UNITED NATIONS UNIVERSITY
UNU-WIDER

WIDER Working Paper 2022/164

Determinants of gender gaps in youth employment in urban Mozambique

Jana Bischler,[1] Eva-Maria Egger,[2] Paul Jasper,[1] Ivan Manhique[3]

December 2022

Abstract: In this study, we explore the correlates of the employment gender gap among urban youth in Mozambique. Young people are confronted with simultaneous decisions about education, work and family life influenced by social norms around gender roles. Using data from a panel of individuals in 2017 and 2020, aged between 15–25 years in 2017, that covers information on education, employment, fertility, social life, gender norms and more, we observe an increase of 10 percentage points in the raw employment gender gap over time to the disadvantage of young women. Exploiting the longitudinal nature of our data, we apply two methods to assess the main correlates of this gap, an Oaxaca-Blinder decomposition on first-differenced data and a data-driven individual-level fixed-effects LASSO approach. Both analyses reveal that young women face a significant trade-off between work and time spent with reproductive activities and that the labour market seems to reward better education only for men.

Key words: gender gap, employment, decomposition, LASSO, Mozambique

JEL classification: J13, J16, J71, O55

Acknowledgements: The authors thank Sam Jones, Janneke Pieters, and Cecilia Poggi for their constructive comments that greatly improved this paper. They are grateful for the feedback from participants at the IGM-CEEG seminar in Maputo, a UNU-WIDER brown bag seminar in Helsinki, and a seminar at the MUVA association in Maputo. The data collection would not have been successful without the excellent work of Andres Arau and the members of the Associação de Nutrição e Segurança Alimentar (ANSA).

---

[1] Oxford Policy Management, Oxford; [2] UNU-WIDER Maputo, [3] independent researcher. Corresponding author: Paul.Jasper@opml.co.uk

This study has been prepared within the UNU-WIDER project Inclusive growth in Mozambique – scaling up research and capacity implemented in collaboration between UNU-WIDER, University of Copenhagen, University Eduardo Mondlane, and the Mozambican Ministry of Economy and Finance. The project is financed through specific programme contributions by the governments of Finland and Norway.

Information and requests: publications@wider.unu.edu

ISSN 1798-7237 ISBN 978-92-9267-297-3

https://doi.org/10.35188/UNU-WIDER/2022/297-3

Typescript prepared by Lorraine Telfer-Taivainen.

United Nations University World Institute for Development Economics Research provides economic analysis and policy advice with the aim of promoting sustainable and equitable development. The Institute began operations in 1985 in Helsinki, Finland, as the first research and training centre of the United Nations University. Today it is a unique blend of think tank, research institute, and UN agency—providing a range of services from policy advice to governments as well as freely available original research.

The Institute is funded through income from an endowment fund with additional contributions to its work programme from Finland and Sweden, as well as earmarked contributions for specific projects from a variety of donors.

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

1 Introduction

It is often acknowledged that increasing women's participation in the labour market and improved access to quality employment is a direct and indirect driver for economic growth and development; see e.g., Klasen (1999) and Verick (2014). Channels via which positive effects of improved labour market outcomes of women might materialize include additional household income (Galor and Weil, 1996), increases in human capital (Klasen and Lamanna 2009), and positive demographic effects (Behrman and Gonalons-Pons 2020; Bloom et al. 2003; Bongaarts et al. 2019). Improved labour market outcomes for women also lead to an increase in women's agency and decision-making power (Annan et al. 2019; Kabeer 2008) and, as a result, to potential improvements in children's nutrition, health, and education outcomes (Chari et al. 2017; Perez-Alvarez and Favara 2020; Quisumbing 2003).

Yet, low female labour force participation rates persist in many countries (Klasen 2019). To understand determinants of low participation rates and gender gaps in labour market outcomes in low- and middle-income countries (LMICs), and to develop effective policy responses, an analysis of dynamics early in the economic lives of young people is important. This is mainly due to two reasons: First, gender differences in labour market outcomes often materialise in youth when decisions around work coincide with marriage and first fertility choices. The differences are sticky, preventing young women from achieving their potential later in life and leading to long-term reductions in women's wellbeing and welfare (Fox 2019). Second, nearly 1 billion of the 1.2 billion youth aged 15–24 worldwide live in developing countries, placing them at the heart of the debate on sustainable development (IFAD 2019). While the population in the rest of the world is, or will soon be, aging, sub-Saharan Africa (SSA) has the highest projected growth rate in youth population, which presents an unprecedented opportunity if countries find effective ways to address the youth employment challenge and manage the demographic transition well (Bandiera et al. 2022; Filmer and Fox 2014). Bandiera et al. (2022) document that compared to youth in other continents, African youth are equally likely to work but less likely to be paid and to work in a salaried job because job growth is limited. While the educational gender gap has declined drastically in the continent, the employment gap much less so (Mariara et al. 2018).

Mozambique is no exception. According to the World Bank development indicators, about 20 per cent of Mozambicans are 15–24 years old, and an increasing proportion of the population lives in urban areas (37 per cent in 2020 with an annual growth rate of more than 4 per cent). On a national level, in 2019 about 80 per cent of the population are active in the agricultural subsistence sector, many of them women. However, subsistence farming is not considered employment (Gaddis et al. 2020; ILO 2013). Employment in the non-subsistence sector has been rising since the 1990s but there remains a significant gap between male and female employment rates. In 2014/15, 12.6 per cent of working-age women and 30 per cent of working-age men in Mozambique were employed in the non-subsistence sector. Among them, 31 per cent of the women worked as unpaid family worker compared to 14 per cent of men (Gradín and Tarp 2019).

Our study focusses on gender gaps in labour market outcomes among urban youth in Mozambique. We use a uniquely rich panel dataset that tracks a sample of young individuals (aged 15–25 in 2017) in two of Mozambique's largest cities between 2017 and 2020. We analyse how gender differentials in non-subsistence employment rates among them change over time and identify factors that are associated with these changes. More specifically, we are interested in assessing how factors that are related to changes in employment status vary between male and female study participants. We first apply a classical regression and decomposition approach to test the relevance of determinants established in the literature for our setting. Then we employ a data-driven approach using Least Absolute Shrinkage and Selection Operator (LASSO) panel regression, inspired by Belloni et al. (2016), to allow for previously not considered factors that are related to employment status changes to emerge.

In our analysis, we show that young men who are part of our study are significantly more likely to be employed than young women and that this difference increases significantly by 10 percentage points between the period 2017–20, which is mainly driven by the youngest cohort of men transitioning into employment over this period. This is confirmed by our decomposition analysis, where we find a gender employment gap of 11 percentage points. This gap is almost fully explained by the unobserved component, meaning unobserved differences in preferences between men and women or discrimination that women face. For men, education is associated with improvements in employment prospects, while for women it is not. In contrast, young women seem to face a significant trade-off between spending time at work or at home with reproductive work, while for men this does not seem to matter. We do not find significant associations in terms of our indicators of social norm attitudes. The LASSO analysis confirms the previous findings. Further, it reveals that there is a positive and significant association between savings group membership and employment for young women. These results are robust to attrition corrections, changes in regression specifications to allow for alternative definitions of social norms, and the choice of counterfactuals in the decomposition analysis.

Our study's contribution to the existing literature is three-fold. First, our longitudinal survey data — which includes specific questions on social norms behaviours and attitudes towards women's employment — allows us to investigate the relevance of a multiplicity of dimensions identified in the literature in a panel data context, i.e., controlling for unobservable time-invariant characteristics among respondents. Second, we implement an innovative data-driven approach to robustly identify significant time-varying predictors of employment that the literature so far might not have paid attention to. Third, we specifically focus on employment outcomes early in the economic lives of young urban Mozambicans. This allows us to provide evidence on factors that drive early labour market differentials among young men and women, which in turn might lead to persistent labour market inequalities later in life for an increasingly important cohort of urban Mozambicans.

The remainder of this paper is structured as follows. In section 2 we provide an overview of the literature on the drivers of gender gaps in labour market outcomes in LMICs. Section 3 introduces the MUVA urban youth panel data that forms the basis of our analyses. We then describe our main estimation approaches in section 4. In section 5, we present and discuss our results. Section 6 concludes with a discussion of limitations to the study and implications of our work.

2 Drivers of gender gaps in labour market outcomes in LMIC

A widely researched theory on the drivers of female labour force participation is the feminization U-shape hypothesis (Boserup et al. 2013; Goldin, 1995). According to this hypothesis, female labour market participation is high in poor countries where women are engaged in subsistence activities out of necessity. As economies develop and move towards male-dominated industrial jobs, female labour force participation falls until it rises again as women's education levels improve, fertility rates fall, and women respond to a growing demand in the service sector. While there is some evidence that this U-shape hypothesis holds (Jayachandran 2020), it is important to note that not all labour markets and economies follow this pattern (Klasen 2019; Verick 2014). For example, Gaddis and Klasen (2014) found little empirical support for the feminization U-shape hypothesis in current developing countries and Idowu and Omowumi (2019) argue that in African countries the relationship between female labour force participation and economic development follows an inverted U-shaped relationship instead.

Besides the level of economic development and structure of the economy, the global literature identifies a range of overlapping factors that explain the gap between men and women in terms of employment outcomes (Verick 2014). This literature has identified a variety of different channels that affect these gender gaps, with a recent focus on three key ones: differences in human capital and skills between men and women; social norms; and discrepancies in access to and use of information and communication technologies (ICT). We briefly review each one of these in turn here.

One of the main factors that is commonly identified as driving employment outcomes is human capital accumulation and skills that workers acquire either in school or on the job (Becker 1962; Mincer 1974; Schultz 1961). Consequently, disparities in the levels of education or skills between men and women are often found to also have significant explanatory power for the differences in labour market outcomes between men and women (Cazes and Verick 2013; Rebollo-Sanz and Rica 2020). For instance, for Mozambique, Gradín and Tarp (2019) use a Oaxaca-Blinder decomposition to show that disparities in human capital, measured by educational attainment, literacy levels, and Portuguese proficiency, is one of two drivers explaining the gap in non-subsistence employment between men and women in Mozambique.

There is also broad agreement in the literature that social or cultural norms, loosely understood as society's informal rules about one's appropriate behaviour, can affect employment outcomes for men and women differentially (Jayachandran 2020). For example, Maxwell and Wozney (2021) found that in the US, norms about work and home explain about 60 per cent of the wage gap between men and women and in China, Xiao and Asadullah (2020) estimated that gender-related community norms account for 41 per cent of the unexplained differences in male and female labour force participation. These social norms can take a variety of different forms. For instance, norms about the types of jobs that are appropriate for women could restrict their access to employment (Boudet et al. 2013). Similarly, in some contexts, both the real risk of sexual harassment at work and perceptions around it and safety more generally can lead to lower levels of female participation in the economy (Chakraborty et al. 2018; Siddique 2018). In other contexts, norms that constrain women's mobility can negatively affect their effective participation in economic activities (Field et al. 2010; Jayachandran 2020). Another important group of behaviours driven by gender norms identified in the literature relate to unpaid care or reproductive work, which is often considered to be the domain of women and hence can limit the time they have available for participation in the formal, paid labour market (Charmes 2019; Clark et al. 2019). In a similar vein, studies have identified home chores and child bearing (Arsalan et al. 2019) and marriage (Lee et al. 2008) as factors constraining female labour force participation. In fact, Gradín and Tarp (2019) find that the second main driver explaining the employment gap between men and women in Mozambique is marriage. The common theme across these studies is that social norms constrain women's participation in the labour force either directly — e.g., by stigmatising certain types of jobs — or indirectly by defining how women ought to behave under certain conditions — e.g., when being married or having children — which then limits their ability to work. Most of the above literature describes the discussed factors as 'discrimination against women'. However, as discussed by Oaxaca (2007) and Neumark (2018), some of the observed gender gaps could also arise from differences in tastes or preferences of the individuals, for example the desired number of children or the types of job they would like to be employed in. We believe, however, that these preferences often are shaped or influenced by the cultural environment and the prevailing social norms that individuals face.

Finally, access to and use of ICT is often found to have positive impacts on labour market outcomes across developing countries (Bahia et al. 2021; Eyike Mbongo 2019; Hjort and Poulsen 2019; Tshukudu 2019). At the same time, numerous studies have shown that women all over the world are less likely to access and use ICTs than men. The gender digital divide is estimated to be the largest in SSA (Antonio and Tuffley 2014). Closing this digital divide has consequently been shown to effectively improve women's labour market outcomes. For example, Nikulin (2017) and Valberg (2020) have shown that ICT access contributed to narrowing the gender gap in labour force participation by increasing female labour force participation but not impacting male labour force participation. In another example, internet access has also been found to lead to women using the internet for job search in Jordan (Viollaz and Winkler 2021). In sum, equal access to and knowledge of ICT is seen as an important condition in modern economies for equal participation in the labour market.

As described above, our study adds to this literature by making use of a panel dataset collected among youth in urban Mozambique that allows us to investigate the relative importance of these channels at the same time. Making use of automated variable selection procedures also allows us to add robustness to these findings, making sure that factors identified as significant are not just driven by researcher, in this case our, discretion and priors. We describe the data and our methods in more detail in the following sections.

3 The MUVA urban youth panel data

3.1 The longitudinal survey

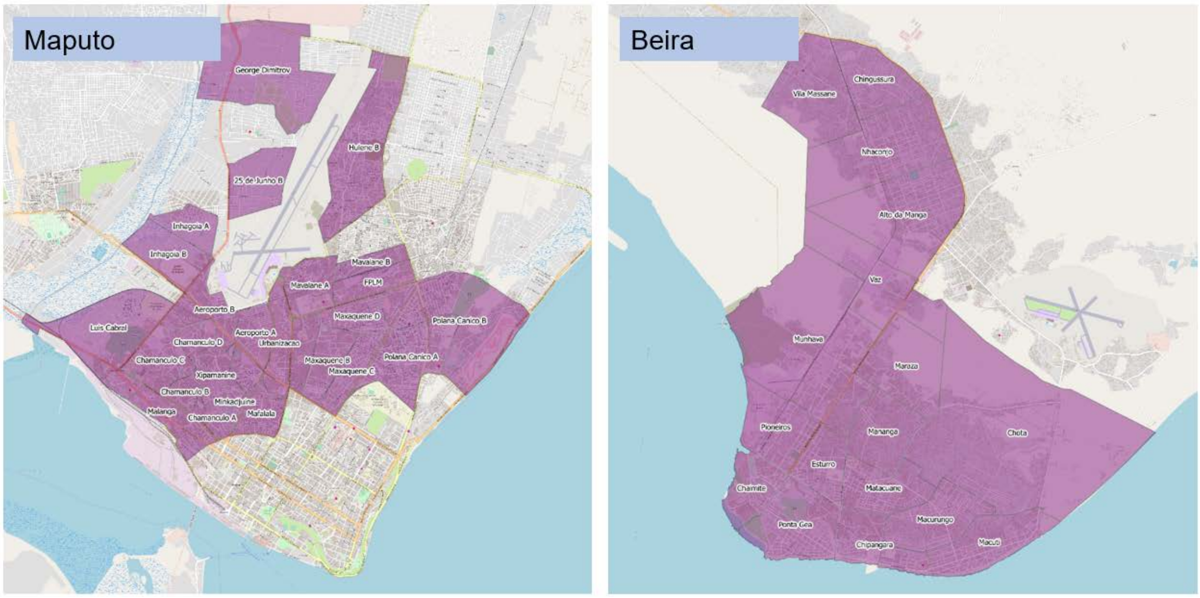
This study draws on a longitudinal survey of youth and young adults aged 15–25 in 2017 in two of the largest cities of Mozambique, Beira and Maputo. The survey was originally started in 2017 by MUVA,[1] an organization that promotes women's economic empowerment among youth in urban areas in Mozambique. Its aim was to capture the challenges young female Mozambicans from low socio-economic backgrounds living in urban areas face in their lives compared to young men. It covers a range of questions about household characteristics, education, employment, fertility, financial inclusion, time use, social norms, social capital, and decision making.

