Motherhood and Female Labor Supply: Causal Evidence from Central Asia

By

Dilnovoz Abdurazzakova

Submitted to

Central European University

Department of Economics and Business

In partial fulfilment of the requirements for the degree of Master of Arts in Economics

Supervisor: Professor Andrea Weber

Vienna, Austria

2021

Abstract:

This paper estimates the impact of having more children on women's labor force participation in Central Asia using all available waves of standard Demographic Health Surveys conducted between 1995 and 2017. The two stage least squares method using infertility, son preference and having twin birth as instrument variables is applied to estimate the causal impact. By using OLS method, I find statistical significant negative association between fertility rate and mother's current labor supply, employment and occupation type. Once potential endogenity issue associated with fertility decision is solved, having more children does not show any impact on female occupation and employment type in Central Asia. It is significantly decreasing only female's current working probability. Estimated negative effect is strong among females living in rich households and urban regions. However, if estimation is restricted to mothers who has at least one or two children, it is also becoming insignificant factor to explain mother's current labor supply. Hence, I can state that becoming a mother decreases female's labor supply significantly but, once female entered to motherhood, marginal change in the number of children does not effect on their working probability significantly.

Acknowledgements

First, I would like to express my sincere gratitude to my research supervisor Professor Andrea Weber for giving me the great opportunity, and providing precious guidance throughout the writing this research project. Moreover, I would like to thank all CEU professors for their interesting and precise lectures that motivated and encouraged me to become a better professional in my sphere. Lastly, I have to mention that I am extremely grateful to my friends, parents and siblings for their constant support and love.

Table of contents

List of figures

List of tables

Introduction

There are a number of previous researchers study the causal effect of fertility rate on female's labor force participation. Although the negative impact of having more children on female labor supply is quite robust across studies conducted in developed economies, its positive and no causal effects are also found in developing countries and context specific explanations are given. Theoretical models find both negative and positive effect justifiable as having more children requires substantial amount of both monetary and time investments from parents. In families where monetary resources are limited, every additional child makes harder to meet daily expenses and motivate mothers to increase their labor supply in order to cover these costs. On the other hand, having more children demands additional time for child care activities and increases mother's time allocation over household tasks by decreasing their labor supply ability especially in countries where females are primary homemakers and males are usual breadwinners. Therefore, it is quite hard to predict the effect of children on maternal employment in both monetary and time constraint binding countries.

In this project I study the causal effect of having more children on mother's employment for Central Asia–the region which has not been studied yet. Central Asian countries are usually associated with high fertility rates and major labor markets inequalities (Mogilevskii, 2020). A number of reports by international organizations indicated that strengthening patriarchal norms during post-Soviet transition period and decreasing government support to preschool childcare facilities contributed to worsening gender inequality in labor market in these countries (International Labour Organization, 2017; Khitarishvili, 2016a, b; Mee, 2001; Somach and Rubin, 2010). To illustrate, Khitarishvili (2016a) mentioned that gender wage gap increases during female primary childbearing years in Central Asia and 61 percent of inactive females reasoned domestic responsibilites for not participating to labor market in Tajikistan. Although these reports signal that having more children decreases Central Asian female labor supply, there is no empirical research conducted in this region yet to estimate the causal impact of having more children on parental labor supply.

Empirical studies that analyzed the causal impact of fertility rate on mother's labor supply, mostly used IV regression estimation due to apparent endogenous nature of chosen interest variable due to reverse causality and omitted variable bias. One of the commonly given example for omitted variable bias in literature is the ambition of the mothers. Usually more ambitious females are more career oriented and prefer to have fewer children. Under this scenario OLS method can overestimate the effect of fertility rate on mother's employment. Thus, to mitigate the potential endogenity problem, I also apply 2SLS regression method to estimate Local Average Treatment Effect (LATE).

I measure female labor supply in three different ways: her current working status, her employment type and occupation type based on skill requirement. Infertility shock, having twin children at first and second births and son preference are applied as instrument variables to estimate the causal effect. I assume they are exogenous random shock to the number of living children female has and cannot impact directly to the mother's labor force participation similar to previous literature.

Using OLS regression I find strong negative association between the number of children and all measurement of female labor supply except for low-skilled occupation type. However, once endogenity issue is mitigated using different instrument variables, having more children only decreased women's current working probability -the effect being stronger for females living in urban regions and in rich wealth quantile households. This negative effect is not observed when women who do not have any children are removed from the sample. On the other hand, I estimate insignificant impact of the number of children on mother's employment and occupation type in the last 12 months' period regardless of using all females or only mothers with children in the analyses.

This paper makes two main contributions to the literature: To the best of my knowledge, it is the first paper studying the causal impact of fertility rate on female labor supply for Central Asia countries. Disaggregated regression estimation of this research paper can be useful for future policy implementations in this region. Secondly, this research paper includes a wide range of the instrument variables while previous literature used only one or two type of instrument variables focusing on restricted sample of females. Exploiting a number of instrument variables for different group of females boosted external validity of the current study and its comparability to wide range of the previous literature outcomes.

The rest of the paper is organized as follows: the next section gives background information about Central Asia. Section 3 describes data used in the analyses followed by the econometric method of the estimation. After that I provide regression results and discuss the estimated outcome. Finally, I draw conclusions and provide insights for further research analyses.

Background information

Central Asia includes 5 developing countries (Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan and Uzbekistan) that are hugely impacted by Soviet Union centralized political system and currently all are in some stages of moving to market economy. These countries share similar cultural, traditional background and share similar historical civilization. According to Mogilevskii, (2020) and Khitarishvili, (2016a) during Soviet Union period gender equality in employment were improved noticeably in this region due to weakening of traditional cultural norms, provided child care facilities and maternal leave opportunities. However, these countries inherited gender segregation in labor market: females are in mainly low paid jobs in education and healthcare sector while males are being busy in industrial activities, high senior positioned jobs. Currently among Central Asian countries, Kazakhstan, which has relatively better gender equality in employment, has the highest ($9812.5) GDP per capita whilst Tajikistan, that shows the worst statistics in terms of gender equality in labor force, has the worst ($870.8) GDP per capita among other Central Asian countries. The other Central Asian countries, Uzbekistan and Turkmenistan does not provide solid data on labor force participation in terms of gender equality. According to the table 1 below, 71.8% of females are economically active in Kazakhstan, and 37% of managerial positions are employed by them. In Kyrgyzstan, under half (47.3%) females participate in labor market, 27.69% of them is part-time workers, while 77.7% of males have been employed, and only 14.47% of them does part-time jobs. Tajikistan, another Central Asian country, shows the worst labor market conditions; just one-thirds (30.7%) of women and near the half (52.3%) men are in the labor force participation mostly explained by substantial labor migration to Russia. Female managers control only 6.6% of local firms; male managers hold 93.6 % of total firms in Tajikistan.

A number of descriptive reports done by UNICEF, ILO and USAID indicated that high fertility rate and strong traditional patriarchal norms can explain this gender inequality in this region as females are mostly busy with domestic and care burdens which in result decrease their time available for professional development (International Labour Organization, 2017; Khitarishvili, 2016a, Somach and Rubin, 2010).

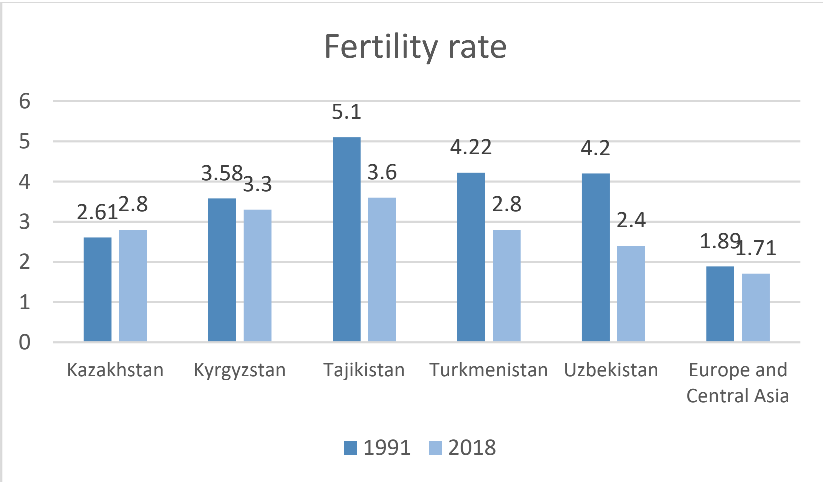
Table 1. Selected economic indicators of Central Asian countries.

