# Child Costs and the Causal Effect of Fertility on Female Labor Supply: An investigation for Indonesia 1993-2008*

Jan Priebe†

October 20, 2010

Over the last two decades Indonesia has experienced a significant decline in fertility rates and substantial increases in the level of education of women. Despite this development female labor force participation rates have remained roughly constant throughout this period. This paper explores the causes for the seeming unresponsiveness of female labor supply to changes in fertility.

The empirical analysis is performed using annual data from the national household survey Susenas for the period 1993-2008. The final sample comprises about 850,000 woman aged 21 to 35 with at least two children. Identification of causal effects builds upon the empirical strategy as outlined in Angrist and Evans (1998).

The results suggest that a considerable share of women in Indonesia works in the labor market in order to finance basic expenditures on their children. Therefore, reductions in fertility rates seem to have led to two opposing effects that contributed to aggregate levels of female labor supply being constant. While some women were more likely to participate in the labor market due to a lower number of children, others might now lack the need to engage in the labor market due to a relaxation in their budget constraint.

**Key words:** Causality, Child Costs, Indonesia, Labor Supply, LATE.

***JEL* codes:** C21, D01, J13, J20.

*I would like to thank the Development Economics Research Group and the Center for Statistics in Göttingen for valuable comments. The paper benefited from discussions with David Blau, Michael Grimm, Stephan Klasen, Matin Qaim and Stefan Sperlich. I am grateful to Alberto Abadie and Markus Frölich for providing some of the programming syntax.

† University of Göttingen, Center for Statistics and Department of Economics, Platz der Göttinger Sieben 3, 37073 Göttingen, Germany. Email: jpriebe@uni-goettingen.de.

# 1. Introduction

Over the last three decades developing countries have experienced a continuous decline in fertility rates accompanied by a steady rise in female education levels. However, despite this development and in contrast to similar experiences from the 20th century in developed countries (Coleman and Pencavel, 1993; Goldin, 1990) female labor supply has stagnated or even decreased in several African and most Asian countries during this period.(ILO, 2008; Mammen and Paxson, 2000).

This paper aims at providing an explanation for the observed pattern that focuses on the impact of fertility and child costs on female labor supply in Indonesia. Recent studies on the U.S. (Blau and Robins, 1988; Conelly, 1992; Kimmel, 1998; Gelbach, 2002) have tended to exclusively stress the role of child care costs on female labor supply. With higher child care costs effectively meaning lower net-wages of women these studies typically find that high child care costs tend to lead to lower female labor supply.

However, children do not only impose costs in terms of time devoted to them but as well in terms of direct costs such as food, clothing, education or health expenditures. A Norman Parish Priest writing in prerevolutionary France noted the following of workmen:

> As young men they work and when by their work they got themselves decent clothing and something to pay the wedding costs, they marry, raise a first child, have much trouble in raising two, and if a third comes along their work is no longer enough for food and the expense. At such a time they do not hesitate to take up the beggar's staff and take to the road. (quoted in William Doyle 1989, p.14)

Although the effects of children these days are less dramatic in developed countries, this is less clear for developing countries in which substantial parts of the population struggle to cover its basic needs. Thus, direct costs on children might induce a negative income effect that is sufficiently large to drive some women into the labor force. A decline in fertility rates can therefore lead to a very heterogeneous impact on female labor supply. While some women might find it easier to participate in the labor market, others might lack the need to engage in the labor market due to a relaxation in their budget constraint. If both effects compensate each other or if an equal share of women is affected by one of the two effects than this helps to explain why there has been no rise in the female labor force participation despite a substantial decline in fertility over the last two decades.

Given the empirical difficulties to accurately quantify the cost of children (Deaton and Muellbauer, 1986) I present in this paper indirect descriptive and causal evidence for the claim that the need to finance the cost of children is an important determinant of female labor supply. Since children are likely to be relatively more expensive for poorer households than for richer households, with poorer households being at the same time less likely to be able to finance child costs out of savings or access to capital markets, the direct child cost constraint seems more relevant for former ones. As a consequence one would

expect the causal effect of fertility on female labor supply to be less negative or even positive for less well-off women.

The empirical set-up and the econometric identification strategy for the causal effects are borrowed from Angrist and Evans (1998). Using an unprecedented large sample on about 850,000 women aged 21 to 35 with at least two children for the period 1993-2008 I find a positive and statistically significant effect of fertility on female labor supply for women who are more likely to be poorer, less educated or live in the rural areas of the country. Moreover, this effect is particularly pronounced in the crisis period 1998-2000 in which women had to cope with a substantial decline in real household incomes. In addition, examining the period 1998-2000 is interesting because the economic crisis was largely unanticipated by households in Indonesia. Therefore, women in the sample were not able to immediately adjust their fertility levels as response to the crisis. In this context I find that women in general increased their labor supply during the crisis with women with higher fertility levels showing an even stronger response.

In this regard the existing study adds new important insights into the functioning of labor markets in developing countries and into the relationship of fertility and female labor supply. While recent research has emphasized that female labor supply in developing countries can be counter-cyclical in contrast to developed countries (Bhalotra and Umaña-Aponte, 2010; Miller and Urdinola, 2010), this paper argues that likewise the effect of fertility on female labor supply can be quite different between developed and developing countries. Furthermore, this paper presents separate causal evidence on the relationship between fertility and labor supply for different time periods and different regional aggregation (rural/urban). A result that is currently unavailable for other developing countries.

The results are of interest to both labor economists and policy makers. In particular, the results demonstrate that a decline in fertility rates does not necessarily imply that female labor supply on the aggregate level will increase. Furthermore, the findings suggest that increases in female labor force participation rates cannot unambiguously be interpreted as improvements in labor market access for women since a substantial share of women might be drawn rather 'involuntarily' into the labor market. In addition, the results indicate that social norms related to the expectation that women with children stay at home play a much larger role in Indonesia than one would expect by comparing labor force participation rates between men and women or between women with different fertility levels.

The paper is structured as follows: Section 2 presents a simple theoretical framework that allows to link child care costs, direct child costs and fertility to female labor supply. Section 3 explains the data set and construction of core variables used in this paper. Section 4 provides background information on the Indonesian context and discusses descriptive empirical evidence. Section 5 outlines the identification strategy of the causal effect of fertility on female labor supply and comprises the econometric analysis, while section 6 presents robustness checks and extensions to the main specification. Section 7 summarizes and concludes.

## 2. Theoretical Framework

The model presented in this section is intended to serve as a framework for cross-section estimation of the effect of child costs and fertility on female labor supply. Dynamic considerations are ignored for simplicity. Child costs are separated into child care costs and direct child costs with former ones being modeled as the time devoted to child care as in Blau and Robins (1988), Conelly (1992) and Kimmel (1998). The utility function U for a representative woman is defined over consumption c, leisure d, fertility n interacted with a child quality parameter q and is assumed to be given by:

$$U(c,d,n) = \log(c) + \alpha \log(d) + \beta \log(qn) \tag{1}$$

For simplicity I assume a logarithmic functional form. The weight on consumption is normalized to 1. The relative weights of leisure and children in utility ($\alpha$ and $\beta$) are assumed to be > 0. Likewise, the child quality parameter (q) is assumed to be positive and monotonically increasing in the level of variable direct child costs devoted to each child.[1]

Total time available to a woman is normalized to 1 and can be divided between leisure time d, child care time b*n and working time l. The time allocated to children is assumed to be linear in the number of children with $b > 0$.

$$1 = d + bn + l \tag{2}$$

The consumption possibilities of a woman are limited to the amount of exogenous income available (E), by the amount of income she earns, the wage w times the working time l, and the direct cost of children. Exogenous income can be thought of comprising husband's earnings (labor supply of husbands is likely to be inelastic to women's labor supply and fertility level n in a developing country context (Pencavel, 1987)), remittances and alike.

Direct child costs are further separated into fixed and variable costs. Fixed child costs (cs) refer to the subsistence requirement for children and are assumed to be linear in the number of children. Moreover, cs is assumed to be the same for all children and to present a lower bound that is above 0. In addition, to necessary expenditures on child subsistence, women might want to add further expenditures on children in line with individual or societal preferences on what might be spent on children (variable child costs). Since wealthier women in practice tend to spend more on children in absolute terms, variable child costs are assumed to increase with disposable household income net of income derived from women's labor supply.[2] Therefore, consumption possibilities are limited

[1] As explained later on direct child costs are separated into variable and fixed child costs. Since fixed child costs are related to the subsistance level of children these are assumed to not affect the quality of children. For simplicity reasons we abstract from time devoted to child care having an effect on child quality. However, the main results of this section do not change if child quality is allowed to depend on women's time devoted to child rearing.

[2] Variable child costs are modeled to increase proportionally with disposable income net of women's wages. However, results do not depend on this specific functional form. Likewise, the main results in this section do not alter if it is assumed that at the margin time devoted to child care (b) is decreasing in n.

to:

$$c = E + wl - csn - cv(E - csn) \tag{3}$$

with cv referring to the share of disposal income net of women's income (E-csn) that is spent on children.

Both constraints as outlined above are assumed to be binding. However, if they are regarded as inequality constraints the fact that consumption and leisure time are always desirable will make them binding under maximization. Given the constraints the utility function can be re-written as:

$$V(E, l, n) = \log\left(E + wl - csn - cv(E - csn)\right) + \alpha \log\left(1 - bn - l\right) + \beta \log\left(qn\right) \tag{4}$$

with $\alpha > 0$, $\beta > 0$, $q > 0$ $w > 0$, $cs > 0$, $b > 0$, $0 < cv < 1$ and $0 \leq l \leq 1$.

The first order condition for an interior maximum with respect to l is given by[3]:

$$\frac{dV}{dl} = \frac{w}{E + wl - csn - cv(E - csn)} - \frac{\alpha}{1 - bn - l} = 0 \tag{5}$$

Given a fixed number of children the optimal labor supply is:

$$l^* = \frac{w - \alpha E - bwn + \alpha(csn + cv(E - csn))}{(1 + \alpha)w} \tag{6}$$

Equations below show the response of optimal labor supply l* given exogenous changes in direct child costs, fertility and exogenous income.

$$\frac{dl^*}{dcs} = \frac{\alpha n(1 - cv)}{(1 + \alpha)w} > 0 \tag{7}$$

$$\frac{dl^*}{dcv} = \frac{\alpha(E - csn)}{(1 + \alpha)w} \geq 0 \tag{8}$$

$$\frac{dl^*}{dn} = \frac{\alpha cs(1 - cv) - bw}{(1 + \alpha)w} \gtreqless 0 \tag{9}$$

$$\frac{dl^*}{dE} = \frac{\alpha(cv - 1)}{(1 + \alpha)w} < 0 \tag{10}$$

Assumption 1:
If $E < csn$ then women have to become engaged in the labor market in any case since exogenous income is not sufficient to finance subsistence expenditures on children (cs). Since cs is the same for all mothers, only poorer women will be affected by this constraint. Given a developing country context poor households

[3]Since the objective function is quasiconcave equation 5 provides the solution for a constrained maximum. Furthermore, as can be easily shown, the corresponding Bordered Hessian matrix fulfils the second-order sufficient conditions for a maximum.

might find it very difficult to finance basic necessities in terms of food and non-food items.[4] Therefore, it seems plausible to expect that a certain share of poorer women is directly affected by this constraint. The more interesting case arises when is assumed that $E \geq csn$ (Assumption 1). If $E \geq cvn$ then, as shown in equations 6 and 9, female labor supply with respect to fertility depends on the sign of $\alpha cs(1-cv)-bw$. Thus, the effect of children on female labor supply depends on the net income effect that arises from direct child costs (cs(1-cv) and the opportunity cost of women associated with the time devoted to child rearing (bw). With direct child costs being sufficiently large (or increasing with n being held constant as in equations 7 and 8) higher fertility is likely to result in higher female labor supply. Likewise, given an exogenous decrease in n, female labor supply is likely to fall if cs(1-cv) is large enough.

