Estimating the causal effect of fertility on women's employment in Africa using twins

Introduction

Promotion of family planning and ensuring access to preferred contraceptive methods is essential to increase the well-being and autonomy of women and to support the health and development of communities. Family planning benefits are related to health issues such as the prevention of pregnancy-related health risks for women, the reduction of infant mortality, and the prevention of HIV/AIDS. Additional advantages are that family planning might enhance women's empowerment, children's education, and reduce adolescent pregnancies and population growth (Cleland et al., 2006; Longwe & Smits, 2012; Singh & Darroch, 2012).

A major channel through which family planning might contribute to women's empowerment is by providing opportunities for women to participate in paid employment (Canning & Schultz, 2012; Besamusca, et al., 2015; Engelhardt & Prskawetz, 2004). Women's labor force participation is considered to be of critical importance for gender equality, as well as for the living standard, dependency burden, and saving patterns of households (Anderson & Eswaran, 2009; Buvinic et al., 2009; Fallon & Lucas, 2002; Kritz & Makinwa-Adebusoye, 1999; Yount & Li, 2009). Ever since the pioneering work of Mincer (1962), women's labor force participation has been studied extensively in both developed and developing countries (Bloom et al., 2009; Mammen & Paxson, 2000; Steiber & Haas, 2012; Besamusca et al., 2015). According to these studies, women in developing countries are mostly involved in non-market activities, at home, in the family business, or in other informal sector work, although a pronounced increase in the contribution of women to modern sector employment activities has also been noticed (Gundüz-Hosgör & Smits, 2008; Besamusca et al., 2015). The latter is partly due to the advances made in females' educational attainment and the expansion of the market economy (Tandrayen-Ragoobur et al., 2011).

The presence of young children in the household is seen as one of the most important factors explaining variation in women's labor market participation (Spierings et al., 2010; Contreras & Plaza, 2010; Besamusca et al., 2015; Browning, 1992). In most cultures, women are considered the prime suppliers of household care needs, which increases with the presence of children (Maume, 2006; Moghadam, 2004). Empirical studies in developed countries generally find a negative relationship between fertility and women's labor force participation (Smits et al. , 1996; Ahn & Mira, 2002; Boushey, 2008; Mishra et al., 2010; Michaud & Tatsiramos, 2011; Angrist & Evans, 1998). In developing countries, there is less consistent evidence of a negative effect of the number of children on women's labor force participation (Agüero & Marks, 2008; Benefo & Pillai, 2003; Cruces & Galiani, 2007; Orbeta, 2005). Although many studies find a negative relationship, Agüero and Marks (2008) report an insignificant effect and Nanfosso et al. (2010) even a positive effect of fertility on women's employment.

An important question in the literature is whether unemployed mothers with young children at home would have entered the labor market if they had less or no children (e.g. Agüero & Marks, 2008; Steiber & Haas, 2012; Besamusca et al., 2015; Rosenzweig & Wolpin, 1980a). Any existing association could as well be the result of a reverse causal effect, with women with a job having a lower propensity to have children. Alternatively, a selection effect could be present; women who want to have more children, or have planned to have their children within a short period, might differ from women with less children in that they have a lower preference for paid work at the moment the survey is conducted.

The ideal way to determine the causal effect of fertility on female employment would be to organize an experiment in which women are randomly assigned to conditions with different numbers of (young) children. Differences in employment levels between the groups, with women

with more children having lower employment levels, would then provide sound evidence for the existence of a negative causal effect of the number of young children on women's employment. Such an experiment is obviously impossible, but there exists a useful alternative; the natural experiment. If as a result of a random shock some women would get more children than they expected to have, a comparison of these women with those who did not experience this shock would be similar to the comparison suggested in the experiment described earlier.

In this paper, the causal effect of fertility on women's employment is studied using such a natural experiment. We follow the pioneering work of Rosenzweig and Wolpin (1980a and 1980b) and Angrist and Evans (1998) in using twins as instrument. In the literature on fertility and female labor force participation (FLFP) other instruments have also been used, in particular same sex siblings (Angrist & Evans, 1998). The assumption is then that parents want to have at least one child of the other sex and thus the probability of a third child increases if the first two are of the same sex. This phenomenon is probably more relevant in developed countries than in developing countries, where fertility is still relatively high .

Studies correcting for the possibility of endogeneity between the number of children and FLFP often report a negative effect of young children on FLFP. A negative relation between the number of young children and FLFP is found for the United States (Rosenzweig & Wolpin, 1980b; Vere, 2007), the United Kingdom (Sprague, 1988), Norway (Black et al., 2005), and various other West-European countries (Del Boca et al., 2009 ; Michaud & Tatsiramos, 2011). It is also found for some countries from other continents, such as Argentina and Mexico (Cruses & Galiani, 2007), and the Phillipines (Orbeta, 2005). However, for other non-Western countries the negative relationship is not confirmed. For Peru, Guatamala, Colombia, Bolivia, Nicaragua and the Dominican Republic the negative relation found under Ordinary Least Squares (OLS)

becomes insignificant when self-reported infertility is used as the instrumental variable (Agüero & Marks, 2008). For urban Cameroon even a positive effect is reported (Nanfosso & Zamo-Akono, 2010). These results raise the question whether the negative relationship between fertility and FLFP is also found in Sub-Saharan Africa, when twins are used as instrument.

Sub-Sahara Africa is a very interesting region in this respect, because it has much within region variation in women's fertility and their labor force participation. Women's fertility shows a declining trend but the differences between the countries are large. In the latest wave of the Demographic and Health Survey (DHS) the lowest number of children per woman is 3.3 (Lesotho), whereas the highest are with 6.3 (Chad) and 6.1 (Mali) almost twice as large. Moreover, it is well known that the factors determining women's fertility can differ a lot, even between two adjacent countries (Benefo & Schultz, 1996). Female labor force participation in Sub-Saharan countries increased from an average of 58% in the years around 2000 to almost 70% in the period 2009-2013.[1] Despite this general upward trend, the differences within the continent remain large. Around 2010, the FLFP ranged from less than 40% in Senegal to more than 80% in Tanzania and Mozambique.