	Kazakhstan			**Kyrgyzstan**			**Tajikistan**		
	female	male	f/m	female	male	f/m	female	male	f/m
Labor force participation rate, %	71.8	81.7	0.88	47.3	77.7	0.61	30.7	52.3	0.59
Estimated earned income, int'l $ 1,000	18.5	31.3	0.59	3.4	7.2	0.47	2.1	4.7	0.45
Legislators, senior officials and managers, %	37	63	0.59	37.8	62.2	0.61	14.8	85.2	0.17
Workers employed part-time, % of employed people	10.8	6.44	1.68	27.69	14.47	1.91	n/a	n/a	n/a
Firms with female top managers, % firms	26	74	0.35	32.9	67.1	0.49	6.6	93.4	0.07
GDP per capita in 2019 (in current USD)	9812.5			1309.5			870.8		

Source: WEF, 2021

Although the overall fertility rate decreased noticeably in this region since 1991 except Kazakhstan, it is still higher than its aggregates- Europe and Central Asia region. Fertility rate of Uzbekistan declined by 1.8, which is the sharpest decrease among the countries, and thereby showing the lowest rate. The rate in Tajikistan continuous to be highest, 5.1 in the year of 1991, and 3.6 in the year of 2018.

Figure 1. Fertility rate among Central Asian countries



Source: World Bank indicators, 2021

Literature review

A number of researchers consider the recent noticeable decrease in fertility rate in developing countries as one of the important reasons of the increase in female employment. For example, Cáceres-Delpiano (2012) found that having more children decreases a mother's labor force participation based on a sample of 40 low- and middle –income countries. Similar negative impact of fertility was also observed in Africa and developed economies. (Aaronson et al., 2017; de Jong, Smits, & Longwe, 2017). However, some papers conclude that having more children does not explain female participation in the labor force (Agüero & Marks, 2008, 2011; Majbouri, 2019). Moreover, Ajefu, (2019), Priebe (2010) and Trako (2018) estimated that having more children actually increases mother's labor supply in Nigeria, Indonesia rural regions and Albania. Thus the impact of having more children on a mother's labor force participation is quite heterogeneous and conditioned on some factors like family size, economy development status, social norms and labor market conditions. Most commonly used instrument variables to estimate the causal impact of fertility rate are parental preference for either mixed sex sibling composition or specifically choosing one gender over another (son preference), twinning at first/second birth and fertility shocks. To illustrate, mixed gender composition preference introduced by Angrist & Evans (1998) later also applied by Cruces & Galiani, (2007) for Latin America and by Schmieder (2020) for Mexico to estimate the effect of having three and more children on maternal labor supply in those region . However, in other developing countries where female faces systematic discrimination or when son provides support to old-aged parents, the presence of daughters are considered positive shock to fertility and instead son preference is used as instrument variable (Azimi, 2015; Chun & Oh, 2002; Daouli, Demoussis, & Giannakopoulos, 2009; Van Der Stoep, 2008). Even though gender of the children is quite randomly assigned and balanced across samples, its estimated effect is quite restricted to the mothers who have two or more children. Another commonly used instrument variable- having twin at first birth is introduced by Rosenzweig & Wolpin (1980) and later applied by a number of researchers for both developing and developed countries (Aaronson et al., 2017; Ajefu, 2019; Cáceres-Delpiano, 2012; de Jong et al., 2017; Majbouri, 2019; Vere, 2011). Twinning at first birth allows to study the effect of having two and more children on parental labor supply covering more portion of females than mixed sex sibling composition instrument choice. However, as having twin is quite rare occurred exogenous shock to the fertility rate, it is quite data demanding comparing to other instrument variable choices. And hence it is hard to do disaggregated analyses using this instrument variable- comparing the effect of having more children for different group of females. Lastly, another instrument variable that is pioneered relatively recently by Agüero & Marks (2008, 2011) is infertility shock- losing ability of conceiving any more children due to some medical reasons. Even though this instrument is less convincible in terms of its random distribution across females, it allows to include women who do not have any children in the analyses and make analyses among broader range of females.

Methodology

Data and sample characteristics

In this paper, I use all available Demographic Health Surveys (DHS) conducted in Central Asia in different period of time starting from 1995 till 2017 available at https://dhsprogram.com/. DHS provides approximately 400 publicly available cross sectional surveys for more than 90 developing countries surveyed extensively from females aged 15-49. Previous DHS surveys conducted in Africa and South Asia were used to study similar research questions as they provide essential variables for estimation. I make pooled cross sectional data analyses by appending all available 7 datasets. Pooling all these surveys allows me to increase the sample size in my estimation and see labor force participation over time and across regions.

I restrict sample size to woman whose age is between 18-44 and age of the oldest child is less than 18 if she has any similar to previous studies. It will reassure all children are present in the household and women whose ages are between 18-44 are most likely to have ability to give birth and to have young children in their family. Consequently, I have 30802 respondents in my final sample group (Appendix Table 1) and table 2 column 1 provides descriptive statistics of these respondents.

Labor force participation of these women is measured as dummy variables in 5 different ways based on their current working status, their occupation type based on skill requirement and employment type in the last 12 months. 32.42 percent of respondents were working at the time of the survey and approximately 40 percent were working in some type of occupation group in the last 12 months. DHS surveys provide occupation groups of the respondents that can be categorized to high skilled versus low skilled occupations. 5 occupation groups are measured as high skilled occupation: professional/technical/managerial, clerical, sales, services and skilled manual. Low skilled occupation groups include 4 occupation groups: agricultural- self employed, agricultural- employee, household and domestic and unskilled manual. Around 30 percent of the respondents are employed in high skilled occupation while 10 percent are busy with low skilled occupation in the last 12 months. Lastly, employment type of the respondents is identified based on question asking for whom she is working for: for family members, for someone else or for herself. If respondents are working for family members or for herself, they are categorized as self employed (1) otherwise 0. If respondents are working for someone else, they are categorized as wage-employee. This question is missing in Kazakhstan DHS survey conducted in 1999 and hence this survey is excluded in the analyses of employment type.

On average respondents in the survey had around two living children and one children who are under 5 years old. Gender of the firstborn and second born children are quite balanced and around 0.7 percent of the females had twin at either first or second birth. Infertility status of the women is identified based on two questions given in DHS survey: female is considered infertile if she pointed it as a reason for not using contraception or if she said she is unable to have any more children to the question to her desire to have more children: 4.7 percent of the females in the survey are infertile according to the table 2.

Table 2. Descriptive statistics of respondents

	All respondents	Sample 1	Sample 2	Sample 3
Labor force participation				
Currently working	0.3242	0.2718	0.3282	0.3291
	(0.4681)	(0.4449)	(0.4696)	(0.4699)
High skilled occupation	0.2954	0.2504	0.3055	0.2893
	(0.4562)	(0.4332)	(0.4606)	(0.4535)
Low skilled occupation	0.0933	0.0851	0.0984	0.1112
	(0.2909)	(0.2791)	(0.2979)	(0.3144)