The survey consists of three rounds of data collection. The first round was collected in 2017 and covered 3,300 individuals. The targeted geographic areas were densely populated, low-income, inner-city areas based on the 2007 Population and Housing Census. As a result, the sample was representative of all inner-city neighbourhoods of Beira, but only a sub-set of these in Maputo.[2] These neighbourhoods are highlighted in purple in the maps of Beira and Maputo in Figure 1. In a first stage, census enumeration areas (EA) were randomly selected from these neighbourhoods. In a second stage, all households with at least one member between the ages of 15–25 years were listed. In each EA, 16 households were randomly sampled from the listed households. If a household had more than one member within the target age range, an individual was randomly chosen among them as the respondent. A more detailed description of the sampling methodology can be found in Arau et al. (2018).

[1] See muvamoz.co.mz for more detail.

[2] In Maputo, the most affluent municipal district, KaMpfumo, was excluded from the sample frame given MUVA's focus on poor urban neighbourhoods.

Figure 1: Areas covered by the MUVA urban youth panel in Maputo and Beira



Source: Arau et al. (2018).

In 2018, a second round of data collection was conducted, following up with a randomly selected sub-sample of 1,600 respondents of the original survey explicitly stratified by city, and implicitly by gender, education, neighbourhood, and enumeration area; see Allen et al. (2020). This round only covered a limited set of the original questions and focused on measuring young people's functional literacy and numeracy skills. Two years later, in late 2020,[3] a third round of the survey was conducted aiming to reinterview as many young participants as possible from the sample of 1,600 youth in the second round. Where necessary and when individuals from the primary sample were not available, other participants from the first round, but who were not followed-up in the second round, were interviewed as replacement respondents. The survey managed to successfully track a total of 1,195 participants in this third round, resulting in an attrition rate of 25.3 per cent. In section 5.3, we observe that, on average, attritors are more likely to be slightly older women, with more children, and more decision-making power about their own mobility from slightly smaller households. Fewer of them are single. This could fit a profile of young women moving away from their baseline household to another location outside of our sample.[4]

Table 1: Sample composition of the MUVA urban youth panel by gender and city

	Beira		Maputo		Total	
	%	N	%	N	%	N
Female	55.2	353	56.9	316	56.0	669
Male	44.8	287	43.1	239	44.0	526
Total	100.0	640	100.0	555	100.0	1,195

Source: authors' calculations using MUVA Urban Youth Survey data.

[3] The survey was conducted during the COVID-19 pandemic, carefully respecting recommended safety measures. Government restrictions were in place throughout the duration of the survey, equal for all respondents.

[4] We discuss the implications of attrition for our results in section 5.3.

The present study focusses on the complete panel of 1,195 participants, for whom data is available from the first round in 2017 and the third round in 2020.This sample includes 56 per cent women (N=669) and 44 per cent men (N=526). Table 1 presents the sample size and proportions by gender and city.

3.2 Variables used in this study

The main labour market outcome used in this study is employment, following the ILO (1982) resolution on employment statistics. This refers to economic activities performed in the seven days prior to the implementation of the survey interview and categorises anyone as 'employed' if they did any remunerated or non-remunerated work, either as employed worker or in self-employment. Importantly, this also includes unpaid work in a family business.[5] 'Any work' means having worked for at least an hour in the seven-day reference period. Finally, this definition also categorised people as employed who stated that they did not work in the seven days prior to the survey, but who had employed, self-employed or unpaid work in a family business that they would definitely return to in the following weeks. In Figure 2 we present the key estimates of this indicator for our panel. In order to provide some context to these estimates, we also present estimates related to some other measures of employment or work in section 3.3.

3.2.1 Covariates used in the theory-driven analysis

As will be described in section 4, we implement two separate estimation strategies in this study that deal with covariate selection differently. For our first estimation strategy, which we call the theory-driven analysis, we select key covariates to be included in our analysis based on the channels identified in the literature review presented in section 2. We describe the selected variables in Table 2 and provide a summary here.

To capture the education channel, we include a measure of the highest grade completed by the survey respondent in our analysis. To capture the broad channel of social norms, we include several different variables: First, we include the hours spent on reproductive work by the individual on the day prior to the survey. Second, we include a variable that indicates whether the individual can decide about their movements outside of the home themselves. Third, we measure social norms around women in leadership positions using two variables, by asking about whether the respondent themselves approve of women in leadership position and whether they think that others in the neighbourhood do. The first variable aims to directly extract views on women in leadership held by survey respondents whereas the second variable exposes what respondents perceive as the social norm in their relevant reference group. An individual might privately agree to women in leadership but still acknowledge that the prevailing opinion on this matter is different. These two complementary variables can thus have different influences on the employment outcome of young people.

Fourth, to capture individuals' social capital, we include a variable that measures whether they form part of social groups, other than religious groups, or not, such as e.g., savings groups or political groups, as we hypothesise that these provide a measure for how well individuals are connected to others in the local community. Fifth, we include the number of children alive in our analysis as it controls for the trade-off between employment, fertility decisions, and care work that individuals face and might be different for men and women due to social norms around child

[5] Subsistence agriculture is not considered employment according to the ILO 2013 revision of labour statistics (ILO 2013). Given that our sample is urban, we do not have anyone working in this sector.

rearing responsibilities (Doepke et al. 2022).[6] Finally, to capture norms around marriage, we include a binary indicator for whether the individual is married or not. Together, these variables cover a broad range of different ways in which social norms can affect young peoples' lives. To capture the third channel that affects employment emphasised above, access and use of ICT, we include a variable on whether an individual frequently uses a computer or not.

Table 2: Covariates used in theory-driven analysis

Variable name	Definition
Highest grade completed	Highest level of education the respondent achieved at the time of each survey round, measured in years.
Single	Whether the respondent is single at the time of each survey round, or not (i.e. in a marital union or currently married). (0=no, 1=yes)
Number of children still alive	The number of respondent's children alive at the time of each survey round.
Number of hours spent per day on reproductive work	The number of hours the respondent spent on reproductive work on the day prior to the interview. This includes domestic chores, looking after children, or looking after the sick and elderly.
Make decisions about my movement alone	Whether the respondent in general makes decisions about their movement alone or not. (0=no, 1=yes)
Frequently uses a computer	Whether the respondent reports using a computer frequently, which means at least once a week. (0=no, 1=yes)
Member of a social group (excl. church)	Whether the respondent reports being member of a social group (credits/savings group, political group, community group, others). (0=no, 1=yes)
Do others approve of women in leadership	Whether the respondent thinks that more than the majority of people in their neighbourhood would approve of a woman being selected for the leadership of an organization (professional, school, political, community organization). (0=no, 1=yes)
Would approve of women in leadership	Whether the respondent would approve of a woman being selected for a leadership position in an organization. (0=no, 1=yes)
Simple Poverty Score Card 2014/2015 poverty score.	Household wealth score estimated using a simple poverty scorecard approach. This covers asset ownership and characteristics of households, presented in (Schreiner 2017).
Age	Age of the respondent at the time of each survey round, in years.
Household size	The number of individuals who are members of the household the respondent is also a member of at the time of each survey round.
Household composition: sex ratio	The ratio of female household members to total household members. A value of one means that all household members are female.

Source: authors' description using MUVA Urban Youth Survey data.

[6] Our main analyses are robust to alternative specifications, where we include other variables that capture norms around child-bearing that respondents might hold, such as e.g. the number of children they consider to be ideal for women to have. See Appendix A2.

In addition, we include a set of control variables to cover some basic socio-demographic characteristics of study participants: the age of the respondent, the city of residence, the household size, and its proportion of female members. Lastly, we define a poverty score at household level based on a poverty score card approach (see Arau et al. 2018 for details).

3.2.2 Covariates used in the LASSO analysis

The idea behind using LASSO for our second estimation procedure is that it allows us to test for significant relationships between our key outcome variable and a potentially large set of covariates (see section 4.2 for more details). Our survey questionnaire includes 199 questions which, after cleaning and indicator creation translates into a dataset of over 450 variables. This includes, however, variables that cannot be used in our analysis for a variety of different reasons. For example, some variables will be defined for certain sub-populations of respondents only or will be directly related to the employment status and type of work done by the individual, i.e., our outcome variable. We therefore implemented a step-by-step data wrangling procedure to reach a final set of variables to be used in our LASSO analysis:

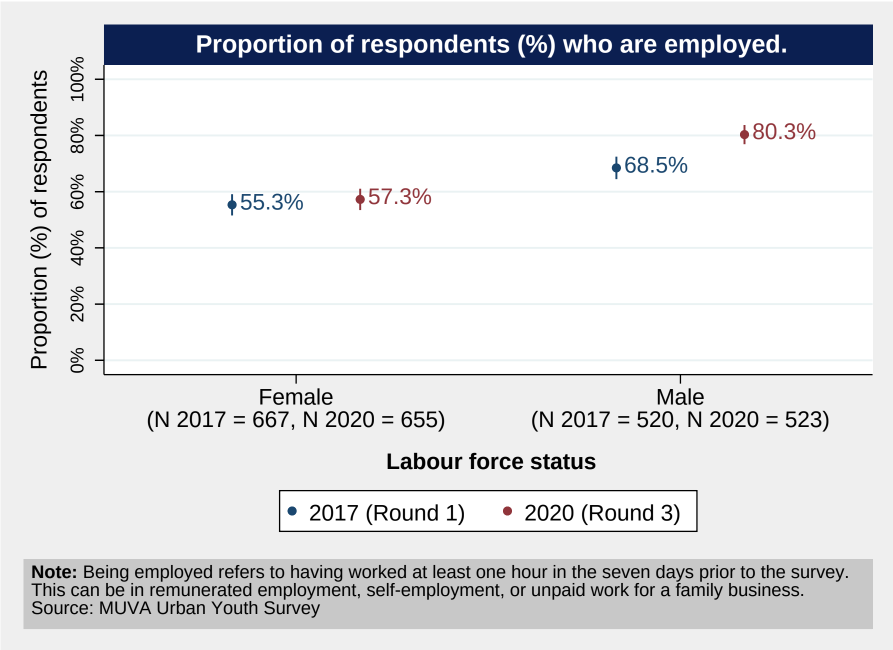
In a first step, we either recoded (where possible) or removed variables that were defined for a sub-population of survey respondents only. For example, if a question was asked only to enrolled individuals, the non-enrolled category was labelled as such. In a second step, we removed survey information variables, such as e.g., the enumeration area or neighbourhood identifier. Third, in order to prevent introducing artificial relationships between the outcome and covariates, we removed any variable that by construction was related to the employment status of the survey respondent. Finally, because any missing values in one variable included in the LASSO regressions would affect the numbers of observations that the model runs on, we exclude any variables that have more than 5 per cent missing values. At the end of this process, we reached a list of 149 binary, categorical, and continuous variables that are included in the LASSO analysis. These include the covariates selected for our theory-driven analysis. The full list of variables is presented in Appendix A1.

3.3 Summary statistics

3.3.1 The gender employment gap over time

This section provides a summary of the main sample characteristics related to the gender employment gap. In Figure 2, we illustrate the gender gap in employment and how it evolves over the survey rounds. The figure shows that there is a large employment gap between young men and women, which is increasing over time. While employment rates have barely moved for female survey participants (from about 55 per cent in 2017 to 57 per cent in 2020, a change that is not statistically significant), they have increased significantly for young men, from 68 per cent in 2017 to 80 per cent in 2020. The increase in the raw gender gap over time, therefore, is 10 percentage points. Most of this increase is due to a significant increase in employment among men who were aged 15–18 years in Round 1 (2017) (see appendix A1).

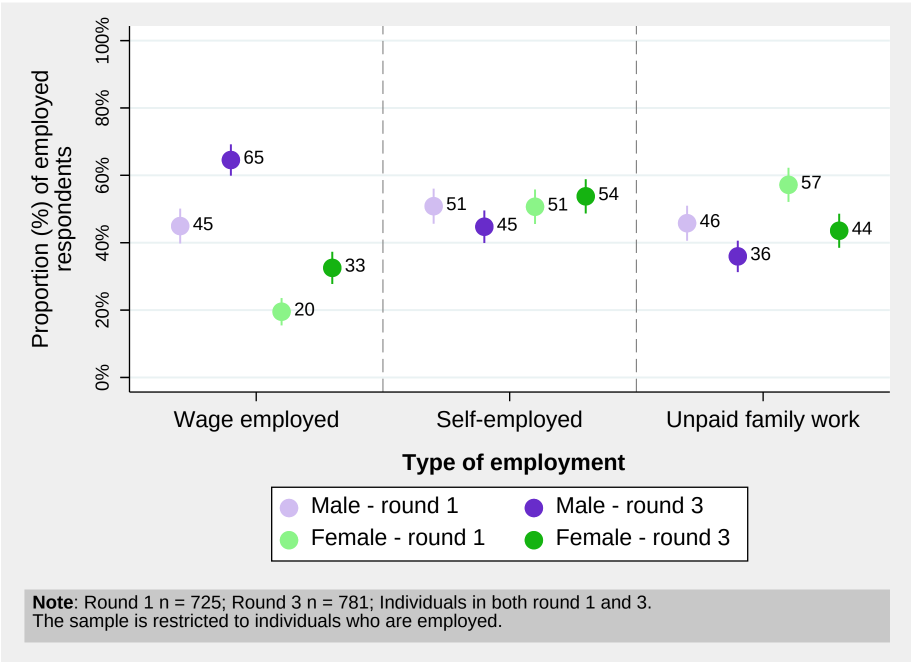
Figure 2: Employment rate of young women and men, by survey wave (%)



Source: authors' calculations using MUVA Urban Youth Survey data.

It is also worth noting that the employment rate in Round 3 (2020) is higher than in Round 1 (2017) despite the COVID-19 pandemic affecting the Mozambican economy in 2020. However, it is important to note that this employment indicator does not capture any loss in income due to reduced hours or revenues during the pandemic (Egger et al. 2021).

Figure 3: Employment type by gender and survey round



Source: authors' calculations using MUVA Urban Youth Survey data.

Figure 4: Occupation by gender and survey round

Note: Round 1 n = 725; Round 3 n = 795; Individuals in both round 1 and 3.
* The sample is restricted to individuals who are employed.
** This figure only includes occupation groups with at least one observation.

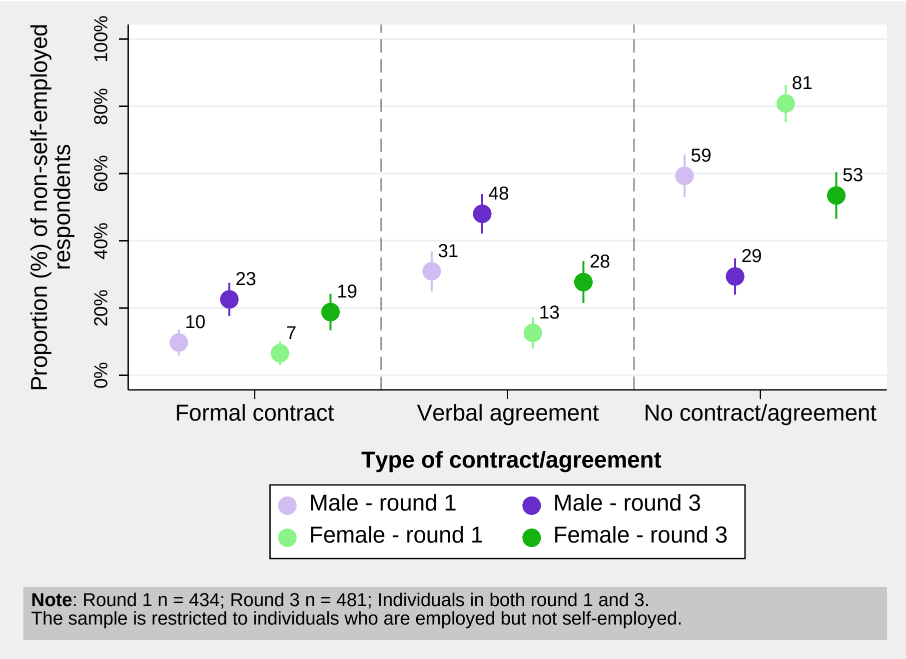
Source: authors' calculations using MUVA Urban Youth Survey data.