The central aim of our paper is to investigate whether, despite great diversity in Africa, a causal relation between fertility rates and female labor force participation can be observed, as was found for other parts of the world (see the references above). To make this possible we will control for the diversity between Sub-Saharan regions in the strongest possible way by using fixed effects at the level of 217 sub-national areas.

We address the reverse causality problem by means of an Instrumental Variable estimation in which twins are used as an instrument. In many studies the number of twins is low, both in relative and absolute sense. Rosenzweig and Wolpin (1980a), for example, have 44 twins

[1] Data are downloaded from the website of the International Labour Organization at June 21, 2016.

out of 2939 households, which is 1,5 percent, and Schultz (2008, p. 3272) reports that in many studies twins form less than one percent of the number of pregnancies. We are in the fortunate position that in our dataset both the number (4,863) and percentage (about two percent) of twins is high, which enhances the precision of the Instrumental Variable (IV) estimate. In Sub-Saharan Africa the natural twinning rate is higher than in any other part of the world (Smits & Monden, 2011) and twin births resulting from artificial reproductive techniques are rare. We use a newly constructed database including information on more than 220,000 women living in 217 sub-national areas (provinces, departments, districts) of 24 Sub-Saharan African countries. To our knowledge, the present study has the highest number of twins of any study that uses twins for estimating causal effects.

Besides by using twin births as an instrument, several other characteristics of the present study will be helpful in determining a causal relationship. First, we control for major confounders at the household level and include fixed effects dummies at the level of the sub-national areas. Second, as women with twins have been pregnant at least once, for reason of comparability and in order to increase the likelihood of a causal effect, we restrict our analysis to women who are similar in this respect (i.e. who have at least one living child). Third, given that women may differ with respect to their demand for family planning depending on their marital status, we include only women who are married or living together with a partner. Fourth, children are not always a burden for their parents. Older children can help their parents by doing household chores or working in the family business (Webbink, Smits & De Jong, 2012). We therefore focus our analysis on children under age six.

The rest of the paper is outlined as follows. In the next section, we present the theoretical framework to be used in our analysis and motivate the included factors. After that, we discuss the

data and methods used. The result section starts with the first stage regression, followed by the second stage model and the interaction analysis. Finally, we draw conclusions and discuss policy implications.

Theoretical Framework

Figure 1 shows the different groups of factors that are included in our analytical model and their expected direction of influence. The explanatory factors belong to one of three groups: family planning factors, household and context factors and interactions with the context.

Insert Figure 1 here.

Family planning factors

The effects of the family planning factors on women's labor force participation are indicated by arrow A in Figure 1. The family planning outcome we are interested in is the number of children requiring extensive care (say below age six). The relationship between this fertility outcome and employment is central to this paper, because the presence of young children who demand significant child care may deter the mother from working outside the home. As a control factor, we also include an indicator for pregnancy of the mother.

Household and context factors

At the household level, age, education, marital status, husband's education and occupation, and the presence of adult women (Arrow B), have been known for long to be important determinants of women's employment. Regarding age, women's employment in developed countries is known to peak before the onset of childbearing, to decline during the child-rearing period and to start

increasing again a few years after childbearing (Brewster & Rindfuss, 2000; Smits et al., 1996; Baah-Boateng et al., 2013). Many empirical studies have found that better educational attainment of women leads to greater labor force participation and increased productivity (Buvinic at al. 2009, p. 354). Although in some cases there is only a small or non-linear relationship between the level of education and women's employment, in general education has a positive effect on the labor force participation of both married and single women (Gundüz-Hosgör & Smits, 2008; Spierings et al., 2010; Yakubu, 2010).

The sign of the influence of marital status on female labor market participation is ambiguous (Angrist, 2002). Some studies have shown that married women are less likely to participate in the labor force than those who are single, divorced/separated or widowed (Benefo & Pillai, 2003; Ntuli & Wittenberg, 2013). The sexual division of labor within the unit of a married couple predicts that a man with higher earnings through a higher level of education and occupation will allow his wife to reduce her market work and to focus more on housework (Devereux, 2004; Kalenkoski et al., 2009). Empirical evidence of the opposite has also been found (Mon, 2000). We circumvent this ambiguity by focusing on women with a partner.

Finally, as mentioned in the introduction, women's employment is influenced by the context in which people live (arrow C in Figure 1). People living in rural and less developed areas are less likely to be active in the labor market compared to those living in urban and developed areas (Benefo & Pillai, 2003; Günduz-Hosgör & Smits, 2008; Ntuli & Wittenberg, 2013). In this study, we will therefore control for urbanization of place of living. Besides urbanization, many other context factors may play a role, including labor market opportunities, religion and culture. Some of these factors are available in our dataset but many of them are not

observed. To control for the effect of these factors, fixed effects dummies at the level of 217 districts will be included in all our analyses.

Interactions

The effects of family planning outcomes on women's labor force participation need not be everywhere the same. We perform an interaction analysis in which we study to what extent these effects depend on characteristics of the household and of the context in which the household lives (as shown by arrows D in Figure 1). Two opposite hypotheses are tested. First, it is possible that women who suffer most from negative family planning outcomes are those in the weakest situations, hence poor women, women with little education and women living in rural areas, with less job opportunities. Having several young children might give these women fewer possibilities to work for pay. For women under better circumstances the number of young children would matter less, as they have more resources at hand to solve child care problems.

On the other hand, it is also possible that women living in more developed areas suffer most, as they tend to have more demanding jobs and may miss the extended kinship network that might take care of the children. By including interactions between the family planning outcome and education and urbanization in our model, we aim to find out which alternative is mostly in line with the situation of African women. A related question is whether the possible problems related to the number of young children are growing bigger or becoming smaller over time. As we have data for two points in time, we will be able to answer this question empirically.