Wage employed	0.2899	0.2431	0.3057	0.3024
	(0.4537)	(0.4290)	(0.4607)	(0.4593)
Self-employed	0.0896	0.0838	0.0884	0.0917
	(0.2856)	(0.2772)	(0.2839)	(0.2886)
Fertility rate				
Number of living children	1.6835	1.7518	2.3877	2.8995
	(1.4914)	(1.3913)	(1.2137)	(1.0219)
Number of children under 5	1.1326	1.3408	1.3815	1.4864
	(1.2334)	(1.3082)	(1.2367)	(1.2909)
Infertile	0.0471	0.0909	0.0460	0.0416
	(0.2119)	(0.2875)	(0.2094)	(0.1997)
Have twin at 1st birth	0.0076	0.0082	0.0073	0.0086
	(0.0868)	(0.0900)	(0.0853)	(0.0922)
Have daughter at 1st birth	0.4891	0.4978	0.4877	0.4876
	(0.4999)	(0.5000)	(0.4999)	(0.4999)
Have twin at 2nd birth	0.0071	0.0059	0.0070	0.0068
	(0.0837)	(0.0766)	(0.0833)	(0.0822)
Have daughter at 2nd birth	0.4894	0.4976	0.4889	0.4832
	(0.4999)	(0.5000)	(0.4999)	(0.4997)
Individual characteristics				
Age	27.8626	28.1539	29.9348	31.2137
	(6.6471)	(6.3325)	(5.9113)	(5.3469)
BMI score	23.4207	23.7165	23.9108	24.2290
	(4.4100)	(4.4875)	(4.5348)	(4.6186)
Higher education	0.2617	0.2277	0.2339	0.2169
	(0.4396)	(0.4194)	(0.4233)	(0.4121)
Had miscarriage	0.0944	0.1014	0.1217	0.1286
	(0.2924)	(0.3019)	(0.3269)	(0.3347)
Visited health facility last 12 months	0.5641	0.6371	0.6421	0.6358
	(0.4959)	(0.4808)	(0.4794)	(0.4812)
Age at first cohabitation	20.4018	20.4645	20.2984	20.1655
	(3.2280)	(3.3598)	(3.0965)	(2.9091)
Married	0.7206	0.8538	0.9064	0.9430
	(0.4487)	(0.3533)	(0.2912)	(0.2318)
Household level factors				
Number of household members	6.3972	6.7527	6.5662	6.8127
	(3.1950)	(3.4726)	(3.1935)	(3.1164)
Urbanity status	0.4205	0.3730	0.4034	0.3780
	(0.4937)	(0.4836)	(0.4906)	(0.4849)
Wealth status: poorest	0.1723	0.1779	0.1774	0.1893
	(0.3777)	(0.3825)	(0.3820)	(0.3917)
poorer	0.1679	0.1813	0.1735	0.1820
	(0.3738)	(0.3853)	(0.3786)	(0.3859)
middle	0.1803	0.1930	0.1841	0.1879

	(0.3845)	(0.3947)	(0.3876)	(0.3907)
richer	0.2014	0.2010	0.2048	0.1978
	(0.4010)	(0.4008)	(0.4035)	(0.3984)
richest	0.2780	0.2468	0.2603	0.2429
	(0.4480)	(0.4312)	(0.4388)	(0.4289)
Observations	30802	15074	21718	15866

mean coefficients; Standard errors are in parentheses

To estimate the causal effect of fertility rate on female's labor supply, further restrictions are introduced based on instrument variable choice. In Sample 1 infertility shock is used as instrument variable. As it is observable only for females who are not using contraceptives I remove females who are using contraceptives or sterilized and who never had sexual encounter from the sample similar to Aguero and Marks (2008, 2010). In sample 2, all respondents are restricted to mothers who has at least one child. In this sample gender of the first child and having twin at first birth are used as instrument variables. As these women decided already to have a child, they are comparable to mothers who have at least one child. Lastly in sample 3, respondents are restricted to mothers who have at least two children and having twin at second birth and gender composition of first two children are used as instrument variables. These women already decided to have two children at least and hence they are comparable to women who have two or more children. Household level factors are quite similar across different sample of groups and whole sample. While individual level factors are slightly different by construct of the sample. In the samples, respondents are older and slightly heavier (higher BMI index) and more portion of them are married.

Estimation model

If the number of children is randomly assigned, then the difference in the mean outcomes of labor force participation of women who have more children and women who have few or no children would give us average treatment effect (ATE). However, as the number of children is endogenously determined by respondents, I use 2SLS regression estimation using infertility, having twin and son preference as instrument variables to estimate local average treatment effect (LATE).

First stage of estimation can be formulated in the following way:

$$N_i = X_i'\delta + \gamma I_i + \tau + \varphi + \varepsilon_i$$

Where N is the number of living children female *i* has at the time of survey. X_i' is a set of relevant control variables including age of the respondent, health status, education level, household wealth quantile, number of household members and urbanity status that can impact on their labor supply decision. I is dummy variable equals to one if the respondent *i* has infertility problem for sample 1; or gender of the firstborn or having twin at first birth for sample 2; or gender composition of firstborn and second born or having twin at second birth for sample 3. Ideally I assume these instrument variables are exogenous random shocks to the number of children female has and it effects to her labor supply through only by decreasing/increasing their fertility rate. If conditional independence, exclusion and monotonicity assumptions are not violated, second stage of the equation provides LATE.

Further predicted number of children is applied to the second stage of the estimation where our interest (LATE) is β coefficient. It shows the effect of having more children to the compliers - subpopulation that responds to a change in the value of the instrument as having more children is arbitrarily heterogeneous among respondents.

$$LFP_i = X_i'\alpha + \beta N_i + \tau + \varphi + \eta_i$$

LFP_i is labor force participation of *i* respondent that is measured as dummy variable. It equals to one if respondent is working at the time of survey otherwise equals to zero. Further it measures the type of employment and occupation she is busy with in the last 12 months' period. In both stage of the equation region τ and time φ fixed effects are included to control time and regional unobservable differences. Errors terms are clustered at household level since respondents living in the same household will share some similar characteristics like family background, household factors and wealth status and hence error terms of these respondents can be correlated with each other.

Identification assumptions

In order to use 2SLS to find LATE, four main assumptions are necessary to hold. Below I discuss the relevance and validity of the chosen instrument variables for the case of Central Asia.

Similar to previous literature, I apply some empirical tests that can show any violation of the assumptions.

Non trivial first stage assumption specify that chosen instrument variables effect the number of children female has. First stage of the IV estimation can show how these variables are effecting to the number of children with and without control variable sets given in table 3. According to the table, infertility indeed decreasing the number of children female has by 0.73 point and the having twin at first and second birth and son preference are increasing fertility rate statistically significantly. According to given in table 3 model 2, OLS coefficients of chosen instrument variables do not change noticeably when control variables are included to the model showing they are not sensitive to observed characteristics of the respondents.

Table 3. First stage of IV estimation. Relevance of the instrument

	Sample 1	Sample 2		Sample 3	
Dependent variable: Number of children	Infertility	Had twin at 1st birth	Had girl at 1st birth	Had twin at 2nd birth	Had girl at 1st and 2nd birth
Model 1					
Infertile	-0.734^{***}				
	(0.037)				
Had twin at 1st birth		0.272^{***}			
		(0.076)			
Had girl at 1st birth			0.097^{***}		
			(0.012)		
Had twin at 2nd birth				0.545^{***}	
				(0.072)	
One boy one girl					
Girls					0.236^{***}
					(0.017)
Boys					-0.024
					(0.016)
Observations	15074	21718	21718	15866	15866
Adjusted R^2	0.32	0.45	0.45	0.32	0.33
Model 2					
Infertile	-0.776^{***}				
	(0.037)				
Had twin at 1st birth		0.264^{***}			
		(0.078)			
Had girl at 1st birth			0.108^{***}		
			(0.012)		
Had twin at 2nd birth				0.610^{***}	

				(0.073)	
One boy one girl					
Girls					0.236***
					(0.017)
Boys					-0.033**
					(0.015)
Observations	14486	20382	20382	15019	15019
Adjusted R^2	0.39	0.53	0.53	0.39	0.40

Source: DHS surveys; Model 1 includes age, age-squared and age at first birth polynomial form, survey fixed effect. Model 2 additionally includes education level, BMI index, marital status, had miscarriage dummy, number of household members, wealth quantile and urbanity status; Standard errors are in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Independence assumption indicate that having twin, infertility status and gender of the children are randomly distributed among respondents and females cannot influence on their probability of receiving it. Having twin children and gender of the children are considered quite randomly distributed historically. However, as medicine is developing significantly, these days some medical advancements are possible that can increase the probability of having twin at birth or having gender selective birth. However, Central Asian countries are still low, low-middle income countries- most of the females cannot afford these medical procedures yet and hence it is highly unlikely that they can influence on their probability of having twin or gender of their children. Another vital factor that can violate the independence assumption is age of the female at birth. Medical literature showed that being older increases the probability of having twin. Therefore, following the previous studies, I control for a third-order polynomial of mother's age at first birth to solve this problem. Table 4 can show independence assumption to certain extent as it shows significance of observed characteristics of female to the probability of facing infertility issue, having twin and having gender selective birth. Based on adjusted-R square values these factors are not explaining the chosen instruments, except infertility shock. The probability of being infertile is affected by age, BMI score, education level, wealth and urbanity status. Medical literature indicates that females' infertility probability increases with their age, weight and experience of miscarriages (Aguero and Marks, 2008, 2010). Furthermore, as this measurement is self-informed by respondents, the effect of wealth and urbanity status can be observed due to access to the information. Wealthier households have higher access to advanced medical services and find out their infertility status. Additionally, as most recent survey has more infertile women comparing to earlier ones, it also explains better access to the medical services over time. Thus we cannot conclude that infertility shock is random among respondents. Therefore, I will control these variables in my IV estimation to reassure conditional independence of this instrument variable.