To provide additional context, we present supplementary statistics on other employment-related indicators in the following graphs and summarise key findings here. First, if they work, the young women in our sample are less commonly wage employed and more likely to do unremunerated family work or be self-employed (Figure 3), a finding that holds across survey rounds. This finding also generally holds when disaggregating this analysis by age group, although younger women (aged 15–18 years in 2017) are significantly less likely to be wage employed than older women (aged 19–25 years in 2020) in both survey rounds (see appendix A1). Second, a large proportion of young people work in low skill occupations (Figure 4) and this tendency is even stronger for women. Third, young working women in our sample are less likely to have a formal contract or verbal agreement, highlighting that a larger proportion of young women than men work in the informal sector and/or engage in irregular work (Figure 5).

Because the period we observe for respondents and the trends over time could be associated with important transitions from school to employment for young men and women, we also look at respondents' current school enrolment and their employment status. Recognizing that respondents can be both employed and enrolled in school at the same time, Figure 6 shows the distribution of respondents across the four possible combinations of both indicators at each survey round for both male and female respondents separately. In essence, these figures show how the employment rates presented in Figure 2 can be disaggregated when also taking current school enrolment status of respondents into account.

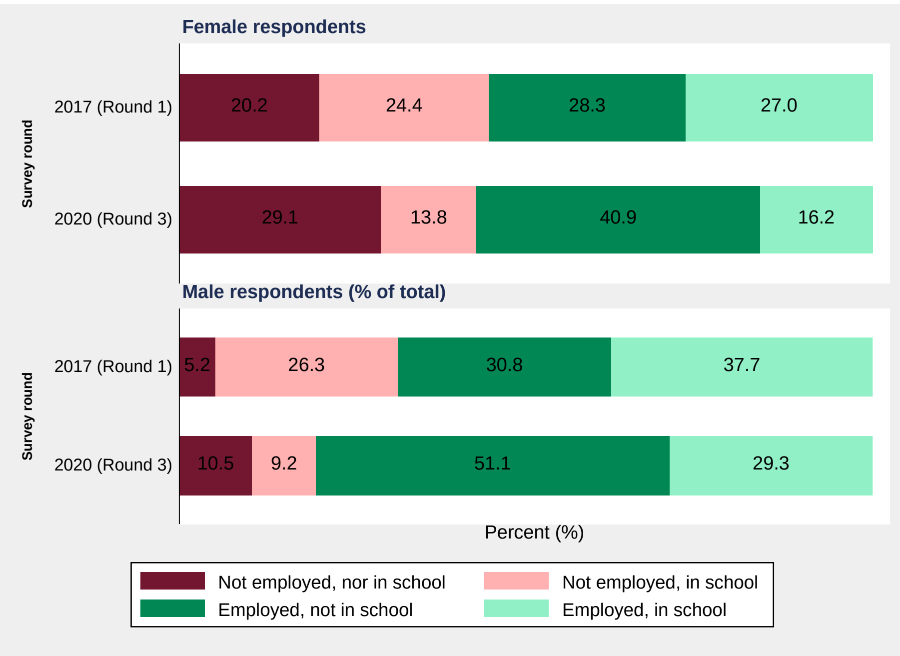
A few key insights can be derived from these graphs. First, in both 2017 and 2020 and both among male and female respondents, individuals were likely to both be employed and work: this proportion varies between 16–38 per cent across the different sub-groups and underlines the fact that in Mozambique the transition from education to work is messy and can involve both being employed and at school at the same time for many individuals. Second, however, the proportion of respondents enrolled in school, whether they are employed at the same time or not, drops from 2017 to 2020 for both men and women, indicating that, indeed, this is a period where many respondents stop studying. Third, most strikingly, the graphs indicate that the proportion of men who were employed and not in school increased by about 20 percentage points between 2017–20, roughly twice the increase seen among female respondents (about 10 percentage points). At the same time, the proportion of men who were not employed but in school decreased by 17 percentage points for men and 11 percentage points for women. Similarly, the percentage point decrease of respondents who were both employed and in school between 2017–20 is roughly the same among male and female respondents. Together, this indicates that men seemed to have been more likely to make the transition away from education into employment than women.

Figure 5: Contract type by gender and survey round



Source: authors' calculations using MUVA Urban Youth Survey data.

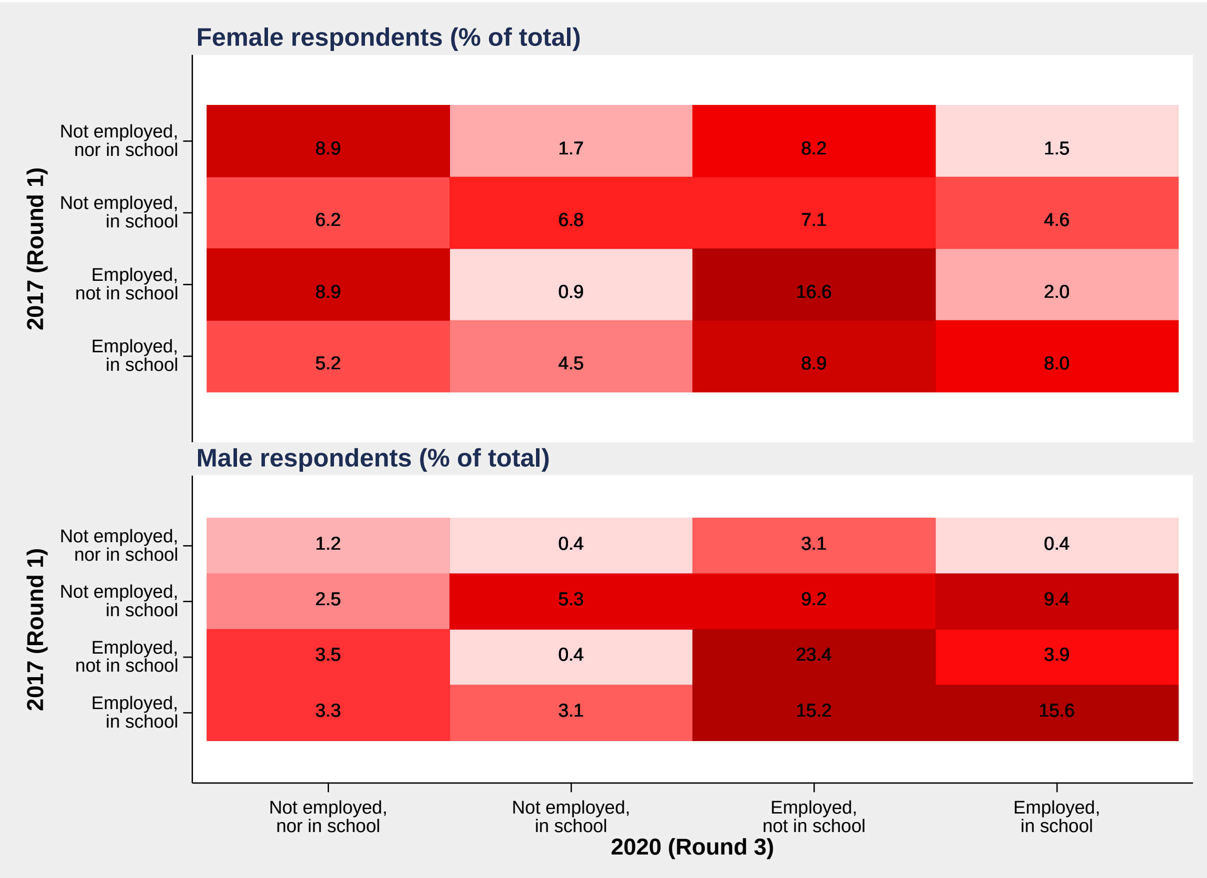
Figure 6: Employment and school transition



Source: authors' calculations using MUVA Urban Youth Survey data.

To investigate this further, we look at transitions between the different combinations of employment and school enrolment among survey respondents in Figure 7. This transition matrix shows how both male and female respondents move from one category of the employment and schooling combination to another, or not, between 2017 (y-axis) and 2020 (x-axis). Each cell represents one transition, with percentages of the total number of female (top panel) or male respondents (bottom panel) displayed.

Figure 7: School to work transition for female and male respondents



Source: authors' calculations using MUVA Urban Youth Survey data.

Three key insights can be derived from this analysis. First, among male respondents, the bottom right-hand quadrant composed of the four cells of individuals who were employed in both 2017 and 2020 together makes up over 50 per cent of the total, which means that most male respondents remained employed and transitioned within the employment categories between the period 2017–20.

Second, still among male respondents, 19 per cent transitioned from not being employed but in school to employment (male panel, second row, right two cells). About half of these (9 per cent), were not in school in 2020 anymore. The equivalent percentages among women were 12 per cent and 5 per cent. This confirms the finding that men were more likely to transition away from education to employment than women.

Third, and in addition to the above findings, the transition matrix for women (top panel) shows that they are more likely to stay unemployed (the top left-hand quadrant of the 'unemployed' cells makes up 24 per cent of the total compared to 9 per cent for men), but also more likely to drop out of employment (the bottom left-hand quadrant of the four cells that indicate this transition make up 20 per cent of the total for women compared to 10 per cent for men). Interestingly, women were also more likely to transition into employment between 2017–20 if they were not employed, nor in in 2017 (10 per cent vs 5 per cent for men).

Together, these findings show that the trends observed in Figure 6 stem from the fact that, on average, female respondents tended to find it more difficult to transition into employment between 2017–20, even when leaving school, and that any gains in employment that might be observed among women were offset by the fact that they also tended to transition into unemployment more easily than men.

3.3.2 Background characteristics of survey participants

Table 3 provides an overview of the main characteristics for young women and men in each survey round, how they differ between those two groups, and how they change over time. We present the mean estimate of each indicator disaggregated by round and gender in columns labelled as (1), (2), (3), and (4). Changes between the survey rounds for men are shown in column (5). The equivalent for women is shown in column (6). The difference-in-difference estimate is given in the last column. This estimate indicates whether trends over time varied significantly between men and women. We describe these statistics by focussing first on levels and how they compare between men and women in each survey round. We then discuss trends over time for men and women. In a final step, we look at how these trends varied between men and women.

First, in terms of levels and differences between men and women, the results in the table show that both in 2017 and 2020 young men and women in our sample are significantly different across a range of outcomes: In both rounds, men are on average more educated, they are more likely to be single and have fewer children than women. In both 2017 and 2020, young men spend on average more than three hours less on reproductive work per day than young women. They are also more likely to be able to make decisions about their movement on their own. Women are significantly less likely, in both waves, to use a computer frequently. They are also less likely to be a member of a social group in the first round of the survey, although this is no longer the case three years later in 2020. In both years, men and women have relatively similar views about whether *others* would approve of women in leadership. However, men are less likely than women in both rounds to approve of women in leadership positions. As measured by our poverty scorecard, young women, on average, live in relatively poorer households and are more likely to live in households with a larger proportion of female to male members. However, in both rounds, young men and women live in households with a similar average number of members. The average age of young men and women across the two rounds is also comparable, although the distribution changes slightly. In 2020, the proportion of younger women (aged 15–18) is slightly larger than younger men, while there are slightly less men aged 19–25 years than women in 2020. This indicates some selective attrition, which we analyse in section 5.3.

With respect to trends between 2017–20, both men and women improved their educational status as they got older. Fewer men and women are single by the second round, while the number of children increased significantly in both groups. Both groups increased their decision-making power about their own movements but the gap between men and women found in 2017 persists. Both groups are more likely to think that others approve of women in leadership in 2020, but hardly changed their own views. Poverty scores indicate that household wealth for both men and women has significantly decreased between 2017–20, while the average household size increased, and the proportion of women in their households decreased significantly. However, across some indicators, the trends between men and women also differ. Women significantly increased the number of hours spent on reproductive work by 2020, while the number of hours men spent on

Table 3: Characteristics by gender in each survey wave and differences between gender and waves

	2017 – Round 1						2020 – Round 3						Change from 2017 to 2020					
Indicator Name	**Male Estimate**		**Female Estimate**			**Difference**	**Male Estimate**		**Female Estimate**			**Difference**	**Male ((3)-(1))**		**Female ((4)-(2))**			**Double difference**
	(1)	SE	(2)	SE	(2)-(1)		(3)	SE	(4)	SE	(4)-(3)		(5)		(6)		(6)-(5)	
Highest grade completed	9.21	0.11	8.81	0.10	-0.400	***	10.24	0.14	9.65	0.11	-0.585	***	1.025	***	0.84	***	-0.185	
Single	0.94	0.01	0.72	0.02	-0.218	***	0.87	0.02	0.68	0.02	-0.187	***	-0.076	***	-0.045	***	0.031	*
Number of children still alive	1.12	0.02	1.56	0.03	0.439	***	1.25	0.03	1.83	0.04	0.584	***	0.124	***	0.269	***	0.145	***
Number of hours spent per day on reproductive work	1.86	0.09	5.02	0.14	3.159	***	2.01	0.10	5.56	0.15	3.548	***	0.149		0.539	***	0.39	**
Make decisions about my movement alone	0.59	0.02	0.31	0.02	-0.283	***	0.78	0.02	0.46	0.02	-0.321	***	0.183	***	0.145	***	-0.038	
Frequently uses a computer	0.44	0.02	0.19	0.02	-0.247	***	0.43	0.02	0.14	0.01	-0.290	***	-0.012		-0.054	***	-0.042	
Member of social group (excl. church)	0.30	0.02	0.23	0.02	-0.071	***	0.32	0.02	0.34	0.02	0.017		0.015		0.103	***	0.088	***
Do others approve of women in leadership	0.38	0.02	0.44	0.02	0.057	*	0.51	0.02	0.51	0.02	-0.004		0.131	***	0.07	***	-0.061	
Would approve of women in leadership	0.85	0.02	0.92	0.01	0.078	***	0.88	0.01	0.94	0.01	0.059	***	0.038	*	0.019		-0.019	
Simple Poverty Score Card 2014/2015	48.14	0.72	45.65	0.57	-2.497	***	46.31	0.61	44.25	0.56	-2.062	***	-1.828	***	-1.393	***	0.435	
Age	19.33	0.12	19.54	0.13	0.219		22.31	0.13	22.54	0.13	0.238		2.981	***	3	***	0.019	
Aged 15–18	*0.440*	*0.020*	*0.420*	*0.020*	*-0.018*		*0.08*	*0.01*	*0.12*	*0.01*	*0.042*	**	*-0.368*	***	*-0.307*	***	*0.061*	**
Aged 19–25	*0.560*	*0.020*	*0.580*	*0.020*	*0.018*		*0.74*	*0.02*	*0.68*	*0.02*	*-0.066*	**	*0.187*	***	*0.102*	***	*-0.085*	**
Aged 26–35	*0*	*0*	*0*	*0*	*0*		*0.18*	*0.02*	*0.20*	*0.02*	*0.024*		*0.181*	***	*0.205*	***	*0.024*	
Household size	5.84	0.13	5.90	0.11	0.057		6.40	0.13	6.42	0.12	0.026		0.555	***	0.525	***	-0.03	
Household composition: sex ratio	0.40	0.01	0.59	0.01	0.194	***	0.35	0.01	0.54	0.01	0.190	***	-0.049	***	-0.053	***	-0.004	
Overall no. of observations	**526**		**669**				**526**		**669**				**1,052**		**1,338**		**2,390**	

Note: SE = Standard Error of the mean estimate. Asterisks represent level of statistical significance of t-test/chi-squared test of difference in means: *** for p<=0.01; ** for p<=0.05 ; * for p<=0.1. Asterisks in columns (5) and (6) indicate whether the change in means between survey rounds for the sub-sample of men and women, respectively, is significant. The overall N will vary by indicator, the numbers given in the last row are for the complete sample.

