms for the results.

We estimate IV models of the following form:

(2) $\textit{Any factory job}_{i,t1} = \alpha Y_{i,t0} + \beta Treatment_i + \gamma X_{i,t0} + \delta List_i + \epsilon_{it}$;

(3) $Y_{i,t1} = \alpha Y_{i,t0} + \beta \ \textit{Predicted(Any factory job)}_{i,t1} + \gamma X_{i,t0} + \delta List_i + \epsilon_{it}$

That is, we predict having had any factory job since baseline with the randomization and use the predicted values for in the second stage to calculate the local average treatment effect of having had a factory job on fertility. We will also instrument *Months factory job* with *Treatment* and conduct a similar analysis.

4.3 Exploratory analyses of already collected data

We will use previous waves of data to dig into mechanisms and to deepen the descriptive aspects of the paper. In particular we will use the previous waves to show the effects on employment, time use, and earnings over time. We think this will be particularly valuable as the differences in e.g. employment or share of household earnings are likely to be considerably smaller in the last round of the survey, while the effects on fertility over the last 8 years are likely mediated by previous differences over time.

5 Power

We hope to interview as many as possible of the initial sample and we will invest heavily in tracking for this purpose, including a phone survey to individuals we can not meet in person. Somewhat conservatively we will anyway assume that we only reach 800 individuals, with around half from the treatment group. At the conventional level of significance of 0.05 and a power of 0.8, our sample size would allow for a minimum detectable effect of 0.2 standard deviations. These calculations do not take into account the potential gains in precision from including the covariates in the estimation.

We will also adjust the p-values for the fact that we are testing the impact on two main outcomes. Using the procedure of Benjamini and Hochberg (1995) the critical p-value would be 0.025 for the one with the lowest p-value (0.05* 1/2, which is the same as a Bonferroni correction). The corresponding minimum detectable effect after accounting for multiple hypothesis testing is still around 0.2 standard deviations.

6 Addressing survey attrition and non-response

We will not manage to reach all the respondents initially sampled. We will check whether non-response is correlated with treatment. If there is a statistically significant difference in non-response (controlling for the strata variables), we will follow Kling, Liebman and Katz (2007)'s correction. We will obtain lower bounds of the treatment effect by replacing missing observations in the treatment (control) arms by the corresponding arm's mean value minus (plus) 0.05, 0.10 and 0.20 standard deviations of the control group. Upper bounds of the treatment effects will be constructed in a symmetrical way.

7 Variables with limited variation

To limit noise caused by variables with minimal variation, questions for which 95 percent of observations have the same value within the relevant sample will be omitted from the analysis and will not be included in any indicators or hypothesis tests. In the event that omission decisions result in the exclusion of all constituent variables for an indicator, the indicator will be not be calculated. If this happens for one of our main outcomes we will not use it as a main outcome and we will not adjust for multiple testing for that variable.

8 Archive

The pre-analysis plan is archived at the registry for randomized controlled trials in economics held by The American Economic Association: https://www.socialscienceregistry.org/ before the endline data is collected.

References

Aalen, Lovise, Andreas Kotsadam, Janneke Pieters and Espen Villanger. 2023. "Jobs and Political Participation - Evidence from a Field Experiment in Ethiopia." The Journal of Politics (ja):null.
URL: *https://doi.org/10.1086/726929*

Aaronson, Daniel, Fabian Lange and Bhashkar Mazumder. 2014. "Fertility Transitions along the Extensive and Intensive Margins." American Economic Review 104(11):3701–24.
URL: *https://www.aeaweb.org/articles?id=10.1257/aer.104.11.3701*

Ager, Philipp, Benedikt Herz and Markus Brueckner. 2020. "Structural Change and the Fertility Transition." The Review of Economics and Statistics 102(4):806–822.
URL: *https://doi.org/10.1162/rest_a_00851*

Ashraf, Nava, Erica Field, Alessandra Voena and Roberta Ziparo. 2022. "Gendered Spheres of Learning and Household Decision Making over Fertility." Working Paper .