Table 4. OLS regression estimation. Randomness of the instrument

	Sample 1	Sample 2		Sample 3	
	Infertility	Had twin at 1st birth	Had girl at 1st birth	Had twin at 2nd birth	Had girl at 1st and 2nd birth
Age	-0.032***	-0.000	0.002	0.002	-0.003
	(0.004)	(0.001)	(0.006)	(0.001)	(0.008)
Age squared	0.001***	0.000	-0.000	-0.000*	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
BMI score	0.002***	0.000*	-0.000	0.000	0.001
	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)
Had miscarriage	0.003	-0.001	0.021*	-0.001	-0.004
	(0.008)	(0.002)	(0.011)	(0.002)	(0.013)
Married	0.040***	-0.001	-0.023*	-0.006	-0.001
	(0.007)	(0.002)	(0.013)	(0.004)	(0.018)
Visited health facility last 12 months	-0.006	0.001	0.011	-0.002	0.022**
	(0.005)	(0.001)	(0.008)	(0.002)	(0.009)
Higher education	-0.015**	0.003*	-0.007	-0.000	0.012
	(0.006)	(0.002)	(0.009)	(0.002)	(0.011)
Wealth status: poorest					
Poorer	0.008	-0.001	-0.002	0.006**	-0.009
	(0.007)	(0.002)	(0.012)	(0.002)	(0.013)
Middle	0.007	-0.001	-0.010	0.002	-0.017
	(0.007)	(0.002)	(0.012)	(0.002)	(0.013)
Richer	0.003	0.000	-0.005	0.001	0.014
	(0.007)	(0.002)	(0.012)	(0.002)	(0.014)
Richest	0.022**	-0.000	-0.006	0.005**	-0.013
	(0.009)	(0.002)	(0.014)	(0.003)	(0.016)
Urbanity status	0.013**	-0.002	0.004	-0.002	-0.001
	(0.007)	(0.002)	(0.010)	(0.002)	(0.012)
Age of respondent at 1st birth		-0.021**	0.014	0.013	0.083
		(0.009)	(0.054)	(0.012)	(0.083)
Age_at_first_squared		0.001**	-0.001	-0.000	-0.003
		(0.000)	(0.002)	(0.001)	(0.003)
Age_at_first_polynomial		-0.000***	0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)	(0.000)
Kazakhstan 1995					

Kazakhstan 1999	0.025	-0.000	-0.011	0.003	-0.001
	(0.016)	(0.003)	(0.019)	(0.005)	(0.024)
Kyrgyzstan 1997	0.044^{***}	0.000	-0.009	-0.001	0.028
	(0.013)	(0.003)	(0.016)	(0.003)	(0.020)
Kyrgyzstan 2012	-0.011	-0.004	-0.012	-0.001	0.026
	(0.010)	(0.002)	(0.015)	(0.003)	(0.018)
Tajikistan 2012	0.003	0.000	-0.010	-0.004	0.022
	(0.010)	(0.002)	(0.014)	(0.003)	(0.017)
Tajikistan 2017	0.037^{***}	0.002	0.006	0.001	0.007
	(0.010)	(0.002)	(0.014)	(0.003)	(0.017)
Uzbekistan 1996	0.003	0.001	0.001	-0.002	0.017
	(0.012)	(0.003)	(0.016)	(0.003)	(0.019)
Constant	0.297^{***}	0.164^{**}	0.391	-0.138	-0.197
	(0.053)	(0.077)	(0.445)	(0.095)	(0.656)
Observations	14481	20360	20360	15000	15000
Adjusted R^2	0.07	0.00	-0.00	0.00	0.00

Source: DHS surveys; Standard errors are in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Exclusion assumption of identification strategy requires that infertility, having twin and sex of the child can impact of female labor force participation only through by decreasing or increasing fertility rate. There are some papers questioning this assumption for having twin and son preference instrument variables: it can decrease the cost of child raising per child as twins grow at the same time and requires less attention comparing to having two separate children subsequently. Furthermore, having the same gender children comparing to mixed gender composition also can be less costly upbringing due to reusing similar resources. Another arising concern can happen if families distribute the resources in discriminatory manner: more resources toward sons comparing to daughters. Unfortunately, I could not compare the household expenditures of families with twin comparing to without twin as well as females who have the same gender children comparing to the ones with both gender as it is not given in the dataset. Previous literature found that household expenditures are not varying significantly between these groups of females and I also assume that it is the case for Central Asia as these countries are also low income developing countries.

Lastly, monotonicity assumption requires that infertility, having twin and son preference should provide similar effect across different group of females. Appendix figure 1 presents the empirical evidence for infertility shock and son preference instrument variables are indeed decreasing and increasing the number of children females have across different age groups. As having twin is not quite frequent among respondents, I cannot construct this figure for twin instrument. Furthermore, some literature mentioned that son preference can vary across different group of females and these households may be systematically different from child sex indifferent households in terms of traditional view. More traditional households prefer to have more sons and also may not tolerate female labor force participation. Therefore, in appendix table 2 I tested the effect of having two girls on the number of children among different group of females. I estimate almost similar OLS coefficients for different group of females indicating that son preference is quite homogenous across different group of females.

Result

Table 5 provides OLS and 2SLS regression estimations of the number of children on the mother's currently working status as well as her employment and occupation type in the last 12 months in Central Asia across different sample groups of females using different instrument variables. OLS regression estimated among a whole sample of females are presented in first row to show general association between fertility rate and female labor supply. Sample 1 excludes females who never had sexual encounter and who were using contraceptives or sterilized. Sample 2 and 3 includes mothers who have at least one and two children respectively. In all estimations, I control for (country#year) survey fixed effect along with control variables-respondent's age in quadratic form, BMI score, education level, marital status, number of household members excluding her and her children, household wealth quantile and urbanity status. Age at first birth is included as polynomial form among females who have at least one or two children (sample 2 and sample 3). Robust standard errors are applied in all estimation.

All standard OLS models estimate that there is a negative significant association between the number of children and a woman working status except for low skilled occupation group. More specifically, having one more child is correlated with around 3.6 percent point decrease in mother's currently working probability in whole sample and association is strongest for high skilled occupation type (4.2 percent). Whereas having one more child is associated with 1 percent point increase in the probability of working low skilled occupation.

Since OLS model provides simple correlation estimation between fertility rate and mother's employment, I move to IV regression estimation to estimate the causal impact. According to second stage of IV regression estimation, having more children decreases women's current working probability among females in sample 1. However, among mothers who have at least one and two children given in sample 2 and 3, using different instrument variables show no effect of children on mother's current employment status. Furthermore, once endogenity issue is solved using IV regression method, the number of children become an insignificant factor to explain females' employment and occupation type in the last 12 months across different sample of females except group of females who has at least one children. When variation in the number of children is increasing due to having twin at the first birth, having more children decreases mother's probability of working low skilled occupation by 17.3 percent point.