Source: authors' calculations using MUVA Urban Youth Survey data.

reproductive work remained stable. The proportion of respondents who use a computer frequently decreased for both groups but significantly so only for women. This may be related to the place of computer usage and COVID-19 prevention measures which saw schools in Mozambique closed between March 2020 and February 2021. In 2017, around 40 per cent of young people that use a computer frequently, reported doing so outside their house. Young women are more likely than men to use a computer at school, while young men are more likely to report using it at someone else's house (Arau et al. 2018). Finally, in contrast to men, women also increased their participation in social groups between 2017–20.

The last column in Table 3 indicates that, with the exception of differential changes in age-group composition, these trends have only varied significantly between women and men in four cases. First, the proportion of female singles has decreased slightly less than the proportion of male singles. As mentioned above, men were more likely to be single in both waves, which means that the two groups resemble each other more on this indicator in 2020 than in 2017. Women have quite significantly increased the number of children they have at a faster rate than men. Perhaps relatedly, they have also had to increase the time they spend on reproductive work by significantly more than men. In both of these cases, this means that the difference between men and women that was already large and significant in 2017 has increased even further. In 2020, women spent on average 5.6 hours on reproductive work, men only 2. Finally, women have caught up with respect to social group membership — the difference in 2017 has disappeared in 2020. Together with results presented in section 3.3.1, these differences in trends are a good indication for how the lives of young men and women in our sample has changed — in quite significantly different ways — during the period we covered with our survey. While some of the changes may be attributed to a 'coming of age' effect as the respondents' age range changed from 15–25 years in the first round to 18–28 years in the third round, others should also be interpreted in the light of COVID-19 measures and impacts that this group experienced in 2020.

4 Estimation approach

The goal of this study is to identify the main determinants of the gender gap in labour market outcomes among young urban Mozambicans, as captured by our survey data and presented in the previous sections. The rich information in this data enables us to pursue two approaches: first, a theory-driven estimation in which we include potential determinants based on our review of the relevant literature presented in section 2. Second, a data-driven approach in which we allow an automated algorithm to identify the most relevant predictors of the gender gap. We then compare and discuss the results from both estimation approaches.

4.1 Theory-driven analysis of the gender gap

We apply a regression-based decomposition of the gender difference in changes in employment between 2017–20 using variables we identified as relevant in the literature review. The decomposition analysis allows us to assess whether the gender gap is driven by differences in characteristics or in coefficients, meaning differences in the behavioural response to a given characteristic. Further, we can observe which characteristics or responses are the main drivers of differences.

In a fixed-effects regression framework, we first specify the probability of being employed at any given time to linearly depend on both a set of observable characteristics of every individual and a time-invariant unobservable fixed effect:

$$P_{i,t}^{g} = \boldsymbol{\beta}\mathbf{X}_{i,t}^{g} + \mu_i + \varepsilon_{i,t}. \quad (1)$$

That is, the employment status, *P*, of individual *i* of gender *g* (f for female, m for male) in period *t* (2017 or 2020) depends on observable time-varying individual characteristics, $\mathbf{X}$, individual unobservable fixed effects, μ_i, and an idiosyncratic error term, ε. We first estimate this relationship using a standard linear fixed effect regression approach, clustering standard errors at the enumeration area level from 2017, which was the original primary sampling unit in our survey. With only two time periods this estimation is equivalent to a first-difference model. By including individual fixed effects in this set up, we control for any time-invariant unobservable characteristics that could influence the change in working status, such as motivation or ability, between 2017–20. Our coefficient estimates ($\widehat{\boldsymbol{\beta}}$) reflect how changes in observable characteristics are associated with a change in employment outcomes among study participants. We estimate this relationship separately for male and female participants to identify the variables that matter for each sub-group separately as well.

It is important to acknowledge that while including individual fixed effects in this estimation allows us to control for endogeneity due to unobservable time-invariant factors, it does not overcome the problem of potential endogeneity due to reverse causality. For example, a significant increase in computer use might be associated with an increase in employment rates because using a computer might support an individual's employment prospects or, alternatively, simply because having a new job might require increased computer use. From our perspective, this implies careful interpretation of our results — not implying definite evidence for causal relationships but rather associations — while acknowledging that controlling for unobservable factors by exploiting the panel structure of our data is a definite improvement over other studies of the gender employment gap that use cross-sectional data; e.g. Xiao (2020) and Gradín and Tarp (2019).

4.1.1 Decomposition

Next, to assess whether changes in observable characteristics have the same influence on the working status for men and women, we apply a Oaxaca-Blinder decomposition (1973). While the canonical decomposition employs linear regression to a single cross-section, we have two waves of a panel, which means that we can exploit variation across time as well as cross-sectional variation and apply the decomposition *after* dealing with the fixed effect component in (1). To do so, we simply take the first difference of the outcome and observable characteristics and then perform the decomposition on the resulting set-up, where $\boldsymbol{\Delta}$ stands for the first-difference transformation of the data:

$$\Delta P_i^{g} = \boldsymbol{\beta}\boldsymbol{\Delta}\mathbf{X}_i^{g} + \Delta\varepsilon_i. \quad (2)$$

Expressed in conditional expectations, the change in the female (male) probability to be working is defined as:

$$\Delta P_i^{g} = E_{\beta^g}\left(\Delta P_i^{g} \middle| \boldsymbol{\Delta}\mathbf{X}_i^{g}\right). \quad (3)$$

There exist various approaches to specify the decomposition. Neumark (2004) and Oaxaca and Ransom (1994) suggest a pooled sample approach that assumes an underlying non-discriminatory coefficient vector, $\boldsymbol{\beta}^*$, allowing for (positive and negative) discrimination toward both groups, men and women. The gender gap in employment, *G*, could then be expressed as follows:

$$G = \left\{E_{\beta^*}\left(\Delta P_i^{f} \middle| \boldsymbol{\Delta}\mathbf{X}_i^{f}\right) - E_{\beta^*}\left(\Delta P_{i,t}^{m} \middle| \boldsymbol{\Delta}\mathbf{X}_i^{m}\right)\right\} + \left\{E_{\beta^f}\left(\Delta P_{i,t}^{m} \middle| \boldsymbol{\Delta}\mathbf{X}_i^{m}\right) - E_{\beta^*}\left(\Delta P_{i,t}^{m} \middle| \boldsymbol{\Delta}\mathbf{X}_i^{m}\right)\right\} + \left\{E_{\beta^*}\left(\Delta P_{i,t}^{m} \middle| \boldsymbol{\Delta}\mathbf{X}_i^{m}\right) - E_{\beta^m}\left(\Delta P_{i,t}^{m} \middle| \boldsymbol{\Delta}\mathbf{X}_i^{m}\right)\right\}. \quad (4)$$

The first component provides the explained part of the gender gap. In other words, it measures how much of the gap is due to differences in changes in women's and men's observable characteristics ($\boldsymbol{\Delta}\mathbf{X}_i^g$). For example, if relatively more young men than young women increased their education between survey waves, this could explain why relatively more young men gained an employment. The second and third part represent the so-called 'discrimination' component, one from a female perspective, the other from the male perspective. They measure how much of the gap is because the same change in a given characteristic is associated with different conditional probabilities to work. For example, if young women who gain a high school education between the survey waves are less likely to also get a job than young men with the same additional education, this cannot be explained by their equal change in education. Instead, this might be due to discrimination or other unobserved factors. In a simple cross-sectional model, such unobserved factors could drive the so-called 'unexplained' or 'discrimination' component of the decomposition. In our application, we can exploit the panel nature of the data to overcome this concern and argue that the fixed effects (or taking first differences) absorb at least the time-invariant unobservables.

In the estimation approach, we follow Jann (2008) and Elder et al. (2010), who suggest that one should include a group dummy in the pooled regression to reduce omitted variable bias. As Elder et al. (2010) illustrate, leaving out the group dummy would overestimate the role of the explained component in the decomposition. To reconcile the various options of decomposing the employment gap, our main analysis will apply the pooled sample including a gender dummy. We apply other approaches to assess robustness of findings to other specifications and present results in the robustness section. The main insights do not change and the small differences we find are in line with the discussions of Lee (2015) and Elder et al. (2010).

We acknowledge that using the term 'discrimination' might overstate the role of actual discrimination against women and underestimate the role of individual preferences (Neumark 2018; Oaxaca 2007). Yet, we also consider the latter to be shaped by social norms and the broader cultural environment that individuals face, so that it would be difficult to fully disentangle preferences from discrimination within the same domain. For example, decisions about, or the ability to make decisions related to, fertility choices are likely a result of both. We carefully interpret results in light of this ambiguity. Furthermore, we extend our analysis by a data-driven approach to overcome some of the constraints when using pre-selected variables from observational data.

4.2 Automated selection of covariates using LASSO

Given the large set of potential covariates that could make up $\mathbf{X_i}$, we also implement an automated covariate selection approach, in which we allow the initial set of covariates to be significantly larger than in the approach presented in the previous section. We follow the approach suggested in Belloni et al. (2016) for this panel setting. This was specifically developed to identify variables of relevance in high-dimensional contexts, i.e., in situations where there are a potentially large set of covariates that could be used to explain variation in the outcome of interest. In our case, as described in section 3.2.2, our survey instrument includes almost 200 questions which translates into over 450 variables. After implementing a set of data management steps described in that section, we end up with a set of 149 potential explanatory variables among which LASSO allows us to select the relevant ones. See Appendix A3 for a full list of these variables.

More specifically, we use the implementation developed for Stata in Ahrens et al. (2020) to implement LASSO fixed-effects regressions in our case. Generally, a LASSO regression works by solving a so-called 'penalised minimization' problem, in which, contrary to a standard OLS regression, not just the residual sum-of-squares (RSS) is minimized, but the RSS is minimised subject to a penalty term formed of the sum of estimated coefficients (Tibshirani 1996). Depending on the penalization chosen, this induces some coefficients to shrink towards zero, i.e. removing these variables from the model. Hence, choosing the right penalty level is crucial for LASSO to perform well. Belloni et al. (2016) adapt this to the panel context, showing that this requires specifying two penalty loadings, rather than only one as is the case in the simple OLS case. Ahrens et al. (2020) show how theory-driven 'rigorous' penalization levels can be chosen for the panel case, which they implement in the 'rlasso' package for Stata and which we use here. It should be noted that the authors emphasise that this approach will typically 'sparse' solutions, i.e., of the many variables that are included in the model a small sub-set will be chosen by this rigorous LASSO.

This means that we estimate equation (1) using a fixed effects transformation and the LASSO approach described in Ahrens et al. (2020), with an initial set of p=149 possible covariates. As before, we cluster standard errors at the EA level and implement these regressions for the male and female sub-sample of study participants first. We also implement a pooled regression in order to be able to compare results.

Importantly, however, we are interested not just in selecting the right set of predictor variables for our outcome of interest (employment status), but we want to estimate how strong this relationship is, i.e., perform post-selection inference. It is important to clarify that this is different from 'classical' inference due to the two-step nature of this procedure: the researcher first looks for important relationships in a first step (e.g., using LASSO) and then wants to assess the strength of this relationship in a second step (e.g., using p-values in a regression). Because of the first step, however, where one has 'looked' at the data already, the threshold for what is significant in the second step needs to be adapted. There is an active literature on how to do so and Bachoc et al. (2019) and, more recently, Zhang et al. (2022) provide an overview. Our reading of this literature is that it is, so far, inconclusive in terms of what the optimal solution for this problem is. On the one hand, some authors suggest different ways of correcting post-selection inference values (confidence intervals, p-values), sometimes depending on the exact selection algorithm (Lee et al. 2016; Berk et al. 2013; Taylor and Tibshirani 2015). On the other hand, other authors claim that some suggested methods are invalid in some contexts (e.g. Bachoc et al. (2019) on Lee et al. (2016)) or that a 'naïve' approach to post-selection inference might work as well as other suggested approaches (Hannes Leeb et al. 2015; Zhao et al. 2017). Zhang et al. (2022) provide the first comprehensive review that systematically compares and discusses different approaches. In our case, we perform 'naïve' post-selection inference for our headline results, selecting a high threshold of $p = 0.01$ as a minimum for anything to be statistically significant.

5 Results and discussion

5.1 Theory-driven approach

Table 4 presents the fixed-effects estimation of the employment status of survey participants, i.e., of (1) specified above, for men and women separately. We present four specifications, adding additional characteristics in each column. Our main results are stable across specifications.[7] While

[7] Note that they are also robust to alternative specifications that check for different ways of testing the relevance of prevalent social norms relating to childbearing and fertility. See Appendix A2.

Table 4 Probability of working for young men and women, fixed effects regression

	Male				Female			
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
	Employed	**Employed**	**Employed**	**Employed**	**Employed**	**Employed**	**Employed**	**Employed**
Age	0.024*	0.024*	0.026*	0.026*	0.011	0.013	0.008	0.007
	(0.012)	(0.012)	(0.012)	(0.011)	(0.012)	(0.011)	(0.012)	(0.012)
Highest grade completed	0.033**	0.034**	0.029**	0.029**	0.001	0.002	0.002	0.002
	(0.010)	(0.010)	(0.010)	(0.010)	(0.012)	(0.011)	(0.011)	(0.011)
Single	-0.021	-0.020	-0.032	-0.033	0.029	0.011	0.027	0.031
	(0.075)	(0.077)	(0.077)	(0.078)	(0.101)	(0.099)	(0.100)	(0.099)
Number of children still alive	0.030	0.036	0.039	0.040	-0.092	-0.070	-0.074	-0.073
	(0.051)	(0.053)	(0.055)	(0.055)	(0.058)	(0.058)	(0.058)	(0.058)
Household size	0.003	0.002	-0.000	0.000	0.028	0.030	0.029	0.030
	(0.020)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.018)	(0.018)
Household composition: sex ratio	-0.035	-0.035	-0.032	-0.036	0.035	0.062	0.053	0.051
	(0.156)	(0.157)	(0.159)	(0.160)	(0.186)	(0.189)	(0.193)	(0.189)
Simple Poverty Score Card 2014/2015	0.005	0.005	0.004	0.004	0.013***	0.012***	0.012***	0.012***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Number of hours spent on reproductive work		-0.006	-0.007	-0.008		-0.026***	-0.021**	-0.022**
		(0.011)	(0.010)	(0.010)		(0.007)	(0.007)	(0.007)
Make decisions about my movement alone			0.012	0.013			-0.024	-0.023
			(0.041)	(0.041)			(0.041)	(0.040)
Frequently uses a computer			0.083	0.085			0.033	0.035
			(0.049)	(0.049)			(0.060)	(0.060)
Member of a social group (excl. church)			0.046	0.043			0.155***	0.156***
			(0.041)	(0.042)			(0.043)	(0.044)
Do others approve of women in leadership				-0.017				-0.005
				(0.035)				(0.044)
Would approve of women in leadership				-0.040				0.075
				(0.053)				(0.071)
Observations	**961**	**961**	**959**	**959**	**1,227**	**1,227**	**1,223**	**1,223**

Notes: standard errors in parentheses are clustered at enumeration area level. Asterisks indicate level of significance: * p<0.05, ** p<0.01, *** p<0.001.

Source: authors' calculations using MUVA Urban Youth Survey data.

not many characteristics seem to play a significant role, we can observe some differences between the role they play for women and men. First, young men who gain additional years of education also improve their employment prospects, and quite significantly so, while for women this does not seem to play a role. Similarly, increases in age seem to be associated with a higher likelihood to be working for men, albeit not for women confirming what we observed in the summary statistics of school-to-work transitions by age-group and gender in section 3.3.

