Online Appendix for
"Rationed Fertility: Treatment Effect Heterogeneity in the Child Quantity-Quality Tradeoff"

Rufei Guo* Junjian Yi† Junsen Zhang‡

June 15, 2020

*Economics and Management School, Wuhan University. Email: rufei_guo@whu.edu.cn.
†Department of Economics, National University of Singapore. Email: junjian@nus.edu.sg.
‡School of Economics, Zhejiang University; Department of Economics, Chinese University of Hong Kong. Email: jszhang@cuhk.edu.hk.

Contents

A1 Full Derivation of the Theory

In this section, we build on Rosenzweig and Wolpin's (1980) model using rationing theory in Neary and Roberts (1980), and formulate a general theory of *exogenous* fertility change from a truly rationed level. We first present and solve a model of rationed fertility. We then derive the effect of rationed fertility on child quality, and follow up with a simulation analysis.

A1.1 The Becker-Lewis Setup

We first consider the three-commodity interactive (q^2) model, as in Becker and Lewis (1973) and Rosenzweig and Wolpin (1980). Parents maximize utility by choosing fertility or child quantity (n), child quality (q), and a composite consumption good (s),

$$\begin{aligned} \max_{n,q,s} \quad & U(n,q,s), \\ \text{subject to} \quad & \pi_n n + \pi_q q + \pi_{nq} nq + p_s s \leq y, \end{aligned} \tag{P1}$$

where y is the monetary income of the family. The price of child quality, i.e., the cost of increasing q by one unit, is

$$p_q = \pi_q + \pi_{nq} n,$$

where π_q represents the costs of "public goods" or "family goods," such as "some aspects of training in the home and the 'handing down' of some clothing" (Becker and Lewis, 1973, p. S283); $\pi_{nq} n$ represents the costs of "private goods," such as tuition fees and health care expenditure, which increases in n. Similarly, the price of child quantity, i.e., the cost of an additional child, is $p_q = \pi_n + \pi_{nq} q$, where π_n represents the "fixed cost" of an additional child, including the cost of giving birth and any necessities to keep the child alive; $\pi_{nq} q$ represents the cost of private goods, which increases in child quality; p_s is the price of the composite good. Solving P1 gives the optimal child quantity (n^o), child quality (q^o), and composite good (s^o).

We define the shadow prices of child quantity (v_n) and quality (v_q), respectively,

$$v_n = \frac{\partial U}{\partial n} / \lambda,$$
$$v_q = \frac{\partial U}{\partial q} / \lambda,$$

where λ is the Lagrangian multiplier of P1, representing the marginal utility of 1 dollar. So v_n and v_q are the marginal utilities of child quantity and quality measured by dollars. We then define the family "total income" in the Beckerian sense, as follows (Becker, 1991):

$$W = v_n n + v_q q + v_s s.$$

When $n = n^o$, $q = q^o$, and $s = s^o$, then $v_n = p_n$, $v_q = p_q$, and $v_s = p_s$; the family achieves the highest income, W^o.

To simplify notation, let $x = (n, q, s)$, $\pi = (\pi_{nq}, \pi_n, \pi_q, p_s)$, $p = (p_n, p_q, p_s)$, and $v = (v_n, v_q, v_s)$. Note that the shadow prices v are the monetary-equivalent utility values of n, q, and s, and the "actual" prices p are the monetary costs of n, q, and s. When there is no restriction on the choice of x, in equilibrium the shadow prices must equal the actual prices ($p = v$).[1]

If $\pi_{nq} > 0$, child quality enters the price of child quantity, and vice versa (Becker and Lewis, 1973). If $\pi_{nq} = 0$, the q^2 model is reduced to the standard non-interactive (q^1) model.

[1] Becker and Lewis (1973) and Rosenzweig and Wolpin (1980) conduct the comparative static analysis only at the unrestricted optimal point, so they do not distinguish p from v. To derive the implication of rationed fertility in the case of p not equal to v, a comparative static analysis off the unrestricted optimal point is necessary. Hence we call p the "actual" price of x to distinguish from the shadow price v of x.

We defined the indirect utility function and uncompensated demand functions as

$$V(\pi, y) = \max_{n,q,s}\{U(n,q,s)|\pi_{nq}nq + \pi_n n + \pi_q q + p_s s \leq y\},$$
$$x(\pi, y) = \arg\max_{n,q,s}\{U(n,q,s)|\pi_{nq}nq + \pi_n n + \pi_q q + p_s s \leq y\},$$

where $x(\pi, y) = (n(\pi, y), q(\pi, y), s(\pi, y))$.

The dual problem of the utility maximization problem in the q^2 model is the expenditure minimization problem,

$$\min_{n,q,s} \quad \pi_{nq}nq + \pi_n n + \pi_q q + p_s s,$$
$$\text{s.t} \quad U(n,q,s) \geq u.$$

We define the expenditure function and compensated demand functions as

$$e(\pi, u) = \min_{n,q,s}\{\pi_{nq}nq + \pi_n n + \pi_q q + p_s s|U(n,q,s) \geq u\},$$
$$x^c(\pi, u) = \arg\min_{n,q,s}\{\pi_{nq}nq + \pi_n n + \pi_q q + p_s s|U(n,q,s) \geq u\},$$

where $x^c(\pi, u) = (n^c(\pi, u), q^c(\pi, u), s^c(\pi, u))$. Throughout the paper, we use superscript "c" to denote compensated demand functions (in the Hicksian sense) in the expenditure minimization problem.

By the duality theorem, $x(\pi, y) = x^c(\pi, V(\pi, y))$. Specifically, we define the optimal fertility level $n^o \equiv n(\pi, y) = n^c(\pi, V(\pi, y))$.