Data

The data used in this study are from the Demographic and Health Surveys (DHS). These are large representative household surveys held since the 1980s in many developing countries (see www.dhsprogram.com). The DHS program is sponsored by USAID and executed in collaboration with national statistical agencies. DHS surveys consist of a household survey in which basic information on all household members is obtained and a women's survey in which all usual resident women aged 16-49 obtain an extensive oral interview. For the purpose of our study we select women who are 18-45 years old, married and have at least one child below the age of six. We selected this group of women because they are in the age group most likely having young children at home. We included all countries for which two waves of the Standard DHS survey were available that contained the necessary variables. Countries with two waves were selected in order to be able to study changes over time. For countries for which more than two standard DHS surveys were available, the two most recent ones at the time of study were used. Our database contains data for 221,556 women from 217 districts of 24 Sub-Saharan African countries as shown in the Appendix Table A1. Of these women, 4,279 had a twin pair among their children under six.

The dependent variable, married women's labor force participation, is measured by a dummy variable indicating whether (1) or not (0) the woman was engaged in non-farm employment in the week before the interview. To measure women's employment, first the following question was asked: "Aside from your own housework, have you done any work in the last seven days?" And if the answer was no: "As you know, some women take up jobs for which they are paid in cash or kind. Others sell things, have a small business or work on the family farm or in the family business. In the last seven days, have you done any of these things or any other work?" The answers to these questions were for the employed women combined with

information on whether or not they were engaged in farming. Women who were employed in farming were included in the non-employed category, as the focus of our study is on women's participation in non-farm work. The reason for this is that for women entering the non-farm labor force generally is a greater step towards economic independence and more likely to be socially recognized than engaging in farm work (Spierings et al., 2010; Gundüz-Hosgör & Smits, 2008). Given that non-farm work generally is more difficult to combine with childcare activities (Brewster & Rindfuss, 2000), we also expect the negative effects of fertility on women's employment to be strongest for non-farm work.

To find out to what extent our findings are sensitive to the way we measured women's labor market participation, we repeated our analyses with three different operationalizations of this variable. In the first one, we removed the women engaged in farm work from the analysis, so that the difference between women engaged in non-farm work and women who reported not to work was studied. In the second one, we compared all employed women, farm and non-farm combined, with the women who reported not to work. In the third one, we compared women employed in farming with women who reported not to work. In all cases, the IV effect was in the same order, direction and level of significance as in the main analysis. The outcomes of these analyses are presented in Table A2 of the Appendix.

Independent variables include family planning outcomes, other household-level factors, urbanization and in the interaction analysis also the average wealth in the district. Fertility is measured by the number of children below age six living in the household. The current age of the woman is in years. Pregnancy prevalence is a dummy variable indicating whether (1) or not (0) the woman was pregnant during the time of the interview. Of the other household-level factors, husband's occupation is measured as (1) farm, (2) lower non-farm and (3) upper non-farm. The

presence of other adult women in the household is a dummy indicating whether (1) or not (0) there were other women aged over 18 living in the household. Education of the women and their husbands is measured by years of schooling. The level of urbanization is measured by a dummy indicating whether (1) or not (0) the household lives in a rural area. To indicate the level of development of the context, we followed Huisman and Smits (2009) and created a district-level index on the basis of six variables available in the household datasets: the percentages of households in a district owning a fridge, a car, a telephone, or a television, and the percentages of households with electricity or running water. Of these characteristics the mean was taken of the standardized values for each district. Given the presence of district dummies in all our models, this index could only be used in the interaction analysis. The time dimension is measured by a dummy called wave, indicating whether the respondent was interviewed in the first (0) or in the second (1) DHS wave of the country. Fixed effects dummies for the 217 districts are included to control for variation in the context. In all analyses the weight factors provided by DHS were used to obtain representative samples for the countries.

Table 1 shows that the average participation rate of women in a non-agricultural job is 27 percent and that the average number of children below six is 1.6. The descriptive statistics further show that two percent of the women had twins among the children under six, which is equivalent to 4,279 twins. The average education level for women in our sample is 3.4 years while that of their husbands is 4.4 years. The mean age of the women is 29 years. We see that on average about half of the husbands were employed in farming and 40 percent had a lower non-farm occupation. At the time of the interview 13 percent of the women were pregnant and 72 percent lived in rural areas. The households have an average of 1.3 adult women.

Insert Table1 about here

Method

The central aim of the analysis is to establish whether there is a causal effect of the presence of young children on women's employment. In order to do so, we estimate an instrumental variables model in order to address the endogeneity of family planning outcomes with regard to women's employment. We specify our general model as follows:

$$WP_i = \beta_0 + \beta_1 FD_i + \beta_2 X_i + \varepsilon_i \quad (1)$$

Where WP is an indicator of a married woman's labor force participation taking value 1 if she works in a non-agricultural job and 0 if not; FD is the endogenous fertility decision indicated by the number of children under six; X is a vector of individual and household characteristics assumed exogenous and ε is the residual. The vector X includes all other explanatory variables mentioned in the previous section and the district fixed effects. The subscript i refers to the fact that these are data for individual women.

The number of children below six is likely to be endogenous, so that merely estimating relation (1) will not inform us about the causality. We assume that *Cov* (X_i, ε_i) =0 and *Cov* (FD_i, ε_i) ≠0. Whether a twin is among the number of young children is used as the exogenous variation in the number of young children. The assumption is that twins represent an unwanted increase in the number of young children, which has no independent effect on a woman's labor supply. The first stage equation of the IV-estimation is given as follows:

$$FD_i = \alpha_0 + \alpha_1 TW_i + \alpha_2 X_i + v_i \quad (2)$$

Where TW is the instrumental variable twins, ν is the residual and X is as discussed above. For each specification, the basic coefficients obtained by the probit model without the number of young children instrumented (equation (1)) are compared with those obtained using the model where the number of young children is instrumented by twins. Exogeneity of the number of young children is tested using the Hausman test. This involves inserting the residuals from the first stage regressions into the original regression. The significance level of the coefficient for the residual constitutes a test for exogeneity of the variable in question. Given a significant negative coefficient with a P-value <0.0001 we rejected the null hypothesis of no endogeneity. The test thus indicates that our fertility variable was indeed endogenous. We also tested whether the occurrence of twins has a direct effect on women's employment and found that the effect was not significant.