For young women, on the other hand, three different variables are significantly associated with employment status: first, becoming a member of a social group, an indicator for social capital, is strongly and significantly associated with a positive change in employment status. Second, an increase in the poverty score, meaning a reduction of poverty in the household in which she lives, is significantly and positively associated with employment. Finally, increasing the hours in reproductive work (care work, domestic chores) is associated with a significant and large drop in the chance to be working for the women of our sample. None of these variables show up as significant in the male-only specification.

Table 5 presents the results of the pooled decomposition. Note that here we consider the changes between baseline and follow-up period, meaning that we are estimating the decomposition on the first-differenced data as specified in section 4.1. As shown in Figure 2, between 2017 and 2020, employment increased significantly for young men but not for women, which is reflected in a gender gap in the employment probability of 11.3 percentage points, which reflects the differences in trends identified in Figure 2. It appears that differences in observable characteristics cannot explain this gender gap. However, the differential increase in time spent on reproductive work is just about significant and to the disadvantage of young women, supporting existent literature that such tasks fall disproportionately to women and are negatively associated with employment prospects. Similarly, the differential increase in social group membership for women indicates the importance of social capital or social connectedness in relation to employment.

The difference in conditional probabilities, or coefficients, is large and significant. This is the so-called 'discriminatory' part of the gender gap. It explains over 90 per cent per cent of the employment gender gap. In detail, this is driven by 'discrimination' against young women who gain additional education. We find that if young women in our sample increased their education the same as young men, they were still significantly less likely to find employment.

Table 5: Decomposition of employment gender gap, pooled sample including gender dummy

	Pooled sample	
Overall decomposition	Estimate	Standard Error
Female	-0.002	(0.027)
Male	0.111***	(0.028)
Difference (*Gender Gap*)	-0.113**	(0.040)
Observables	-0.002	(0.012)
Discrimination	-0.111**	(0.040)
Detailed decomposition		
Observables		
Change in Age	0.000	(0.000)
Change in Highest grade completed	-0.004	(0.003)
Change in Number of hours spent per day on reproductive work	-0.008†	(0.004)
Change in Single	0.000	(0.002)
Change in Number of children still alive	-0.004	(0.006)

Change in Make decisions about my movement alone	0.000	(0.001)
Change in Frequently uses a computer	-0.001	(0.002)
Change in Member of social group (excl. church)	0.012*	(0.005)
Change in Do others approve of women in leadership	0.000	(0.001)
Change in Would approve of women in leadership	-0.000	(0.000)
Change in Simple Poverty Score Card 2014/2015	0.002	(0.005)
Discrimination		
Change in Age	-0.050	(0.046)
Change in Highest grade completed	-0.028†	(0.016)
Change in Number of hours spent per day on reproductive work	-0.003	(0.004)
Change in Single	-0.006	(0.008)
Change in Number of children still alive	-0.018	(0.014)
Change in Make decisions about my movement alone	-0.007	(0.010)
Change in Frequently uses a computer	0.002	(0.003)
Change in Member of social group (excl. church)	0.005	(0.004)
Change in Do others approve of women in leadership	0.003	(0.006)
Change in Would approve of women in leadership	0.002	(0.002)
Change in Simple Poverty Score Card 2014/2015	-0.010	(0.006)
Observations	1,013	

Notes: standard errors in parentheses are clustered at enumeration area level. Asterisks indicate level of significance: † p<0.1, * p<0.05, ** p<0.01, *** p<0.001.

Source: authors' calculations using MUVA Urban Youth Survey data.

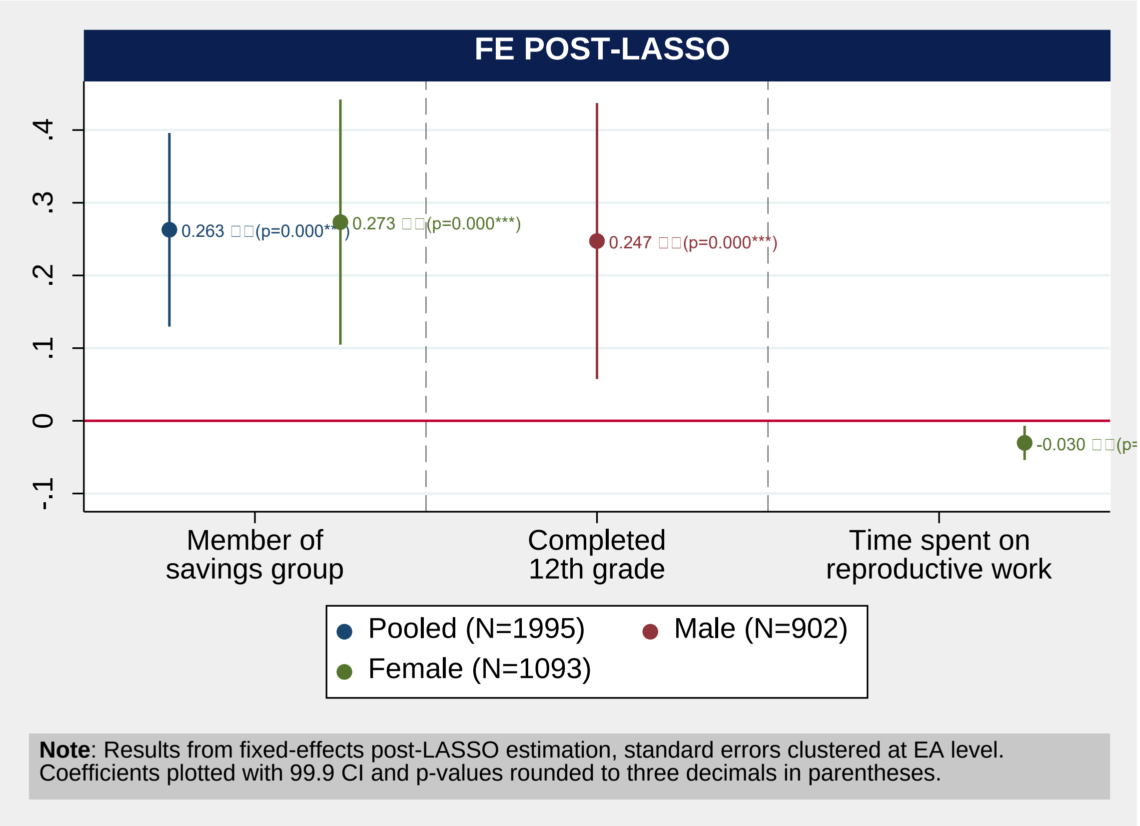
5.2 LASSO and post-LASSO results

We present results from our fixed-effects post-LASSO analysis in Figure 8. As mentioned in section 4.2, these are the results from fixed-effects estimations implemented after 'rigorous' LASSO that take both clustering at the EA level and the fixed-effects structure of equation (1) into account. The results in the below figure hence show the estimated coefficients of the post-LASSO fixed-effects regressions together with — given the post-selection inference issues described in section 4.2 — large confidence intervals of 99.9 per cent. We also show p-values, rounded to three decimals. In the graph, we show results both for a pooled analysis and for two separate regressions for the male and female sub-samples only.

The results can be summarised as follows. First, as expected, the rigorous LASSO approach identifies only a small sub-set of covariates as being relevant predictors for employment status. In fact, for the pooled regression, the approach selects only one predictor of employment status: being a member of a savings group. The post-LASSO analysis indicates that this variable is significantly and positively related to the employment status of survey participants. This indicator is also selected by LASSO and is highly significant in the female-only analysis. In addition, however, LASSO selects the time spent on reproductive work in the female-only analysis, which is significantly and negatively related to employment status for women. Finally, for the male only analysis, LASSO again only selects one variable as being relevant: whether the survey participant completed 12th grade or not. The post-LASSO analysis indicates that this indicator is significantly and positively related to employment for men.

Overall, these findings broadly support the results derived from our theory-driven approach. On the one hand, men's likelihood to be employed is associated with higher education achievements, which is not the case for women. On the other hand, there is strong evidence that women – but not men – benefit from becoming members of a social group, and in particular, savings groups, which seems to be the group membership driving the significant relationship between the more general 'group membership' indicator and employment in the previous section. Similarly, women's employment status is negatively associated to the time they spend on reproductive work, which is not the case for men, again confirming the above results.

Figure 8: Fixed-effects post-LASSO results



Source: authors' calculations using MUVA Urban Youth Survey data.

Two significant relationships identified by our theory-driven analysis are not picked up by this analysis: First, the change in poverty score card values, which the fixed-effects regressions identify to be relevant for women. Second, the age variable, which our fixed-effects regressions identify to be significantly related to employment status for the male sub-sample. Both, however, are not identified to be significant drivers of the gender gap in employment in our Oaxaca-Blinder decomposition. For both, we interpret the inconclusiveness of the evidence to indicate that these relationships are not particularly robust and hence that we cannot report with confidence that these variables are determining or are associated with young people's employment status – whether women or men – in our sample.

5.3 Robustness

5.3.1 Attrition

Following up with respondents from our baseline sample over time provided challenging as young people tend to move out of their homes when becoming adults. Especially in Maputo City, many move to the neighbouring urbanized areas. Our tracking protocol during the field work instructed enumerators to try to find out the new address of those who had moved and follow them there if that new address was within our original sampling areas. The reasoning was that these areas were representative of the Mozambican urban poor and that this approach was feasible within survey budget constraints.

Table 6 presents the mean of baseline characteristics of the main sample containing all individuals who were successfully tracked from baseline to Round 3 (column 1) and of those who were sampled but not encountered in Round 3 (column 2). Column 3 presents the difference in means and asterisks indicate whether this difference is statistically significant.

We observe that, on average, attritors are more likely to be slightly older women, with more children, more decision-making power about their own mobility from slightly smaller households. Fewer of them are single. This could fit a profile of young women moving away from their baseline household.

Table 6: Characteristics of individuals successfully tracked or lost from the baseline to the follow-up survey

	(1)	(2)	(3)
	Tracked	Lost	Difference
	Mean	Mean	in means
Male	44.02	38.40	-5.62*
Employed	61.08	62.61	1.53
Highest grade completed	8.99	8.95	-0.04
Number of children still alive	0.37	0.45	0.08*
Single	82.09	76.29	-5.80**
Frequently uses a computer	29.87	32.39	2.51
Spent time on reproductive work	84.44	86.48	2.04
Make decisions about my movement alone	43.50	49.25	5.74*
Do others approve of women in leadership	41.42	38.56	-2.86
Would approve of women in leadership	88.95	88.81	-0.14
Youth in bairro approve of women in leadership	88.84	89.04	0.19
Member of social group (excl. church)	13.05	14.52	1.47
Age	19.45	19.86	0.42**
City	1.46	1.50	0.03
Household size	5.87	5.36	-0.51***
Household composition: sex ratio	0.51	0.51	0.00
Simple Poverty Score Card 2014/2015	46.74	47.57	0.83
Observations	1,195	599	1,794

Notes: standard errors in parentheses are clustered at enumeration area level. Asterisks indicate level of significance: * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

Source: authors' calculations using MUVA Urban Youth Survey data.

The attrition profile could influence our results primarily due to its influence on the gender imbalance of the sample. If relatively more women than men left the sample, it is skewed towards a more 'male-looking' sample. In our decomposition analysis, this could influence the role of the explained component as the sample will look relatively more like men than it should and thus attribute more influence on these observable gender differences than it would otherwise; see Elder et al. (2010). It could also influence the unexplained part by underestimating the 'discriminatory' effect as those women who left the sample might have done so because of the factors that we identify under the unexplained part.

Table 7: Oaxaca-Blinder decomposition with main sample and full sample, pooled approach

	Female proportion reweighted		Inverse probability weighted	
	Estimate	S.E.	Estimate	S.E.
Overall decomposition				
Female	-0.002	(0.027)	-0.001	(0.028)
Male	0.111***	(0.028)	0.123***	(0.029)
Difference (Gender gap)	-0.113***	(0.040)	-0.124***	(0.040)
Observables	-0.002	(0.012)	0.000	(0.012)
Discrimination	-0.111***	(0.040)	-0.124***	(0.040)
Observables				
Change in Age	0.000	(0.000)	-0.000	(0.000)
Change in Highest grade completed	-0.004	(0.003)	-0.005	(0.003)
Change in Number of hours spent per day on reproductive work	-0.008*	(0.004)	-0.007	(0.004)
Change in Single	0.000	(0.002)	0.001	(0.001)
Change in Number of children still alive	-0.004	(0.006)	-0.005	(0.006)
Change in Make decisions about my movement alone	0.000	(0.001)	0.001	(0.002)
Change in Frequently uses a computer	-0.001	(0.002)	-0.002	(0.002)
Change in Member of social group (excl. church)	0.013**	(0.006)	0.014**	(0.006)
Change in Do others approve of women in leadership	0.000	(0.001)	0.000	(0.001)
Change in Would approve of women in leadership	0.000	(0.000)	0.000	(0.001)
Change in Simple Poverty Score Card 2014/2015	0.002	(0.005)	0.002	(0.004)
Discrimination				
Change in Age	-0.050	(0.046)	-0.053	(0.046)
Change in Highest grade completed	-0.028*	(0.016)	-0.031*	(0.017)
Change in Number of hours spent per day on reproductive work	-0.003	(0.004)	-0.003	(0.004)
Change in Single	-0.006	(0.008)	-0.009	(0.010)
Change in Number of children still alive	-0.018	(0.013)	-0.021*	(0.013)
Change in Make decisions about my movement alone	-0.007	(0.010)	-0.010	(0.012)
Change in Frequently uses a computer	0.002	(0.003)	0.001	(0.002)
Change in Member of social group (excl. church)	0.005	(0.004)	0.006	(0.004)
Change in Do others approve of women in leadership	0.003	(0.006)	0.004	(0.006)
Change in Would approve of women in leadership	0.002	(0.002)	0.002	(0.002)
Change in Simple Poverty Score Card 2014/2015	-0.010	(0.007)	-0.008	(0.006)
Observations	1,013		1,013	

Notes: standard errors in parentheses are clustered at enumeration area level. Asterisks indicate level of significance: * p<0.05, ** p<0.01, *** p<0.001.

Source: authors' calculations using MUVA Urban Youth Survey data.

However, for our main outcome, we cannot directly use a 'non-attritor' sample. Instead, we can think of a sensible bound to test how much our results change if we assume a relatively larger proportion of women in the sample than currently observed (as the attrition analysis indicated that a larger proportion of them left the sample). We can then reweight the female observations in the sample to replicate the 'original' sample without attrition in terms of the female proportion. Applying such a reweighting approach to our case, the female proportion of respondents increases from 56 to 62 per cent, meaning an increase by a factor of 1.107. Thus, for our first robustness check, we assign each woman in the main analysis sample a weight of 1.107 and each man a weight of 1 and re-run the analysis. Table 7 columns 1 and 2 present the resulting decomposition results. We observe that the results remain almost identical to our main results from Table 5.

However, this does not yet consider the difference in other characteristics, such as freedom of movement or number of children, between individuals who remain in our sample and those who drop out between Round 1 and Round 3. These could influence both attrition and the gender gap in employment. To correct for this, we apply an inverse probability weighting approach. First, we estimate the probability to be an attritor based on baseline characteristics. Then we take the inverse of the predicted probability for everyone as weight in the decomposition analysis (Wooldridge 2002). The intuition is that we aim to make the sample look more like the originally sampled baseline sample, including attritors. Results are presented in columns 3 and 4 of Table 5. There are some small differences in the size of the gender gap and of the contribution of each significant characteristic. First, the employment gender gap is a bit larger now, with 12.4 percentage points. The coefficient on spending time in reproductive work is no longer significant in our reweighted sample, but the discrimination for education remains and the number of children now enters significantly. Based on the attrition profile presented in Table 6, those leaving the sample were more likely to be women and with more children. Thus, the 'discriminatory' effect against having more children would be even larger if they had stayed in the main sample. The education result is in line with our findings from the LASSO analysis: men are more likely to be employed with higher education, whereas it does not appear to matter for women.