A1.2 Rationed Fertility

We build on the rationing theory of Tobin and Houthakker (1950) and Neary and Roberts (1980) to extend Rosenzweig and Wolpin (1980), and derive a generalized comparative static analysis of an exogenous fertility change. As in Becker and Lewis (1973), both n

and q are choice variables, so a direct comparative static analysis of the effect of n on q is not feasible. Rather, we focus primarily on the effect of rationed fertility.

To begin with, we fix n at $n = \bar{n}$. We call the q^2 (or q^1) model with fixed fertility ($n = \bar{n}$) the restricted q^2 (or q^1) model. The utility maximization problem in the restricted q^2 model is

$$\begin{aligned} \max_{q,s} \quad & U(\bar{n}, q, s), \\ \text{subject to} \quad & \pi_n \bar{n} + \pi_q q + \pi_{nq} \bar{n} q + p_s s \leq y, \end{aligned} \tag{P2}$$

where $\bar{n}$ is the number of children mandated by some exogenous force. Note that the restriction on or rationing of fertility makes $\bar{n}$ a parameter rather than a choice variable. We define the indirect utility function and uncompensated demand functions in the restricted q^2 model as

$$\tilde{V}(\pi, y, \bar{n}) = \max_{q,s}\{U(\bar{n}, q, s) | \pi_n \bar{n} + (\pi_{nq}\bar{n} + \pi_q) q + p_s s \leq y\},$$
$$\tilde{x}(\pi, y, \bar{n}) = \arg\max_{q,s}\{U(\bar{n}, q, s) | \pi_n \bar{n} + (\pi_{nq}\bar{n} + \pi_q) q + p_s s \leq y\},$$

where a tilde "$\sim$" denotes functions in the restricted model.

The expenditure minimization problem in the restricted q^2 model is

$$\begin{aligned} \min_{q,s} \quad & \pi_n \bar{n} + (\pi_{nq}\bar{n} + \pi_q) q + p_s s, \\ \text{subject to} \quad & U(\bar{n}, q, s) \geq u. \end{aligned}$$

The corresponding expenditure function and compensated demand functions are

$$\tilde{e}(\pi, u, \bar{n}) = \min_{q,s}\{\pi_n \bar{n} + (\pi_{nq}\bar{n} + \pi_q) q + p_s s | U(\bar{n}, q, s) \geq u\},$$
$$\tilde{x^c}(\pi, u, \bar{n}) = \arg\min_{q,s}\{\pi_n \bar{n} + (\pi_{nq}\bar{n} + \pi_q) q + p_s s | U(\bar{n}, q, s) \geq u\}.$$

To derive testable implications, we carry out the analysis in four steps. First, we link the restricted q^2 model and the unrestricted q^2 model. Second, we define two types of fertility changes. Third, we link the restricted q^2 model and the restricted q^1 model. Finally, we decompose the derivatives of q and s with respect to $\bar{n}$ in the restricted q^2 model into the derivatives of q, s, and n with respect to prices in the unrestricted q^1 model.[2]

In the restricted q^2 model, the shadow price of child quantity does not equal p_n if $\bar{n} \neq n^o$. Specifically, if $\bar{n} < n^o$, parents prefer more children, so the shadow price of child quantity is higher than p_n; if $\bar{n} > n^o$, parents prefer fewer children, so the shadow price is lower than p_n. We adjust the fixed price π_n (as a component of p_n) to equate p_n with the shadow price of n, inducing parents to choose $n = \bar{n}$ in the *unrestricted* q^2 model. The *supporting fixed price* $\bar{\pi}_n$ is defined by

$$\bar{n} = n^c(\pi_{-n}, \bar{\pi}_n, u), \tag{A1.1}$$

where $\pi_{-n} = (\pi_{nq}, \pi_q, p_s)$, and $n^c(\pi_{-n}, \bar{\pi}_n, u)$ is the compensated demand function of n in the unrestricted q^2 model (P2). The supporting fixed price $\bar{\pi}_n$ causes parents to choose $n = \bar{n}$ in the expenditure minimization problem in the unrestricted q^2 model. In the following analysis, functions derived from the unrestricted q^2 model are always evaluated at $(\pi_{-n}, \bar{\pi}_n, u)$.

Because the first-order conditions of q and s in the expenditure minimization problem in the restricted q^2 model and those in the unrestricted q^2 model (when $\pi = \bar{\pi}_n$) are the same, we have

$$\tilde{q^c}(\pi, u, \bar{n}) = q^c(\pi_{-n}, \bar{\pi}_n, u), \tag{A1.2}$$

$$\tilde{s^c}(\pi, u, \bar{n}) = s^c(\pi_{-n}, \bar{\pi}_n, u). \tag{A1.3}$$

[2] The decomposition of derivatives in restricted models into derivatives in standard models helps deliver testable implications. This approach was pioneered by Becker and Lewis (1973), who decompose income and price elasticities of q and n in the q^2 model, which they call "observed" elasticities, into income and price elasticities of q and n in the q^1 model, which they call "true" elasticities.

By Eqs. (A1.1), (A1.2), and (A1.3), we connect the expenditure functions in the restricted q^2 model and those in the unrestricted q^2 model,

$$\tilde{e}(\pi, u, \bar{n}) = e(\pi_{-n}, \bar{\pi}_n, u) + (\pi_n - \bar{\pi}_n)\bar{n}. \tag{A1.4}$$

Differentiating Eq. (A1.4) with respect to $\bar{n}$, we have

$$\begin{aligned} \frac{\partial \tilde{e}}{\partial \bar{n}} &= \left(\frac{\partial e}{\partial \pi_n} - \bar{n} \right) \frac{\partial \bar{\pi}_n}{\partial \bar{n}} + (\pi_n - \bar{\pi}_n) \\ &= \pi_n - \bar{\pi}_n. \end{aligned} \tag{A1.5}$$

The second equality in Eq. (A1.5) holds because (i) $\frac{\partial e}{\partial \pi_n} = n^c$ by the envelope theorem (Shephard's lemma) and (ii) $n^c = \bar{n}$ by Eq. (A1.1).