Furthermore, it is important to note that direct child costs constitute in general a larger share of a women's budget the poorer she is.[5] Therefore, female labor supply effects arising from direct child cost considerations seem in general more important for poorer women.[6]

Prediction 1:
Summarizing the discussion above, it seems plausible to expect that the effect of direct child costs on female labor supply seems particularly relevant for poorer women. Therefore, since reliable information on child costs (child care costs and direct child costs) is absent from the later analysis and with income being excluded as a control variable in the later econometric specifications, we would expect the effect of fertility on female labor supply to be less negative or even positive for poorer women compared to richer ones.[7]

[4] The subsistence level might be determined by food but likewise by clothing, education or health expenditures. In Indonesia, for instance, primary education is compulsory with six years of schooling. Partly due to large school construction programs in the mid 1970s (Duflo, 2001) primary school enrollment rates in Indonesia are about 95% across 1993-2008. Thus, schooling costs in terms of school uniforms, books, pencils constitute an important additional component of direct child costs in Indonesia that affects both, poorer and richer women.

[5] As long as the share of income used to finance variable child costs is not increasing in income, this statement holds.

[6] In addition, one might argue that child care cost constraints might in reality even play a smaller role for poorer women since they might be more likely to live in an environment where other care givers are present (smaller b) and face lower wages (w). Therefore, opportunity costs (bw) are likely to be lower for poorer women which underscores the importance in direct child costs for poorer women when making labor supply decisions.

[7] Fertility levels are likely to be determined by exogenous and endogenous factors. In Indonesia where the two-child norm has become widely accepted by a large share of the population (Permana and Westoff, 1999), the exogenous determinant of fertility might be relatively large compared to other developing countries. If fertility is partly determined endogenously and if direct child cost affect particularly poorer women, one might expect poorer woman to have less children. However, in practice it seems plausible that poorer women will not necessarily respond with lower fertility levels. On the one hand, children might be much more important as old age security to poorer women. Moreover, poorer women are more likely to face higher child mortality risks which, as pointed out in Soares and Falc̆ao (2008), might even lead to higher fertility levels. Thus, prediction 1 is likely to hold even when fertility levels are partially determined endogenously.

In the discussed model fertility is assumed to be given exogenously and to be fixed. Since fertility and female labor supply decisions might be made simultaneously the previous considerations are somewhat restrictive. However, even if both decisions are made simultaneously the presented model is nonetheless useful in the Indonesian case. In late 1997 until 2000 Indonesia experienced a severe economic crisis that was largely unanticipated by households in the country. As explained in more detail in section 4 the crisis led to a fall in real wages and job-losses especially among men. Given the immediate effect of the crisis women were not able to respond with adjustments in their fertility level to a fall in exogenous income (E) and to a relative increase in direct child costs (cs). Therefore, n can be plausibly considered to be fixed for this period.

Prediction 2: As indicated by equation 10 one would expect female labor supply to increase during the crisis. Furthermore, if direct child cost considerations play an important role in this setting (share out of total expenditures), than from equation 7 one would expect that women with higher levels of fertility show an even stronger increase in labor supply.

However, since many women do not work, corner solutions are at least potentially a very important issue when studying female labor supply. For instance, if after an increase in E, the constraint $n * cv > E$ is still binding for all women than we won't expect any effect on women's decision to work or not. However, bearing potential corner solutions in mind, the main conclusions and predictions of this section remain valid as long as a sufficiently large share of women is always affected by the constraints as discussed above.

# 3. Data

One of the main goals of this paper is to identify the causal effect of fertility on female labor supply in Indonesia. To accomplish this I adopt the LATE framework as developed in Imbens and Angrist (1994) and Angrist et al. (1996). As explained in more detail in section 5 the empirical instrumentation strategy relies on the sex preference argument introduced in Angrist and Evans (1998). Reliable estimation of the LATE parameter of interest using the sex preference argument typically demands big data sets in order to guarantee that a sufficiently large number of mothers with at least two children is affected by the instrument. As a consequence all relevant empirical studies have focused on micro census data so far. For instance, Angrist and Evans (1998) used census data for the U.S. (1980, 1990), Cruces and Galiani (2007) census data for Argentina (1991) and Mexico (2000), Ebenstein (2009) census data for Taiwan (2000) and the U.S. (2000) and Angrist and Lavy (2011) census data for Israel (1983, 1995).

In contrast to these previous studies, this paper uses Indonesian household survey data from the annual Susenas rounds conducted in July of each year. Susenas is the principal data source to calculate official education, fertility, and poverty statistics in Indonesia. Every year Susenas collects socio-economic data

on about 200,000 households and 1,000,000 individuals and presents, together with the Indian National Sample Surveys, one of the largest household surveys in the world. Given its large sample size the data set seems appropriate to implement the empirical strategy as outlined in section 5. Since 1993 Susenas is representative at the province level. The final data comprises 16 cross-sectional rounds of Susenas from 1993 to 2008 and contains about 850,000 individual observations on women aged 21 to 35 with at least two children.[8]

For Indonesia alternative household data exist. Both the Indonesian Family Life Surveys (IFLS) and Sakernas (Indonesian Labor Force Survey) have been extensively used by researchers, e.g. (Smith et al., 2002; Frankenberg et al., 2003; Thomas et al., 2004). However, the sample size in both surveys is too small to allow for the adopted empirical strategy. Moreover, Sakernas lacks expenditure data that can be used to test some of the instrument assumptions.

# 4. The Indonesian context and descriptive statistics

## The counter-cyclic nature of female labor supply

Official Indonesian labor force statistics and ILO estimates show that female labor force participation rates among the working-age population have been roughly constant throughout the last two decades (According to ILO data (ILO, 2008) 50.4% in 1986 vs. 50.6% in 2005 with a peak of 51.0% in 1999).[9] Data from Sakernas and Susenas seem to confirm these results.

The period analyzed in this paper (1993-2008) can be separated into three distinct time spans namely before, during and after the large financial and economic crisis of 1998-2000. The pre-crisis period is characterized by stable and high economic growth rates. Using data from Sakernas Smith et al. (2002) find that between 1986 and 1997 real wages for men and women increased by about 40% and 60% respectively. While employment of men increased about 3% points during the same period, employment of women slightly fell by about 2% points. It might be that the income effect from the husband's earnings had more than offset the increase in the women's real wage and the narrowing of the gender pay gap in terms of its effect on female labor supply.

The crisis period starts with the second half of 1997 when the effect of the Asian Financial crisis trickled down and set the Indonesian Rupiah under strong pressure. Interest rates quadrupled between July and December 1997 with the Rupiah devaluing about 60% towards the U.S. Dollar. In January 1998 the Indonesian Rupiah collapsed. The financial crisis, was followed by an economic and political crisis. Early on in 1998 subsidees were removed from rice, oil and fuels. Exacerbating the situation a severe drought associated with El Niño depressed agricultural output in the rural areas of the country. Official inflation

[8] A detailed description of Susenas and on the construction of some key variables is provided in appendix A1. Summary statistics on the core variables used in the analysis are given in Table 1.

[9] Labor force data on Indonesia before 1986 is based on a different survey framework which expresses itself in substantially lower labor force participation rates of women in national and ILO labor force statistics.

for 1998 is reported to be 80% with an increase of rice prices by about 120% within a year and food prices doubling between 1997 and 1998 (Frankenberg et al., 2003). In May 1998 Suharto resigned from the presidency after 30 years in power.

However, despite the economic downturn aggregate statistics on labor market participation show a very stable pattern. Smith et al. (2002) report that overall employment for men decreased only slightly by 2% in urban and about 1% in rural areas. In contrast, there was an increase in women's employment (mostly in unpaid family work) due to the need to compensate for losses in the real wage of men during the crisis period. Although aggregate statistics on employment remained stable Smith et al. (2002) show that there was high fluctuation in the labor force with several individuals losing their job without finding a new one and others joining the labor force.

Data from Susenas confirms the counter-cyclical nature of female labor force participation in Indonesia. Figure 1 plots two indicators, work for pay (wage and self-employment) and work (work for pay together with unpaid family labor) separately for men and women.[10] While male employment in both indicators remained remarkably stable over the entire period there has been a relatively strong fluctuation in the female labor force participation. As in Smith et al. (2002) one observes a decrease in the female labor force participation in both indicators for the pre-crisis period with an increase during the crisis. Given that the difference between both indicators lies in the inclusion of unpaid family work, Susenas data suggests that particular unpaid family work has increased. With economic growth resuming in 2000-2001, one observes again a decrease in the female labor force participation after the crisis.[11]

[insert Figure 1]

## Fertility and female labor supply

Before turning to the role of children in the labor force participation of women, it is important to note that Indonesia has seen a substantial decline in fertility levels over the last decades. Partly due to the rigorous implementation of family planning programs and the promotion of the two child norm, the total fertility rate (TFR) has dropped from 5.6 in 1967 to 2.8 in 1997 (Permana and Westoff, 1999). Likewise, the World Development Indicators report that the TFR has further fallen from 2.85 to 2.17 between 1993 and 2008. Susenas data confirms the trend of declining fertility rates. Table 1 shows that the share of women with a third child decreased from 51% to 32% between 1993 and 2008. Furthermore, Figure 3c illustrates that during this time period there has been a steady decline in the share of women with a third birth across the entire age distribution.

[10] Labor force statistics in Figure 1 are based on the sample of women aged 21 to 35 with at least two children and their spouses (if married). See Table 2 for the respective statistics on female labor supply.

[11] Furthermore, Figures 3a and 3b indicate that particularly women in the age range of 27 to 35 years seem to have increased their labor supply during the crisis.

Comparing labor force participation rates of women with two children to those who have a third child (Table 2) for 1993-2008 one observes a slightly higher labor force participation rate of the former one by about 3% points for wage employment and 2% points in the 'work for pay' variable. However, no difference is observed in the overall 'work' variable. As Table 3 reports for the 'work for pay' variable, the difference is slightly larger in urban (3% points) than in rural areas (1% point). Furthermore, Table 4 shows, that the difference seems to partly depend on the wealth situation (measured in household per capita expenditure levels) of a woman.[12] In line with the predictions derived in section 2 one finds that the difference increases with the level of per capita expenditures. While in the poorest quintiles women with more than two children are at least as likely to work as women with two children, women with more than two children are less likely to work by about 2-3 percentage points in the richest expenditure quintile.

The descriptive results can be interpreted in mainly two ways. On the one hand it might be that there is a very small or no effect of fertility on the decision of women to enter the labor force.[13] On the other hand it might be, as argued in section 2, that some mothers enter the labor force in order to finance the costs of children while for others children might pose a constraint to participate in the labor market. If both effects cancel each other off in the overall population then this might explain why we do not observe a strong relationship between fertility and labor force participation on the aggregate level.

Furthermore, the interpretation of the observed relationship demands caution for more reasons. First of all, the bivariate relationship is potentially plagued by issues of reversed causality between fertility and labor supply. Secondly, decisions on fertility and labor force participation might be made simultaneously which makes the interpretation more cumbersome.

## The crisis as an exogenous shock: Child costs, fertility and female labor supply

The crisis period offers the opportunity to obtain evidence on the role that child costs play in the relationship between fertility and female labor supply. First of all, the crisis caused a negative income effect to most households either

[12] Susenas does not provide income or asset data on an annual base. However, Susenas collects expenditure data in each round which is used to update official annual poverty estimates by the Indonesian National Statistical Office. As explained in the appendix detailed household expenditure data is available only in the Module sections of Susenas every three years. Mean and median expenditure levels from the Module section tend to be higher than those from the Core section. However, since the correlation between both measures is very high (> .95) the resulting ranking classification seems reliable. In the following, information on household expenditures per capita is used to assign women into five distinct expenditure quintiles for each of the survey rounds separately. The expenditure figures are deflated by provincial CPIs. Since CPI figures for Indonesia are based on urban price surveys no adjustment for urban/rural differences could be made.

[13] Angrist (2001) finds a difference of about 17 percentage points in the labor force participation rate between mothers with two children and those with a third child using data from the 1980 U.S. census. In comparison the difference in Indonesia is small. Cruces and Galiani (2007) do not report this difference.

through the loss of jobs or the decline in real wages. Second, the crisis was largely unanticipated and therefore both, the decline in real incomes and the rise in relative child costs, could not be anticipated by women. Moreover, once the crisis was in place, woman could not anymore adjust their fertility levels to this event.[14] Thus, fertility can be assumed to be fixed and given exogenously during the crisis period.

