# The Impact of Children on Earnings of Japanese Fathers: Estimate using Exogenous Variation in Family Size

Miho Iwasawa

National Institute of Population and Social Security Research, Tokyo

Akira Motegi

University of Tokyo

September 25, 2014

**Abstract**

Having children have various impacts on parents' life. While raising children cost directly and indirectly, having children may give sufficient incentive to work harder, which results in covering those costs. In this paper, we shed light on the causal relationship between childbearing and father's earnings. Since decision to have another child is likely to be jointly determined with prospective earnings, naïve estimates with OLS of the causal effect of family size on parents' earning should suffer from endogenous bias. To overcome this endogeneity problem, we use the exogenous variation in sex composition of the first two children to construct instrumental variables estimates of the effect of having additional children on father's income. Using large sample surveys on fertility in Japan, we estimated conventional OLS model and two-stage least-square models with the instrument. Naïve OLS estimates show the negative relationship between having a third birth and father's income. However, once we correct for endogeneity with IV estimates, such relationship alters to positive direction especially for less educated fathers.

**Introduction**

One of the drivers for fertility decline is increasing cost of raising children. According to OECD report (2011), a child accounts for approximately 15 to 39% of the budget of a couple without children. In addition to the increase in house prices in the 2000s in OECD countries, higher costs of private education in Japan and Korea are more likely to establish a barrier to fertility. On the other hand, many people feel that raising their children is rewarding experience. In fact, Leibenstein proposed in this work in 1957 that there are three types of utilities derived from having an additional child: the utility of a child as a productive unit, the utility of a child as a source of old-age security, and the utility of the child as a consumption good (Leibenstein 1957). As economy and welfare system are developed, the first two utilities of children have become weaker and only the last one remains an effective motive for having children. Children as consumption goods are fairly costly. On the other hand, people increasingly expect psychological satisfaction or motivation in life from having children. These psychological conditions may affect parental productivity in a positive way. In this paper, we focus on this advantageous aspect of having more children in the case of men. What impact does having children on their father's earnings?

**Background**

It is widely documented that motherhood wage is negatively affected by having children. On the other hand, the opposite is observed to hold for father across developed countries (Harkness and Waldfogel 1999). For parenthood penalty, several explanations have been proposed. Difficulty in reconciliation between work and family and unfavorable treatment for discrimination against parenthood are considered to be possible mechanisms to lower wage of mothers. On the other hand, increasing family size often shows positive relationship with fathers' wage. Increasing productivity

by specialization between father and mother would be one of the plausible explanations (Lundberg and Rose 2002). In addition, we expect that an additional child gives a father mental fulfillment and a feeling of responsibility. This may also increase his productivity. Other than these, Kawaguchi (2005) suggested that there are many Japanese companies that pay child allowance to employees with children as a part of wage. This may also affect positively on father's wage.

However, estimating these relationships between family size and parents' earnings as a framework of causality is often complicated due to the endogeneity problem, which typically arises in analyses using cross-sectional data. Fathers who have good career prospects are more likely to have an additional child. In fact, not a few previous studies using fixed effect models suggested that magnitude of family size effect becomes smaller than the OLS or random effect models without considering unobserved heterogeneity or self-selection process. To overcome this problem, we utilize exogenous variation in family size based of parental preference for sex composition of children following the strategy in Angrist and Evans (1998).

**Endogenous variation in family size**

A significant impact of parental preference for children's sex composition on the number of children has been well documented (Bongaarts and Potter 1983). Since most parents have desire to have at least one boy and one girl, the third birth probability for couples with two boys or two girls tends to be much higher than those with mixed-sex offspring. Sex composition of children is reasonably considered to be randomly assigned in Japan where sex selection is very uncommon. That means that their sex composition is unrelated to any unobserved factors that might affect father's earnings. This property allows us to use the previous-two-children's sex composition as a plausible instrument for the third birth progression among couples with at least two children. We create a dummy variable *same sex* indicating whether the sex of the previous two children is the same (boy-boy or girl-girl).

**Data and methods**

We use the five rounds of the Japanese National Fertility Surveys conducted by the National Institute of Population and Social Security Research from 1992 through 2010. Data for couple surveys include information of married wives under age 50 and their husband. Response rates are about 85%. The sample is limited to the couples in which husband is under age 50 and working having at least two children. Those fathers who did not provide educational level are excluded from the analyses and these restrictions leave us with a sample of 17,724.