Similarly, the indirect utility function in the restricted q^2 model is connected with that in the unrestricted q^2 model,

$$\tilde{V}(\pi, y, \bar{n}) = V(\pi_{-n}, \bar{\pi}_n, y + (\bar{\pi}_n - \pi_n)\bar{n}). \tag{A1.6}$$

Differentiating Eq. (A1.6) with respect to $\bar{n}$, we have

$$\begin{aligned} \frac{\partial \tilde{V}}{\partial \bar{n}} &= \left(\bar{n} + \frac{\partial V / \partial \bar{\pi}_n}{\partial V / \partial y} \right) \frac{\partial \bar{\pi}_n}{\partial \bar{n}} \frac{\partial V}{\partial y} + (\bar{\pi}_n - \pi_n) \frac{\partial V}{\partial y} \\ &= (\bar{\pi}_n - \pi_n) \frac{\partial V}{\partial y}. \end{aligned} \tag{A1.7}$$

The second equality in Eq. (A1.7) holds because (i) $\frac{\partial V / \partial \bar{\pi}_n}{\partial V / \partial y} = -n(\pi_{-n}, \bar{\pi}_n, y + (\bar{\pi}_n - \pi_n)\bar{n})$ by Roy's identity; (ii) $n(\pi_{-n}, \bar{\pi}_n, y + (\bar{\pi}_n - \pi_n)\bar{n}) = n^c(\pi_{-n}, \bar{\pi}_n, u)$ by the duality theorem; and (iii) $n^c = \bar{n}$ by Eq. (A1.1). We assume that the shadow price of income is always larger than zero, namely, $\frac{\partial V}{\partial y} > 0$.

A1.3 Desired and Undesired Fertility Changes

Based on Eqs. (A1.5) and (A1.7), we define desired and undesired fertility changes.

Definition A1 *When $\bar{n} < n^o$, $\bar{\pi}_n - \pi_n > 0$ and $\frac{\partial \tilde{V}}{\partial \bar{n}} > 0$, an increase in $\bar{n}$ is a **desired** fertility increase; a decrease in $\bar{n}$ is an **undesired** fertility decrease.*

Definition A2 *When $\bar{n} > n^o$, $\bar{\pi}_n - \pi_n < 0$ and $\frac{\partial \tilde{V}}{\partial \bar{n}} < 0$, an increase in $\bar{n}$ is an **undesired** fertility increase; a decrease in $\bar{n}$ is a **desired** fertility decrease.*

When fertility is rationed below the unrestricted optimum ($\bar{n} < n^o$), parents prefer more children. The shadow price of child quantity ($v_n = \bar{\pi}_n + \pi_{nq}\tilde{q^c}$) is higher than the actual price ($p_n = \pi_n + \pi_{nq}\tilde{q^c}$). As such, $\bar{\pi}_n - \pi_n = (\bar{\pi}_n + \pi_{nq}\tilde{q^c}) - (\pi_n + \pi_{nq}\tilde{q^c}) = v_n - p_n > 0$. By Eqs. (A1.5) and (A1.7), $\partial\tilde{e}/\partial\bar{n} < 0$ and $\partial\tilde{V}/\partial\bar{n} > 0$. An increase in n reduces the minimal expenditure to achieve a given utility level or raises the maximal utility attainable at a given income level. In this case, we call the fertility increase a desired one. Similarly, when fertility is rationed above the unrestricted optimum ($\bar{n} > n^o$), parents prefer fewer children. The shadow price of child quantity ($\bar{\pi}_n + \pi_{nq}\tilde{q^c}$) is lower than the actual price ($\pi_n + \pi_{nq}\tilde{q^c}$). As such, $\bar{\pi}_n - \pi_n = (\bar{\pi}_n + \pi_{nq}\tilde{q^c}) - (\pi_n + \pi_{nq}\tilde{q^c}) < 0$. By Eqs. (A1.5) and (A1.7), $\partial\tilde{e}/\partial\bar{n} > 0$ and $\partial\tilde{V}/\partial\bar{n} < 0$. An increase in n either raises the minimal expenditure necessary to achieve a given utility level or reduces the maximal utility attainable at a given income level. In this case, we call the fertility increase an undesired one.

Undesired versus desired fertility changes are illustrated in Figure 1. Desired fertility changes move fertility toward the unrestricted optimal fertility level (n^0): Fertility increases at $\bar{n} < n^o$ and reductions at $\bar{n} > n^o$ are desired. In contrast, undesired fertility changes move fertility away from n^0: Fertility increases at $\bar{n} > n^o$ and reductions at $\bar{n} < n^o$ are undesired.

A1.4 Duality and Decomposition

By the duality theorem, we link uncompensated and compensated demand functions:

$$\tilde{q}(\pi, \tilde{e}(\pi, u, \bar{n}), \bar{n}) = \tilde{q^c}(\pi, u, \bar{n}), \tag{A1.8}$$

$$\tilde{s}(\pi, \tilde{e}(\pi, u, \bar{n}), \bar{n}) = \tilde{s^c}(\pi, u, \bar{n}). \tag{A1.9}$$

Differentiating Eqs. (A1.8) and (A1.9) with respect to $\bar{n}$ and invoking Eq. (A1.5), we have

$$\frac{\partial \tilde{q}}{\partial \bar{n}} = \frac{\partial \tilde{q^c}}{\partial \bar{n}} + (\bar{\pi}_n - \pi_n)\frac{\partial \tilde{q}}{\partial y}, \tag{A1.10}$$

$$\frac{\partial \tilde{s}}{\partial \bar{n}} = \frac{\partial \tilde{s^c}}{\partial \bar{n}} + (\bar{\pi}_n - \pi_n)\frac{\partial \tilde{s}}{\partial y}. \tag{A1.11}$$