Results

First step regressions

Table 2 shows the full results for the first stage Ordinary Least Squares (OLS) regressions of the number of young children in the household on the presence of twins under age six. In order for the instrument to be valid, it needs, in addition to its exogeneity with respect to labor force participation, also to have a strong relationship with fertility. The effect of twin presence on the number of young children is highly significant. Its coefficient has a t-statistic which is much higher than 3.3, the minimum level required (see bottom of Table 2). This indicates that the effect of twin births is strong enough to use this factor as an instrumental variable (Carter- Hill et

al., 2012, p. 414). The regression controls for age, square of age, education, husband's education, husband's occupation, number of adult women in the family, whether the respondent is pregnant, urbanization and wave effects. The coefficients have the expected signs. The number of children younger than six increases with a woman's age but this effect becomes less for older women. Higher education of both spouses and non-farm occupation of the husband decrease the number of children under six. This number is also reduced by the fact that the woman is currently being pregnant and the household contains another adult woman. Women in rural areas of a district have more children. The number of young children decreases over time. Context level effects are controlled for by including the district fixed effect dummies in all estimated regressions.

Insert Table 2 here.

Second step model

The standard probit model and two-stage probit model are presented in Table 3. The results of the standard model suggest that the number of children under six has a negative effect on the chances of a woman to participate in non-agricultural work. The regression controls for age, age square, education, husband's education and occupation, pregnancy of the woman, number of adult women in the family and urbanization, which all show the expected effects, and includes fixed effects dummies at the district level.

Insert Table 3 here

The second-stage outcome confirms the standard outcome completely. The coefficient remains strongly significant and has about the same value as in the standard model: having more children under six is negatively associated with women's non-farm employment. Note that the difference in size with the non-instrumented variable is very small. This indicates that twins form a strong instrumental variable in our study. This result is most likely due the high number of twins in our sample, which allows us to make precise estimates. From these findings we may conclude that the number of children under age six negatively influences the chances of African women to work in a non-agricultural job. The multiplicative version of the coefficient of the instrument is exp (-0.061)=0.94. Hence we can conclude that the presence of one additional child under six reduces the odds that the mother participates in the non-farm labor force by six percent. When they have more children, the women may respond by dropping out of the labor market or not engaging in it, or by opting for an agricultural job that keeps them near their home.

Looking briefly at the control variables in the models, we observe that the labor participation function is concave in age, implying that both having children at a very young age and at a relatively old age decreases a woman's ability to engage in non-farm work. The education coefficients for the respondents and their spouses are positive, indicating that more highly educated women and women of more highly educated men tend to work more often outside the agricultural sector. Women with husbands in non-farm occupations are also more likely to be involved in non-farm jobs. Pregnant women work less and women in rural areas participate less in non-agricultural jobs than their urban counterparts. The presence of other adult women in a household consistently shows a negative association with women's employment. This might be due to the fact that extended families are often more traditional households. The wave variable indicates that women's participation in the labor force increases over time.

Interaction effects

The coefficients of the interaction analysis are presented at the bottom of the third column of Table 3. This analysis is important, because it gives an impression of the degree to which and the way in which the associations of fertility with women's employment vary across situations. Given the explorative nature of this analysis, we tested for significant interactions with all variables in our model and included the significant ones in the final interaction model, presented in Table 3, column 3. We observe that the coefficient of the instrumented variable increases somewhat when the interaction terms are included in the model and remains strongly significant. The control factors also keep their significance and direction compared to the IV model without interactions.

There are significant interactions of women's age, education, district wealth and wave with fertility. Hence the fertility effect varies between younger and older women, between women with less or more education, between women living in poorer or wealthier areas and it changes over time. The interaction effect with women's age shows that having more children is stronger negative for older women. The interaction coefficient of education shows that women with more children can profit less from the advantage of being educated. The interaction effect with district wealth shows that women in wealthier areas can more easily overcome the negative effects of children on their employment. The interaction with the wave dummy indicates that over time the negative effect of fertility on women's employment decreases.

Conclusion

In this paper, we study the causal relationship between African women's fertility and their labor force participation in non-agricultural work. The focus is on non-agricultural work, as farm work is often done around the house and is generally better compatible with care for children than non-farm work. The number of children below six years is used as the indicator of fertility. An instrumental variables method is used to address the endogeneity problem that exists between fertility and women's labor force participation. The presence of twins among the children is used as instrumental variable. We also performed an interaction analysis in order to understand to what extent the effects depend on characteristics of the household and of the context in which the household lives. The study is based on more than 220,000 married women with at least one child below the age of six in 24 African countries.

The number of children below age six was found to have a significantly negative effect on the woman's ability to work in the non-farm sector. Our results indicate that the presence of one unexpected additional child under six reduces the odds of non-farm employment of African mothers by six percent. This effect is highly significant, which is probably due to the high number of twins in our dataset. The interaction analysis revealed that the effects of the number of young children on women's non-farm work are more problematic for older women and for women with more years of education. The effects are less problematic in more developed areas and also become less negative over time.

Our study shares with some others (such as Besamusca et al., 2015) the interest in the question whether certain patterns hold worldwide. Although the circumstances can differ considerably between countries, women everywhere go through stages as school-going, transition to adulthood, motherhood and grand-motherhood. Does this also mean that except for local institutions and habits the same factors are relevant for women's decisions everywhere? In

particular, when a women is unexpectedly confronted with the care for young children does that prevent her from entering the labor market? As in many other studies (Angrist & Evans, 1998; Michaud & Tatsiramos, 2011; Agüero & Marks, 2008; Benefo & Pillai, 2003 ; Cruces & Galiani, 2007) for different countries the answer obtained from our analysis on Sub-Saharan African data is yes.