Ashraf, Nava, Erica Field and Jean Lee. 2014. "Household Bargaining and Excess Fertility: An Experimental Study in Zambia." American Economic Review 104(7):2210–37.
URL: *https://www.aeaweb.org/articles?id=10.1257/aer.104.7.2210*

Athey, Susan, Julie Tibshirani and Stefan Wager. 2019. "Generalized random forests." The Annals of Statistics 47(2):1148–1178.

Becker, Gary S. and H. Gregg Lewis. 1973. "On the Interaction between the Quantity and Quality of Children." Journal of Political Economy 81(2, Part 2):S279–S288.
URL: *https://doi.org/10.1086/260166*

Belloni, Alexandre, Victor Chernozhukov and Christian Hansen. 2014. "Inference on treatment effects after selection among high-dimensional controls." The Review of Economic Studies 81(2):608–650.

Benjamini, Yoav and Yosef Hochberg. 1995. "Controlling the False Discovery Rate : A Practical and Powerful Approach to Multiple Testing." Journal of the Royal Statistical Society 57(1):289–300.

Bleakley, Hoyt and Fabian Lange. 2009. "Chronic Disease Burden and the Interaction of Education, Fertility, and Growth." The Review of Economics and Statistics 91(1):52–65.
URL: *https://doi.org/10.1162/rest.91.1.52*

Caldwell, John C. 1978. "A Theory of Fertility: From High Plateau to Destabilization." Population and Development Review 4(4):553–577.

Chernozhukov, Victor, Mert Demirer, Esther Duflo and Ivan Fernandez-Val. 2018. Generic Machine Learning Inference on Heterogeneous Treatment Effects in Randomized Experiments, with an Application to Immunization in India. Technical report National Bureau of Economic Research.

Delecourt, Solène and Anne Fitzpatrick. 2021. "Childcare Matters: Female Business Owners and the Baby-Profit Gap." Management Science 67:4455–4474.

Doepke, Matthias and Fabian Kindermann. 2019. "Bargaining over Babies: Theory, Evidence, and Policy Implications." American Economic Review 109(9):3264–3306.
URL: *https://www.aeaweb.org/articles?id=10.1257/aer.20160328*

Doepke, Matthias and Michèle Tertilt. 2018. "Women's Empowerment, the Gender Gap in Desired Fertility, and Fertility Outcomes in Developing Countries." AEA Papers and Proceedings 108:358–62.
URL: *https://www.aeaweb.org/articles?id=10.1257/pandp.20181085*

Goldin, Claudia. 1992. Understanding the Gender Gap: An Economic History of American Women. Oxford University Press.
URL: *https://EconPapers.repec.org/RePEc:oxp:obooks:9780195072709*

Jones, Larry, Alice Schoonbroodt and Michèle Tertilt. 2011. "Fertility Theories: Can They Explain the Negative Fertility-Income Relationship?" Demography and the Economy .

Kotsadam, Andreas and Espen Villanger. 2022. "Jobs and Intimate Partner Violence - Evidence from a Field Experiment in Ethiopia." Journal of Human Resources .
URL: *https://jhr.uwpress.org/content/early/2022/08/01/jhr.0721-11780R2*

Lim, K. 2019. "Do American mothers use self-employment as a flexible work alternative?" Review of Economics of the Household 17:805–842.

Rasul, Imran. 2008. "Household Bargaining over Fertility: Theory and Evidence from Malaysia." Journal of Development Economics 86(2):215 – 241.
URL: *http://www.sciencedirect.com/science/article/pii/S0304387807000247*

Rossi, Pauline and Mathilde Godard. 2022. "The Old-Age Security Motive for Fertility: Evidence from the Extension of Social Pensions in Namibia." American Economic Journal: Economic Policy 14(4):488–518.
URL: *https://www.aeaweb.org/articles?id=10.1257/pol.20200466*

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

Zipfel, Céline. 2023. "The Demand Side of Africa's Demographic Transition: Desired Fertility, Wealth, and Jobs." Mimeo .