The model presented in section 2 assigned direct child costs (cs) the role to present a fixed minimum level of expenditures per child. Given the severe economic crisis some families, particularly poorer ones, might had to cut even into these basic expenditures. For instance, families might decide to postpone health-checks or buying new clothes for their children or take children out of primary school. If these cuts were large enough to off-set the experienced loss of real incomes during the crisis than one might not see any effect of the crisis on female labor supply. However, this is not observed by the empirical evidence at hand. Frankenberg et al. (2003) and Thomas et al. (2004) report that households responded to the crisis by reducing expenditures on durables and selling assets but not on basic necessities related to children. Furthermore, they find that school attendance remained stable during the crisis period with households being reluctant to cut in education expenditures on children. Expenditure patterns as observed from the Susenas Module supports this. Table 5 reports total expenditure and expenditure shares on food, clothing, education and health.[15] Between 1996 and 1999 the average food share of household expenditures increased about 5% points while the share of household clothing and health expenditures declined by about 0.2% and 0.3% points respectively. Regarding expenditures on children there seems to have been only small reductions in the education expenditure share by about 0.5% points with clothing expenditures even slightly increasing. Considering the strong increase in the share of food expenditures, the small decrease in child education expenditures and the evidence from previous studies, it seems that household on average did not try to fully compensate income losses and increases in food prices with reductions in direct child costs.

Given that the number of children is roughly fixed at the time of the crisis, we would expect that both, the negative income effect as well as the increases in the relative cost of children, would increase the propensity of some mothers to work. Moreover, Prediction 2 from section 2 states that mothers with more children are expected to show an even stronger reaction to the crisis. Table 2 support this point of view. In all three indicators mothers with a third child show a stronger response (in relative and absolute terms) to the crisis as mothers with two children. Moreover, Tables 3 and 4 suggest that this pattern is observed in both, rural and urban labor markets and across all wealth levels.[16]

[14] An exception might be woman who were pregnant during the crisis or women with infants. However, neither abortions nor infanticide is a common practice in Indonesia. Moreover, as described in the appendix woman with children below the age of 1 were excluded from the analysis.

[15] Susenas allows only for clothing and education expenditures to be separated between child and adult expenditures.

[16] Temporary migration from urban to rural areas was occasionally reported in Indonesia for

The finding above is not only useful as evidence for the existence of an effect of child costs on the relationship between fertility and female labor supply but points as well to sample selection effects. Thus, given the stronger participation of women with higher fertility during the crisis period than in the pre and post-crisis period one can expect the causal effect of fertility on female labor force participation being less negative or more positive during the crisis.

# 5. Identification of the causal effect

The interpretation of the relationship between fertility and female labor supply is plagued by the difficulty to convincingly extract cause and effect relationships from the observed correlations. In order to disentangle the causal mechanism linking fertility and female labor supply I adopt an empirical strategy as introduced by Angrist and Evans (1998) that uses child sex preferences as an instrument in the first stage regressions within the Local Average Treatment Effect (LATE) framework as introduced in Imbens and Angrist (1994).

## 5.1. The sex preference instrument

The identification strategy of Angrist and Evans (1998) exploits parental preferences for a mixed sibling sex composition as an instrument for fertility. Since parents of same-sex siblings are significantly more likely to have an additional child, and the sex mix is virtually randomly assigned, an indicator variable for whether the sex of the second child matches the sex of the first child provides a plausible instrument for further childbearing among women with at least two children. In terms of relevance of the instrument Angrist and Evans (1998) find that mothers whose first two children are of the same sex are more likely to have a third child by about 6% points. Likewise, Cruces and Galiani (2007) report a difference of about 3.5% and 3.8% for Argentina and Mexico.

Table 6 depicts birth composition figures for Indonesia based on the pooled Susenas sample. As expected from the demographic literature the likelihood of having a boy is slightly higher than that of girls. Moreover, the sex of the first child does not seem to have a strong effect on the decision of having a second child.[17] Column 3 shows the share of mothers who have a third child (out of those with two children). The data shows that mothers whose first two children are either two boys or two girls are about 3% points more likely to have a third child. The difference is statistically significant at the 1% level. Therefore, it seems that the sex composition of the first two children has an

the crisis period. Therefore, household composition effects might eventually influence the interpretation of results. However, as reported in Table 3 figures on average household sizes in rural and urban areas do not exhibit any unusual movement which suggest that the results are not much affected by this issue.

[17] In contrast to other Asian countries son preference is not a very pronounced phenomena in Indonesia. More importantly, son preference does not translate into discrimination against females in terms of sex-selective abortions, fertility stopping rules, female infanticide or discrimination in the allocation of household resources (Levine and Kevane, 2003). Therefore, the sample on women with at least two children does not seem to suffer from sample selection problems caused by child sex composition issues.

effect on mother's fertility decisions in the Indonesian context, although the difference is not as large as observed for the U.S..[18]

Although the sex composition of the first two children is virtually randomly assigned this does not completely rule out that there might be a direct effect of the instrument on mothers' labor force participation. For instance, Rosenzweig and Wolpin (2000) demonstrate for the Indian context that the child sex composition of the first two children leads to secular effects that might question the exogeneity of the instrument from the second stage. Using expenditure data Rosenzweig and Wolpin (2000) find that same sex siblings are related to lower expenditures per child. They attribute this effect to "hand-me-down" savings, which are more likely to arise when there are children of the same sex in the household for items such as clothing and footwear. Table 7 reports expenditure differences for mothers with two boys, two girls, both children from the same sex and mixed sibling sex composition for Indonesia. The statistics show that there is no difference in the expenditure behavior among the four groups considered. Therefore, the instrument seems not to be related to any indirect income effects that might question its exogeneity (at least not to testable sources of influence).

## 5.2. Econometric Framework: LATE

Identification of the causal effect of fertility on female labor supply rests on an instrumental variable approach that allows for heterogeneity in the treatment effect (fertility) on the response variable (female labor force participation) and selection on unobservables. In this context identification of the average treatment effect and the average effect on the treated often runs into difficulties (Imbens and Wooldridge, 2009). However, for the case of a binary treatment variable and a binary instrument Imbens and Angrist (1994) have shown that the average treatment effect for the subpopulation affected by the instrument can be consistently estimated relying on standard instrumental variable (IV) techniques (2SLS). To better understand for which subgroup the average treatment effect can be consistently estimated it is useful to classify individuals i into the following categories (G) as in Angrist et al. (1996).

$$G_i = \begin{cases} \text{never-taker} & \text{if } W_i(0) = W_i(1) = 0 \\ \text{complier} & \text{if } W_i(0) = 0, W_i(1) = 1 \\ \text{defier} & \text{if } W_i(0) = 1, W_i(1) = 0 \\ \text{always-taker} & \text{if } W_i(0) = W_i(1) = 1 \end{cases} \tag{11}$$

where $W_i$ refers to the actual level of treatment received (fertility) while $W_i(0)$ and $W_i(1)$ denote the level of treatment received if the instrument (child sex composition) takes on the values 0 and 1 respectively.

Invoking the assumption of monotonicity in the effect of the instrument on the treatment variable ($W_i(1) \geq W_i(0)$) the presence of defiers is ruled out.

[18] Despite declining fertility rates during 1993-2008 the difference of about 3% is approximately the same when looking at each single round of Susenas. The difference fluctuates between 2.7 and 3.3. Results for each single year can be obtained from the author.

Imbens and Angrist (1994) show that under this assumption the average causal effect of treatment ($\tau$) can be estimated for the sub-group of compliers which makes it a Local Average Treatment Effect.

The LATE estimand can then be written as:

$$\tau_{LATE} = \frac{E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]}{E[W_i|Z_i = 1] - E[W_i|Z_i = 0]} \tag{12}$$

where $Y_i$ refers to whether a mother participates in the labor market (0/1), $W_i$ refers to the endogenous/treatment variable (More than 2 children, (0/1)) and $Z_i$ to the instrument (Same sex, (0,1).
Although the original LATE framework outlined above applies to causal models without covariates, the LATE parameter of interest can still be estimated with 2SLS when covariates are included.[19]

Since the dependent variable (work for pay) is binary with covariates being included 2SLS might not give the best approximation of the Conditional Expectation Function (CEF). If covariates are dummy variables Angrist (2001) and Angrist and Pischke (2010) take the position that 2SLS is not less appropriate for binary dependent variables since the CEF can be parameterized as linear using a saturated model regardless of the support of the dependent variable. Although I will largely stick to 2SLS in this article, I report as well results when using a semi-parametric type of approximation (Abadie, 2003) and a non-parametric approximation (Fröhlich, 2007) for the CEF.[20]

## 5.3. Discussion of results

As in Angrist and Evans (1998), Cruces and Galiani (2007) and Ebenstein (2009) this study focuses on mothers with at least two children who are between 21 and 35 years old at the time of the survey. The dependent variable discussed in this article is 'work for pay' which has been the core dependent variable in previous studies.[21] To reduce the likelihood of bias in the LATE parameter of interest I pool the different rounds into three time periods (pre-crisis, crisis, and post-crisis).[22] The pre-crisis period comprises the rounds of 1993-1997, the crisis period includes the rounds 1998-2000, and the post-crisis period comprises

[19] Covariates might be included because the conditional independence assumption and the exclusion restrictions underlying IV estimation are more likely to be valid after conditioning. Moreover, 2SLS estimates might be more precise if conditioning on covariates reduces some of the variability in the dependent variable (Angrist and Pischke, 2009).

[20] Furthermore, since 2SLS can be biased towards OLS if the instrument is weak, so called Split Sample IV estimates (Angrist and Krueger, 1995) are presented.

[21] Estimates for the case that the dependent variable is 'wage employment' or 'work' can be obtained from the author. The main results of this article however do not change when considering these two variables. Likewise, the results do not change substantially when using the age group 18 to 35 or 21 to 45.

[22] The bias might arise from the instrument being weak or the number of mothers affected by the instrument being too small. Moreover, pooling makes the sample size more comparable to those studies using census data. In addition, when increasing the number of instruments in section 6 the likelihood of bias in the 2SLS estimate increases (Bound et al., 1995). Pooling the data helps to mitigate this problem.

the rounds 2001-2008.[23] The instrument used is 'Same sex' which is a binary variable that takes the value 1 if the first two children are either two boys or two girls. In this section it is assumed that the effect of the instrument on the endogenous variable (More than two children) remains constant over time. Therefore, there is exactly one exclusive instrument used in the first step. A limited set of covariates was included as explained in Table 8.

The first stage estimates in Table 8 show that women with same sex children are about 3 percentage points more likely to have a third child in a model with covariates. This relationship seems robust over time with a slightly stronger effect observed in urban than in rural areas. Moreover, in each case the relationship is statistically significant. The corresponding F-statistics of the first stage relationship are all clearly above 10 which indicates that the instrument does not seem to be weak (Staiger and Stock, 1997).

OLS estimates of the causal effect indicate that there is a small negative, statistically significant, relationship between having a third child and a mother's labor supply. The point estimates for the pooled sample and the three subperiods range from -3.5 percentage points to -4.4 percentage points. These estimates are much smaller than those reported in Angrist and Evans (1998) for the U.S. (-18 percentage points) and Cruces and Galiani (2007) for Argentina/Mexico (-9.5 percentage points).[24] With respect to the urban-rural dichotomy one observes a more negative (stronger effect) in urban areas with a difference of about 2 percentage points.

The IV estimates are much smaller than the corresponding OLS estimates and become even positive with all IV approaches providing highly similar results. On the national level the estimates for the entire period and the crisis period are statistically significant with parameter estimates of 4.3 and 9.0 percentage points respectively ('saturated 2SLS model').[25] Examining the results for urban and rural areas separately one finds that estimates in both areas become less negative. However, the national results seem to be largely driven by the strong positive effects found in rural areas. While urban estimates remain at least slightly negative, they are all statistically insignificant. In contrast, rural estimates are highly positive and statistically significant. Given that poverty is substantially higher in rural areas than in urban areas this result is line with Prediction 1 from section 2 that particularly poorer woman are likely to participate in the labor market in order to being able to finance direct child costs.[26]

[23] The core crisis period is 1998 and 1999. Since (food price) inflation was still very high in 2000 this round is assigned to the crisis period as well.