Using this data we estimate our models with ordinary least-squares (OLS) and two-stage least-squares (2SLS). With a simple OLS model, log of father's last-year income is accounted for by the presence of the third child. For a 2SLS equation, we use same covariates using *same sex* dummy as an instrument for the exogeneous fertility variable. Other covariates in the models are year from the second childbirth, region of residence, survey round, father's age, father's age-squared, father's

education, father's employment status, father's occupation, father's firm size, and mother's working status.

Besides relationship between children and father's earnings in the entire sample, we investigate how the impact of children on father's earnings varies with father's socio-economic status or working conditions. As a provisional analysis, we estimate separate models with the same covariates for fathers with high school graduate and junior high school graduate, and for fathers graduated from vocational school or college/university, respectively.

**Results**

The first panel (A) in Table 1 shows the result of the OLS and the 2SLS estimates for all fathers. We show only the results for the instrument variable and the main explanatory variable. The right column is for the OLS model of father's income. The coefficient for presence of a third birth shows negative direction and is significantly different from zero at 10% level. This indicates that fathers having three children tend to earn less money than those with two children.

Let us move to the result for the 2SLS model. In the 1$^{st}$ stage estimation, we use *same sex* as an exogenous variation in family size. As we expected, *same sex* has a significant positive effect on the presence of a third child. The F-statistics from the first-stage estimation is above 50 and well beyond 10, the result which ensures the predictive power of the instrument. Looking at the 2SLS estimates, we found that the negative direction in the presence of a third birth disappears. This suggests that negative relationship observed in the OLS model is a spurious one caused by the selection process that fathers with less earning power are more likely to have a third child.

The panel B and C give the comparison between the results for low-educated fathers and highly educated fathers, respectively. For low-educated fathers, we got an interesting result. With the OLS model, a presence of a third birth child does not show any significant association with father's income. However, once we correct for endogeneity of fertility, significant opposite sign at 10% level is found. That is, having an additional child may increase father's income.

**Conclusion**

Our powerful instrument of the sex composition of the first two children reflecting parents' sex balance preference allows us to examine the causal relationship between childbirth and father's earnings. The difference in the results between naïve OLS and IV estimates suggests that observed negative correlation between childbearing and father's income in the former is in fact spurious, which is largely accounted for by the selection process; fathers with less economic resources are more likely to have additional children. This is opposite to the theoretical expectation assuming that sufficient income encourages a father to have more children. However, if we assume their relatively low opportunity costs of children and relatively high cost burden of contraception, high fertility of low income households would be reasonable.

We also found heterogeneity in the effect of additional childbearing on father's income. The result from the separate IV regression for less-educated fathers shows positive causal effect of additional childbirth on father's income. This finding points out the possibility that sensitivity of father's earnings to increased children may vary with their own characteristics such as labor attitude or working condition, which should be examined in further analyses with alternative specifications.

* Permission to use the data on National Fertility Surveys was obtained through the National Institute of Population and Social Security Research on the basis of the Statistics Act (Act No. 53 of 2007), Article 32.

**References**

Angrist, J.D. and N. E. William (1998) "Children and Their Par-ents' Labour Supply: Evidence from Exogeneous Variation in Family Size." *The American Economic Review* 88(3):450-477

Bongaarts, J., and R. G. Potter (1983) *Fertility, Biology, and Behavior: An Analysis of the Proximate Determinants*, New York: Academic Press.

Harkness, S. and J. Waldfogel (1999) *The Family Gap in Pay: Evidence from Seven Industrialised Countries*. Center for Analysis of Social Exclusion. CASEpaper No.29.

Hirvonen, L. (2009) *The Effect of Children on Earnings Using Exogenous variation in Family Size: Swedish Evidence*, Working Paper 2 /2009, Swedish Institute for Social Research, Stockholm University.

Kawaguchi, A. (2005) "Kekkon to Syussan ha Danjyo no Chingin ni Donoyona Eikyo wo Oyobosite Irunoka." Nihon Rodo Kenkyu Zassi, No.535.