Following Becker and Lewis (1973) and Rosenzweig and Wolpin (1980), we further decompose $\frac{\partial \tilde{q}}{\partial \bar{n}}$ and $\frac{\partial \tilde{s}}{\partial \bar{n}}$ into derivatives in the q^1 model. The expenditure minimization problem in the restricted q^1 model is

$$\begin{aligned} \min_{q,s} \quad & p_n^* \bar{n} + p_q^* q + p_s^* s, \\ \text{subject to} \quad & U(\bar{n}, q, s) \geq u, \end{aligned} \tag{P3}$$

where p_n^*, p_q^*, and p_s^* denote the actual prices in the non-interactive (q^1) model. We use the superscript "$*$" to denote functions in the q^1 model. The corresponding expenditure function and compensated demand functions are

$$\tilde{e^*}(p^*, u, \bar{n}) = \min_{q,s}\{p_n^* \bar{n} + p_q^* q + p_s^* s | U(\bar{n}, q, s) \geq u\},$$

$$\tilde{x^{*c}}(p^*, u, \bar{n}) = \arg\min_{q,s}\{p_n^* \bar{n} + p_q^* q + p_s^* s | U(\bar{n}, q, s) \geq u\},$$

where $p^* = (p_n^*, p_q^*, p_s^*)$ for notational brevity.

If $p_q^* = \pi_{nq}\bar{n} + \pi_q$, $p_n^* = \pi_n$, and $p_s^* = p_s$, the expenditure minimization problem in the restricted q^2 model (P2) is equivalent to that in the restricted q^1 model (P3). Hence the compensated demands of q and s in the two models are equal:

$$\tilde{q}^c(\pi, u, \bar{n}) = \tilde{q^{*c}}(\pi^*, u, \bar{n}), \tag{A1.12}$$

$$\tilde{s}^c(\pi, u, \bar{n}) = \tilde{s^{*c}}(\pi^*, u, \bar{n}). \tag{A1.13}$$

Differentiating Eqs. (A1.12) and (A1.13) with respect to $\bar{n}$,

$$\frac{\partial \tilde{q}^c}{\partial \bar{n}} = \pi_{nq} \frac{\partial \tilde{q^{*c}}}{\partial p_q^*} + \frac{\partial \tilde{q^{*c}}}{\partial \bar{n}}, \tag{A1.14}$$

$$\frac{\partial \tilde{s}^c}{\partial \bar{n}} = \pi_{nq} \frac{\partial \tilde{s^{*c}}}{\partial p_q^*} + \frac{\partial \tilde{s^{*c}}}{\partial \bar{n}}. \tag{A1.15}$$

If we further consider the utility maximization problem in the restricted q^1 model, we have

$$\frac{\partial \tilde{q}}{\partial \bar{n}} = \pi_{nq} \frac{\partial \tilde{q^*}}{\partial p_q^*} + \frac{\partial \tilde{q^*}}{\partial \bar{n}}, \tag{A1.16}$$

$$\frac{\partial \tilde{s}}{\partial \bar{n}} = \pi_{nq} \frac{\partial \tilde{s^*}}{\partial p_q^*} + \frac{\partial \tilde{s^*}}{\partial \bar{n}}, \tag{A1.17}$$

$$\frac{\partial \tilde{q}}{\partial y} = \frac{\partial \tilde{q^*}}{\partial y}, \tag{A1.18}$$

$$\frac{\partial \tilde{s}}{\partial y} = \frac{\partial \tilde{s^*}}{\partial y}, \tag{A1.19}$$

where $\tilde{q^*}$ and $\tilde{s^*}$ are uncompensated demand functions of q and s in the restricted q^1 model when $p_q^* = \pi_{nq}\bar{n} + \pi_q$, $p_n^* = \pi_n$, $p_s^* = p_s$. Substituting Eqs. (A1.14) and (A1.18) into Eq. (A1.10), and Eqs. (A1.15) and (A1.19) into Eq. (A1.11), we have

$$\frac{\partial \tilde{q}}{\partial \bar{n}} = \pi_{nq} \frac{\partial \tilde{q^{*c}}}{\partial p_q^*} + \frac{\partial \tilde{q^{*c}}}{\partial \bar{n}} + (v_n - p_n) \frac{\partial \tilde{q^*}}{\partial y}, \tag{A1.20}$$

$$\frac{\partial \tilde{s}}{\partial \bar{n}} = \pi_{nq} \frac{\partial \tilde{s^{*c}}}{\partial p_q^*} + \frac{\partial \tilde{s^{*c}}}{\partial \bar{n}} + (v_n - p_n) \frac{\partial \tilde{s^*}}{\partial y}. \tag{A1.21}$$

Applying the Neary and Roberts's (1980, pp. 32-34) Eqs. (19), (24), and (29), we decompose derivatives of q and s with respect to $\bar{n}$ in the restricted q^1 model to derivatives in the unrestricted q^1 model,

$$\frac{\partial \tilde{q^{*c}}}{\partial \bar{n}} = \frac{\partial q^{*c}}{\partial p_n^*} \left(\frac{\partial n^{*c}}{\partial p_n^*} \right)^{-1}, \tag{A1.22}$$

$$\frac{\partial \tilde{q^{*c}}}{\partial p_q^*} = \frac{\partial q^{*c}}{\partial p_q^*} - \left(\frac{\partial n^{*c}}{\partial p_q^*} \right)^2 \left(\frac{\partial n^{*c}}{\partial p_n^*} \right)^{-1}, \tag{A1.23}$$

$$\frac{\partial \tilde{q^*}}{\partial y} = \frac{\partial q^*}{\partial y} - \frac{\partial \tilde{q^{*c}}}{\partial \bar{n}} \frac{\partial n^*}{\partial y}. \tag{A1.24}$$