5.3.2 Decomposition: change the counterfactual in the decomposition

Table 8 presents the results of three different counterfactual scenarios to decompose the gender gap in employment, compared to our headline results presented in Table 5. This exercise also addresses the concern of attrition bias discussed above. The top part of the table presents the overall decomposition of the gap into its explained and discriminatory part. There is no significant difference in terms of estimates across these different scenarios — whether we consider a female (1) or a male (2) counterfactual, or a pooled one (3) that excludes a gender dummy. All estimates fall within less than one standard error of each other and our headline findings. As before, the discriminatory component is the significant one across the different models.

Looking at the detailed decomposition, for the female and male counterfactuals respectively, the following key findings need to be highlighted. First, for the female counterfactual, i.e., if all men were facing the same conditional probabilities as women, none of the observable characteristics are identified as significant. However, in the discriminatory part, social group membership enters significantly and positively, meaning that similar changes in social group membership are more significantly related to employment for women than for men. For the male counterfactual, i.e., if all women were facing the same conditional probabilities as men, two observable characteristics are identified as significant: changes in the number of hours spent on reproductive work and the change in social group membership. The sign of the social group membership coefficient is positive, which means that increases in this variable are related to increases in female employment. On the other hand, increases in the hours spent on reproductive work are related to decreases in female employment. In the discriminatory part, only education enters significantly and negatively and with a larger coefficient than in the female counterfactual scenario.

Table 8: Oaxaca-Blinder decomposition of probability of working, female or male counterfactual, pooled sample excluding gender dummy

	Female coefficient		Male coefficient		Pooled sample excl. gender	
	(1)		(2)		(3)	
Overall decomposition	Estimate	SE	Estimate	SE	Estimate	SE
Female	-0.002	(0.027)	-0.002	(0.027)	-0.002	(0.027)
Male	0.111***	(0.028)	0.111***	(0.028)	0.111***	(0.028)
Difference (*Gender Gap*)	-0.113**	(0.040)	-0.113**	(0.040)	-0.113**	(0.040)
Observables	-0.002	(0.013)	0.004	(0.016)	-0.007	(0.012)
Discrimination	-0.110**	(0.039)	-0.117**	(0.041)	-0.105**	(0.038)
Detailed decomposition						
Observables						
Change in Age	0.001	(0.001)	0.000	(0.000)	0.000	(0.001)
Change in Highest grade completed	-0.008	(0.005)	-0.001	(0.003)	-0.004	(0.003)
Change in Number of hours spent per day on reproductive work	-0.004	(0.005)	-0.009†	(0.005)	-0.008†	(0.005)
Change in Single	-0.001	(0.003)	0.003	(0.003)	-0.000	(0.002)
Change in Number of children still alive	0.006	(0.008)	-0.009	(0.008)	-0.006	(0.006)
Change in Make decisions about my movement alone	-0.001	(0.002)	0.001	(0.002)	0.000	(0.001)
Change in Frequently uses a computer	-0.002	(0.003)	-0.001	(0.002)	-0.001	(0.002)
Change in Member of social group (excl. church)	0.005	(0.005)	0.017*	(0.008)	0.011*	(0.005)
Change in Do others approve of women in leadership	0.001	(0.002)	0.000	(0.002)	0.000	(0.001)
Change in Would approve of women in leadership	-0.000	(0.001)	0.000	(0.001)	-0.000	(0.000)
Change in Simple Poverty Score Card 2014/2015	0.001	(0.003)	0.003	(0.007)	0.002	(0.005)
Discrimination						
Change in Age	-0.051	(0.046)	-0.050	(0.046)	-0.051	(0.046)
Change in Highest grade completed	-0.024†	(0.014)	-0.031†	(0.018)	-0.028†	(0.016)
Change in Number of hours spent per day on reproductive work	-0.007	(0.008)	-0.001	(0.002)	-0.002	(0.004)
Change in Single	-0.005	(0.006)	-0.009	(0.010)	-0.006	(0.008)
Change in Number of children still alive	-0.028	(0.021)	-0.013	(0.010)	-0.016	(0.014)
Change in Make decisions about my movement alone	-0.006	(0.008)	-0.007	(0.011)	-0.007	(0.009)
Change in Frequently uses a computer	0.002	(0.004)	0.001	(0.002)	0.002	(0.003)
Change in Member of social group (excl. church)	0.013†	(0.007)	0.000	(0.003)	0.006†	(0.004)
Change in Do others approve of women in leadership	0.002	(0.005)	0.003	(0.007)	0.003	(0.006)
Change in Would approve of women in leadership	0.002	(0.002)	0.002	(0.003)	0.002	(0.002)
Change in Simple Poverty Score Card 2014/2015	-0.009	(0.006)	-0.011	(0.007)	-0.010	(0.006)
Observations	1,013		1,013		1,013	

Notes: standard errors in parentheses are clustered at enumeration area level. Asterisks indicate level of significance: † p<0.1, * p<0.05, ** p<0.01, *** p<0.001.

Source: authors' calculations using MUVA Urban Youth Survey data.

Lastly, in column (3) we also present the decomposition of a pooled sample excluding the gender dummy. This specification is prone to omitted variable bias overestimating the contribution of the explained component (Elder et al. 2010). In our case, the results suggest that the explained (observable) part remains insignificant and like the other estimations, although indeed largest in absolute values. The results of the detailed decomposition are almost identical to our main approach (pooled sample incl. gender dummy).

6 Conclusion and policy implications

We use a unique longitudinal dataset that tracks a cohort of young, urban Mozambicans over a period of three years between 2017 and 2020 to analyse how their employment changes over time and, in particular, how trends vary between male and female respondents. Our results show that employment outcomes differ significantly between young men and women in our survey and that these differences increase from about 13 percentage points in 2017 to 23 percentage points in 2020, a large and significant increase of over 75 per cent. Most of this increase is due to increases in employment among men in the youngest cohort of our sample (aged 15–18 years in 2017). Further, men in our sample increasingly transition away from some form of schooling into employment, which is not the case for women. This finding indicates that gaps in employment outcomes between men and women in urban Mozambique materialise early in their economic lives and that these gaps tend to increase over time, as young people transition into adulthood.

Our results indicate that these differences in trends in employment rates can be explained in part by observable differences in the lives and characteristics of young people but also, to a seemingly larger extent, by differences in how the labour market rewards or punishes certain characteristics, depending on whether you are a man or a woman. For instance, we find that being a member of a social group — especially a savings group — is significantly positively related to employment for women but not for men. In addition, we find very strong evidence that young women spend significantly more time on reproductive work than men, that this gap increases further as they get older and that that this is negatively related to employment rates among women. Hence, women are less likely to be employed over time while also being more likely to spend more time on domestic chores or caring for children and relatives, compared to men. In addition, we find strong evidence that achieving higher levels of education – particularly finishing secondary school – is significantly and positively related to employment for men, but not for women. This means that our analysis indicates that even if women were to achieve the same level of schooling as men, they would still be less likely to be employed, all else being equal. In fact, our decomposition results indicate that this 'discriminatory' effect explains much of the divergence in employment trends between men and women in our sample. Similarly, Jones et al. (2020, 2021) find that even young Mozambican women in urban areas with higher education, such as technical/vocational training or university degrees, have more difficulties finding an employment after their graduation than their male counterparts. Jointly these results indicate constraints on the demand side of the Mozambican labour market with potentially some discrimination against (young) women.

Overall, these results indicate that employment gaps among youth in urban Mozambique are partially driven by norms that dominate young women's lives. On the one hand, young women are expected to spend more time on domestic work than men, which harms their employment prospects. On the other hand, even if they achieve similar education levels as men, they are not rewarded in a similar way as men by positive employment prospects. Our results are in line with Gradín and Tarp (2019), concerning the important role of education and social norms for the gender employment gap in Mozambique. The findings lend additional strength to the argument of organizations like MUVA that aim to foster women's economic empowerment and strive to achieve gender parity in employment not only by improving women's 'observable' conditions, for example by working on getting more young women into education for longer, but also by focussing on social norms around women's participations in the labour force. Tackling these norms is imperative to achieve their goals.

We acknowledge that our study has some limitations that future research could address. First, our findings are not representative for urban Mozambique overall. While Beira and Maputo are among the four biggest cities in the country according to the 2017 census, there are large young urban populations in other cities as well, such as for example Matola and Nampula. It is possible, therefore, that our insights would vary if conducted using similar survey data from those places. Second, our main analysis only considered overall employment due to sample size concerns, not differentiating between wage employment, self-employment, and unremunerated employment. It is likely that the determinants of the gender gap might vary depending on the employment type that one looks at. Third, despite the longitudinal nature of our data, we cannot fully overcome endogeneity concerns. Some variables, such as the use of computers, might be a result of the employment status and not its driver. Future research that employs experimental approaches could resolve such concerns.

References

Ahrens, A., Hansen C.B., and Schaffer M.E. (2020). 'lassopack: Model selection and prediction with regularized regression in Stata'. *The Stata Journal*, 20(1): 176–235. https://doi.org/10.1177/1536867X20909697

Allen, R., J. Bischler, J., and Jasper, P. (2020). Assessing Everyday Literacy and Numeracy Skills in Young Adults in Mozambique. *MUVA Technical Report.*

Annan, J., Donald, A., Goldstein, M., Gonzalez Martinez, P., and Koolwal, G. (2019). 'Taking Power: Women's Empowerment and Household Well-Being in Sub-Saharan Africa'. Working Paper. World Bank. https://doi.org/10.1596/1813-9450-9034

Antonio, A., and Tuffley, D. (2014). 'The Gender Digital Divide in Developing Countries'. *Future Internet,* 6(4): Article 4. https://doi.org/10.3390/fi6040673

Arau, A., Bhalla, G., Bischler, J., Gronbaek, A., and Jasper, P. (2018). *MUVA Urban Youth Survey*. OPM. https://muvamoz.co.mz/wp-content/uploads/2019/08/MUVA-Urban-Youth-Survey-Report_final.pdf

Arsalan, A., Tschirley, D., Dolislager, M., Reardon, T., Fox, L., Liverpool-Tasie, L. S., and Christine, S. (2019). IFAD RESEARCH SERIES 43, 'Youth Agrifood System Employment in Developing Countries: A Gender-Differentiated Spatial Approach'. https://doi.org/10.2139/ssrn.3520554

Bachoc, F., Leeb, H., and Pötscher, B.M. (2019). 'Valid confidence intervals for post-model-selection predictors'. *Annals of Statistics*, 47(3): 1475–504. https://doi.org/10.1214/18-AOS1721

Bahia, K., Castells, P., Cruz, G., Masaki, T., Rodríguez-Castelán, C., and Sanfelice, V. (2021). 'Mobile Broadband Internet, Poverty and Labor Outcomes in Tanzania', (SSRN Scholarly Paper ID 3921510). Social Science Research Network. https://doi.org/10.2139/ssrn.3921510

Bandiera, O., Elsayed, A., Smurra, A., and Zipfel, C. (2022). 'Young Adults and Labor Markets in Africa'. *Journal of Economic Perspectives*, 36(1): Article 1. https://doi.org/10.1257/jep.36.1.81

Becker, G.S. (1962). 'Investment in Human Capital: A Theoretical Analysis. In Investment in Human Beings', *Journal of Political Economy*, Vol. LXX, No. 5, Part 2: 9–49). https://doi.org/10.1086/258724

Behrman, J., and Gonalons-Pons, P. (2020). 'Women's employment and fertility in a global perspective (1960–2015)', *Demographic Research*, 43: 707–44. https://doi.org/10.4054/DemRes.2020.43.25

Belloni, A., Chernozhukov, V., Hansen, C., and Kozbur, D. (2016). 'Inference in High-Dimensional Panel Models With an Application to Gun Control', *Journal of Business and Economic Statistics*, 34(4): 590–605. https://doi.org/10.1080/07350015.2015.1102733

Berk, R.., L. Brown, A. Buja, K. Zhang, and L. Zhao (2013). 'Valid post-selection inference', *The Annals of Statistics*, 41(2): 802–37. https://doi.org/10.1214/12-AOS1077

Bloom, D. E., Canning, D., and Sevilla, J. (2003). 'The demographic dividend: A new perspective on the economic consequences of population change'. Rand. https://doi.org/10.7249/MR1274

Bongaarts, J., Blanc, A. K., and McCarthy, K. J. (2019). 'The links between women's employment and children at home: Variations in low- and middle-income countries by world region', *Population Studies*, 73(2): 149–63. https://doi.org/10.1080/00324728.2019.1581896

Boserup, E., Tan, S.F., and Toulmin, C. (2013). *Women's Role in Economic Development*. Abingdon-on-Thames: Routledge. https://doi.org/10.4324/9781315065892

Boudet, A.M.M., Petesch, P., and Turk, C. (2013). *On norms and agency: Conversations about gender equality with women and men in 20 countries*. Washington DC: World Bank Publications.

Cazes, S., and Verick, S. (2013). *The labour markets of emerging economies: Has growth translated into more and better jobs?* London: Palgrave Macmillan. https://doi.org/10.1057/9781137325358

Chakraborty, T., Mukherjee, A., Rachapalli, S. R., and Saha, S. (2018). 'Stigma of sexual violence and women's decision to work', *World Development*, 103: 226–38. https://doi.org/10.1016/j.worlddev.2017.10.031

Chari, A. V., Heath, R., Maertens, A., and Fatima, F. (2017). 'The causal effect of maternal age at marriage on child wellbeing: Evidence from India', *Journal of Development Economics*, 127: 42–55. https://doi.org/10.1016/j.jdeveco.2017.02.002

Charmes, J. (2019). The Unpaid Care Work and the Labour Market. An analysis of time use data based on the latest World Compilation of Time-use Surveys. International Labour Office: Geneva.

Clark, S., Kabiru, C.W., Laszlo, S., and Muthuri, S. (2019). 'The Impact of Childcare on Poor Urban Women's Economic Empowerment in Africa', *Demography*, 56(4): 1247–72. https://doi.org/10.1007/s13524-019-00793-3

Doepke, M., Hannusch, A., Kindermann, F., and Tertilt, M. (2022). 'The Economics of Fertility: A New Era'. NBER Working Paper 29948. https://doi.org/10.3386/w29948

Egger, E.-M., Arslan, A., and Zucchini, E. (2021). 'Does connectivity reduce gender gaps in off-farm employment? Evidence from 12 low- and middle-income countries' WIDER Working Paper 3/2021 UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/2021/937-2

Elder, T.E., Goddeeris, J.H., and Haider, S. J. (2010). 'Unexplained gaps and Oaxaca–Blinder decompositions', *Labour Economics*, 17(1): 284–90. https://doi.org/10.1016/j.labeco.2009.11.002

Eyike Mbongo, L. de B. (2019). 'ICT diffusion and employment in Africa'. *Economics Bulletin*, 39(1): 521–32.

Field, E., Jayachandran, S., and Pande, R. (2010). 'Do traditional institutions constrain female entrepreneurship? A field experiment on business training in India', *American Economic Review,* 100(2): 125–29. https://doi.org/10.1257/aer.100.2.125

Filmer, D., and Fox, L. (2014). *Youth employment in sub-Saharan Africa*. Washington, DC: World Bank Publications. https://doi.org/10.1596/978-1-4648-0107-5

Fox, L. (2019). 'Economic participation of rural youth: What matters?', IFAD RESEARCH SERIES 46. https://doi.org/10.2139/ssrn.3521170

Gaddis, I., and Klasen, S. (2014). 'Economic development, structural change, and women's labor force participation', *Journal of Population Economics,* 27(3): 639–81. https://doi.org/10.1007/s00148-013-0488-2

Gaddis, I., Oseni, G., Palacios-Lopez, A., and Pieters, J. (2020). Who Is Employed? Evidence from Sub-Saharan Africa on Redefining Employment', *Journal of African Economies*. https://doi.org/10.1093/jae/ejac021

Galor, O., and Weil, D.N. (1996). 'The Gender Gap, Fertility, and Growth'. *The American Economic Review*, 86(3): 374–87.