[24] Given that the bivariate relationship as discussed in section 4 was already much weaker and that only a reduced set of covariates is included this result was to be expected.

[25] This difference looks relatively large. However, even in Angrist and Evans (1998) the IV estimates are about 6 percentage points lower than that of OLS. Therefore, a magnitude of 8 percentage points at the national level does not seem to be implausible.

[26] Given that the urbanization rate in Indonesia (about 40 percent) is substantially lower than in Argentina (about 90 percent) or Mexico (about 80 percent) this might partly help to explain why Cruces and Galiani (2007) find a small negative effect of fertility on female labor supply on the national level in these two countries. However, the relative small employment difference among women with two and more than two children in urban

These results above support the argument that the direct cost of children is an important reason why women might join the labor market in Indonesia, at least when looking at the national and rural level. Table 8 provides further evidence for this claim. Independently of whether OLS/IV techniques or the National/Urban/Rural subset is considered, the effect of having a third child is always less negative/more positive in the crisis period than in the pre- or post-crisis period. This results seems consistent with the considerations in sections 2 and 4 that children are likely to be relatively more costly during the crisis for a large part of the population (either due to declining real incomes of the family or direct increases in the cost of children) and that women with more than two children were relatively more likely to self-select into the labor market during the crisis than women with two children (Prediction 2). Since the estimates for the urban labor market follow the same pattern over time, this might point to the circumstance that child costs play a similar but less important role in the urban labor market, despite the respective coefficients being slightly negative and statistically insignificant.

# 6. Robustness Checks and Extensions

## Additional Instruments

Estimation of models in section 5 was based on the assumption that the effect of the same sex composition of the first two children on fertility (having more than two children) is constant. This subsection relaxes this assumption allowing the effect to vary (A) between having two girls or two boys as the first two children and (B) between different survey rounds and different provinces. The first case might be relevant if parents show a different fertility response depending on whether two daughters or two sons were born while the second case is useful if there are regional differences in the preference for a mixed sibling sex composition or changes in this preference over time. In the second case the same sex instrument is interacted with the 25 province and 16 time dummies.[27]

On the one hand, if the additional instruments help to explain part of the variation in fertility responses than the LATE parameter of interest can be estimated more efficiently. On the other, it is likely that in case (B) several instruments are rather weak. In the just identified IV models discussed before 2SLS is approximately median unbiased. However, in overidentified IV models with potentially weak instruments 2SLS is likely to be biased towards OLS. The bias decreases when the number of observations increases. However, as discussed in Bound et al. (1995) the bias can be substantial even in large samples.

An often used alternative to 2SLS in this case is estimation with Limited Information Maximum Likelihood (LIML) which is approximately median-unbiased for overidentified constant effect models and which provides a finite sample bias

Indonesia suggests that other factors might play a role as well.

[27] Therefore these models include about 25*16=400 additional instruments. Due to multicollinearity reasons the final number is slightly lower.

reduction compared to 2SLS. However, LIML estimates can have large dispersions when the instruments are weak (Hahn et al., 2004) and be inconsistent in the case of heteroskedastic data (Bekker and van der Ploeg, 2005; Chao and Swanson, 2005). A further alternative in the case of multiple weak instruments are Jackknife Instrumental Variable Estimators (JIVE) as in Phillips and Hale (1977) and Angrist et al. (1999) which are more likely to be unbiased but which face efficiency problems.

To address the problems that can arise in the context of multiple weak instruments and heteroskedastic data $\tau$ and the respective standard errors are estimated with a jackknife version of LIML developed in Hausman et al. (2009). The estimator is called 'HFUL' by the authors and seems to perform well in terms of bias, consistency and efficiency independently of whether the data is heteroskedastic or not and whether the instruments are weak or not.[28]

Results for 2SLS, LIML and HFUL are shown in Table 9. Comparing the estimates between using same sex as an instrument with those from introducing two girls and two boys as separate instruments shows that coefficients and standard errors are very similar. Therefore, it seems that splitting the instrument 'same sex' into 'two boys' and 'two girls' does not change the previous results. Likewise, interacting the same sex instrument with year dummies provides results similar to those obtained before independently of whether 2SLS, LIML or HFUL is considered. In contrast, results based on the interaction between same sex with year and province dummies provide different results. While the 2SLS estimates becomes smaller in magnitude, the LIML and HFUL estimates tend to become larger compared to the previous specifications. Nonetheless, all three procedures still report statistically significant positive effects for the same coefficients as before. However, the specification with this much larger set of additional instrumental variables results in a strong drop in the F-statistics and even less precise estimates.

Given these results it seems fine to stick with the basic model from the previous section which assumed a constant effect of the sibling sex composition on fertility and that does not explicitly separate between two boys and two girls.

## Heterogeneity in the Impact of Children on Female Labor Supply

In the previous sections model specifications were discussed that allowed the effect of fertility on female labor supply to vary across time (three time periods) and location (rural/urban). Likewise, it seems plausible that the effect might vary by other factors. For example, if child cost considerations are one important reason in Indonesia why women join the labor market than one would expect the effect of fertility on the decision to work to depend on the income or wealth level of the mother. Consequently, one would expect the effect to be less negative/more positive among poorer mothers (Prediction 1, section 2).

In order to examine whether the effect of fertility on female labor supply depends on the wealth level of the mother, woman are pooled together corre-

[28] Programming of the 'HFUL' estimator is done as outlined in section 2 of Hausman et al. (2009).

sponding to its grouping into a specific expenditure quintile, e.g. all women in quintile 1 (lowest expenditure quintile) are pooled together.[29] Table 10 provides descriptive statistics and 2SLS estimates on the five different quintiles. As one would expect women in the poorer quintiles are more likely to be younger, to have a third child, to have less years of formal schooling and to live in rural areas. The presented 2SLS estimates which are again based on the specifications related to the 'saturated model' show a much stronger positive effect in the two lowest expenditure quintiles. Moreover, the effect of having a third child on female labor supply is only statistically significant for these two quintiles. These results are in line with the reasoning above and seem to support the view that higher fertility associated with higher total child costs are an important factor that influences the labor supply decision of a large share of women in Indonesia.[30]

An alternative proxy for the income and wealth status is the education level of the mother. Thus, women can be assigned into different groups according to their level of education. Table 11 depicts descriptive statistics and 2SLS estimates for three different education groups (less than completed primary, completed primary, completed secondary) according to the highest schooling degree obtained.[31] The 2SLS coefficient for women in the lowest education tercile has a strong positive sign and is statistically significant. In contrast the respective coefficient for the other two terciles is statistically insignificant. Therefore, again one finds that there is a positive effect of fertility on female labor supply which in line with previous reasoning is only found among the less educated/economically less well of mothers.

Furthermore, the effect of fertility on female labor supply is likely to vary with the age of a child. On the one hand, when children are young the need to care for a child restricts the opportunities of a mother to work for pay. In line with this reasoning, empirical studies (Blau and Robins, 1988; Conelly, 1992; Kimmel, 1998) tend to find that mothers' are more likely to work the older their children are. On the other hand, the age of a child does not only lead to substitution effects with respect to time needed for child care but as well to direct income effects since older children tend to be more expensive than younger ones. For instance, the older a child gets, the more calories it will need. In addition, older children might be simply more expensive compared to younger ones because parents will at some point need to finance education expenditures or to buy more quality clothing for their offspring. As outlined in section 2, both, the

[29] Women are grouped into expenditure quintiles as described in section 4.

[30] To a certain degree the classification of women into expenditure quintiles introduces selection effects since women who find paid work might end up in higher quintiles. For instance, among women with at least three children those who are able to find paid work might be more likely to end up in quintile 3 than those who are not able to find work and which therefore classified into quintile 2. In this particular case controlling for the selection effect would even lead to a stronger positive effect in the 2SLS estimate for women in the lower expenditure quintiles.

[31] The classification is based on the consideration to have an approximately equal number of observations in each education group. Results do not change much if five education categories are used or if women with tertiary education are excluded or if women are assigned into groups according to completed years of education.

direction of the substitution and the income effect as discussed above, implies that direct child cost considerations become more important the older children are. Hence, one would expect an even more positive (less negative) effect of fertility on female labor supply with increasing age of a mother's child.

Table 12 presents separate estimates for the causal effect of fertility on female labor supply for mothers whose youngest child is either below the age of six, between six and ten years or older than ten years. As expected from the discussion above, the positive effect of fertility on female labor supply increases for women whose youngest child is older than five years. However, the estimated causal effect is largest for women whose youngest child is between six and ten years (.068) and decreases again for women with children above the age of ten (0.33). Taking into account that children above the age of ten are more likely to already contribute to the income of parents, particularly among poorer families or in rural areas, while at the same time they finish their formal education (e.g. finish primary school), the results indicate that for a certain share of the population net direct costs of children might fall after a certain age which as a consequence tend to relax the parent's budget constraint.

## Married Women's Labor Supply

So far the discussion has focused on women between the age of 21 and 35 with at least two children in which the labor supply decision was made independently of other adult members of the household. However, in practice such labor supply decisions are likely to depend on the characteristics of these other members as well.

Several empirical and theoretical models have focused on the role of women's unearned income, especially the income of the husband, on female labor supply. In Indonesia marriage is usually seen as a prerequisite for having children. Thus, not surprisingly, the share of married women in the Susenas sample amounts to 98% and remains approximately constant over the period 1993-2008 (See Table 1). Therefore, the previous results are likely to hold as well for the married women sub-sample. Nonetheless, reducing the analysis to sub-sample of married provides the empirical advantage of including the husband's characteristics as additional control variables that might have a direct effect of women's labor force decision and which might help to estimate the effect of fertility on female labor supply more precisely.[32] Since husband's income data is not available in Susenas this study uses the husband's labor force participation information (work for pay) and educational attainment as a proxy for the earning position of the husband.

In addition to women's unearned income, the own earnings potential might influence women's labor supply. Again, since the data does not allow to condition on women's earnings, educational characteristics are used instead.[33]

[32] The virtual random assignment of the same sex instrument makes it very likely that the reduced-form regressions considered before have a meaningful causal interpretation. As stated in section 5 identification of the LATE parameter does not require to condition on other covariates.

[33] Note that at least two of the three variables, women's educational attainment and the hus-

Table 13 presents IV estimates of the effect of fertility on female labor supply for the married sub-sample. As expected the baseline model specified like in Table 8 provides estimates that are highly similar between the previous sample and the married women sub-sample. Moreover, the IV estimates are remarkably insensitive to the list of additional covariates, including the potentially endogenous schooling variables. Although the effect of fertility on female labor supply decreases when including the husband's work status together with the educational variables of both parents, it still remains of comparable sign and magnitude as in the baseline specification. Hence, the effect of fertility on female labor supply remains strongly positive on the national and urban level with peak values for the crisis period. Likewise, the effects are statistically significant (at least at the 10% level) for the same coefficients are before (National-All, National-Crisis and all rural specifications). Thus, the previous results seem to be robust to the inclusion of these additional covariates.

## Alternative Instruments

An alternative equally plausible instrument for the effect of fertility on labor supply is multiple births with women who experience a multiple (second) birth having more children on average than women who don't. The twin status of a child as an instrument for fertility has been introduced by Rosenzweig and Wolpin (1980) and has been widely used in the context of estimating the causal effect of fertility on educational or labor market outcomes (Bronars and Grogger, 1994; Angrist and Evans, 1998; Jacobsen et al., 1999).