Table 5. Main regression estimation

All respondents	Working	Self-employed	Wage-employed	High- skilled	Low-skilled
OLS					
Number of living children	-0.036***	-0.004**	-0.027***	-0.042***	0.010***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Observations	28851	27307	27307	28864	28864
Adjusted R^2	0.17	0.02	0.26	0.26	0.08
Sample 1					
2SLS-infertility					
Number of living children	-0.048***	-0.004	-0.012	-0.019	0.008
	(0.017)	(0.012)	(0.016)	(0.015)	(0.011)
First stage F-value	586.06	606.58	606.58	586.77	586.77
Observations	14477	14003	14003	14486	14486
Sample 2					
2SLS -Had daughter at 1st birth					
Number of living children	0.052	0.012	0.071	0.024	0.059
	(0.055)	(0.036)	(0.051)	(0.050)	(0.038)
First stage F-value	1116.24	1160.09	1160.09	1116.86	1116.86
2SLS-Had twin at 1st birth					
Number of living children	-0.128	-0.059	0.047	0.133	-0.173*
	(0.136)	(0.079)	(0.128)	(0.140)	(0.092)
First stage F-value	1018.88	1052.06	1052.06	1019.44	1019.44
Observations	20375	19261	19261	20382	20382
Sample 3					
2SLS –Had 2 girls at 1st and 2nd birth					
Number of living children	0.014	-0.006	0.034	0.020	0.009
	(0.033)	(0.022)	(0.031)	(0.030)	(0.024)
First stage F-value	466.12	481.20	481.20	466.58	466.58

2SLS -Had twin at 2nd birth					
Number of living children	0.076	0.027	0.003	-0.056	0.076
	(0.073)	(0.052)	(0.065)	(0.063)	(0.054)
First stage F-value	441.37	453.51	453.51	441.73	441.73
Observations	15013	14311	14311	15019	15019

Source: DHS surveys; all estimation includes age, age-squared, education level, BMI index, marital status, had miscarriage, wealth quantile, number of household members, urbanity status and survey fixed effect.; Sample 2 and 3 additionally includes age at first birth at polynomial form; Kazakhstan DHS 1999 is not included for wage employment and self-employment analyses due to data availability issue; Standard errors are in parentheses; * p < 0.1, ** $p < 0.05$, *** $p < 0.01$

Estimated negative effect of having more children on women's current working probability and being employed in low skilled occupation are robust to control variable set which can indicate chosen instrument variables are not sensitive to the given control variable set and less likely suffer from omitted variable bias (Appendix table 3). Moreover, the effect of the number of children is the strongest when children under 6 are used as fertility rate rather than total number of children (Appendix table 4). Among females in sample 1 having one more children under 6 in the household decreases the current working probability of mothers by 8.6 percent point. Lastly among married females given in appendix table 5, I consistently find that the number of children decreases the women's working probability and having more children decreases the probability of working in low skilled occupation among mothers who has at least one child when twin is instrumented in the model.

Disaggregated analyses

The effect of the number of children on maternal labor supply conditioned on many external factors and varied noticeably from one group of females to other. Thus, I estimate the effect of having more children for different group of females and compare the outcome: urban versus rural regions, nuclear versus extended families and rich versus poor wealth quantile households. In disaggregated regression analyses I do not use having twin at first and second birth as instrument variables since twinning is not balanced across groups due to very few number of respondents with twins. Furthermore, disaggregated analyses are estimated among only married females who are most part of the sample as they are main representative of the population and living conditions of not married respondents can be quite different than married ones.

Usually, urban and rural females face different labor demand conditions and estimating the effect for these regions separately allows to study it under different labor market conditions. According to OLS estimations given in Table 6, negative correlation between the number of children and mother's labor force participation in urban region is stronger comparing to in rural region: having one more child is correlated with around 5.2 percent and 1.6 percent point decrease in mother's current working probability in urban and rural regions respectively. 2SLS regression estimation showed that actually the negative effect of having more children is even stronger for urban region as the coefficient increased to 9.4 percent while it is insignificant factor for rural females' labor supply. However, estimation among only mothers who have at least one and two children given in sample 2 and 3 showed no effect of having more children on female labor supply in both regions.

Studying the effect of having more children for females who live in nuclear families (approximately 40% of the sample) versus extended families (approximately 60% of the sample) can be informative as mothers living in extended families can get external help from other household members in their child care and domestic activities. On the other hand, they may have extra homecare burdens due to extra household members. While having more children decreases current labor supply of mothers from both family types, it is stronger for women living in nuclear families.

Lastly estimating the effect of fertility rate for females living in rich households where monetary constraint is less binding versus poor households could show how household wealth mediates the effect having more children. According to 2SLS estimation using infertility as instrument variable, the number of children decreases current working probability of mothers from rich income households while it is insignificant factor for mothers from poor wealth quantile. Furthermore, individual level disaggregated estimation showed that the negative effect of having more children on current working probability is stronger for younger mothers and for mothers who has higher education (Appendix table 6).

Table 6 . Disaggregated regression estimation among different group of married females

	Working	Self-employed	Wage-employed	High-skilled	Low-skilled	Working	Self-employed	Wage-employed	High-skilled	Low-skilled
Living place			**Urban**					**Rural**		
OLS	-0.052***	-0.006**	-0.031***	-0.038***	0.000	-0.016***	0.000	-0.016***	-0.029***	0.012***
	(0.005)	(0.003)	(0.005)	(0.005)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)
Observations	7830	7271	7271	7834	7834	13137	12708	12708	13143	13143
Adjusted R^2	0.20	0.02	0.28	0.25	0.03	0.14	0.02	0.28	0.25	0.08
Sample 1										
2SLS-infertility instrument	-0.094***	-0.038*	-0.015	-0.036	-0.002	-0.026	0.019	-0.007	-0.006	0.014
	(0.029)	(0.020)	(0.027)	(0.027)	(0.012)	(0.021)	(0.014)	(0.019)	(0.016)	(0.016)
Observations	4205	4038	4038	4208	4208	8268	8124	8124	8273	8273
Sample 2										
2SLS -Had daughter at 1st birth	0.037	0.022	0.029	0.006	0.029	0.041	0.042	0.034	0.027	0.057
	(0.130)	(0.074)	(0.113)	(0.128)	(0.060)	(0.056)	(0.040)	(0.053)	(0.048)	(0.045)
Observations	6938	6434	6434	6941	6941	11632	11224	11224	11635	11635
Sample 3										
2SLS –Had 2 Girls at 1st and 2nd birth	0.030	0.035	0.017	0.070	-0.021	-0.018	-0.012	0.009	-0.021	0.014
	(0.056)	(0.036)	(0.052)	(0.055)	(0.025)	(0.039)	(0.027)	(0.036)	(0.033)	(0.033)
Observations	5121	4813	4813	5123	5123	9074	8759	8759	9077	9077
Family type			**Nuclear**					**Extended**		
OLS	-0.040***	-0.004	-0.027***	-0.044***	0.011***	-0.024***	-0.003	-0.018***	-0.028***	0.007***
	(0.004)	(0.003)	(0.004)	(0.004)	(0.003)	(0.004)	(0.002)	(0.003)	(0.003)	(0.003)
Observations	8304	7628	7628	8308	8308	12663	12351	12351	12669	12669
Adjusted R^2	0.16	0.02	0.27	0.24	0.10	0.14	0.02	0.27	0.27	0.06
Sample 1										
2SLS-infertility	-0.065**	0.007	-0.027	-0.031	0.018	-0.049**	-0.014	0.001	-0.016	0.006
	(0.030)	(0.020)	(0.029)	(0.026)	(0.019)	(0.021)	(0.014)	(0.018)	(0.017)	(0.014)
Observations	3887	3707	3707	3890	3890	8586	8455	8455	8591	8591
Sample 2										
2SLS -Had daughter at 1st birth	0.089	0.121**	0.042	0.071	0.102**	-0.003	-0.059	0.044	-0.012	-0.008
	(0.072)	(0.048)	(0.064)	(0.067)	(0.049)	(0.081)	(0.057)	(0.078)	(0.072)	(0.057)
Observations	7978	7336	7336	7982	7982	10592	10322	10322	10594	10594
Sample 3										
2SLS –Had 2 Girls at	-0.002	0.047	-0.010	-0.017	0.040	0.003	-0.036	0.042	0.049	-0.038