As in these other studies, our conclusion is conditional on having given birth to twins. This conditionality forms a limitation of studies using Instrumental Variables for identifying causal patterns (Deaton, 2010). The causal effect is only established for the small group identified by the Instrumental Variable, in our case the women who take care of young children among which is a twin. The advantages of our study are that due to the high number of twins the estimator is very precise and that the high number of fixed effects for sub-national differences effectively controls for a multitude of differences between regions. However, it cannot be precluded that factors related to getting twins, like maternal health problems affect our outcomes as well.

This study confirms that twins hamper women for entering the labor force and makes this claim stronger than it was before. To what extent these findings are relevant for policy making remains, however, to be seen. Policy makers cannot prevent women from getting twins. However, they can facilitate childcare arrangements that help women to combine work with childcare and improve public health for mothers with twins. Whether such measures will actually increase the labor force participation of women with twins remains however to be seen. For drawing more policy relevant conclusions, a longitudinal analysis is required in which women's labor force participation is among others explained by exogenous policy measures.

References

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

Ahn, N., & Mira, P. (2002). A note on the changing relationship between fertility and female employment rates in developed countries. *Journal of Population Economics, 15* (4), 667–682.

Anderson, S, & Eswaran, M. (2009). What determines female autonomy? Evidence from Bangladesh. *Journal of Development Economics, 90* (2), 179–91.

Angrist, J. (2002). How do sex ratios affect marriage and labor markets? Evidence from America´s 2nd generation. *The Quarterly Journal of Economics, 117* (3), 997-1038.

Angrist, J., & Evans, W. (1998). Children and their parents' labor supply: Evidence from exogenous variation in family size. *American Economic Review, 88* (3), 450-77.

Baah-Boateng, W., Nketiah-Amponsah, E., & Frempong, R. (2013). The effect of fertility and education on female labour force participation in Ghana. *Ghanaian Journal of Economics, 1*, 119-137.

Benefo, .D., & Schultz, T.P., (1996). Fertility and child mortality in Cote d' Ivoire and Ghana. *World Bank Economic Review, 10* (1), 123-158.

Benefo, K. D.,& Pillai, V.K. (2003). Determinants of women non- family work in Ghana and Zimbabwe. *Canadian Studies of Population, 30* (2), 389-406.

Besamusca, J., Tijdens, K., Keune, M., & Steinmetz, S. (2015). Working women worldwide. Age effects in female labor force participation in 117 countries. *World Development, 74*, 123-141.

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

Bloom, D. E., Canning, D., Fink, G., & Finlay, J. E. (2009). Fertility, female labor force participation, and the demographic dividend. *Journal of Economic Growth, 14* (1), 79-101.

Brewster, K., & Rindfuss, R. (2000). Fertility and women's employment in industrialized nations. *Annual Review of Sociology, 26*, 271-296.

Browning, M. (1992). Children and household economic behavior. *Journal of Economic Literature, 30* (3), 1,434-1,475.

Boushey, H., (2008). 'Opting out?' The effect of children on women's employment in the United States. *Feminist Economics, 14* (1), 1-36.

Buvinic, M., Das Gupta, M., & Casabonne, U. (2009). Gender, poverty and demography: An overview. *The World Bank Economic Review, 29* (3), 347-369.

Carter Hill, R., Griffiths, W. E., & Lim, G. C. (2012). *Principles of Econometrics.* Hoboken, NJ: Wiley & Sons, Fourth Ed

Canning, D., & Schultz, P. (2012). The economic consequences of reproductive health and family planning. *The Lancet, 380* (9837), 14–20.

Cleland, J., Bernstein, S., Ezeh, A., Faundes, A., Glasier, A., & Innis, J. (2006). Family planning: The unfinished agenda. *The Lancet: Sexual and Reproductive Health, 368* (9549), 1810-1827.

Contreras, D., & Plaza, G. (2010). Cultural factors in women's labor force participation in Chile. *Feminist Economics, 16* (2), 27-46.

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

Deaton, A. (2010). Instruments, randomization, and learning about development. *Journal of Economic Literature, 48* (2), 424-455.

Del Boca, D., S. Pasqua, & Pronzato, C. (2009). Motherhood and market work decisions in institutional context: a European perspective. *Oxford Economic Papers, 61*, i147-i171.

Devereux, P. (2004). Changes in relative wages and family labor supply. *Journal of Human Resources, 39* (3), 696-722.

Engelhardt, H., & Prskawetz, A. (2004). On the changing correlation between fertility and female employment over space and time. *European Journal of Population/Revue Europeenne De Demographie, 20* (1),35–62.

Fallon, P.R., & Lucas, R.E.B. (2002). The impact of crises on labor markets, household incomes, and poverty: A review of evidence. *World Bank Research Observer, 17* (1), 21-45.

Gundüz-Hosgör, A., & Smits, J. (2008). Variation in labour market participation of married women in Turkey. *Women Studies International Forum, 31*, 104–117.

Huisman, J & Smits, J. (2009). Effects of household- and district-level factors on primary school enrollment in 30 developing countries. *World Development, 37* (1), 179-193.

Kalenkoski, S., Ribar, D. S., & Stratton, L. S. (2009). The influence of wages on parents' allocation of time to child care and market work in the United Kingdom. *Journal of Population Economics, 22* (2), 399-419.

Kritz, M.M. & Makinwa-Adebusoye, P. (1999). Determinants of women's decision-making authority in Nigeria: The ethnic dimension. *Sociological Forum, 14* (3), 399–424.

Longwe, A & Smits, J. (2012). Family planning outcomes and primary school attendance in Sub-Saharan Africa. *Studies in Family Planning, 43* (2),127-134.

Mammen, K., & Paxson, C. (2000).Women's work and economic development. *Journal of Economic Perspectives, 14* (4), 141–164.

