Fertility, Child Gender, and Parental Migration Decision: Evidence from One Child Policy in China

Zibin Huang [1] Lin Lin [2] Junsen Zhang [2]

[1]University of Rochester

[2]The Chinese University of Hong Kong

October 30, 2020

Introduction: Motivation

- One Child Policy
- Huge change in China's demography
  Low population growth + Sex ratio imbalance
- Our question: how this change in demography affects family decisions? (migration decision, labor supply, divorce etc.)

Introduction: Motivation

Ebenstein(2010)

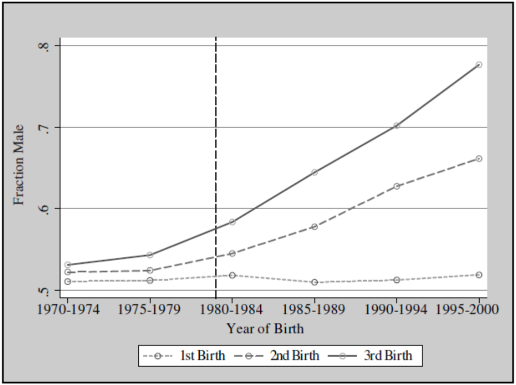

Figure 1
Male Fraction of Births Following Daughters in China
Source: China census 1982–2000. Sample restricted to mothers ages 21–40. See technical appendix at author's website for details. Vertical line indicates year of introduction of China's One Child Policy (1979).

Introduction: Literature Review

- Children number effect on family decisions: Oliveira(2016), Mincer(1978), Gemici(2006), Jena(2006), Conley(2004), Lee(2003), Adda, Dustmann and Stevens(2017)
- Children gender effect on family decisions: Dahl and Moretti(2008), Behrman(1988), Ichino, Lindstrom and Viviano(2014), Li and Yi(2015)
  Exclusion restriction problem: IV can affect both number and gender of children
- One Child Policy: Garcia(2017), Ebenstein(2010), Chen, Li and Meng(2013), Wei and Zhang(2010), Huang(2016), Huang, Lei and Zhao(2016), Wang, Guo and Zhang(2007), Wang(2005), Qian(2009), Rosenzweig and Zhang(2009), Li and Zhang(2016)
  None considers number and gender effect at the same time.

Introduction: Main Results and Contribution

- We establish a semi-parametric method to decompose number effect and gender effect of OCP on family decisions
- Applying to migration, we show that both number of children and boys will increase migration probability at a similar magnitude
- A misuse of IV method and a misinterpretation of the reduced form One Child Policy effect may mask the real effects

Introduction: Main Results and Contribution

- This is the first attempt to deal with the identification difficulty when both the number and the gender of children will affect the family decision
- This is the first attempt to inspect OCP's effect on parental migration
- This is the first study departing from the traditional linear model and discussing OCP effect in a heterogenous treatment effect framework

Data

- China National Population Census 2010 (Census 2010)
- Household level survey over the Mainland China (Excluding MC, HK and TW)
  0.35% random sample of the households
- Only consider Rural Household Heads' migration

Traditional IV: Identification

Consider a traditional OLS first:

$$y_{il} = {}_0 + NumChild_{il}\ {}_1 + X_{il}\ {}_4 + C_{il}\ {}_5 + {}_l + u_{il}$$

i: household; l: home location
y_{il} is migration choice; $NumChild_{il}$ is the number of children; X_{il} is the household control; C_{il} is the male/female head control; ${}_l$ is the original/home location dummy.

Traditional IV: IV Identification

$$y_{il} = \ {}_0 + NumChild_{il}\ {}_1 + X_{il}\ {}_4 + C_{il}\ {}_5 + \ {}_l + u_{il}$$

- Endogeneity of fertility choice
- Migrants are those more open to modern values, so less likely to have more children.
- Correlation between unobserved migration preference with fertility decision.

Traditional IV: Identification

- IV: First child gender. (1.5 Child Policy)

	First Boy	First Girl
Second Birth		p

- First gender (1) Random; (2) Affect number of children.
- Area offering the 1.5 Child Policy to all rural people: 19 out of 31 provinces, accounting for 71.3% of GDP and 74.9% population.