1st and 2nd birth										
	(0.046)	(0.031)	(0.042)	(0.042)	(0.033)	(0.046)	(0.031)	(0.042)	(0.040)	(0.033)
Observations	6847	6393	6393	6850	6850	7348	7179	7179	7350	7350
Wealth status			**Poor**					**Rich**		
OLS	-0.014***	0.002	-0.011***	-0.024***	0.015***	-0.047***	-0.008**	-0.033***	-0.042***	0.002
	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)	(0.002)
Observations	7591	7289	7289	7596	7596	9371	8869	8869	9376	9376
Adjusted R^2	0.13	0.03	0.26	0.23	0.08	0.20	0.02	0.29	0.27	0.04
Sample 1										
2SLS –infertility	-0.026	0.016	0.004	0.017	0.005	-0.069**	-0.014	-0.027	-0.033	0.006
	(0.027)	(0.018)	(0.024)	(0.019)	(0.022)	(0.027)	(0.018)	(0.025)	(0.025)	(0.013)
Observations	4676	4571	4571	4681	4681	5286	5152	5152	5289	5289
Sample 2										
2SLS -Had daughter at 1st birth	-0.001	0.002	0.018	-0.083	0.078	0.090	0.021	0.089	0.122	0.011
	(0.070)	(0.050)	(0.070)	(0.058)	(0.060)	(0.093)	(0.055)	(0.081)	(0.092)	(0.046)
Observations	6767	6482	6482	6769	6769	8280	7823	7823	8284	8284
Sample 3										
2SLS –Had 2 Girls at 1st and 2nd birth	0.011	-0.022	0.011	-0.044	0.012	-0.019	0.017	0.002	0.052	-0.023
	(0.048)	(0.033)	(0.045)	(0.039)	(0.041)	(0.049)	(0.032)	(0.045)	(0.048)	(0.025)
Observations	5347	5125	5125	5349	5349	6116	5844	5844	6119	6119

Source: DHS surveys; all estimation includes age, age-squared, education level, BMI index, husband age, had miscarriage, and survey fixed effect; Sample 2 and 3 additionally includes age at first birth at polynomial form; Kazakhstan DHS 1999 is not included for wage employment and self-employment analyses due to data availability issue; Standard errors are in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Conclusion

In this paper I estimate the impact of having more children on women's labor force participation using all waves of cross sectional DHS conducted between 1995 and 2017. Two stage least squares method using infertility shock, having twin children at first/second birth and son preference as instrument variables are applied to estimate the causal impact. Standard OLS estimation showed statistical significant negative correlation between the number of children and mother's current working status, employment and occupation type. Once the endogeinity issue was solved using IV-regression method, having more children did not show any impact on female employment and occupation type. Results were robust to regional labor market differences as well as respondents age and educational level. It is significantly decreasing only female's current working probability. Estimated negative effect is strong among females living in rich households and urban regions. However, if estimation is restricted to only mothers who has at least one or two children, it is also becoming insignificant factor to explain mother's current labor supply. Hence, I can state that becoming a mother decreases female's labor supply significantly but, once female entered to motherhood, marginal change in the number of children does not effect on their working probability significantly. Exploiting a number of instrument variables for different group of females boosted external validity of the current study and its comparability to wide range of the previous literature outcomes.

This study can be extended in a number of directions. Firstly, due to data availability issue, female labor force participation is measured in a quite basic level. It would be better if future studies can distinguish high paid jobs from low paid jobs as well as half time jobs from full time jobs as they are more informative for policy implications. Secondly, even though using a number of instrument variables improved external validity of the estimated outcome, due to data finding problem internal validity of the instrument variables are not tested fully.

Reference

Aaronson, D., Dehejia, R. H., Jordan, A., Pop-Eleches, C., Samii, C., & Schulze, K. (2017). The Effect of Fertility on Mothers' Labor Supply over the Last Two Centuries. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2915334

Agüero, J. M., & Marks, M. S. (2008). Motherhood and female labor force participation: Evidence from infertility shocks. *American Economic Review*, *98*(2), 500–504. https://doi.org/10.1257/aer.98.2.500

Agüero, J. M., & Marks, M. S. (2011). Motherhood and female labor supply in the developing world: Evidence from infertility shocks. *Journal of Human Resources*, *46*(4), 800–826. https://doi.org/10.3368/jhr.46.4.800

Ajefu, J. B. (2019). Does having children affect women's entrepreneurship decision? Evidence from Nigeria. *Review of Economics of the Household*, *17*(3), 843–860. https://doi.org/10.1007/s11150-019-09453-2

Angrist, J. D., & Evans, W. N. (1998). Children and Their Parents ' Labor Supply : Evidence from Exogenous Variation in Family Size Author (s): Joshua D . Angrist and William N . Evans Source : The American Economic Review , Vol . 88 , No . 3 (Jun ., 1998), pp. *American Economic Association*, *88*(3), 450–477.

Azimi, E. (2015). The effect of children on female labor force participation in urban Iran. *IZA Journal of Labor and Development*, *4*(1). https://doi.org/10.1186/s40175-015-0030-x

Cáceres-Delpiano, J. (2012). Can We Still Learn Something From the Relationship Between Fertility and Mother's Employment? Evidence From Developing Countries. *Demography*, *49*(1), 151–174. https://doi.org/10.1007/s13524-011-0076-6

Chun, H., & Oh, J. (2002). An instrumental variable estimate of the effect of fertility on the labour force participation of married women. *Applied Economics Letters*, *9*(10), 631–634. https://doi.org/10.1080/13504850110117850

Cruces, G., & Galiani, S. (2007). Fertility and female labor supply in Latin America: New causal evidence. *Labour Economics*, *14*(3), 565–573. https://doi.org/10.1016/j.labeco.2005.10.006

Daouli, J., Demoussis, M., & Giannakopoulos, N. (2009). Sibling-sex composition and its effects on fertility and labor supply of Greek mothers. *Economics Letters*, *102*(3), 189–191. https://doi.org/10.1016/j.econlet.2009.01.002

de Jong, E., Smits, J., & Longwe, A. (2017). Estimating the Causal Effect of Fertility on Women's Employment in Africa Using Twins. *World Development*, *90*, 360–368. https://doi.org/10.1016/j.worlddev.2016.10.012

International Labour Organization (2017) Women in business and management: Gaining momentum in Eastern Europe and Central Asia. Bureau for Employers' Activities (ACT/EMP)

Khitarishvili, T. (2016a) Gender inequalities in labour markets in Central Asia. UNDP/ILO conference on Employment, Trade and Human Development in Central Asia

Khitarishvili, T. (2016b) Gender Dimensions of Inequality in the Countries of Central Asia, South Caucasus, and Western CIS*. Working paper N 858, Levy Economics Institute of Bard College

Majbouri, M. (2019). Twins, family size and female labour force participation in Iran. *Applied Economics*, *51*(4), 387–397. https://doi.org/10.1080/00036846.2018.1497853

Mee, W. (2001). Country Briefing Paper: Women in the Republic of Uzbekistan

Mogilevskii, R. (2020). Labour Market and Technological Development in Central Asia. Working paper, University of Central Asia.

Priebe, J. (2010). *Labor Supply : An investigation for Indonesia 1993-2008*.

Rosenzweig, M. R., & Wolpin, K. I. (1980). Life-Cycle Labor Supply and Fertility: Causal Inferences from Household Models. *Journal of Political Economy*, *88*(2), 328–348. https://doi.org/10.1086/260868

Schmieder, J. (2020). Fertility as a Driver of Maternal Employment. *SSRN Electronic Journal*, (13496). https://doi.org/10.2139/ssrn.3661170

Somach, S and Rubin, D. (2010). Gender Assessment USAID/Central Asian Republics. United States Agency for International Development

Trako, I. (2018). Fertility and Parental Labor-Force Participation: New Evidence from a Developing Country in the Balkans. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2671618

Van Der Stoep, G. (2008). *Childbearing and labour force participation in South Africa: sibling composition as an identification strategy? Childbearing and labour force participation in South Africa: sibling composition as an identification strategy?* (52908). Retrieved from http://mpra.ub.uni-muenchen.de/52908/

Vere, J. P. (2011). Fertility and parents' labour supply: New evidence from US census data: Winner of the OEP prize for best paper on women and work. *Oxford Economic Papers, 63*(2), 211–231. https://doi.org/10.1093/oep/gpr003

WEF (2021). The Report and an interactive data platform are available at http://reports.weforum.org/globalgender-gap-report-2021/dataexplorer.