As for the comparative static analysis with respect to s, we have

$$\frac{\partial \tilde{s^{*c}}}{\partial \bar{n}} = \frac{\partial s^{*c}}{\partial p_n^*} \left(\frac{\partial n^{*c}}{\partial p_n^*} \right)^{-1}, \tag{A1.25}$$

$$\frac{\partial \tilde{s^{*c}}}{\partial p_q^*} = \frac{\partial s^{*c}}{\partial p_q^*} - \frac{\partial n^{*c}}{\partial p_s^*} \frac{\partial n^{*c}}{\partial p_q^*} \left(\frac{\partial n^{*c}}{\partial p_n^*} \right)^{-1}, \tag{A1.26}$$

$$\frac{\partial \tilde{s^*}}{\partial y} = \frac{\partial s^*}{\partial y} - \frac{\partial \tilde{s^{*c}}}{\partial \bar{n}} \frac{\partial n^*}{\partial y}. \tag{A1.27}$$

Substituting Eq. (A1.24) into Eq. (A1.20), and Eq. (A1.27) into Eq. (A1.21), we have

$$\frac{\partial \tilde{q}}{\partial \bar{n}} = \pi_{nq} \frac{\partial \tilde{q^{*c}}}{\partial p_q^*} + (1 - \alpha_\Delta \epsilon_{n^* \cdot y}) \frac{\partial \tilde{q^{*c}}}{\partial \bar{n}} + (v_n - p_n) \frac{\partial q^*}{\partial y}, \tag{A1.28}$$

$$\frac{\partial \tilde{s}}{\partial \bar{n}} = \pi_{nq} \frac{\partial \tilde{s^{*c}}}{\partial p_q^*} + (1 - \alpha_\Delta \epsilon_{n^* \cdot y}) \frac{\partial \tilde{s^{*c}}}{\partial \bar{n}} + (v_n - p_n) \frac{\partial s^*}{\partial y}, \tag{A1.29}$$

where $\alpha_\Delta = \frac{(\bar{\pi}_n - \pi_n)\bar{n}}{y}$ and $\epsilon_{n^* \cdot y} = \frac{\partial n^*}{\partial y} \frac{y}{n^*}$. Note that α_Δ is the share of compensating income change,[3] $(\bar{\pi}_n - \pi_n)\bar{n}$, out of total monetary income y. $\alpha_\Delta > 0$ if child quantity is rationed below the unrestricted optimum ($\bar{n} < n^o$), and $\alpha_\Delta < 0$ if child quantity is rationed above ($\bar{n} > n^o$). The income elasticity of child quantity is given by $\epsilon_{n^* \cdot y}$. If $\bar{n} = n^o$, $\bar{\pi}_n = \pi_n$.

[3] Note that $(\bar{\pi}_n - \pi_n)\bar{n}$ appears in Eq. (A1.6). When π_n is adjusted to the supporting fixed price $\bar{\pi}_n$, income y should be adjusted to $y + (\bar{\pi}_n - \pi_n)\bar{n}$ to induce the unrestricted household to choose $\bar{n}$.

Eqs. (A1.28) and (A1.29) are reduced to Eqs. (18) and (19) in Rosenzweig and Wolpin (1980, p. 231).[4] Importantly, Eq. (A1.28) enables us to derive the effect of $\bar{n}$, i.e., rationed fertility, on q, even though a comparative static analysis of the effect of n is not possible. Eq. (A1.28) is Eq. (2) in the main text.

A1.5 Theoretical Extensions on Parental Responses

We extend the model in Section 2 to incorporate parental time allocation. Parents care about children's average human capital (quality), the number of children, their own consumption, and leisure. Parents' utility function is $u = u(h, n, c, l)$, where h is the average human capital of children, c is parental consumption, and l is parental leisure. To enhance a child's human capital h, parents can either increase child expenditure q, or allocate more time to home tutorials t, i.e. $h = h(q, t)$. We use T to denote total parental time. Hence $d = T - t - l$ represents parental labor supply.

The parents' utility maximization problem is

$$\begin{aligned} \max_{q,t,c,l} \quad & u = u(h, c, l, n), \\ \text{subject to} \quad & h = h(q, t), \\ & c + nq + \pi_n n + \pi_q q + w(t + l) = wT + y, \end{aligned} \tag{A1.30}$$

where $U(q, t, c, l, n) = u(h(q, t), c, l, n)$ is the reduced-form utility function. Here, q no longer represents child quality; instead, q represents parental investment in each child.[5]

[4] Rosenzweig and Wolpin (1980) are the first to decompose derivatives of q and s with respect to $\bar{n}$ in the restricted q^2 model into derivatives of q, s, and n with respect to prices in the unrestricted q^1 model. They note that the method of decomposing derivatives in a restricted model into derivatives in an unrestricted model is analogous to the one used in rationing theory by Tobin and Houthakker (1950), which enables them to conduct a comparative static analysis only at the unrestricted optimal fertility level ($\bar{n} = n^o$). As a generalization of Tobin and Houthakker's (1950) rationing theory, Neary and Roberts (1980) apply duality techniques to evaluate functions both at and off the unrestricted optimum. We adopt their duality techniques to extend Rosenzweig and Wolpin's (1980) analysis. In this manner, we are able to evaluate derivatives of q and s with respect to $\bar{n}$ both at and off the unrestricted optimal fertility level (i.e., $\bar{n} = n^o$ and $\bar{n} \neq n^o$), respectively.