Goldin, C. (1995). Investment in Women's Human Capital and Economic Development'. In T.P. Schultz (ed) *Investment in Women's Human Capital*. Chicago: University of Chicago Press.

Gradín, C., and Tarp, F. (2019). 'Gender Inequality in Employment in Mozambique', *South African Journal of Economics*, 87(2): 180–99. https://doi.org/10.1111/saje.12220

Hannes, L., B.M. Pötscher, and K. (2015). 'On Various Confidence Intervals Post-Model-Selection', *Statistical Science,* 30(2): 216–27. https://doi.org/10.1214/14-STS507

Hjort, J., and Poulsen, J. (2019). 'The Arrival of Fast Internet and Employment in Africa', *American Economic Review*, 109(3): 1032–79. https://doi.org/10.1257/aer.20161385

Idowu, O.O., and Owoeye, T. (2019). 'Female Labour Force Participation in African Countries: An Empirical Analysis', *Indian Journal of Human Development*, 13(3): 278–93. https://doi.org/10.1177/0973703019895234

IFAD (2019). *Creating opportunities for rural youth*. IFAD 2019 Rural Development Report. https://www.ifad.org/documents/38714170/41190221/RDR2019_Overview_e_W.pdf/699560f2-d02e-16b8-4281-596d4c9be25a

ILO (1982). *Resolution Concerning Statistics of the Economically Active Population, Employment, Unemployment and Underemployment*, Adopted by the Thirteenth International Conference of Labour Statisticians. October. Geneva: ILO.

ILO (2013). Resolution Concerning Statistics of Work, Employment and Labour Underutilization. Geneva: ILO. http://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/normativeinstrument/wcms_230304.pdf

Jann, B. (2008). 'The Blinder–Oaxaca Decomposition for Linear Regression Models. *The Stata Journal*, 8(4): 453–479. https://doi.org/10.1177/1536867X0800800401

Jayachandran, S. (2020). Social Norms as a Barrier to Women's Employment in Developing Countries (SSRN Scholarly Paper ID 3637751). Social Science Research Network. https://doi.org/10.2139/ssrn.3637751

Jones, S., Santos, R., and Xirinda, G. (2020). *Final report on the survey on the education-employment transitions of university finalists*. Relatório Final do Inquérito à Transição Ensino-Emprego dos Finalistas Universitários. UNU-WIDER, University of Copenhagen, University of Eduardo Mondlane.

Jones, S., Santos, R., and Xirinda, G. (2021). Survey on the school-to-work transition of technical and vocational training graduates in Mozambique. Final report. Helsinki: UNU-WIDER.

Kabeer, N. (2008). 'Paid work, women's empowerment and gender justice: Critical pathways of social change'. Monograph No. 3. Institute of Development Studies. http://www.pathwaysofempowerment.org/

Klasen, S. (1999). 'Does gender inequality reduce growth and development: Evidence from cross-country regressions'. https://doi.org/10.5282/UBM/EPUB.1602

Klasen, S. (2019). 'What Explains Uneven Female Labor Force Participation Levels and Trends in Developing Countries?', *World Bank Research Observer,* 34(2): 161–197. https://doi.org/10.1093/wbro/lkz005

Klasen, S., and Lamanna, F. (2009). 'The Impact of Gender Inequality in Education and Employment on Economic Growth: New Evidence for a Panel of Countries'. *Feminist Economics*, 15(3): 91–132. https://doi.org/10.1080/13545700902893106

Lee, B.S., Jang, S., and Sarkar, J. (2008). 'Women's labor force participation and marriage: The case of Korea'. *Journal of Asian Economics*, 19(2): 138–154. https://doi.org/10.1016/j.asieco.2007.12.012

Lee, M.-J. (2015). 'Reference parameters in Blinder-Oaxaca decomposition: Pooled-sample versus intercept-shift approaches', *Journal of Economic Inequality*, 13(1): 69–82. https://doi.org/10.1007/s10888-014-9293-8

Lee, J.D., DL. Sun, Yuekai Sun, and Jonathan E. Taylor. (2016). 'Exact post-selection inference, with application to the lasso', *The Annals of Statistics*, 44(3): 907–27. https://doi.org/10.1214/15-AOS1371

Mariara, J., McKay, A., Newell, A., and Rienzo, C. (2018). 'Gender gaps in the path to adulthood for young females and males in six African countries from the 1990s to the 2010s', *IZA Journal of Development and Migration*, 8(1): Article 1. https://doi.org/10.1186/s40176-018-0124-8

Maxwell, N. L., and Wozny, N. (2021). 'Gender Gaps in Time Use and Labor Market Outcomes: What's Norms Got to Do with it?', *Journal of Labor Research*, 42(1): 56–77. https://doi.org/10.1007/s12122-020-09306-3

Mincer, J.A. (1974). *Schooling and Earnings. In Schooling, Experience, and Earnings* . NBER. https://www.nber.org/books-and-chapters/schooling-experience-and-earnings/schooling-and-earnings.

Neumark, D. (2004). Employers' discriminatory behavior and the estimation of wage discrimination. In *Sex Differences in Labor Markets*. Abingdon-on-Thames: Routledge.

Neumark, D. (2018). 'Experimental Research on Labor Market Discrimination', *Journal of Economic Literature*, 56(3): Article 3. https://doi.org/10.1257/jel.20161309

Nikulin, D. (2017). 'The Impact of ICTs on Women's Economic Empowerment' In H. Kaur, E. Lechman, and A. Marszk (eds) *Catalyzing Development through ICT Adoption: The Developing World Experience*. New York: Springer. https://doi.org/10.1007/978-3-319-56523-1_2

Oaxaca, R.L. (1973). Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review*, 14, 693–709. https://doi.org/10.2307/2525981

Oaxaca, R.L. (2007). 'The challenge of measuring labor market discrimination against women', *Swedish Economic Policy Review*, 14(1): Article 1.

Oaxaca, R.L., and Ransom, M.R. (1994). 'On discrimination and the decomposition of wage differentials', *Journal of Econometrics*, 61(1): 5–21. https://doi.org/10.1016/0304-4076(94)90074-4

Perez-Alvarez, M., and Favara, M. (2020). 'Early Motherhood and Offspring Human Capital in India', CSAE Working Paper Series No. 2020/15. Centre for the Study of African Economies, University of Oxford. https://ideas.repec.org/p/csa/wpaper/2020-15.html.

Quisumbing, A.R. (2003). Household decisions, gender, and development: A synthesis of recent research. Household Decisions, Gender, and Development: A Synthesis of Recent Research. https://www.cabdirect.org/cabdirect/abstract/20043095514.

Rebollo-Sanz, Y.F., and Rica, S. de la. (2020). 'From gender gaps in skills to gender gaps in wages: Evidence from PIAAC', Working Paper 20.09. Universidad Pablo de Olavide, Department of Economics. https://ideas.repec.org/p/pab/wpaper/20.09.html.

Schreiner, M. (2017). Simple Poverty Scorecard® Poverty-Assessment Tool, Urban Maputo and Beira. MUVA MEL Team and Mark Schreiner.

Schultz, T.W. (1961). 'Investment in Human Capital',*American Economic Review*, 51(1): 1–17.

Siddique, Z. (2018). 'Violence and female labor supply', IZA Discussion Paper No. 11874. http://dx.doi.org/10.2139/ssrn.3273713

Taylor, J., and Tibshirani, R.J. (2015). *Statistical learning and selective inference. Proceedings of the National Academy of Sciences,* 112(25): 7629. https://doi.org/10.1073/pnas.1507583112

Tibshirani, R. (1996). 'Regression shrinkage and selection via the lasso', *Journal of the Royal Statistical Society: Series B* (Methodological): 58(1): 267–88. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Tshukudu, M. (2019). 'Impacts of Access to ICTs on Employment Status in Botswana', Working Paper No. 68. Botswana Institute for Development Policy Analysis. https://EconPapers.repec.org/RePEc:bid:wpaper:68.

Valberg, S. (2020). 'ICT, Gender, and the Labor Market: A Cross-Country Analysis'. In D. Maiti, F. Castellacci, and A. Melchior (eds) *Digitalisation and Development: Issues for India and Beyond.* New York: Springer. https://doi.org/10.1007/978-981-13-9996-1_15

Verick, S. (2014). 'Female labor force participation in developing countries', *IZA World of Labor.* https://doi.org/10.15185/izawol.87

Viollaz, M., and Winkler, H. (2021). 'Does the Internet Reduce Gender Gaps? The Case of Jordan', *Journal of Development Studies*, 0(0): 1–18. https://doi.org/10.1080/00220388.2021.1965127

Wooldridge, J.M. (2002). 'Inverse probability weighted M-estimators for sample selection, attrition, and stratification', *Portuguese Economic Journal*, 1(2): 117–39. https://doi.org/10.1007/s10258-002-0008-x

Xiao, S., and Asadullah, M.N. (2020). 'Social Norms and Gender Differences in Labor Force Participation in China'. *Feminist Economics*, 26(4): 114–48. https://doi.org/10.1080/13545701.2020.1758337

Zhang, D., Khalili, A., and Asgharian, M. (2022). 'Post-model-selection inference in linear regression models: An integrated review', *Statistics Surveys,* 16: 86–136. https://doi.org/10.1214/22-SS135

Zhao, S., Witten, D., and Shojaie, A. (2017). 'In Defense of the Indefensible: A Very Naive Approach to High-Dimensional Inference',. arXiv. https://doi.org/10.48550/ARXIV.1705.05543

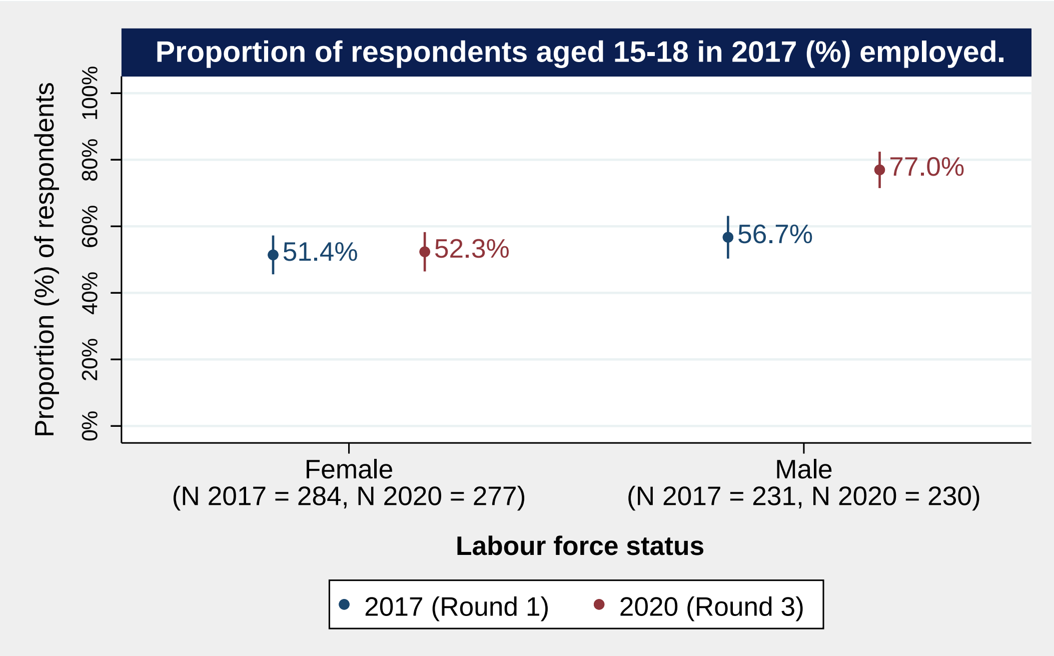
Appendices

A1: Employment gap by age categories

To provide some more background information on the employment transition that survey respondents experienced between 2017–20, we disaggregate the results presented in Figure 2 by whether respondents were 15–18 years old in 2017 (Figure 9) or 19–25 (Figure 10). The main objective of this analysis is to assess whether the broad trends presented in our main analysis vary by age group.

The figures below show that, broadly, the main observed trends hold up: the gender employment gap increases for both age groups over time, although this increase varies in size. For the younger age group (15–18 years old in 2017), the gap increases from an insignificant 5 per cent to 25 per cent, i.e., by twenty percentage points. For the older age group (19–25 in 2017), it increases from 20 per cent to 22 per cent, i.e., by only two percentage points. In essence, most of the increase in the gender gap observed in the aggregate analysis in section 3.3.1 is driven by changes in employment status among men in the younger age group, who catch up with their older peers. The same does not hold true for the female survey respondents in our study.

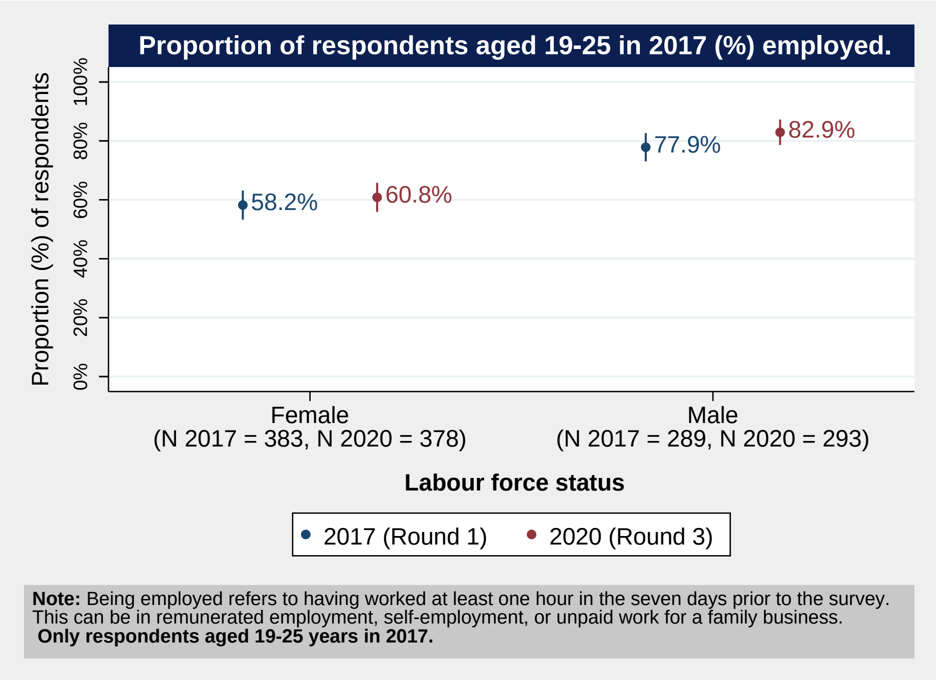
Figure 9: Employment rate of young women and men aged 15–18 in 2017, by survey wave (%)



Note: Being employed refers to having worked at least one hour in the seven days prior to the survey. This can be in remunerated employment, self-employment, or unpaid work for a family business.
Only respondents aged 15-18 years in 2017.

Source: authors' calculations using MUVA Urban Youth Survey data.