Although the use of second multiple births (2nd birth is a twin birth) as an instrument for fertility is not completely free of criticism, it can provide an useful additional robustness check for several reasons. First of all, it is often the case that different instruments provide different results. If one obtains the same results when using twin status as an instrument suggests that the previous results are not specific to the same sex instrument. Secondly, the twin instrument automatically fulfills the monotonicity assumption of the LATE framework since 'non-compliers' cannot exist. Therefore, it is less sensitive to criticism that aims at the lack of a structural approach in the LATE framework to estimate the parameter of interest.[34] Thirdly, the group of 'never-takers' does not exist. Therefore, using twin status as an instrument identifies the causal effect of a different 'group' namely the treatment on the non-treated as pointed out in Angrist and Lavy (2011).

Unfortunately, Susenas does not collect information on twinning. Angrist and Evans (1998) face a similar problem with U.S. census data. Due to the lack of a direct identifier they use information on the year and quarter of birth of children to construct a twin variable retrospectively. Since Susenas collects only the age in years of a child, twin status is assigned to children who have the

band's labor supply, are potentially endogenous because they might be partly determined by fertility. For this reason, women's educational attainment was excluded from the main set of estimates.

[34] See Deaton (2010), Heckman (2010) and Imbens (2010) for an overview on recent discussions.

same age in years. As shown in appendix 1 this induces a considerable amount measurement error in the twin variable.

Results from using the constructed twin variable as instrument for fertility are presented in Table 14. Although the overall level of the effect of fertility on female labor supply is now smaller, the overall pattern (national, urban/rural, pre-crisis/crisis/post-crisis) remains largely the same as before. However, only the estimate for the national sample over the entire period remains significant at a 5% level. Given the amount of measurement error inherent in the constructed twin variable this circumstance is likely to have contributed to the more imprecise estimates than before. In general, the twin estimates seem to be in line with the results obtained in the previous sections.

# 7. Conclusion

The causal estimates reported in this paper suggest that fertility has a positive impact on female labor supply in Indonesia. In particular, the coefficient on fertility implies that an increase in fertility by one child increases female labor supply by about 4 percentage points (on the national level for the entire period), a result that is both, economically and statistically significant.

In order to shed some light on the mechanism that leads to this surprising finding, a simple theoretical model is derived that links the cost of children with fertility and female labor supply. Although child costs are not observed in the data the model provided testable empirical predictions. In line with these predictions I find that the effect of fertility on female labor supply is even more positive (9 percentage points) during the period of the economic and financial crisis (1998-2000) and that almost the entire overall effect can be attributed to labor market responses by woman which are poorer, less educated and who live in the rural areas of the country. These two separate pieces of evidence lend support to the claim that the need to finance the cost of children, more precisely direct child cost in contrast to child care costs, lead a considerable share of women into the labor force in Indonesia.

The findings of this paper are important for various reasons. First of all, the results confirm the common view in labor economics that children are an important determinant of women's labor supply. However, in contrast to findings for developed countries, children might not only restrict the opportunity of woman to participate in the labor market but as well create an incentive to search for employment. Especially in developing countries in which households act under tight budget constraints with a large share of income being devoted to food and other basic necessities, the later constraint seems to matter.

The dual role that children play in the labor supply decision of woman helps to explain why female labor force participation rates have stagnated over the last two decades in Indonesia despite a significant drop in fertility rates and substantial increases in women's years of schooling. With fertility levels falling some women might have found it easier to participate in the labor market while an equal share of women now might lack the incentive to engage in the labor market due to a relaxation in their budget constraint. Furthermore, given sim-

ilar developments of declining fertility, increasing levels of education and stagnating or even falling levels of female labor supply in other developing countries, particularly in Africa and Asia, over the last three decades, the findings of this study contribute in understanding more global patterns in female labor supply that go beyond the Indonesian context.

Second, the results suggest that similar policy interventions are likely to have very different effects in developed and developing countries. In the U.S. and other Western countries much attention has been given to the role of child care costs/subsidies, including direct provision of public pre-school, on female labor supply (Blau and Robins, 1988; Conelly, 1992; Kimmel, 1998; Gelbach, 2002). However, in the Indonesian case child care cost considerations seem to play a much less important role. Therefore, the overall effect of such interventions on female labor supply is likely to be rather small in the country. In contrast, policies that affect the direct costs of children, such as the free provision of school meals or school uniforms, or simply a fixed amount of money a woman receives per child, might lead to a relatively large response in overall female labor supply in Indonesia. If the support provided is meaningful enough to relax the budget constraint of a sufficiently large share of women, than we would expect significant crowding out effects with respect to female labor supply with overall labor supply falling particularly among less well-off women.

## References


**Abadie, A.**, "Semiparametric Instrumental Variable Estimation of Treatment Response Model," *Journal of Econometrics*, 2003, *113* (2), 231–263.

**Angrist, J.D.**, "Estimation of Limited Dependent Variable Models with Dummy Endogenous Regressors: Simple Strategies for Empirical Practice," *Journal of Business and Economic Statistics*, 2001, *19* (1), 2–28.

_ **and A.B. Krueger**, "Split Sample Instrumental Variable Estimates of the Return to Schooling," *Journal of Business and Economic Statistics*, 1995, *13* (2), 225–235.

_ **and J. Pischke**, *Mostly Harmless Econometrics: An Empiricist's Companion*, Oxfordshire:Princeton University Press, 2009.

_ **and** _ , "The Credibility Revolution in Empirical Economics: How better Research Design is Taking the Con out of Econometrics," *Journal of Economic Perspectives*, 2010, *24* (2), 3–30.

_ **and V. Lavy**, "Multiple Experiments for the Causal Link between the Quantity and Quality of Children," *Journal of Labor Economics*, 2011, *forthcoming.*

_ **and W.N. Evans**, "Children and Their Parents Labor Supply: Evidence from Exogenous Variation in Family Size," *American Economic Review*, 1998, *88* (3), 450–477.

_ **, G.W. Imbens, and A.B. Krueger**, "Jackknife Instrumental Variables Estimation," *Journal of Applied Econometrics*, 1999, *14* (1), 57–67.

_ **,** _ **, and D. Rubin**, "Identification of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association*, 1996, *91* (434), 444–455.

**Bekker, P.A. and J. van der Ploeg**, "Instrumental Variable Estimation Based on Grouped Data," *Statistica Neerlandica*, 2005, *59* (3), 239–267.

**Bhalotra, S. and M. Umaña-Aponte**, "The Dynamics of Women's Labor Supply in Developing Countries," *IZA Discussion Paper No. 4879*, 2010.

**Blau, D.M. and P.K. Robins**, "Child-Care Costs and Family Labor Supply," *The Review of Economics and Statistics*, 1988, *70* (3), 374–381.

**Bound, J., D.A. Jaeger, and R.M. Baker**, "Problem with Instrumental Variable Estimation when the Correlation between the Instrument and the Endogeneous Explanatory Variable is Weak," *Journal of the American Statistical Association*, 1995, *90* (430), 443–450.

**Bronars, S.G. and J.T. Grogger**, "The Economic Consequences of Unwed Motherhood: Using Twin Births as Natural Experiment," *American Economic Review*, 1994, *84* (5), 1141–1156.

**Chao, J. and N. Swanson**, "Consistent Estimation with a Large Number of Weak Instruments," *Econometrica*, 2005, *73* (5), 1673–1692.

**Coleman, M.T. and J. Pencavel**, "Trends in Market Work Behavior of Women since 1940," *Industrial and Labor Relations Review*, 1993, *46* (4), 653–676.

**Conelly, R.**, "The Effect of Child Care Costs on Married's Women Labor Force Participation," *The Review of Economics and Statistics*, 1992, *74* (1), 83–90.

**Cruces, G. and S. Galiani**, "Fertility and Female Labor Supply in Latin America: New Causal Evidence," *Labour Economics*, 2007, *14* (3), 565–573.

**Deaton, A.**, "Instruments, Randomization, and Learning about Development," *Journal of Economic Literature*, 2010, *48* (2), 424–455.

_ **and J. Muellbauer**, "On Measuring Child Costs: With Applications to Poor Countries," *Journal of Political Economy*, 1986, *94* (4), 720–744.

**Duflo, E.**, "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment," *American Economic Review*, 2001, *91* (4), 795–813.

**Ebenstein, A.**, "When is the Local Average Treatment close to the Average: Evidence from Fertiliy and Labor Supply," *Journal of Human Resources*, 2009, *44* (4), 955–975.

**Frankenberg, E., J. P. Smith, and D. Thomas**, "Economic Shocks, Wealth and Welfare," *Journal of Human Resources*, 2003, *38* (2), 280–321.

**Fröhlich, M.**, "Nonparametric IV estimation of Local Average Treatment Effects with Covariates," *Journal of Econometrics*, 2007, *139* (1), 35–75.

**Gelbach, J.B.**, "Public Schooling for Young Children and Maternal Labor Supply," *American Economic Review*, 2002, *92* (1), 307–322.

**Goldin, C.**, *Understanding the Gender Gap*, New York: Oxford University Press, 1990.

**Hahn, J., J.A. Hausman, and G.M. Kuersteiner**, "Estimation with Weak Instruments: Accuracy of Higher-order Bias and MSE Approximations," *Econometrics Journal*, 2004, *7* (1), 1–2.

**Hausman, J.A., W.K. Newey, T. Woutersen, J. Chao, and N. Swanson**, "Instrumental Variable Estimation with Heteroskedasticity and Many Instruments," *Working Paper*, 2009, *Link: http://www.econ.jhu.edu/people/woutersen/hetiv09-09-2.pdf.*

**Heckman, J.J.**, "Building Bridges Between Structural and Program Evaluation Approaches to Evaluating Policy," *Journal of Economic Literature*, 2010, *48* (2), 356–398.

**ILO**, "Economically Active Population Estimates and Projection 1980-2020 (5th edition, Revision 2008)," *The International Labor Organization of the United Nations*, 2008, *48*.

**Imbens, G.W.**, "Better LATE Than Nothing: Some Comments on Deaton(2009) and Heckman and Urzua(2009)," *Journal of Economic Literature*, 2010, *48* (2), 399–423.

_ **and J.D. Angrist**, "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 1994, *62* (2), 467–475.

_ **and J.M. Wooldridge**, "Recent Developments in the Econometrics of Program Evaluation," *Journal of Economic Literature*, 2009, *47* (1), 5–86.

**Jacobsen, J.P., J.L. Rosenbloom, and J.W. Pearce III**, "Using Twin Births as a Natural Experiment," *Journal of Human Resources*, 1999, *34* (3), 449–474.

**Kimmel, J.**, "Child Care Costs as a Barrier to Employment for Single and Married Mothers," *The Review of Economics and Statistics*, 1998, *80* (2), 287–299.

**Levine, D. and M. Kevane**, "Are Investments in Daughters Lower When Daughters Move Away? Evidence from Indonesia," *World Development*, 2003, *31* (6), 1065–1084.

**Mammen, K. and C. Paxson**, "Women's Work and Economic Development," *Journal of Economic Perspectives*, 2000, *14* (4), 141–164.

**Miller, G. and B.P. Urdinola**, "Cyclicity, Mortality, and the Value of Time: The Case of Coffee Price Fluctuations and Child Survival in Columbia," *Journal of Political Economy*, 2010, *118* (1), 113–155.

**Pencavel, J.**, "Labor Supply of Men: A Survey," in O.C. Ashenfelter and R. Layard, eds., *Handbook of Labor Economics Vol.1*, Elsevier/North-Holland 1987, pp. 3–102.

**Permana, I.B. and C.W. Westoff**, "The Two Child Norm in Indonesia," *Macro International Technical Report to the DHS*, 1999.

**Phillips, G.D.A. and C. Hale**, "The Bias of Instrumental Variable Estimators of Simultaneous Equation Systems," *International Economic Review*, 1977, *18* (1), 219–228.

**Rosenzweig, M.R. and K. Wolpin**, "Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment," *Econometrica*, 1980, *48* (1), 227–240.

_ **and** _ , "Natural 'Natural Experiments' in Economics," *Journal of Economic Literature*, 2000, *38* (4), 827–874.

**Smith, J.P., D. Thomas, E. Frankenberg, K. Beegle, and G. Teruel**, "Wages, employment and economic shocks: Evidence from Indonesia," *Journal of Population Economics*, 2002, *15* (1), 161–193.