[5] $\pi_{nq} = 1$ by construction. Child human capital h can also be multidimensional. Heckman (2007) emphasizes the importance of noncognitive skills as a form of human capital.

Meanwhile, c, t, and l are components of composite good s in the three-commodity model of Section 2.

Differentiating the budget constraint with respect to $\bar{n}$, and multiplying both sides by $\frac{1}{wT+y}$, we have

$$\epsilon_{cn} b_c + b_n + \epsilon_{qn} b_q + \epsilon_{(t+l)n} b_{t+l} = 0, \tag{A1.31}$$

where $\epsilon_{kn} = \frac{\partial k}{\partial n} \frac{1}{k}$, $b_k = \frac{p_k k}{wT+y}$, $\forall k = c, n, q, t+l$. Note that $p_c = 1$, $p_n = q + \pi_n$, $p_q = n + \pi_q$, and $p_{t+l} = w$; ϵ_{kn} is the semi-elasticity of k with respect to $\bar{n}$, $\forall k = c, q, t+l$; and b_k is the budget share of k, $\forall k = c, n, q, t+l$.[6]

If $\epsilon_{qn} = 0$, then $\epsilon_{cn} b_c + \epsilon_{(t+l)n} b_{t+l} = -b_n < 0$. Then either $\epsilon_{cn} < 0$ or $\epsilon_{(t+l)n} < 0$.[7] The theory implies that in response to an exogenous fertility increase, parents will either reduce self-consumption or increase labor supply, or both, in order to maintain expenditure per child.

A1.6 A Parametric Simulation on Parental Consumption

We use the parametric model in Section 2.4 to simulate parental consumption response to rationed fertility. Figure A1a depicts optimal parental consumption s against rationed fertility $\bar{n}$ for $\bar{n} \in [1, 5]$. Note that s monotonically declines as $\bar{n}$ increases from $\bar{n} = 1$ to $\bar{n} = 5$. The decline of s is stronger at lower $\bar{n}$ than at higher $\bar{n}$. A higher $\bar{n}$ uses up a larger portion of parental income, leaving less leeway for parents to reduce s.

As in Section 2.4, we decompose the marginal effects of rationed fertility $\bar{n}$ on parental consumption s using Eq. (A1.29). We use the semi-elasticity form to ensure that the results are insensitive to the unit of measurement of s. Figure A1b shows that the total effect is decomposed into three parts. The price effect is positive, and it remains sizable for the whole range of $\bar{n}$. The substitution effect is negative; its magnitude declines with $\bar{n}$ and approaches zero for large $\bar{n}$. The rationing income effect is cancelled out by the

[6] $\epsilon_{(t+l)n} = \frac{b_t}{b_{t+l}} \epsilon_{tn} + \frac{b_l}{b_{t+l}} \epsilon_{ln}$. Note that $\epsilon_{ky} = \frac{\partial k}{\partial y} \frac{y}{k}$ is the income elasticity of k.

[7] $\epsilon_{(t+l)n} < 0$ means $\epsilon_{dn} > 0$ since $d = T - t - l$.

rationing price effect for large $\bar{n}$. Overall, the total effect appears to follow the path of the substitution effect.

A2 Full Derivation of the Econometrics

A2.1 Setup

For purposes of illustration, we assume a pool of three types of mothers ($i \in A, B, C$), as shown in Table 1. The shares of the three types of mothers are P_A, P_B, and P_C (column (5)), which are unobservable to researchers. Without twinning at the second birth ($Z_i = 0$) and the birth-control policy ($X_i = 0$), the three types of mothers can achieve their optimal fertility levels in P1, which are two, three, and three (column (1)).

We consider the differences between twinning-induced fertility increases without and with the policy. Without the policy ($X_i = 0$), twinning at the second birth shifts the realized fertility of type-A mothers from two to three, but does not affect the realized fertility of type-B and type-C mothers (column (2)). In this case, type-A mothers are "compliers" of twinning; type-B and type-C mothers are "always-takers" of twinning.[8]

We assume that the birth-control policy targets two children per family. Because type-A mothers' optimal fertility is two, the policy does not affect the realized fertility of type-A mothers (column (3)). The policy rations fertility of type-B mothers, reducing type-B mothers' fertility from three to two. Type-C mothers do not comply with the policy, and realize three children regardless of the policy.

Under the policy ($X_i = 1$), twinning shifts the realized fertility from two to three for both type-A and type-B mothers (column (4)). In this case, both type-A and type-B mothers are compliers of twinning; type-C mothers remain always takers of twinning.

In this simple scenario, we observe two types of fertility changes. Without the policy, the twinning shifts fertility of type-A mothers from the optimal two to an undesired three (the arrow between columns (1) and (2) in Table 1). Black, Devereux, and Salvanes (2010) consider twinning to be an "unexpected" and "unplanned" shock to fertility. Similarly,

[8] In the terminology of Imbens and Angrist (1994) and Angrist, Imbens, and Rubin (1996), compliers are individuals whose treatment status is affected by the instrument, and always-takers are individuals who are treated irrespective of whether the instrument is switched on or off.

Mogstad and Wiswall (2016, p. 173) conclude that "twin births increase the number of siblings beyond the desired family size." Thus, the twinning-induced fertility increase for type-A mothers is undesired. By contrast, because the policy rations the fertility level of type-B mothers at two, twinning helps type-B mothers circumvent the rationing and achieve the optimal fertility of three, which represents a desired fertility increase (the arrow between columns (3) and (4) in Table 1).[9]

A2.2 Assumptions

We adopt the framework from the literature on the treatment effect. The indicator of three children, D_i, is the realized treatment status of mother i ($D_i = 0, 1$). Denote D_{zi} as the potential treatment status of mother i when $Z_i = z$ ($z = 0, 1$). We have $D_i = D_{0i} + (D_{1i} - D_{0i}) \cdot Z_i$. As Y_i is the realized outcome, we further denote $Y_i(d, z)$ as the potential outcome when $D_i = d$ and $Z_i = z$.