Appendix

Table 1. All available surveys conducted in this region

Country	*Year*	*N of respondents in the sample*
Kazakhstan	1995	2575
Kazakhstan	1999	3252
Kyrgyzstan	1997	2653
Kyrgyzstan	2012	5291
Tajikistan	2012	6495
Tajikistan	2017	7488
Uzbekistan	1996	3048
In pooled sample		30802

Source: DHS surveys.

Table 2. Son preference across different group of females. The effect of having two girls on number of children female has.

	Urban	Rural	Rich	Poor	Higher education	No higher education	Working	Not working	Beating justified	Beating not justified
One girl, one boy –base category										
Girls	0.228***	0.244***	0.243***	0.252***	0.193***	0.258***	0.203***	0.258***	0.271***	0.227***
	(0.027)	(0.022)	(0.025)	(0.029)	(0.032)	(0.020)	(0.031)	(0.021)	(0.025)	(0.034)
Boys	-0.030	-0.042**	-0.025	-0.055**	-0.000	-0.034*	-0.019	-0.036*	-0.016	-0.044
	(0.024)	(0.020)	(0.021)	(0.027)	(0.030)	(0.018)	(0.029)	(0.019)	(0.024)	(0.030)
Control variables										
Age	0.346***	0.387***	0.352***	0.431***	0.331***	0.384***	0.333***	0.406***	0.427***	0.387***
	(0.020)	(0.018)	(0.018)	(0.023)	(0.025)	(0.016)	(0.026)	(0.017)	(0.020)	(0.025)
Age-squared	-0.004***	-0.004***	-0.004***	-0.005***	-0.004***	-0.004***	-0.003***	-0.004***	-0.005***	-0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
BMI score	-0.001	-0.006***	0.002	-0.008***	0.003	-0.009***	-0.010***	-0.005**	-0.002	0.006*
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)
Had miscarriage	0.018	-0.080***	-0.012	-0.091**	-0.017	-0.048*	-0.004	-0.063**	-0.072*	-0.136***
	(0.034)	(0.028)	(0.031)	(0.037)	(0.037)	(0.027)	(0.036)	(0.028)	(0.040)	(0.046)
Married	0.489***	0.692***	0.502***	0.701***	0.555***	0.657***	0.592***	0.547***	0.664***	0.672***
	(0.038)	(0.042)	(0.037)	(0.051)	(0.051)	(0.034)	(0.037)	(0.046)	(0.053)	(0.053)
Age at 1st birth	0.167	0.188	0.129	0.329	0.334	0.300*	0.175	0.372**	0.306	0.229
	(0.182)	(0.175)	(0.162)	(0.232)	(0.240)	(0.154)	(0.228)	(0.159)	(0.197)	(0.220)
Age at 1st squared	-0.013*	-0.016**	-0.012*	-0.023**	-0.019**	-0.020***	-0.014	-0.024***	-0.021***	-0.018**
	(0.007)	(0.007)	(0.006)	(0.009)	(0.009)	(0.006)	(0.009)	(0.006)	(0.008)	(0.009)
Age at 1st polynomial	0.000**	0.000***	0.000**	0.000***	0.000**	0.000***	0.000*	0.000***	0.000***	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Survey Fixed	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

effect										
Constant	-4.390***	-4.575***	-4.328***	-6.140***	-6.051***	-5.660***	-4.140**	-6.246***	-6.160***	-5.065***
	(1.470)	(1.406)	(1.313)	(1.873)	(2.020)	(1.229)	(1.853)	(1.272)	(1.596)	(1.792)
Observations	5555	9464	6587	5594	3272	11747	4883	10130	6253	3702
Adjusted R^2	0.30	0.42	0.32	0.45	0.30	0.38	0.30	0.41	0.43	0.35

Source: DHS surveys; Standard errors are in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3. Main regression estimation without control variables.

	Working	Self-employed	Wage-employed	High- skilled	Low-skilled
OLS					
Number of living children	-0.072***	-0.010***	-0.051***	-0.079***	0.014***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Observations	30786	27550	27550	30802	30802
Adjusted R^2	0.14	0.02	0.22	0.17	0.03
Sample 1					
2SLS-infertility					
Number of living children	-0.030*	0.001	0.007	-0.006	0.011
	(0.018)	(0.012)	(0.017)	(0.017)	(0.011)
Observations	15064	14098	14098	15074	15074
Sample 2					
2SLS -Had daughter at 1st birth					
Number of living children	0.098	0.016	0.077	0.047	0.072*
	(0.065)	(0.039)	(0.057)	(0.061)	(0.041)
2SLS-Had twin at 1st birth					
Number of living children	-0.052	-0.064	0.129	0.198	-0.175**
	(0.134)	(0.084)	(0.145)	(0.153)	(0.086)
Observations	21708	19410	19410	21718	21718
Sample 3					
2SLS –Had 2 girls at 1st and 2nd birth					
Number of living children	0.024	-0.006	0.040	0.026	0.008
	(0.035)	(0.022)	(0.032)	(0.033)	(0.024)
2SLS -Had twin at 2nd birth					
Number of living children	0.103	0.031	0.019	-0.011	0.083
	(0.085)	(0.055)	(0.072)	(0.078)	(0.061)
Observations	15857	14407	14407	15866	15866

Source: DHS surveys; all estimation includes age, age-squared and survey fixed effect.; Sample 2 and 3 additionally includes age at first birth at polynomial form; Kazakhstan DHS 1999 is not included for wage employment and self-employment analyses due to data availability issue; Kazakhstan DHS 1999 is not included for wage employment and self-employment analyses due to data availability issue; Standard errors are in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4. Main regression estimation using young children

	Working	Self-employed	Wage-employed	High- skilled	Low-skilled
OLS					
Number of living children under 5	-0.033***	-0.008***	-0.012***	-0.021***	0.001
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Observations	28851	27307	27307	28864	28864
Adjusted R^2	0.17	0.02	0.26	0.25	0.08
Sample 1					
2SLS-infertility					
Number of living children under 5	-0.086***	-0.007	-0.020	-0.034	0.014
	(0.030)	(0.021)	(0.027)	(0.026)	(0.020)
Observations	14477	14003	14003	14486	14486
Sample 2					
2SLS -Had daughter at 1st birth					
Number of living children under 5	0.117	0.026	0.156	0.054	0.134
	(0.128)	(0.080)	(0.118)	(0.113)	(0.090)
2SLS-Had twin at 1st birth					
Number of living children	-0.430	-0.195	0.156	0.445	-0.581
	(0.622)	(0.316)	(0.437)	(0.623)	(0.643)
Observations	20375	19261	19261	20382	20382
Sample 3					
2SLS –Had 2 girls at 1st and 2nd birth					
Number of living children under 5	0.011	-0.020	0.055	0.028	0.010
	(0.060)	(0.040)	(0.055)	(0.054)	(0.043)
2SLS -Had twin at 2nd birth					
Number of living children under 5	0.250	0.086	0.010	-0.184	0.250
	(0.281)	(0.172)	(0.207)	(0.230)	(0.225)
Observations	15013	14311	14311	15019	15019

Source: DHS surveys; all estimation includes age, age-squared, education level, BMI index, marital status, had miscarriage, wealth quantile, number of household members, urbanity status and survey fixed effect.; Sample 2 and 3 additionally includes age at first birth at polynomial form; Kazakhstan DHS 1999 is not included for wage employment and self-employment analyses due to data availability issue; Standard errors are in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5. Main regression estimation among married females

Among married respondents	Working	Self-employed	Wage-employed	High- skilled	Low-skilled
OLS					
Number of living children	-0.032***	-0.003*	-0.022***	-0.035***	0.009***