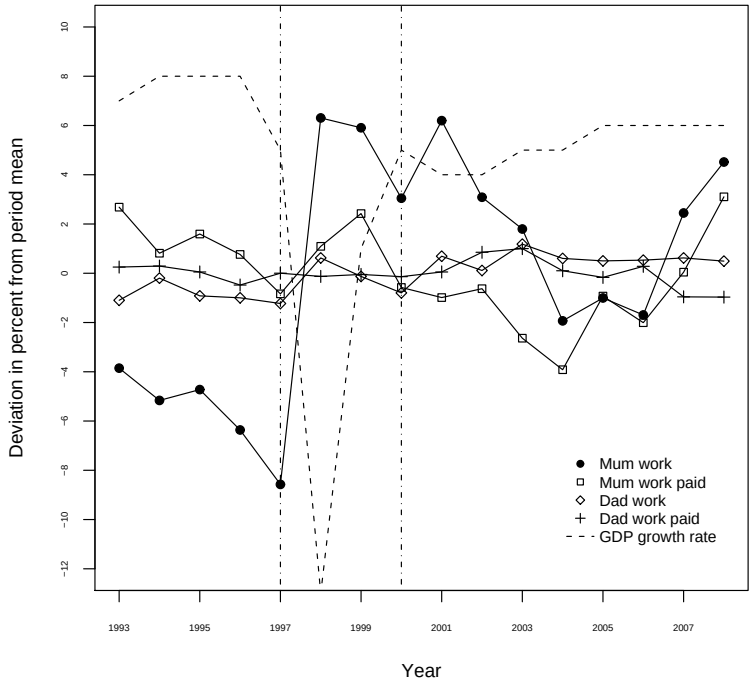
**Soares, R.R. and B.L.S. Falc̃ao**, "The Demographic Transition and the Sexual Division of Labor," *Journal of Political Economy*, 2008, *116* (6), 1058–1104.

**Staiger, D. and J.H. Stock**, "Instrumental Variable Regressions with Weak Instruments," *Econometrica*, 1997, *65* (3), 557–586.

**Thomas, D., K. Beegle, E. Frankenberg, B. Sikoki, J. Strauss, and G. Teruel**, "Education in a crisis," *Journal of Development Economics*, 2004, *74* (1), 53–85.

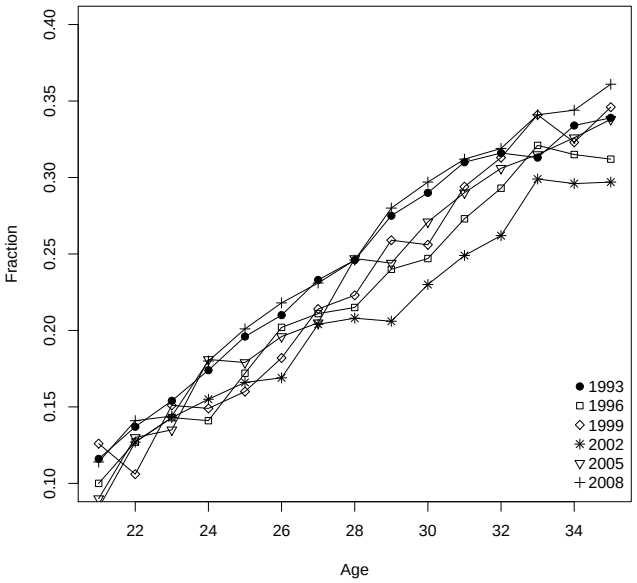
# 8. Figures and Tables

Figure 1: Deviations from period average: Indonesia 1993-2008, Sample: Women aged 21-35 with 2 or more children and their spouses

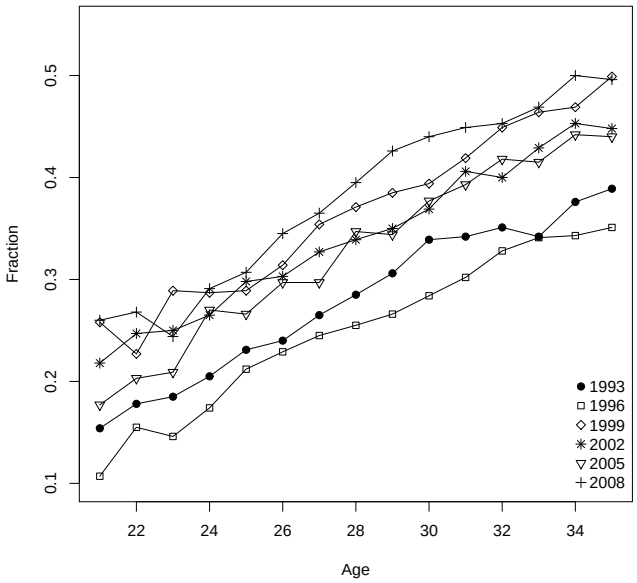



*Source: Susenas and IMF International Financial Statistics Database.*

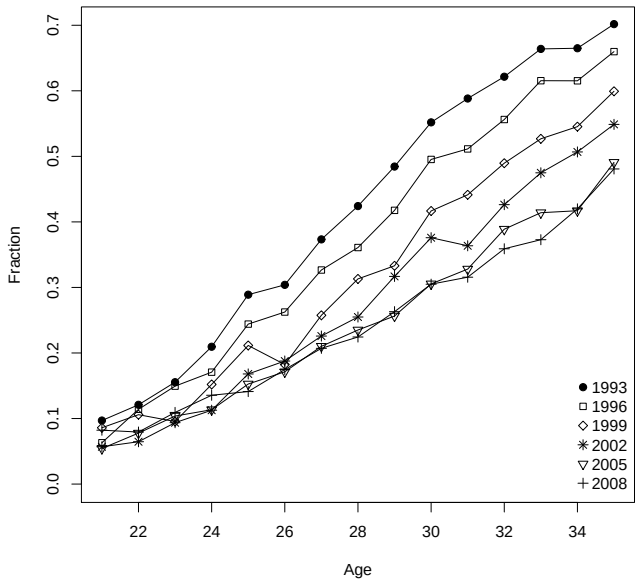
Figure 2: Means of Fertility and Labor Market Outcomes by Age, Women Aged 21-35 with 2 or more children




(a) Worked for pay last month




(b) Worked last month




(c) More than 2 children

Table 1: Descriptive Statistics
(Women Aged 21-35 with at least two children)

| Round | Age | Age at first birth | % with 3rd child | % Rural | HH size | Years of Education Mother | Years of Education Father | % Married |
|---|---|---|---|---|---|---|---|---|
| 1993 | 29.9 | 19.9 | 51 | 69 | 5.3 | 5.5 | 6.4 | 97 |
| 1994 | 30.0 | 20.0 | 49 | 68 | 5.2 | 5.9 | 6.7 | 97 |
| 1995 | 30.0 | 20.0 | 48 | 67 | 5.2 | 5.9 | 6.9 | 97 |
| 1996 | 30.1 | 20.1 | 47 | 68 | 5.2 | 6.1 | 6.9 | 98 |
| 1997 | 30.1 | 20.1 | 45 | 66 | 5.1 | 6.5 | 7.0 | 98 |
| 1998 | 30.2 | 20.2 | 43 | 65 | 5.0 | 6.6 | 7.1 | 98 |
| 1999 | 30.3 | 20.3 | 41 | 64 | 5.0 | 6.8 | 7.2 | 98 |
| 2000 | 30.4 | 20.4 | 39 | 61 | 4.9 | 6.9 | 7.3 | 98 |
| 2001 | 30.2 | 20.5 | 37 | 59 | 4.9 | 7.1 | 7.5 | 98 |
| 2002 | 30.4 | 20.4 | 37 | 58 | 4.8 | 7.4 | 7.8 | 98 |
| 2003 | 30.4 | 20.5 | 36 | 61 | 4.7 | 7.3 | 7.8 | 98 |
| 2004 | 30.4 | 20.6 | 34 | 60 | 4.9 | 7.6 | 8.0 | 98 |
| 2005 | 30.6 | 20.6 | 33 | 58 | 4.9 | 7.7 | 8.0 | 98 |
| 2006 | 30.6 | 20.5 | 32 | 60 | 4.8 | 7.8 | 8.1 | 98 |
| 2007 | 30.5 | 20.6 | 33 | 60 | 4.9 | 8.0 | 8.1 | 98 |
| 2008 | 30.6 | 20.8 | 32 | 56 | 4.9 | 8.0 | 8.1 | 98 |
| All | 30.3 | 20.3 | 40 | 63 | 5.0 | 6.9 | 7.4 | 98 |

*Note: Survey weights are applied.*

Table 2: Female labor force participation
(Women Aged 21-35 with at least two children)

| Round | Mothers with two children | | | Mothers with a third child | | |
|---|---|---|---|---|---|---|
| | Wage employed | Work for pay | Work | Wage employed | Work for pay | Work |
| 1993 | 0.13 | 0.28 | 0.32 | 0.09 | 0.26 | 0.31 |
| 1994 | 0.12 | 0.27 | 0.30 | 0.09 | 0.25 | 0.30 |
| 1995 | 0.13 | 0.27 | 0.31 | 0.09 | 0.25 | 0.30 |
| 1996 | 0.11 | 0.27 | 0.29 | 0.08 | 0.24 | 0.28 |
| 1997 | 0.12 | 0.25 | 0.27 | 0.08 | 0.23 | 0.26 |
| 1998 | 0.12 | 0.27 | 0.41 | 0.08 | 0.24 | 0.41 |
| 1999 | 0.13 | 0.27 | 0.41 | 0.09 | 0.27 | 0.41 |
| 2000 | 0.11 | 0.25 | 0.38 | 0.09 | 0.25 | 0.38 |
| 2001 | 0.09 | 0.24 | 0.41 | 0.06 | 0.23 | 0.42 |
| 2002 | 0.10 | 0.25 | 0.38 | 0.07 | 0.23 | 0.38 |
| 2003 | 0.08 | 0.22 | 0.37 | 0.06 | 0.21 | 0.37 |
| 2004 | 0.08 | 0.21 | 0.33 | 0.06 | 0.20 | 0.34 |
| 2005 | 0.12 | 0.25 | 0.35 | 0.08 | 0.23 | 0.35 |
| 2006 | 0.09 | 0.23 | 0.33 | 0.05 | 0.22 | 0.33 |
| 2007 | 0.12 | 0.27 | 0.38 | 0.08 | 0.25 | 0.35 |
| 2008 | 0.13 | 0.30 | 0.41 | 0.08 | 0.27 | 0.39 |
| All | 0.11 | 0.26 | 0.36 | 0.08 | 0.24 | 0.36 |

*Note: 'Work for pay' consists of wage and self-employment while 'Work' includes in addition unpaid family work. Survey weights are applied.*

Table 3: Female labor force participation by rural/urban
(Women Aged 21-35 with at least two children)

| Round | Rural | | | Urban | | |
|---|---|---|---|---|---|---|
| | HH size | M2 Work for pay | M3 Work for pay | HH size | M2 Work for pay | M3 Work for pay |
| 1993 | 5.2 | 0.27 | 0.26 | 5.4 | 0.29 | 0.26 |
| 1994 | 5.2 | 0.26 | 0.25 | 5.3 | 0.28 | 0.25 |
| 1995 | 5.1 | 0.26 | 0.25 | 5.3 | 0.29 | 0.26 |
| 1996 | 5.1 | 0.26 | 0.24 | 5.3 | 0.28 | 0.24 |
| 1997 | 5.1 | 0.24 | 0.22 | 5.2 | 0.26 | 0.21 |
| 1998 | 5.0 | 0.26 | 0.25 | 5.1 | 0.27 | 0.25 |
| 1999 | 5.0 | 0.26 | 0.26 | 5.0 | 0.30 | 0.28 |
| 2000 | 4.9 | 0.24 | 0.22 | 5.0 | 0.26 | 0.25 |
| 2001 | 4.9 | 0.22 | 0.21 | 4.9 | 0.27 | 0.25 |
| 2002 | 4.8 | 0.23 | 0.22 | 4.8 | 0.27 | 0.24 |
| 2003 | 4.7 | 0.21 | 0.21 | 4.7 | 0.24 | 0.22 |
| 2004 | 4.9 | 0.20 | 0.20 | 4.9 | 0.23 | 0.21 |
| 2005 | 4.9 | 0.24 | 0.22 | 4.9 | 0.27 | 0.25 |
| 2006 | 4.8 | 0.22 | 0.20 | 4.8 | 0.26 | 0.24 |
| 2007 | 4.9 | 0.26 | 0.23 | 5.0 | 0.31 | 0.28 |
| 2008 | 4.9 | 0.29 | 0.27 | 5.0 | 0.33 | 0.29 |
| All | 5.0 | 0.25 | 0.24 | 5.0 | 0.28 | 0.25 |