We make the three standard assumptions on the independence, exclusion, and monotonicity of twinning (Z_i).

Assumption 1 *Conditional on policy exposure X_i, twinning status Z_i is independent of the potential outcomes.* $\{Y_i(d, z), D_{zi}\}_{\forall z \in \{0,1\}, \forall d \in \{0,1\}} \perp Z_i \Big|_{X_i}$.

Assumption 2 *Conditional on policy exposure X_i, twinning status Z_i affects Y_i only via D_i.* $Y_i(d, 0) = Y_i(d, 1) = Y_i(d) = Y_{di} \Big|_{X_i}, \forall d \in \{0, 1\}$

Assumption 3 *Twinning status Z_i monotonically shifts D_i for everyone. $D_{1i} \geq D_{oi}, \forall i$.*

The monotonicity assumption automatically holds in our setting, because mothers with twins at the second birth have at least three children. We also require the "relevancy" condition, $\mathbb{E}[D_{1i} - D_{0i} | X_i] > 0$, which automatically satisfies for the twin instrument.

[9] Two more changes are present in Table 1. Given the policy ($X_i = 1$), twinning shifts the fertility of type-A mothers from the optimal two to an undesired three, which also represents an undesired increase. Given non-twinning ($Z_i = 0$), the policy rations fertility of type-B mothers, and reduces the fertility from three to two, which represents an undesired decrease.

We make two additional assumptions on the policy (X_i).

Assumption 4 *The policy X_i does not change mothers' types.* $\mathbf{Pr}(i \in S|X_i = x) = \mathbf{Pr}(i \in S) = P_S, \forall x \in \{0,1\}, \forall S = A, B, C.$

Assumption 5 *The policy X_i can be excluded from the average treatment effect of D_i on Y_i for each type of mother.* $\mathbb{E}[Y_{1i} - Y_{0i}|i \in S, X_i = x] = \mathbb{E}[Y_{1i} - Y_{0i}|i \in S], \forall x \in \{0,1\}, \forall S = A, B, C.$

Assumption 4 holds by definition. Assumption 5, which is similar to assumption A4 in Hull (2018), states that the policy does not change the average treatment effect for each type of mother.

A2.3 Fertility Equation

To distinguish between the two types of fertility increases, we consider the fertility equation

$$D_i = \alpha_0 + \alpha_1 Z_i + \alpha_2 Z_i \cdot X_i + \alpha_3 X_i + v_i,$$

where D_i is an indicator equal to one if mother i has three children, and zero otherwise; v_i represents idiosyncratic fertility preference shocks.

For mothers who are not exposed to the birth-control policy ($X_i = 0$), the effect of twinning on fertility is

$$\begin{aligned}
&\mathbb{E}[D_i|Z_i = 1, X_i = 0] - \mathbb{E}[D_i|Z_i = 0, X_i = 0] \\
=&\mathbb{E}[D_{1i} - D_{0i}|X_i = 0] \\
=&\mathbf{Pr}(D_{1i} > D_{0i}|X_i = 0) \\
=&\mathbf{Pr}(i \in A|X_i = 0) \\
=&\mathbf{Pr}(i \in A) \\
=&P_A,
\end{aligned}$$

where the first equality uses A1, the second equality uses A3, and the third equality uses A4. The compliers are type-A mothers, who desire two children and experience undesired fertility increases because of twinning. We have, $\alpha_1 = P_A$.

For mothers exposed to the birth-control policy ($X_i = 1$), the effect of twinning on fertility is

$$\begin{aligned}
&\mathbb{E}[D_i|Z_i = 1, X_i = 1] - \mathbb{E}[D_i|Z_i = 0, X_i = 1] \\
=&\mathbb{E}[D_{1i} - D_{0i}|X_i = 1] \\
=&\mathbf{Pr}(D_{1i} > D_{0i}|X_i = 1) \\
=&\mathbf{Pr}(D_{0i} = 0|X_i = 1) \\
=&\mathbf{Pr}(i \in A|X_i = 1) + \mathbf{Pr}(i \in B|X_i = 1) \\
=&\mathbf{Pr}(i \in A) + \mathbf{Pr}(i \in B) \\
=&P_A + P_B,
\end{aligned}$$

where the first equality uses A1, the second equality uses A3, and the fourth equality uses A4. When $X_i = 1$, twinning induces an undesired fertility increase with probability P_A, and a desired fertility increase with probability P_B. We have, $\alpha_1 + \alpha_2 = P_A + P_B$, and $\alpha_2 = P_B$.

The effect of the birth-control policy on the fertility of mothers without twins is

$$\begin{aligned}
&\mathbb{E}[D_i|Z_i = 0, X_i = 1] - \mathbb{E}[D_i|Z_i = 0, X_i = 0] \\
=&\mathbb{E}[D_{0i}|X_i = 1] - \mathbb{E}[D_{0i}|X_i = 0] \\
=&\mathbf{Pr}(i \in C|X_i = 1) - \{\mathbf{Pr}(i \in C|X_i = 0) + \mathbf{Pr}(i \in B|X_i = 0)\} \\
=&-\mathbf{Pr}(i \in B) \\
=&-P_B,
\end{aligned}$$

where the first equality uses A1 and the third equality uses A4. We have, $\alpha_3 = -P_B$.

A2.4 Child-quality Equation

We then estimate QQ effects for the two types of fertility changes by considering the child-quality equation:

$$Y_i = \rho_0 + \rho_1 Z_i + \rho_2 Z_i \cdot X_i + \rho_3 X_i + \varepsilon_i. \tag{A2.1}$$

where Y_i is the quality of the child of mother i and ε_i is the idiosyncratic shocks to child quality.