Traditional IV: Results

Table: First Stage Results for Traditional IV

Dependent Variable	numchi (Number of Children in HH)		
Method	OLS	OLS	OLS
fcsex(=1 if girl)	0.465***	0.411***	0.417***
	(0.0155)	(0.0114)	(0.0112)
Household Controls	No	Yes	Yes
Head Personal Controls	No	Yes	Yes
Home Place Fixed Effect	No	No	Yes
Prob>F	0.000	0.000	0.000
Number of Observations	64095	64095	64095

Traditional IV: Results

Table: IV Results for Migration

Dependent Variable	Both	Male	Female
Method	2SLS	2SLS	2SLS
numchi	-0.00499	0.00607	-0.00861
	(0.00595)	(0.00658)	(0.00691)
Household Controls	Yes	Yes	Yes
Head Personal Controls	Yes	Yes	Yes
Home Place Fixed Effect	Yes	Yes	Yes
Number of Observations	64095	64095	64095

Traditional IV: IV Results

- Tiny effect, not significant, undetermined signs
- But what if children gender will affect migration decision?
  First child as girl)
  (1) ″ Number of children) ? Migration
  (2) # Number of boys) ? Migration
- Traditional IV identification will fail!
- The reduced form effect will be a composition of number effect and gender effect.

Number and Gender Effect: Identification without Gender Selection

- Consider a non-parametric migration and fertility decision model as follows:

$$y_i = f(c_i; b_i; X_i; u_i)$$
$$c_i = g(z_i; X_i; e_i)$$

y_i is the family decision (migration choice, labor supply etc.); c_i is the number of children; b_i is the number of boys; z_i is the gender of first child; u_i is the unobservable in migration decision, u_i ? z_i; e_i is the unobservable in fertility decision, e_i ? z_i; $\text{Cov}(u_i, e_i) \neq 0$ and $\text{Cov}(c_i, u_i) \neq 0$.

Number and Gender Effect: Identification without Gender Selection

Define three types of household(people)

- Alway takers: $c_i(z_i = 1) = c_i(z_i = 0) = 2$ Always Two Children
- Never takers: $c_i(z_i = 1) = c_i(z_i = 0) = 1$ Always One Child
- Compliers: $c_i(z_i = 1) = 2; c_i(z_i = 0) = 1$ Two Children if First is Girl

Number and Gender E ect: Identi cation without Gende
Selection

(z,b,c) determines a speci c node

Number and Gender Effect: Identification without Gende Selection

- Denote $y_{zbc} = E[y_i | (z_i = z; b_i = b; c_i = c)]$, $y_{icb} = f(c_i = c; b_i = b; X_i; u_i)$. y_{zbc} is the conditional expectation of outcome at a specific node, y_{icb} is the potential family decision (migration decision) for household i if it has c children and b boys.
- y_{zbc} known from the data
- y_{icb} unknown potential outcomes need to be identified
- Random gender $E[y_{icb} | z_i; b_i; c_i; Type_i] = E[y_{icb} | Type_i]$

Number and Gender Effect: Identification without Gende Selection

Treatment Effects we want:

$$
\begin{aligned}
FLATE &= E[y_{i21}|Co] - E[y_{i11}|Co] \\
BLATE_{A1} &= E[y_{i22}|A] - E[y_{i21}|A] \\
BLATE_{A2} &= E[y_{i21}|A] - E[y_{i20}|A] \\
BLATE_{Co} &= E[y_{i21}|Co] - E[y_{i20}|Co] \\
BLATE_{N} &= E[y_{i11}|N] - E[y_{i10}|N]
\end{aligned}
$$

FLATE: Number of children effect; BLATE: Number of boy effect

Number and Gender E ect: Identi cation without Gende Selection

(z,b,c) determines a speci c node

Number and Gender E ect: Identi cation without Gende Selection

- Households with same type are comparable across nodes.