Leibenstein, H. (1957) Economic Backwardness and Economic Growth. New York: Wiley.

Lundberg, S. and E. Rose (2002)"The Effects of Sons and Daughters on Men's Labor Supply and Wages." *Review of Economics and Statistics* 84(2): 251-268.

OECD (2011) *Doing Better for Families*, OECD Publishing.

## Table 1: OLS and 2LSL estimates of father's income

### A: All fathers

| Covariates | 2SLS | | | | OLS | |
|---|---|---|---|---|---|---|
| | 1st stage [Presence of a third child] | | 2nd stage [Log of father's incme] | | [Log of father's incme] | |
| | Coefficient | Robust standard error | Coefficient | Robust standard error | Coefficient | Robust standard error |
| Presence of a third child | | | 0.131 | 0.149 | -0.013 | 0.007 # |
| Same sex (IV) | 0.042 | 0.007 ** | | | | |
| N | 17,724 | | 17,724 | | 17,724 | |
| R-squared | 0.090 | | 0.307 | | 0.323 | |
| Adjusted R-squared | 0.088 | | | | | |
| F-statistics | 54.83 (37, 17,686) | | | | 273.7 (37, 17,686) | |
| Tests of endogeneity | | | | | | |
| Ho: variables are exogenous | | | | | | |
| Robust score chi2(1) | | | 0.963 (p=0.3266) | | | |
| Robust regression F-statistics(1, 17,685) | | | 0.960 (p=0.3271) | | | |

** p<0.01 * p<0.05 # p<0.1

Other covariates are year from the second child birth, region, survey round, father's age, father's age-squared, father's education, father's employment status, father's occupation, father's firm size, and mother's working status.

### B: Low-educated fathers

| Covariates | 2SLS | | | | OLS | |
|---|---|---|---|---|---|---|
| | 1st stage [Presence of a third child] | | 2nd stage [Log of father's incme] | | [Log of father's incme] | |
| | Coefficient | Robust standard error | Coefficient | Robust standard error | Coefficient | Robust standard error |
| Presence of a third child | | | 0.446 | 0.263 # | -0.016 | 0.010 |
| Same sex (IV) | 0.037 | 0.009 ** | | | | |
| N | 9,761 | | 9,761 | | 9,761 | |
| R-squared | 0.082 | | 0.085 | | 0.257 | |
| Adjusted R-squared | 0.078 | | | | | |
| F-statistics | 29.84 (34, 9,726) | | | | 114.49 (34, 9,726) | |
| Tests of endogeneity | | | | | | |
| Ho: variables are exogenous | | | | | | |
| Robust score chi2(1) | | | 3.799 (p=0.0513) | | | |
| Robust regression F-statistics(1, 9,725) | | | 3.787 (p=0.0517) | | | |

** p<0.01 * p<0.05 # p<0.1

Other covariates are year from the second child birth, region, survey round, father's age, father's age-squared, father's education, father's employment status, father's occupation, father's firm size, and mother's working status.

### C: Highly-educated fathers

| Covariates | 2SLS | | | | OLS | |
|---|---|---|---|---|---|---|
| | 1st stage [Presence of a third child] | | 2nd stage [Log of father's incme] | | [Log of father's incme] | |
| | Coefficient | Robust standard error | Coefficient | Robust standard error | Coefficient | Robust standard error |
| Presence of a third child | | | -0.199 | 0.186 | -0.014 | 0.010 |
| Same sex (IV) | 0.047 | 0.010 ** | | | | |
| N | 7,963 | | 7,963 | | 7,963 | |
| R-squared | 0.106 | | 0.298 | | 0.327 | |
| Adjusted R-squared | 0.102 | | | | | |
| F-statistics | 30.52 (35, 7,927) | | | | 132.14 (35, 7,927) | |
| Tests of endogeneity | | | | | | |
| Ho: variables are exogenous | | | | | | |
| Robust score chi2(1) | | | 1.026 (p=0.3112) | | | |
| Robust regression F-statistics(1, 7,926) | | | 1.021 (p=0.3123) | | | |

** p<0.01 * p<0.05 # p<0.1

Other covariates are year from the second child birth, region, survey round, father's age, father's age-squared, father's education, father's employment status, father's occupation, father's firm size, and mother's working status.