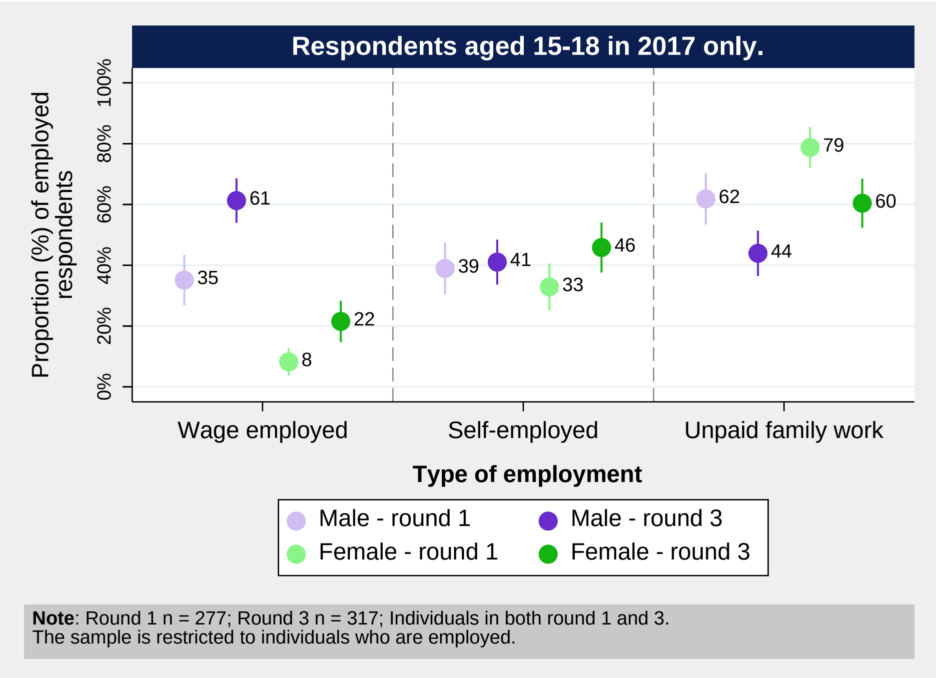
Figure 10: Employment rate of young women and men aged 19–25 in 2017, by survey wave (%)



Source: authors' calculations using MUVA Urban Youth Survey data.

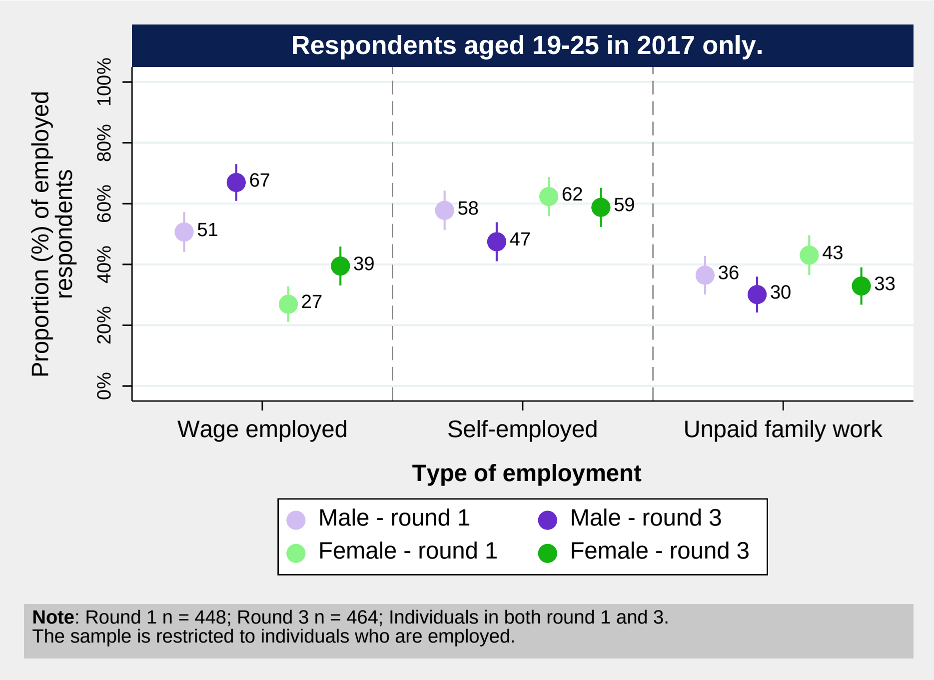
We also disaggregate the results on employment type (Figure 3) by age group in Figure 11 and Figure 12. Again, our main finding that women are much less likely to be wage employed than men holds for both age groups in both survey rounds.

Figure 11: Employment type by gender and survey round, for respondents aged 15–18 in 2017.



Source: authors' calculations using MUVA Urban Youth Survey data.

Figure 12: Employment type by gender and survey wave, for respondents aged 19–25 in 2017.



Source: authors' calculations using MUVA Urban Youth Survey data.

A2: Fixed-effects regressions using alternative indicators for norms relating to childbearing

Table 9: Probability of working for young men and women, fixed effects regressions, alternative norms specifications

	Female			Male	
	(1)	(2)	(3)	(4)	(5)
	Employed	Employed	Employed	Employed	Employed
Age	0.007	0.006	0.000	0.028*	0.029*
	(0.013)	(0.012)	(0.012)	(0.012)	(0.011)
Highest grade completed	-0.006	-0.007	-0.006	0.030**	0.029**
	(0.012)	(0.012)	(0.012)	(0.011)	(0.011)
Single	0.032	0.029	0.054	-0.009	-0.037
	(0.111)	(0.110)	(0.109)	(0.079)	(0.069)
Number of children still alive	-0.080	-0.078		0.045	
	(0.055)	(0.054)		(0.056)	
Household size	0.034	0.034	0.032	-0.007	-0.006
	(0.021)	(0.021)	(0.021)	(0.019)	(0.019)
Household composition: sex ratio	0.111	0.139	0.143	-0.022	-0.024
	(0.187)	(0.189)	(0.189)	(0.163)	(0.163)
Simple Poverty Score Card 2014/2015	0.012***	0.012***	0.012***	0.004	0.004
	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)
Number of hours spent on reproductive work	-0.024**	-0.024**	-0.026***	-0.007	-0.006
	(0.007)	(0.008)	(0.007)	(0.011)	(0.010)
Make decisions about my movement alone	-0.021	-0.023	-0.023	0.010	0.008
	(0.044)	(0.045)	(0.044)	(0.042)	(0.042)
Frequently uses a computer	-0.008	-0.008	-0.009	0.082	0.082
	(0.065)	(0.066)	(0.065)	(0.049)	(0.049)
Member of a social group (excl. church)	0.162***	0.165***	0.163***	0.045	0.045
	(0.042)	(0.043)	(0.044)	(0.043)	(0.043)
Do others approve of women in leadership	0.007	0.014	0.011	-0.011	-0.012
	(0.047)	(0.048)	(0.048)	(0.036)	(0.036)
Would approve of women in leadership	0.063	0.058	0.061	-0.039	-0.038
	(0.076)	(0.077)	(0.078)	(0.054)	(0.054)
Ideal number of children I would like to have	0.039	0.040	0.040		
	(0.040)	(0.039)	(0.039)		
Ideal age for a woman to have her first child		0.010	0.011	-0.005	-0.006
		(0.011)	(0.011)	(0.010)	(0.010)
Ideal age for a man to have his first child		0.002	0.002	-0.001	-0.001
		(0.008)	(0.008)	(0.007)	(0.007)
Observations	1135	1126	1126	954	954

Source: authors' calculations using MUVA Urban Youth Survey data.

Our main theory-based specification includes the number of children the survey respondents have and who are still alive as covariate to capture the social norms individuals might face with regards to childbearing (see section 3.2.1 and Table 4). To assess whether results change when we include other variables in our specifications that capture social norms around the number of children women should have and the age at which individuals should have their first child, we include these our fixed effects regressions and present results in Table 9. Note that the variable 'ideal number of children I would like to have' is only asked to female respondents, and hence not included in specifications (4) and (5). The table shows that results from our main specifications in Table 4 are robust to these alternative specifications.

A3: Full set of variables used in LASSO specifications

In Table 10, we present the full list of 149 variables included in LASSO regressions for which we present results in section 5.2. This is the set of variables remaining after implementing the data management steps described in that section. Note that all categorical variables are included in the regression as transformed binary variables, i.e., each category transformed into a 1/0 dummy.

Table 10: Full list of variables included in LASSO regressions

Variable name	N	Type
Poverty score card (2008) indicator 9: does the HH own a bicycle, motorcycle, or car.	2390	categorical
Poverty score card (2008) indicator 8: does the HH own a radio.	2390	categorical
Poverty score card (2014) indicator 10: does the HH own a fridge.	2390	categorical
Who makes decisions about large expenses?	2374	categorical
Who makes decisions about small household expenses?	2378	categorical
Who makes decisions about movement of respondent?	2383	categorical
What are your educational aspirations? (combined to levels)	2389	categorical
What are your educational aspirations? (grades)	2389	categorical
What was your attendance record in the previous academic year?	2373	categorical
Are you currently enrolled at school?	2387	categorical
What is the highest education level you were enrolled in?	2378	categorical
What is the year in which you left school? (enrolled coded as 'enrolled')	2356	categorical
Age at birth of first child.	2389	categorical
Do you live with all children that you have given birth to/fathered? (no children coded as such)	2350	categorical
Is your child alive? (no children coded as such)	2390	categorical
Is your child at home? (no children coded as such)	2350	categorical
Child bearing status? (men coded as such)	2381	categorical
Year of birth of oldest child.	2389	categorical
Ideal number of children? (men coded as such, question not asked to them)	2295	categorical
Marital status	2389	categorical
Age when first married/entered marital union.	2386	categorical
All pregnancies led to birth or was there a miscarriage/abortion? (men and never pregnant coded as such)	2381	categorical
Number of children still alive.	2390	categorical
Number of children still at home.	2350	categorical
Parenthood was planned. (never pregnant/fathered child coded as such)	2341	categorical
Number of pregnancies. (men and never pregnant coded as such)	2381	categorical
Does respondent live in annex?	2390	categorical
HH source of energy for cooking.	2390	categorical
HH main source of drinking water.	2390	categorical
HH main material of floor.	2390	categorical
HH energy source for lighting.	2390	categorical
HH type of toilet.	2389	categorical
HH main material of exterior walls.	2390	categorical
What's the MAIN activity you use a computer for?	2388	categorical
What's the main location of computer use?	2388	categorical

Phone ownership	2388	categorical
How often are women selected into leadership positions here?	2388	categorical
Would you ever like to be selected for leadership of an organisation?	2388	categorical
In your opinion, how many people around here approve of a woman being selected for the leadership of an organisation?	2388	categorical
Would you approve or disapprove if a woman around here was selected for leadership of an organisation?	2388	categorical
Number of births/children fathered - categories.	2390	categorical
Poverty score card (2014) indicator 3 and 4: can household head read or write.	2390	categorical
Age of respondent.	2390	continuous
Simple Poverty Score Card 2008, final score.	2339	continuous
Probability of HHs lying under 1.25USD/day 2005 PPP line (2008 SPS).	2339	continuous
Probability of HHs lying under 2.5USD/day 2005 PPP line (2008 SPS).	2339	continuous
Age dependency ratio in the household.	2373	continuous
Respondent's age when starting school.	2334	continuous
Ideal age for a woman to have a child.	2370	continuous
Ideal age for a man to have a child.	2367	continuous
Number of births/children fathered.	2390	continuous
Highest grade of schooling completed.	2361	continuous
Number of dependents in the household.	2373	continuous
Number of female household members.	2373	continuous
Household composition: sex ratio.	2373	continuous
Household size.	2373	continuous
Time spent caring for children (hours per day).	2390	continuous
Time spent classroom study (hours per day).	2390	continuous
Time spent domestic chores (hours per day).	2390	continuous
Time spent hygiene (hours per day).	2390	continuous
Time spent caring for the ill and elderly (hours per day).	2390	continuous
Time spent leisure outside the home (hours per day).	2390	continuous
Time spent leisure at home (hours per day).	2390	continuous
Time spent other activities (hours per day).	2390	continuous
Time spent sleeping and relaxing (hours per day).	2390	continuous
Time spent studying at home (hours per day).	2390	continuous
Time spent traveling (hours per day).	2390	continuous
Number of hours spent per day on reproductive work.	2390	continuous
Would approve of women in leadership?	2390	binary
Gender: male or female?	2390	binary
Decision maker on large expenses when youth not involved: other.	2374	binary
Decision maker on large expenses when youth not involved: parent.	2374	binary
Decision maker on large expenses when youth not involved: partner.	2374	binary
Involved in large household expenses? (Yes/No)	2374	binary
Decision maker on small expenses when youth not involved: other.	2378	binary
Decision maker on small expenses when youth not involved: parent.	2378	binary
Decision maker on small expenses when youth not involved: partner.	2378	binary
Involved in small household expenses? (Yes/No)	2378	binary

Main reason for missing school: Illness.	2373	binary
Main reason for missing school: Illness of a family member.	2373	binary
Main reason for missing school: Household chores.	2373	binary
Main reason for missing school: Childcare.	2373	binary
Main reason for missing school: Biscate/emprego.	2373	binary
Main reason for missing school: Leisure activities.	2373	binary
Main reason for missing school: No money for transport/other education related expenses.	2373	binary
Main reason for missing school: Menstruation.	2373	binary
Late for school.	2373	binary
Main reason for missing school: Other.	2373	binary
Has completed at least 10th grade (Secondary 1).	2361	binary
Has completed at least 12th grade (Secondary 2).	2361	binary
Has completed at least primary school.	2361	binary
Has completed at least one year of university.	2361	binary
Lives in female headed household.	2373	binary
Had a child when being under the age of 18.	2390	binary
Ever given birth/fathered a child?	2390	binary
Knows where to get information about family planning or delaying pregnancy.	2385	binary
Is currently married.	2389	binary
Was married under the age of 18.	2389	binary
Currently pregnant?	2380	binary
Ever pregnant?	2381	binary
Is there anyone that contributes with cash to your living expenses?	2389	binary
My partner contributes to my monthly expenditure.	2389	binary
People that contribute to monthly living expenses: Mother.	2389	binary
People that contribute to monthly living expenses: boyfriend/girlfriend.	2389	binary
People that contribute to monthly living expenses: another family member.	2389	binary
People that contribute to monthly living expenses: siblings, uncle/aunt, grandpa.	2390	binary
People that contribute to monthly living expenses: Father.	2389	binary
People that contribute to monthly living expenses: Husband/wife.	2389	binary
People that contribute to monthly living expenses: HH (if different from others).	2389	binary
People that contribute to monthly living expenses: Brother/sister.	2389	binary
People that contribute to monthly living expenses: Uncle/aunt.	2389	binary
People that contribute to monthly living expenses: Grandfather/grandmother.	2389	binary
People that contribute to monthly living expenses: Another family member (in MZ).	2389	binary
People that contribute to monthly living expenses: Another family member who live abroad.	2389	binary
People that contribute to monthly living expenses: Other (specify).	2389	binary
Financially independent?	2389	binary
Assets owned by household: A television.	2390	binary
Assets owned by household: A non-electric iron.	2390	binary
Assets owned by household: An electric iron.	2390	binary
Assets owned by household: A clock.	2390	binary
Assets owned by household: A bed or a cot.	2390	binary

Assets owned by household: Cell phone.	2390	binary
Respondent is the household head.	2369	binary
Household head worked in past 7 days.	2366	binary
Respondent lives alone.	2373	binary
Frequently uses a computer.	2388	binary
Belongs to any social group.	2388	binary
Membership of groups: Church group.	2388	binary
Membership of groups: Credit/savings group.	2388	binary
Membership of groups: Political group.	2388	binary
Membership of groups: Community group.	2388	binary
Membership of groups: Other (specify).	2388	binary
Frequency women selected into leadership. (Never/rarely vs sometimes/often)	2388	binary
Would like a leadership position one day?	2388	binary
Do others approve of women in leadership?	2390	binary
Spent time on: childcare.	2388	binary
Spent time on: class room.	2388	binary
Spent time on: domestic chores.	2388	binary
Spent time on: hygiene.	2388	binary
Spent time on: taking care of ill and elderly.	2388	binary
Spent time on: leisure outside home.	2388	binary
Spent time on: leisure home.	2388	binary
Spent time on: other.	2388	binary
Spent time on: sleeping.	2388	binary
Spent time on: studying home.	2388	binary
Spent time on: traveling.	2388	binary
Member of social group (excl. church).	2388	binary
Spent time on reproductive work?	2390	binary
Did the individual move of residence between survey waves?	2390	binary

Source: authors' descriptions using MUVA Urban Youth Survey data.