$$
\begin{aligned}
E[y_i \mid z_i = 0; b_i = 1; c_i = 2] &= E[y_{i21} \mid A] \\
E[y_i \mid z_i = 0; b_i = 2; c_i = 2] &= E[y_{i22} \mid A] \\
E[y_i \mid z_i = 0; b_i = 1; c_i = 1] &= P(N \mid z_i = 0; b_i = 1; c_i = 1)\, E[y_{i11} \mid N] \\
&\quad + P(Co \mid z_i = 0; b_i = 1; c_i = 1)\, E[y_{i11} \mid Co] \\
E[y_i \mid z_i = 1; b_i = 0; c_i = 1] &= E[y_{i10} \mid N] \\
E[y_i \mid z_i = 1; b_i = 0; c_i = 2] &= P(A \mid z_i = 1; b_i = 0; c_i = 2)\, E[y_{i20} \mid A] \\
&\quad + P(Co \mid z_i = 1; b_i = 0; c_i = 2)\, E[y_{i20} \mid Co] \\
E[y_i \mid z_i = 1; b_i = 1; c_i = 2] &= P(A \mid z_i = 1; b_i = 1; c_i = 2)\, E[y_{i21} \mid A] \\
&\quad + P(Co \mid z_i = 1; b_i = 1; c_i = 2)\, E[y_{i21} \mid Co]
\end{aligned}
$$

Number and Gender Effect: Identification without Gender Selection

Assumption 1

The expectation of boy's partial effect on migration decision is the same for from no boy to one boy, and from one boy to two boys, conditional on always-taker group. That is,
$E[y_{i22}|A] - E[y_{i21}|A] = E[y_{i21}|A] - E[y_{i20}|A]$.

- Some linearity of boy effect for always-taker

Number and Gender Effect: Identification without Gende Selection

Assumption 2

The expectation of boy's partial effect on migration decision is the same conditional on never-taker and always-taker group. That is, $E[y_{11}|N] - E[y_{10}|N] = E[y_{22}|A] - E[y_{21}|A]$.

- Some homogeneity of boy effect for always-taker and never-taker

Number and Gender E ect: Identi cation without Gende Selection

Proposition 1

If Assumption 1 and Assumption 2 hold, we can identify Local Average Treatment E ect of an additional child as:

$$FLATE = \frac{(n_A + n_C)y_{112} \quad (n_A + n_N)y_{012} \quad (n_N + n_C)y_{011} + n_N(y_{022} + y_{101})}{n_C}$$

Proposition 2

We can identify boys' premium in a ecting the family decision for di erent types of people as follows:

(1) $BLATE_{AT} = y_{022} \quad y_{012}$ for Always-Taker

(2) If Assumption 1 holds,

$BLATE_{Co} = \frac{(n_A + n_C)(y_{112} \quad y_{102}) \quad n_A(y_{022} \quad y_{012})}{n_C}$ for Compliers

(3) If both Assumption 1 and Assumption 2 hold,

$BLATE_N = BLATE_{AT} = y_{022} \quad y_{012}$ for Never-Taker

Number and Gender E ect: Identi cation with Gender Selection

Number and Gender E ect: Identi cation with Gender Selection

- Additional identi cation di culty 1: Process of gender selection is unobservable and not random.
- Need assumption for gender selection:
  (1) No gender selection for Always takers with rst child as boy;
  (2) Gender selection is exogenous to the family decision after controlling for X. That is, no sorting after controlling for X.

Number and Gender E ect: Identi cation with Gender Selection

- Assume that the gender of second child when rst is girl will be determined by:

$$1(\text{Boy})_i = h(X_i;\ \ _i)$$

On left hand side the indicator function equals to one if the second birth of household i is a boy; on the right hand side X is observed controls;   is unobserved factors to selection "success" such as the luck of not being spotted by the government, or the willingness to bear the abortion cost.

Number and Gender E ect: Identi cation with Gender Selection

Assumption 3

Always-takers do not implement gender selection when they have a rst-born boy, that is, $P_{A_0} = P_0$. Compliers and always-takers may implement gender-based abortion at the second birth if their rst child is a girl. However after controlling for observables, conditional on types, their sucess of gender selection depends on exogenous factors (to migration decision) that are independent of u_i, i.e., $_i$? u_ijA and $_i$? u_ijC.

- Exogeneous selection after controlling X in terms of the family decision (Selection on observables)
- More relaxed than traditional selection on observable assumption
- Discussion about the failure of A3: Strengthen our conclusion

Number and Gender E ect: Identi cation with Gender Selection

Number and Gender E ect: Identi cation with Gender Selection

- Additional identi cation di culty 2: Need to identify P_{A_1}
- Method: Matching or Set Identi cation
- Matching:
  (1) Use some observable X_A, identify probability of being an always-taker, i.e. $P(A|X_A)$, using the upper part of the tree.
  (2) Assign $P(A|X_A)$ to each household at (z=1,b=1,c=2) and (z=1,b=0,c=2). Sum them up and take the proportion.
  (3) Divide predicted number of Always takers at (1,1,2) by predicted number of Always takers at (1,1,2) and (1,0,2).