Maume, D. J. (2006).Gender difference in restricting work efforts because of family responsibilities. *Journal of Marriage and Family, 68*, 859-869.

Michaud, P.-C. & Tatsiramos, K. (2011). Fertility and female employment dynamics in Europe: the effect of using alternative econometric modeling assumptions. *Journal of Applied Econometrics, 26* (4), 641-668.

Mincer, J. (1962). Labor force participation of married women: A study of labor supply. In: H.G. Lewis. (Ed.), *Aspects of Labor Economics* (pp. 63-101). Princeton: Princeton University Press.

Mishra, V., Nielsen, I., & Smyth, R. (2010). On the relationship between female labour force participation and fertility in G7 countries: Evidence from panel co-integration and Granger causality. *Empirical Economics, 38* (2), 361–372.

Moghadam, V. M. (2004). *Towards gender equality in the Arab/Middle East region: Islam, culture, and feminist activism*. Human Development Reports. New York: United Nations Development Programme.

Mon, M. (2000). Determinants of female labour force participation in Burma: An empirical Analysis of socio-economic survey data. *ABAC Journal, 21* (1).

Nanfosso, R.A.T. & Zamo-Akono C.M. (2010). Fertility, health and female labour force participation. *International Business Research, 3* (2), 136-156.

Ntuli, M., & Wittenberg, M. (2013). Determinants of black women's labour force

participation in post-apartheid South Africa. *Journal of African Economics, 22* (3), (347-374. doi:10.1093/jae/ejs039.

Orbeta, A. (2005). Children and the labor force participation and earnings of parents in the Philippines. *Philippine Journal of Development, 32* (1), 19–52.

Pettit, B., & Hook, J. (2005). The structure of women's employment in comparative perspective. *Social Forces, 84* (2), 779-801

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

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

Schultz, T.P. (2008). Population policies, fertility, women's human capital, and child quality. In T.P. Schultz & J. Strauss (Eds.). *Handbook of Development Economics*, 4 (pp. 3249-3303). Amsterdam: Elsevier,.

Singh, S., & Darroch, J.E. (2012). *Adding It Up: Costs and Benefits of Contraceptive Services-Estimates for 2012*. New York: Guttmacher Institute and UNFPA.

Smits, J & Monden, C. (2011). Twinning across the developing world. *Plos One 6:* e25239 *doi:10.1371/journal.pone.0025239.*

Smits, J., Ultee, W., & Lammers, J. (1996). Effects of occupational status differences between spouses on the wife's labour force participation and occupational achievement: findings from 12 European countries. *Journal of Marriage and Family, 58*, 101-115.

Spierings, N., Smits. J., & Verloo, M. (2010).Micro and macro-level determinants of women's employment in six Arab countries. *Journal of Marriage and Family, 72*, 1391-1407.

Sprague, A. (1988). Post-war fertility and female labour force participation rates. *The Economic*

Journal 98 (392), 682-700.

Steiber, B., & Haas, N. (2012). State of the art. Advances in explaining women's employment patterns. *Socio-Economic Review, 10* (2), 343–367.

Tandrayen-Ragoobur, V., Ummersingh, S., & Bundhoo, Y. (2011). The power to choose: Women and labour market decisions in Mauritius. *Journal of Emerging Trends in Economics and Management Sciences, 2* (3), 193-205.

Vere, J. P. (2007). "Having it all" no longer: Fertility, female labor supply, and the new life choices of generation x. *Demography, 44* (4), 821-828.

Webbink, E., J. Smits & de Jong, E. (2012). Hidden child labour: Determinants of housework and family business work of children in 16 developing countries. *World Development, 40* (3), 631-642.

Yakubu, A.Y. (2010). Factors influencing female labor force participation in South Africa in 2008. *The African Statistical Journal, 11*, 85-104

Yount, K. M. & Li L. (2009). Women's ''justification'' of domestic violence in Egypt. *Journal of Marriage and Family, 71* (5), 1125–1140.

Table1. Descriptive statistics of the variables included in the analysis (N=221,556).

Variable	Mean	Std Dev.
Proportion employed women	0.27	0.44
Number of children < 6 years	1.60	0.67
Proportion women with twins < 6 years	0.02	0.14
Age	29.12	6.72
Years of education	3.41	4.19
Years of husband's education	4.41	4.73
Occupation husband		
Proportion in farming	0.50	0.40
Proportion in lower non- farm	0.40	0.49
Proportion in upper non- farm	0.11	0.31
Proportion of women currently pregnant	0.13	0.34
number of adult women in the family	1.33	0.80
Proportion of women living in rural area	0.72	0.45

Table 2: First stage OLS regression results with number of children under six as dependent variable.

Predicting:	**Children<6years**
Twins	1.126***
	(0.024)
Age	0.158***
	(0.002)
Age square	-0.003***
	(0.000)
Education	-0.015***
	(0.000)
Husband's education	-0.003***
	(0.000)
Occupation husband	
Lower non- farm	-0.042***
	(0.003)
Upper non- farm	-0.054***
	(0.006)
Currently pregnant	-0.232***
	(0.004)
Number of adult women in the family	-0.025***
	(0.002)
Rural	0.074***
	(0.003)
Wave	-0.030***
	(0.003)
Observations	*221,556*
adjR-square	*0.144*
Test for strength of instrumental variable	
t-statistic	*118.06*

****P value<0.01; ** P value<0.05; *P value<0.1*

Notes: standard errors are shown in parentheses; fixed effects are controlled for at district level in all models.