*Note: M2 refers to women with exactly two children. M3 refers to women with more than 2 children. Survey weights are applied.*

Table 4: Female labor force participation by expenditure quintile
(Women Aged 21-35 with at least two children)

| Round | Quintile 1 | | Quintile 2 | | Quintile 3 | | Quintile 4 | | Quintile 5 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | M2 Work for pay | M3 Work for pay | M2 Work for pay | M3 Work for pay | M2 Work for pay | M3 Work for pay | M2 Work for pay | M3 Work for pay | M2 Work for pay | M3 Work for pay |
| 1993 | 0.26 | 0.26 | 0.27 | 0.27 | 0.27 | 0.26 | 0.28 | 0.28 | 0.30 | 0.27 |
| 1994 | 0.24 | 0.24 | 0.26 | 0.25 | 0.25 | 0.25 | 0.27 | 0.25 | 0.27 | 0.25 |
| 1995 | 0.25 | 0.25 | 0.26 | 0.26 | 0.26 | 0.25 | 0.27 | 0.25 | 0.28 | 0.25 |
| 1996 | 0.23 | 0.23 | 0.24 | 0.24 | 0.27 | 0.26 | 0.27 | 0.25 | 0.28 | 0.24 |
| 1997 | 0.22 | 0.23 | 0.24 | 0.24 | 0.26 | 0.25 | 0.25 | 0.24 | 0.25 | 0.24 |
| 1998 | 0.24 | 0.26 | 0.26 | 0.27 | 0.26 | 0.27 | 0.27 | 0.27 | 0.27 | 0.26 |
| 1999 | 0.24 | 0.25 | 0.27 | 0.28 | 0.27 | 0.28 | 0.28 | 0.28 | 0.29 | 0.27 |
| 2000 | 0.23 | 0.24 | 0.24 | 0.26 | 0.24 | 0.25 | 0.25 | 0.25 | 0.26 | 0.25 |
| 2001 | 0.21 | 0.21 | 0.23 | 0.22 | 0.23 | 0.22 | 0.25 | 0.24 | 0.28 | 0.27 |
| 2002 | 0.22 | 0.22 | 0.23 | 0.24 | 0.24 | 0.23 | 0.24 | 0.23 | 0.28 | 0.26 |
| 2003 | 0.20 | 0.20 | 0.21 | 0.21 | 0.23 | 0.22 | 0.23 | 0.21 | 0.24 | 0.23 |
| 2004 | 0.19 | 0.18 | 0.21 | 0.20 | 0.21 | 0.20 | 0.22 | 0.22 | 0.24 | 0.21 |
| 2005 | 0.22 | 0.22 | 0.22 | 0.22 | 0.23 | 0.22 | 0.24 | 0.24 | 0.26 | 0.23 |
| 2006 | 0.20 | 0.20 | 0.22 | 0.21 | 0.22 | 0.22 | 0.24 | 0.23 | 0.26 | 0.25 |
| 2007 | 0.22 | 0.22 | 0.25 | 0.24 | 0.27 | 0.25 | 0.29 | 0.27 | 0.33 | 0.29 |

*Note: Women are assigned into quintiles based on per capita expenditures. Quintile 1 refers to the poorest quintile. Susenas 2008 did not collect expenditure information in the Core part. Therefore, no statistics for 2008 can be presented. M2 refers to women with exactly two children. M3 refers to women with more than 2 children. Survey weights are applied.*

Table 5: Expenditure patterns over time in Indonesia
(Women Aged 21-35 with at least two children)

| Indicator | 1993 | 1996 | 1999 | 2002 | 2005 | 2008 |
|---|---|---|---|---|---|---|
| Per-Capita expenditure | 139,193 | 165,481 | 166,216 | 189,692 | 213,301 | 231,346 |
| Food (S) | .631 | .634 | .687 | .649 | .596 | .586 |
| Health Expenditures (S) | .013 | .016 | .014 | .019 | .019 | .020 |
| Clothing Expenditures (S) | .054 | .051 | .048 | .048 | .038 | .036 |
| Child Clothing (S) | .011 | .014 | .016 | .014 | .014 | .015 |
| Child Clothing (S.p.c.) | .003 | .004 | .005 | .004 | .004 | .004 |
| Child Education (S) | .021 | .023 | .017 | .019 | .028 | .025 |
| Child Education (S.p.c.) | .007 | .008 | .006 | .007 | .011 | .010 |
| Number of observations | 16,751 | 15,905 | 13,824 | 13,564 | 15,314 | 15,847 |

*Note: Expenditures are monthly and in real terms with base year 2002 using the national CPI. 'S' refers to the share of a particular expenditure item over total expenditures while 'S.p.c.' refers to the respective share divided by the number of children living in the household. Survey weights are applied.*

Table 6: Composition of births
(Women Aged 21-35 with at least two children)

| | % | % with another child | $\chi^2$ test |
|---|---|---|---|
| First boy | 51.8 | 59.7 | |
| First girl | 48.2 | 59.4 | |
| Two boys | 26.8 | 41.8 | 576 |
| Two girls | 23.4 | 41.2 | 428 |
| Same sex | 50.2 | 41.5 | 751 |
| Mixed sex | 49.8 | 38.5 | |

*Note: $\chi^2$ test for contingency tables used based on the pooled Susenas sample. Test statistics are reported against 'Mixed sex'.*

Table 7: Expenditure patterns conditional on child composition in Indonesia
(Women Aged 21-35 with at least two children - Pooled sample)

| Indicator | Mixed (A) | Same sex (B) | Two boys (C) | Two girls (D) | Diff1 (A)-(B) | Diff2 (A)-(C) | Diff3 (A)-(D) | Diff4 (C)-(D) |
|---|---|---|---|---|---|---|---|---|
| Food expenditures (S) | 0.626 | 0.626 | 0.626 | 0.626 | 0.000 | 0.000 | 0.000 | 0.000 |
| Health expenditures (S) | 0.017 | 0.017 | 0.017 | 0.017 | 0.000 | 0.000 | 0.000 | 0.000 |
| Child clothing (S) | 0.014 | 0.014 | 0.014 | 0.014 | 0.000 | 0.000 | 0.000 | 0.000 |
| Child clothing (S.p.c.) | 0.004 | 0.004 | 0.004 | 0.004 | 0.000 | 0.000 | 0.000 | 0.000 |
| Child education (S) | 0.023 | 0.023 | 0.023 | 0.023 | 0.000 | 0.000 | 0.000 | -0.001 |
| Child education (S.p.c.) | 0.008 | 0.008 | 0.008 | 0.008 | 0.000 | 0.000 | 0.000 | 0.000 |

*Note: Expenditures are monthly and in real terms with base year 2002 using the national CPI. 'S' refers to the share of a particular expenditure item over total expenditures while 'S.p.c.' refers to the respective share divided by the number of children living in the household. Survey weights are applied. Each single difference is significant according to appropriate t-test or the Mann-Whitney-U-test/Wilcoxon-ranksum-test.*

Table 8: Estimates of Female Labor Supply Model, Women Aged 21-35 with 2 or more children

Dependent Variable: Mother worked for pay

| Geographical Level | Time Period | 2SLS First Stage (saturated model) | OLS | 2SLS | 2SLS (saturated) | Abadie03 | Fröhlich07 | SSIV | Number of observations |
|---|---|---|---|---|---|---|---|---|---|
| National | All | .031(.002)/4295 | -.040 (.001) | .049(.021) | .043(.020) | .044(.022) | .043(.023) | .039(.023) | 845,978 |
| | Pre-crisis | .030(.003)/1934 | -.044 (.002) | .030(.059) | .027(.057) | .029(.059) | .028(.059) | .018(.060) | 274,171 |
| | Crisis | .031(.003)/962 | -.035(.002) | .097(.044) | .090(.044) | .094(.044) | .092(.045) | .084(.043) | 139,747 |
| | Post-crisis | .031(.002)/2327 | -.038 (.001) | .044(.031) | .041(.030) | .042(.031) | .042(.032) | .037(.033) | 432,060 |
| Urban | All | .035(.003)/1503 | -.055 (.002) | -.029(.036) | -.033(.037) | -.030(.038) | -.031(.038) | -.035(.038) | 275,096 |
| | Pre-crisis | .034(.003)/635 | -.063 (.004) | -.044(.051) | -.047(.052) | -.045(.052) | -.044(.053) | -.048(.053) | 80,482 |
| | Crisis | .035(.003)/343 | -.051 (.004) | -.001(.055) | -.009(.054) | -.006(.056) | -.006(.057) | -.013(.058) | 44,088 |
| | Post-crisis | .035(.003)/819 | -.053(.003) | -.026(.043) | -.030(.044) | -.027(.045) | -.028(.044) | -.031(.045) | 150,526 |
| Rural | All | .028(.002)/3506 | -.034(.001) | .082(.025) | .081(.023) | .081(.024) | .083(.024) | .075(.027) | 570,882 |
| | Pre-crisis | .028(.002)/1399 | -.036(.002) | .062(.031) | .063(.031) | .062(.032) | .063(.032) | .061(.032) | 193,689 |
| | Crisis | .028(.003)/672 | -.033(.003) | .162(.080) | .156(.076) | .157(.079) | .159(.079) | .142(.082) | 95,659 |
| | Post-crisis | .028(.002)/1593 | -.034(.002) | .084(.039) | .080(.038) | .083(.039) | .086(.041) | .068(.040) | 281,534 |

*Note: Each model is estimated with province and year fixed effects. The simple '2SLS model' does not include further covariates. All other models include the following covariates: sex of the first child, age of the mother, age at first birth of the mother, marital status and an urban dummy. In the case of the 'saturated 2SLS model' and the 'SSIV' the age and birthage variables are each divided into five mutually exclusive age categories with an approximately equal number of observations. These two models include in addition all possible interaction effects between age, age at first birth, and the sex of the first child. Therefore, these models are not fully saturated. The Abadie03 model uses a linear first stage to estimate Abadie's Kappa. In the second stage it used a probit approximation as in Angrist (2001) using Matlab code provided by Alberto Abadie. Standard errors for Abadie03 are bootstrapped using 500 replications on 1/10 of the respective sample. The SSIV model follows Angrist and Krueger (1995). Standard errors are in brackets. F-statistics from the first stage are reported after the backslash.*

Table 9: Estimates of Female Labor Supply Model, Women Aged 21-35 with 2 or more children

Dependent Variable: Mother worked for pay

| | Method | 2SLS | 2SLS | 2SLS | 2SLS | LIML | LIML | HFUL | HFUL |
|---|---|---|---|---|---|---|---|---|---|
| | Instrument | Same sex | 2 boys, 2 girls | Same sex | Same sex | Same sex | Same sex | Same sex | Same sex |
| Instrument | Year dummies | | | x | x | x | x | x | x |
| Interacted | Province dummies | | | | x | | x | | x |
| Geographical Level | | | | | | | | | |
| National | First stage F-stat | 4,295 | 4,213 | 3,333 | 487 | 3,333 | 487 | 3,333 | 487 |
| | Sargan/Basman (p-value) | | .092 | .062 | .019 | .062 | .019 | .062 | .019 |
| | Second stage $\tau$ | .043(.020) | .046(.022) | .048(.024) | .021(.012) | .050(.023) | .058(.025) | .048(.023) | .056(.026) |
| Urban | First stage F-stat | 1,503 | 1,476 | 1,162 | 168 | 1,162 | 168 | 1,162 | 168 |
| | Sargan/Basman (p-value) | | .052 | .033 | .009 | .033 | .009 | .033 | .009 |
| | Second stage $\tau$ | -.033(.037) | -.031(.037) | -.018(.036) | .009(.030) | -.026(.038) | .004(.032) | -.031(.040) | .000(0.37) |
| Rural | First stage F-stat | 3,506 | 2,876 | 2,238 | 337 | 2,238 | 337 | 2,238 | 337 |
| | Sargan/Basman (p-value) | | .073 | .049 | .014 | .049 | .014 | .049 | .014 |
| | Second stage $\tau$ | .081(.023) | .083(.025) | .071(.026) | .032(.016) | .078(.027) | .088(.027) | .079(.028) | .084(.033) |