Number and Gender E ect: Identi cation with Gender Selection

Assumption 4

We can observe a set of predetermined features X_A such that b will not provide with more information about being an always-taker after conditioning on it, for households with rst girl and two children, (z=1,c=2). That is,
$P(A|X_A; b; c = 2; z = 1) = P(A|X_A; z = 0)$.

- Assume that we have a good predictor set
- Also try some set identi cation (bound the P_A)

Number and Gender E ect: Identi cation with Gender Selection

Proposition 3

If Assumption 1, 2, 3 and 4 hold simultaneously, we can identify Local Average Treatment E ect of an additional child as:

$$FLATE_x$$

$$= \frac{(P_{A_1 x} n_{Ax} + P_{Cx} n_{Cx}) y_{112x} \quad (P_{A_1 x} n_{Ax} + P_{Cx} n_{Nx}) y_{012x} \quad P_{Cx}(n_{Nx} + n_{Cx}) y_{011x} + P_{Cx} n_{Nx}(y_{022x} + y_{101x})}{P_{Cx} n_{Cx}}$$

$$FLATE = \int_X FLATE_x P(X j C) dx$$

Proposition 4

We can identify boys' premium in a ecting the family decision for di erent types of people as follows:

(1) $BLATE_{AT} = y_{022} \quad y_{012}$

(2) $BLATE_{Co} = \int_X [\frac{y_{112x}(P_{A_1 x} n_{Ax} + P_{Cx} n_{Cx}) \quad P_{A_1 x} n_{Ax} y_{012x}}{P_{Cx} n_{Cx}}$
$\frac{[(1 \quad P_{A_1 x}) n_{Ax} + (1 \quad P_{Cx}) n_{Cx}] y_{102x} \quad (1 \quad P_{A_1 x}) n_{Ax} (2 y_{012x} \quad y_{022x})}{(1 \quad P_{Cx}) n_{Cx}}] P(X j C) dx$

(3) $BLATE_N = BLATE_{AT} = y_{022} \quad y_{012}$

Number and Gender E ect: Identi cation with Gender Selection

- Another challenge: With a lot of candidate Xs at hand, nding a good algorithm to estimate $P(A|X_A)$.
- Machine Learning kicks in: L-1 Regularized Logistic Model

Number and Gender Effect: Identification with Gender Selection

- L-1 Regularized Logistic Model

$$max \sum ln[\frac{e^{x_i^0 \beta}}{1 + e^{x_i^0 \beta}}] \quad \frac{1}{2} \; k \;\; k_1$$

- Shrink the coefficient to avoid overfitting
- Filtering X_A to get those with highest prediction power
- Choosing tunning parameter   by 10 folds cross-validation

Number and Gender Effect: Identification with Gender Selection

Table: Results with Estimated Selection Probability

Dependent Variable	Both	Male	Female
Method	Bootstrap	Bootstrap	Bootstrap
FLATE	0.0133	0.0231*	0.00248
	(0.0103)	(0.0134)	(0.0113)
BLATE	0.0103***	0.00928*	0.00864**
	(0.00369)	(0.00504)	(0.00406)
Number of Observations	56794	56794	56794

Relative to the mean migration rate, 13.71%, 10.62% for both.

Number and Gender Effect: Identification with Gender Selection

Compared with Traditional IV

- Boy effect cannot be ignored!
- Positive and much larger point estimation for number effect
- Traditional IV is contaminated, two effects offset each other
  First child as girl rather than boy)
  (1) ″ Number of children) ″ Migration
  (2) # Number of boys) # Migration

Conclusion

- We construct a semi-parametric method to decompose number effect and gender effect of children on family decisions, and apply it to migration decision in China
- Number of children has positive effect on migration decision. One more child will result in 1.33 percentage points increase in probability of having both male or female head migrate
- The boy effect is at the same level. One more boy will result in 1.03 percentage points increase in probability of having both male or female head migrate
- A warning to previous and future research: Never ignore gender effects!