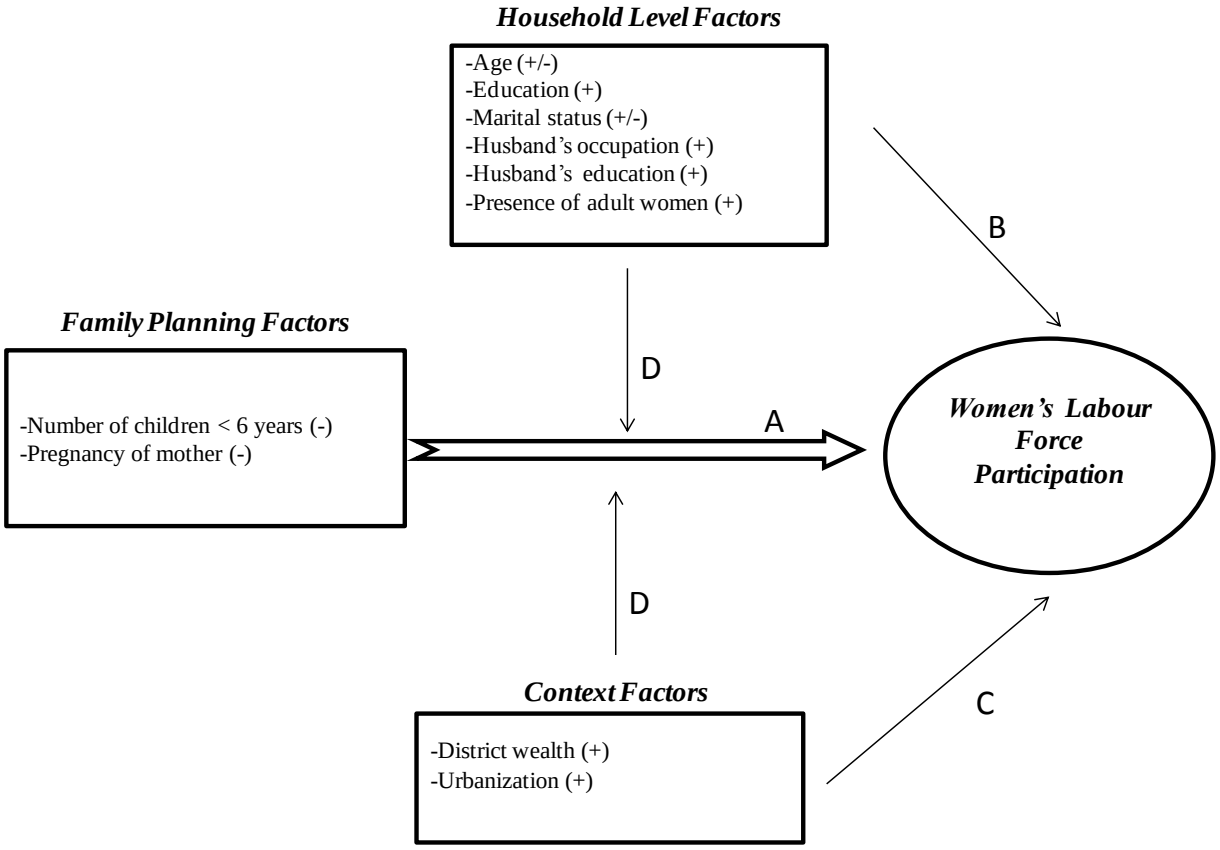
Table3. Original and 2nd stage probit regressions with women's labor force participation as dependent variable (N=221,556)

	Model with children<6years		
	Standard Model	***2nd Stage model***	***with interactions***
Children <6years / Instrumented	-0.060***	-0.061***	-0.107***
	(0.005)	(0.019)	(0.025)
Age	0.106***	0.106***	0.110***
	(0.004)	(0.005)	(0.005)
Age square	-0.001***	-0.001***	-0.002***
	(0.000)	(0.000)	(0.000)
Education	0.039***	0.039***	0.039***
	(0.001)	(0.001)	(0.001)
Husband's education	0.003**	0.003**	0.003**
	(0.001)	(0.001)	(0.001)
Occupation husband			
Lower non- farm	0.0540***	0.540***	0.537***
	(0.008)	(0.008)	(0.008)
Upper non- farm	0.579***	0.578***	0.576***
	(0.012)	(0.012)	(0.012)
Currently pregnant	-0.058***	-0.058***	-0.053***
	(0.009)	(0.010)	(0.010)
Number of adult women in the	-0.025***	-0.025***	-0.022***
family	(0.004)	(0.004)	(0.004)
Rural	-0.365***	-0.365***	-0.369***
	(0.008)	(0.008)	(0.008)
Wave	0.052***	0.051***	0.052***
	(0.007)	(0.007)	(0.007)
Interactions			
Young children(instrumented)*age			-0.017***
			(0.002)
Young children (instrumented)			-0.009***
*education			(0.003)
Young children (instrumented)*district			0.128***
wealth			(0.013)
Young children (instrumented)*wave			0.110***
			(0.025)

****P value<0.01; ** P value<0.05; P value<0.10*

Notes: standard errors are shown in parentheses; dependent variable is women's labor force participation; fixed effects are controlled for at district level in all models.

FIGURE 1 Household and context-level determinants of women labor force participation



Appendix

Table A1: List of countries and years of the DHS waves included in the study

Country	*Wave 1*	*Wave 2*
Benin	2001	2006
Burkina Faso	1998	2003
Cameroon	2004	2011
Chad	1997	2004
Cotedivoire	1994	1999
Eritrea	1995	2002
Ethiopia	2005	2011
Ghana	2003	2008
Guinea	1999	2005
Kenya	2003	2008
Lesotho	2004	2010
Madagascar	2004	2009
Malawi	2004	2010
Mali	2001	2006
Mozambique	1997	2003
Namibia	2000	2006
Niger	1998	2006
Nigeria	2003	2008
Rwanda	2005	2010
Senegal	2005	2011
Tanzania	2004	2010
Uganda	2006	2011
Zambia	2002	2007
Zimbabwe	2006	2011

Table A2. Sensitivity analyses with three different operationalizations of women's labor force participation as dependent variable