	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)
Observations	20967	19979	19979	20977	20977
Adjusted R^2	0.17	0.02	0.29	0.27	0.08
Sample 1					
2SLS-infertility					
Number of living children	-0.053^{***}	-0.006	-0.010	-0.022	0.011
	(0.017)	(0.011)	(0.015)	(0.015)	(0.011)
Observations	12473	12162	12162	12481	12481
Sample 2					
2SLS -Had daughter at 1st birth					
Number of living children	0.052	0.040	0.043	0.037	0.049
	(0.056)	(0.036)	(0.051)	(0.050)	(0.038)
2SLS-Had twin at 1st birth					
Number of living children	-0.165	-0.051	0.066	0.231	-0.249^{*}
	(0.177)	(0.101)	(0.162)	(0.192)	(0.133)
Observations	18570	17658	17658	18576	18576
Sample 3					
2SLS –Had 2 girls at 1st and 2nd birth					
Number of living children	0.001	0.006	0.014	0.015	-0.001
	(0.033)	(0.022)	(0.030)	(0.029)	(0.023)
2SLS -Had twin at 2nd birth					
Number of living children	0.076	0.037	-0.007	-0.077	0.091
	(0.075)	(0.053)	(0.066)	(0.063)	(0.057)
Observations	14195	13572	13572	14200	14200

Source: DHS surveys; all estimation includes age, age-squared, education level, BMI index, husband age, had miscarriage, wealth quantile, number of household members, urbanity status and survey fixed effect.; Sample 2 and 3 additionally includes age at first birth at polynomial form; Kazakhstan DHS 1999 is not included for wage employment and self-employment analyses due to data availability issue; Standard errors are in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

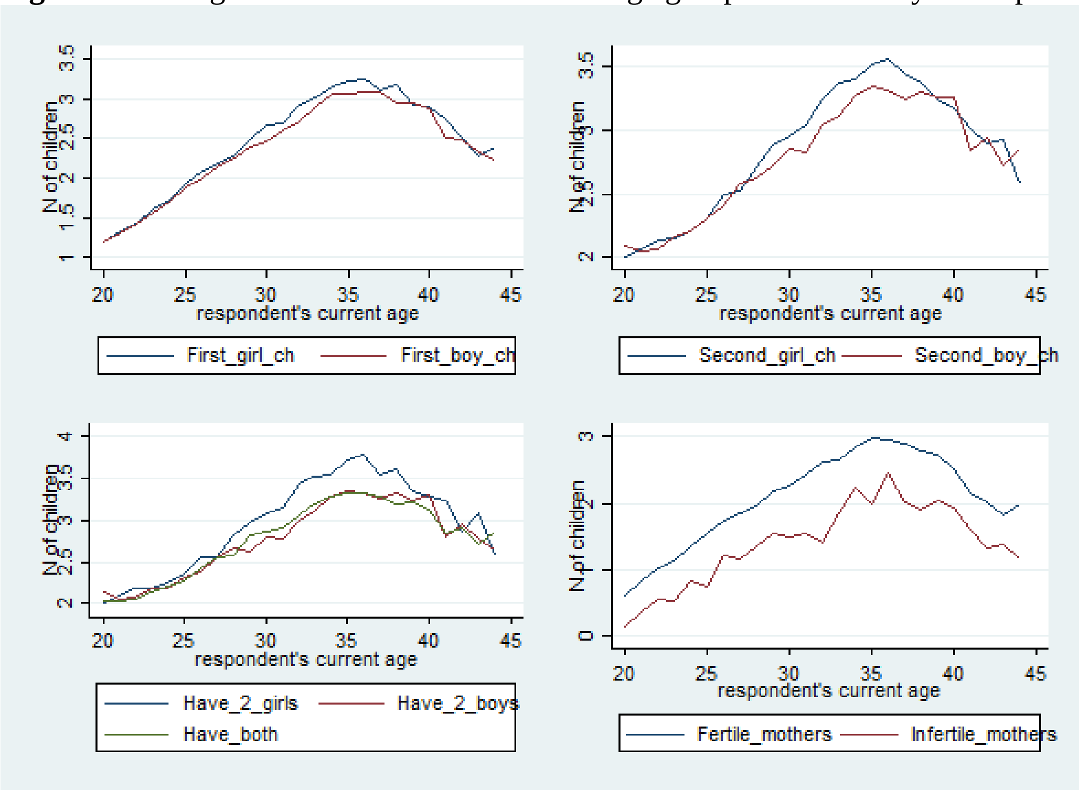
Table 6. Disaggregated regression estimation by age and education among married females

	Working	Self-employed	Wage-employed	High-skilled	Low-skilled	Working	Self-employed	Wage-employed	High-skilled	Low-skilled
Age			**Females <30**					**Females≥30**		
OLS	-0.041^{***}	-0.002	-0.029^{***}	-0.041^{***}	0.008^{***}	-0.026^{***}	-0.004	-0.017^{***}	-0.030^{***}	0.009^{***}
	(0.004)	(0.003)	(0.004)	(0.004)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
Observations	11527	11111	11111	11534	11534	9440	8868	8868	9443	9443
Adjusted R^2	0.10	0.01	0.26	0.23	0.07	0.16	0.01	0.28	0.29	0.08
Sample 1										
2SLS-infertility	-0.091^{***}	-0.018	0.001	-0.033	0.020	-0.039^{*}	-0.001	-0.013	-0.016	0.007
	(0.032)	(0.021)	(0.027)	(0.028)	(0.019)	(0.021)	(0.014)	(0.020)	(0.018)	(0.014)
Observations	8035	7877	7877	8040	8040	4438	4285	4285	4441	4441
Sample 2										
2SLS -Had daughter at 1st birth	-0.041	-0.033	0.045	-0.018	0.036	0.085	0.065^{*}	0.051	0.060	0.055
	(0.130)	(0.085)	(0.124)	(0.123)	(0.091)	(0.057)	(0.038)	(0.051)	(0.050)	(0.038)
Observations	9460	9101	9101	9463	9463	9110	8557	8557	9113	9113

Sample 3										
2SLS –Had 2 Girls at 1st and 2nd birth	0.054	-0.005	0.039	0.042	-0.017	-0.007	0.009	0.012	0.014	0.004
	(0.098)	(0.063)	(0.088)	(0.088)	(0.070)	(0.033)	(0.023)	(0.031)	(0.029)	(0.024)
First stage F-value										
Observations	5901	5720	5720	5903	5903	8294	7852	7852	8297	8297
Education level			**Educated**					**Not educated**		
OLS	-0.056***	-0.006	-0.040***	-0.049***	0.004**	-0.026***	-0.003	-0.017***	-0.031***	0.010***
	(0.007)	(0.004)	(0.007)	(0.007)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)
Observations	4797	4579	4579	4798	4798	16170	15400	15400	16179	16179
Adjusted R^2	0.12	0.01	0.15	0.16	0.01	0.14	0.02	0.30	0.23	0.07
Sample 1										
2SLS-infertility	-0.099*	-0.055	-0.006	-0.065	0.010	-0.046***	0.003	-0.012	-0.015	0.010
	(0.058)	(0.039)	(0.058)	(0.056)	(0.015)	(0.017)	(0.012)	(0.015)	(0.014)	(0.013)
Observations	2680	2625	2625	2680	2680	9793	9537	9537	9801	9801
Sample 2										
2SLS -Had daughter at 1st birth	-0.009	-0.016	0.085	0.082	-0.035	0.068	0.057	0.032	0.036	0.065
	(0.144)	(0.076)	(0.133)	(0.145)	(0.042)	(0.060)	(0.043)	(0.054)	(0.052)	(0.048)
Observations	4228	4033	4033	4229	4229	14342	13625	13625	14347	14347
Sample 3										
2SLS –Had 2 girls at 1st and 2nd birth	0.061	-0.031	0.133	0.068	0.007	-0.013	0.012	-0.010	0.009	-0.007
	(0.103)	(0.059)	(0.106)	(0.100)	(0.032)	(0.033)	(0.024)	(0.030)	(0.028)	(0.027)
Observations	3066	2943	2943	3066	3066	11129	10629	10629	11134	11134

Source: DHS surveys; all estimation includes age, age-squared, BMI index, husband age, had miscarriage, wealth quantile, number of household members, urbanity status and survey fixed effect.; Sample 2 and 3 additionally includes age at first birth at polynomial form; Kazakhstan DHS 1999 is not included for wage employment and self-employment analyses due to data availability issue; Standard errors are in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure 1. Average number of children for each age group: monotonicity assumption



Source: DHS surveys