*Note: 'x' indicates with which dummy variables the instrument is interacted. Variables included are those as in the 'saturated model' in Table 8. Standard errors are presented in parenthesis. In the 2SLS and LIML procedure, clustered standard errors are used.*

Table 10: Statistics on Expenditure Quintiles

| Variables | Expenditure Quintiles Q 1 | Q 2 | Q 3 | Q4 | Q5 |
|---|---|---|---|---|---|
| Age | 29.98 | 30.04 | 30.12 | 30.26 | 30.66 |
| Age at first birth | 20.16 | 20.15 | 20.32 | 20.64 | 21.47 |
| % Rural | .88 | .79 | .69 | .54 | .33 |
| % with 3rd child | .56 | .47 | .41 | .36 | .30 |
| 2SLS estimate | .170(.066) | .116(.064) | .005(.041) | .060(.045) | .039(.038) |
| No. of observations | 157,542 | 157,557 | 157,575 | 157,553 | 157,548 |

*Note: The sample is based on women aged 21-35 with at least two children. No observations for 2008 are included. Q1 refers to the poorest expenditure quintile. 2SLS is based on the 'saturated model' as described in Table 8. Clustered standard errors are presented in parenthesis.*

Table 11: Statistics on Education Terciles

| Variables | Mother's Schooling T1 | T2 | T3 |
|---|---|---|---|
| Age | 30.4 | 29.9 | 30.3 |
| Age at first birth | 19.7 | 19.9 | 21.8 |
| % Rural | .88 | .79 | .48 |
| % with 3rd child | .54 | .42 | .35 |
| % Married | .97 | .98 | .98 |
| 2SLS estimate | .068(.027) | .025(.037) | .049(.031) |
| No. of observations | 226,804 | 319,391 | 299,802 |

*Note: The sample is based on women aged 21-35 with at least two children. T1 refers to all women who have not completed primary schooling. T2 refers to all women who have completed primary schooling and T3 refers to all women who finished secondary schooling. 2SLS is based on the 'saturated model' as described in Table 8. Clustered standard errors are presented in parenthesis.*

Table 12: Estimates of Female Labor Supply

| Child Age | 2SLS National | 2SLS Urban | 2SLS Rural |
|---|---|---|---|
| Below 6 | .025(.018) | -.057(.037) | .070(.038) |
| 6 to 10 | .068(.028) | -.022(.023) | .118(.054) |
| Above 10 | .033(.018) | -.021(.152) | .041(.038) |

*Note: The sample is based on women aged 21-35 with at least two children. Age classification of children is based on age of the youngest child. Variables included are those as in the 'saturated model' in Table 8. Clustered standard errors are presented in parenthesis.*

Table 13: Estimates of Female Labor Supply Model, Women Aged 21-35 with 2 or more children

Dependent Variable: Mother worked for pay

| Variables | Method | 2 SLS | 2SLS | 2SLS | 2SLS | 2SLS | 2SLS |
|---|---|---|---|---|---|---|---|
| | Sample | All | All | Married | Married | Married | Married |
| Basic model (saturated) | | x | x | x | x | x | x |
| Mother's Education | | | x | | x | x | x |
| Husband's Education | | | | | | x | x |
| Husband works for pay | | | | | | | x |
| Geographical Level | Time Period | | | | | | |
| National | All | .049(.021) | .045(.022) | .048(.022) | .044(.020) | .046(.021) | .039(.018) |
| | Pre-crisis | .030(.059) | .026(.058) | .024(.058) | .022(.059) | .023(.062) | .018(.057) |
| | Crisis | .097(.044) | .084(.043) | .094(.044) | .086(.042) | .085(.040) | .075(.039) |
| | Post-crisis | .044(.031) | .036(.035) | .051(.033) | .043(.034) | .045(.035) | .038(.034) |
| Urban | All | -.029(.036) | -.033(.042) | -.031(.039) | -.036(.042) | -.039(.044) | -.042(.047) |
| | Pre-crisis | -.044(.051) | -.053(.061) | -.050(.054) | -.058(.062) | -.061(.065) | -.062(.067) |
| | Crisis | -.001(.055) | -.017(.063) | -.003(.057) | -.013(.065) | -.019(.067) | -.021(.069) |
| | Post-crisis | -.026(.043) | -.023(.050) | -.022(.047) | -.024(.053) | -.023(.054) | -.027(.055) |
| Rural | All | .082(.025) | .078(.027) | .082(.026) | .080(.033) | .078(.034) | .071(.031) |
| | Pre-crisis | .062(.031) | .061(.030) | .068(.033) | .066(.031) | .061(.030) | .058(.029) |
| | Crisis | .162(.080) | .158(.079) | .161(.081) | .158(.080) | .154(.080) | .142(.078) |
| | Post-crisis | .084(.039) | .080(.038) | .082(.040) | .080(.043) | .079(.045) | .077(.045) |

*Note: 'x' indicates which covariates are included. The saturated model is estimated as in Table 8. Education variables are added as dummies for each education category without being interacted with other variables.*

Table 14: Twin estimates of second multiple birth
Dependent Variable: Mother worked for pay

| Method/Sample | 2SLS (National) | 2SLS (Urban) | 2SLS (Rural) |
|---|---|---|---|
| All | .021(.011) | -.013(.015) | .056(.028) |
| Pre-crisis | .002(.024) | -.041(.029) | .023(.017) |
| Crisis | .052(.027) | .017(.023) | .089(.056) |
| Post-crisis | .016(.012) | -.004(.018) | .055(.032) |

*Note: Models are based on the 'saturated model' as described in Table 8. The sample uses women aged 21-35 with at least two children. Clustered standard errors are presented in parenthesis.*

# A. Appendix 1 - Data sources and data construction

The National Socioeconomic Survey 'Susenas' (Survei Sosial Ekonomi Nasional) is conducted annually by the Indonesian National Statistical Institute 'BPS' (Badan Pusat Statistik). The first round of Susenas was collected in 1963-1964. Since 1993, Susenas covers a nationally representative sample typically composed of about 200,000 households. This article uses the rounds of Susenas that are collected in the month of July in each year which are cross-sectional in its nature. Each Susenas survey contains a Core and a supplemental Module questionnaire.

Data from the Core: The Core questionnaire collects data on age, sex, marital status, educational attainment, labor force participation and wage from employment in the principal occupation for each member of the household. Furthermore, the Core questionnaire gathers data on some broad categories of household expenditures and on housing/dwelling conditions.

Data from the Module: The module questionnaire is administered to about 1/4 of the Core sample. Moreover, the topics covered in the Module rotate over a three year period. The rounds of 1993, 1996, 1999, 2002, 2005, and 2008 comprise a detailed income and expenditure section which is used in this article to obtain figures on clothing and schooling expenditures on children.

Linking mothers and children: In contrast to the Census data used in Angrist and Evans (1998), Cruces and Galiani (2007), Ebenstein (2009) and Angrist and Lavy (2011) Susenas data provides a direct mother identifier which is used to link children to the right mothers. Therefore, I do not need to confine the sample to mothers which are the head of the household (matching over the relation to the household head variable) as in Cruces and Galiani (2007) or Ebenstein (2009). As a cross-check I used the reported number of children ever born to a mother. I deleted any mother for whom the number of children in the household did not match the reported number of children ever born. The respective individual observations are reported in Table A.1 below.

Dependent Variable: This paper utilizes 'Work for pay' as the dependent variable in the analyses. The 'Work for pay' variable captures all persons that indicated to typically work as self-employed or employed. In contrast, the 'Work' variable is coded 1 if the person reports to have worked in the last week or indicated to typically work and 0 otherwise. The data indicates that all persons working as self-employed, employed or unpaid family workers are captured in the 'Work' variable. Woman who are engaged in agriculture are typically classified in the data by BPS as unpaid family workers. However, if the main occupation of the household head or the husband is outside agriculture than women working in agriculture are classified as self-employed. Housework is always coded as 0. Mothers who were attending school at the time of the survey and mothers with an own child below the age of 1 or above the age of 18 were excluded from the analysis.

Multiple births: Susenas does not provide an identifier of multiple births. Since quarter of birth information is missing in Susenas I assign each child 'twin status' if it has the same age in years as another child of the same mother. This procedure induces some measurement error in the twin variable. However, since child bearing takes about 9 months followed by 4-6 weeks in which a mother is typically not able to conceive, the overall measurement error is rather low. To obtain an idea of the measurement errors I use information from the six rounds of the Indonesian Demographic and Health Surveys (1987-2007) in which direct twinning information is available (Table A.2). From about 100,000 second births approximately 1% are multiple second births. Using the matching over age in years leads to about .097% of second births births being erroneously classified as twins and .093% being erroneously classified as single births. However, from those being classified as multiple second births about 19% are actually single births.

Table A.1: Sample Sizes: Susenas 1993-2008

| Round | Individuals | Male | Female | Women 20<Age<36 | Women with 2 children 20<Age<36 | Married Women with 2 children 20<Age<36 |
|---|---|---|---|---|---|---|
| 1993 | 903,351 | 449,757 | 453,594 | 115,045 | 57,666 | 56,011 |
| 1994 | 904,793 | 450,153 | 454,640 | 116,342 | 56,529 | 54,897 |
| 1995 | 873,630 | 434,450 | 439,180 | 111,414 | 53,719 | 52,230 |
| 1996 | 897,382 | 445,189 | 452,193 | 114,840 | 54,345 | 52,986 |
| 1997 | 887,265 | 440,751 | 446,514 | 114,068 | 51,912 | 50,573 |
| 1998 | 879,936 | 437,527 | 442,409 | 112,955 | 50,732 | 49,552 |
| 1999 | 864,580 | 431,133 | 433,447 | 109,426 | 47,690 | 46,497 |
| 2000 | 780,121 | 389,418 | 390,723 | 100,794 | 41,325 | 40,344 |
| 2001 | 889,413 | 446,089 | 443,324 | 115,403 | 48,950 | 47,896 |
| 2002 | 862,210 | 431,267 | 430,943 | 112,289 | 45,800 | 44,842 |
| 2003 | 895,427 | 450,578 | 444,849 | 116,855 | 48,131 | 47,075 |
| 2004 | 1,030,250 | 517,017 | 513,233 | 135,967 | 55,729 | 54,471 |
| 2005 | 1,052,091 | 528,467 | 523,624 | 136,738 | 58,176 | 56,950 |
| 2006 | 1,107,594 | 557,186 | 550,408 | 142,433 | 56,132 | 54,884 |
| 2007 | 1,167,019 | 585,106 | 581,913 | 146,329 | 60,920 | 59,457 |
| 2008 | 1,142,675 | 571,874 | 570,801 | 141,803 | 58,222 | 56,942 |
| Total | 15,137,737 | 7,565,962 | 7,571,795 | 1,942,701 | 845,978 | 825,607 |

*Note: Columns 2 to 5 report the number of observations as provided in the official Susenas data set. The last two columns report the number of observations after data cleaning procedures were applied.*

Table A.2: Twin Statistics - Pooled Indonesian DHSs

| # 2nd child | # Twin pairs | Error 1 (%) | Error 2 (%) | Error in Twin=1 (%) |
|---|---|---|---|---|
| 100,147 | 505 | 0.097 | 0.093 | 19.057 |

*Note: Number of twin pairs refers to second multiple birth pairs. Error 1 refers to wrongly being assigned to twin status. Error 2 refers to wrongly being assigned to non-twin status.*