	Model 1[a]	***Model 2***[b]	***Model 3***[c]
Children instrumented	-0.80***	-0.089***	-0.058**
	(0.019)	(0.019)	(0.024)
Age	0.098***	0.135***	0.054***
	(0.005)	(0.006)	(0.006)
Age square	-0.001***	-0.002***	-0.001***
	(0.000)	(0.000)	(0.000)
Education	0.018***	0.034***	-0.010***
	(0.001)	(0.001)	(0.002)
Husband's education	0.007***	0.005***	0.010**
	(0.001)	(0.001)	(0.001)
Occupation husband			
Lower non-farm	-0.167***	0.260***	-0.549***
	(0.008)	(0.009)	(0.010)
Upper non-farm	-0.195***	0.226***	-0.683***
	(0.013)	(0.014)	(0.019)
Currently pregnant	-0.052***	-0.079***	-0.026***
	(0.010)	(0.012)	(0.012)
Number of adult women in the	-0.010**	-0.009*	-0.007
family	(0.004)	(0.005)	(0.006)
Rural	0.180***	-0.149***	0.891***
	(0.008)	(0.009)	(0.013)
Wave	0.035***	0.108***	-0.046***
	(0.006)	(0.008)	(0.008)

****P value<0.01; ** P value<0.05; P value<0.10*

Notes: standard errors are shown in parentheses; fixed effects are controlled for at district level in all models.

[a] *All employed women (farm and non-farm combined) compared to women who reported not to work (n=221556)*

[b] *Women engaged in non-farm work compared to women who reported not to work (n=147928)*

[c] *Women engaged in farm work compared to women who reported not to work (n=158319)*

Note to the editor

Only one reviewer still had two questions. Here is how we treated his/her remarks.

Reviewer

First, both reviewers have drawn the attention on the potential differentials in the effect of fertility of on female labor-force participation in the region across religious groups. Answering that « there is no need to control for variation in religion among countries and regions, as we use a fixed effects design at the level of 217 sub-national regions » is simply not enough as they may be variations across and within the sub-national regions regarding religion composition. Was the effect tested at the household level (like age, education etc.)? If not, why not?

Our answer

It would indeed be nice if we could have controlled for religion at the household level, but this is with the data at hand not well possible. Religion is not a straightforward variable like age or education that can be measured everywhere in years or in years of education completed. Age and education are therefore rather well comparable between situations, whereas religion is not. For example, within the broader religious category of Christianity there are many different denominations, the frequency of which differs considerably between different African countries. Consequently, many of these denominations are not always included in the religion variables. The available religion variables also include all kinds of traditional local religions that are not comparable among countries. Another complicating factor is that no information on the degree of religiosity is available. That means that we cannot distinguish between women who are very religious and women who only in name adhere to a certain religion and are not active members. These issues considerably complicate the use of religion at the household level. At the same time, we know that many regions have a dominant religion and that this is particularly the case in smaller geographic units like sub-national areas. We therefore feel that using fixed effects at the level of sub-national areas is a good alternative for using religion at the household level and that by doing so the effect of religion is to a large extent controlled for.

Reviewer

Second, as pointed out by both reviewers the lack of distinction between formal and informal non-farm labor force participation is questionable and needs, at the very least, to be further discussed. The issue of women's ability to combine production and reproduction is indeed at the heart of the paper and it varies greatly depending on the sector of activity (formal and informal). Like for religion, the information is available in the DHS, which makes the distinction feasible.

Answer

We agree with this reviewer that the operationalization of the dependent variable is an important issue that potentially may affect the outcomes of the study. We therefore have done our best to measure women's labor force participation as well as possible. Given the concerns of this reviewer, we have added Table A2 to the appendix in which our IV analysis is repeated with three operationalizations of the dependent variable, based on different combinations of non-farm work, farm work and not working. The outcomes of these sensitivity tests are substantially the same as those of the main analysis. We therefore feel confident that women's labor force participation is measured well in our study. We do not follow the suggestion of this reviewer to look also at informal and formal labor and the occupational categories in which the employed women are working. The aim of this study is to find out whether the presence of young children influences women's labor force participation and questions about formal/informal work and occupational categories -- although interesting and important -- are beyond the scope of this study.

Text which has been changed

The third paragraph of the section Data has been rewritten (changes in italics)

To find out to what extent our findings *are sensitive* to the way we measured women's labor market participation, we repeated our analyses with *three* different operationalizations of this variable. In the first one, we removed the women engaged in farm work from the analysis, so that the difference between women engaged in non-farm work and women who reported not to work was studied. In the second one, we compared *all employed women, farm and non-farm combined, with* the women who reported not to work. *And in the third one, we compared women employed in farming with women who reported not to work. In all cases, the IV effect was in the same order, direction and level of significance as in the main analysis. The outcomes of these analyses are presented in Table A2 of the Appendix.*

The last two paragraphs of the Conclusion have been changed (changes in italics)

As in *these* other studies, our conclusion is conditional on having given birth to twins. This conditionality forms a limitation of studies using Instrumental Variables for identifying causal patterns (Deaton 2010). The causal effect is only established for the small group

identified by the Instrumental Variable, in our case the women who take care of young children among which is a twin. The advantage of our study is that due to the high number of twins the estimate is very precise and that the high number of fixed effects for sub-national differences effectively controls for a multitude of differences between regions. *However, it cannot be precluded that factors related to getting twins, like maternal health problems affect our outcomes as well.*

This study confirms that twins hamper women for entering the labor force and makes this claim stronger than it was before. To what extent these findings are relevant for policy making remains, however, to be seen. Policy makers cannot prevent women from getting twins. *However, they can facilitate childcare arrangements that help women to combine work with childcare and improve public health for mothers with twins. Whether such measures will actually increase the labor force participation of women with twins remains however to be seen. For drawing more policy relevant conclusions*, a longitudinal analysis *is required* in which women's labor force participation is among others explained by exogenous policy measures.

In addition Table A2 is new

*Conflict of Interest Statement

Estimating the causal effect of fertility on women's employment in Africa using twins

None of the authors has any conflicting interest.