For mothers who are not exposed to the birth-control policy ($X_i = 0$), the effect of twinning on child quality is

$$\begin{aligned}
&\mathbb{E}[Y_i|Z_i = 1, X_i = 0] - \mathbb{E}[Y_i|Z_i = 0, X_i = 0] \\
=& \mathbb{E}[Y_{0i} + (Y_{1i} - Y_{0i})D_{1i}|X_i = 0] - \mathbb{E}[Y_{0i} + (Y_{1i} - Y_{0i})D_{0i}|X_i = 0] \\
=& \mathbb{E}[(Y_{1i} - Y_{0i})(D_{1i} - D_{0i})|X_i = 0] \\
=& \mathbb{E}[Y_{1i} - Y_{0i}|D_{1i} - D_{0i} > 0, X_i = 0] \cdot \mathbf{Pr}(D_{1i} - D_{0i} > 0|X_i = 0) \\
=& \mathbb{E}[Y_{1i} - Y_{0i}|i \in A, X_i = 0] \cdot \mathbf{Pr}(i \in A|X_i = 0) \\
=& \mathbb{E}[Y_{1i} - Y_{0i}|i \in A] \cdot \mathbf{Pr}(i \in A) \\
=& P_A \cdot \beta_A,
\end{aligned}$$

where the first equality uses A1 and A2, the third equality uses A3, and the fifth equality uses A4 and A5. Here $\beta_A = \mathbb{E}[Y_{1i} - Y_{0i}|i \in A]$ is the average treatment effect for type-A mothers who experience undesired fertility increases. We have $\rho_1 = P_A \cdot \beta_A$.

For mothers exposed to the birth-control policy ($X_i = 1$), the effect of twinning on child quality is

$$
\begin{aligned}
&\mathbb{E}[Y_i|Z_i = 1, X_i = 1] - \mathbb{E}[Y_i|Z_i = 0, X_i = 1] \\
=&\,\mathbb{E}[Y_{1i} - Y_{0i}|D_{1i} - D_{0i} > 0, X_i = 1] \cdot \mathbf{Pr}(D_{1i} - D_{0i} > 0|X_i = 1) \\
=&\,\mathbb{E}[Y_{1i} - Y_{0i}|D_{0i} = 0, X_i = 1] \cdot \mathbf{Pr}(D_{0i} = 0|X_i = 1) \\
=&\{\mathbb{E}[Y_{1i} - Y_{0i}|i \in A, X_i = 1] \cdot \mathbf{Pr}(i \in A|D_{0i} = 0, X_i = 1) \\
&+ \mathbb{E}[Y_{1i} - Y_{0i}|i \in B, X_i = 1] \cdot \mathbf{Pr}(i \in B|D_{0i} = 0, X_i = 1)\} \cdot \mathbf{Pr}(D_{0i}|X_i = 1) \\
=&\,\mathbb{E}[Y_{1i} - Y_{0i}|i \in A, X_i = 1] \cdot \mathbf{Pr}(i \in A|X_i = 1) \\
&+ \mathbb{E}[Y_{1i} - Y_{0i}|i \in B, X_i = 1] \cdot \mathbf{Pr}(i \in B|X_i = 1) \\
=&\,\mathbb{E}[Y_{1i} - Y_{0i}|i \in A] \cdot \mathbf{Pr}(i \in A) + \mathbb{E}[Y_{1i} - Y_{0i}|i \in B] \cdot \mathbf{Pr}(i \in B) \\
=&\,P_A \cdot \beta_A + P_B \cdot \beta_B,
\end{aligned}
$$

where the first equality uses A1-A3, and the second-last equality uses A4 and A5. Coefficient $\beta_B = \mathbb{E}[Y_{1i} - Y_{0i}|i \in B]$ is the average treatment effect for type-B mothers, who experience desired fertility increases. We have $\rho_1 + \rho_2 = P_A \cdot \beta_A + P_B \cdot \beta_B$, and $\rho_2 = P_B \cdot \beta_B$.

The effect of the birth-control policy on the child quality of mothers without twins is

$$
\begin{aligned}
&\mathbb{E}[Y_i|Z_i=0,X_i=1]-\mathbb{E}[Y_i|Z_i=0,X_i=0]\\
=&\mathbb{E}[Y_{0i}+(Y_{1i}-Y_{0i})D_{0i}|X_i=1]-\mathbb{E}[Y_{0i}+(Y_{1i}-Y_{0i})D_{0i}|X_i=0]\\
=&\mathbb{E}[Y_{0i}|D_{0i},X_i=1]\cdot\mathbf{Pr}(D_{0i}=0|X_i=1)+\mathbb{E}[Y_{1i}|D_{0i},X_i=1]\cdot\mathbf{Pr}(D_{0i}=1|X_i=1)\\
&-\mathbb{E}[Y_{0i}|D_{0i},X_i=0]\cdot\mathbf{Pr}(D_{0i}=0|X_i=0)-\mathbb{E}[Y_{1i}|D_{0i},X_i=0]\cdot\mathbf{Pr}(D_{0i}=1|X_i=0)\\
=&\mathbb{E}[Y_{0i}|i\in A\cup B,X_i=1]\cdot(P_A+P_B)+\mathbb{E}[Y_{1i}|i\in C,X_i=1]\cdot P_C\\
&-\mathbb{E}[Y_{0i}|i\in A,X_i=0]\cdot P_A-\mathbb{E}[Y_{1i}|i\in B\cup C,X_i=0]\cdot(P_B+P_C)\\
=&\{\mathbb{E}[Y_{0i}|i\in A,X_i=1]-\mathbb{E}[Y_{0i}|i\in A,X_i=0]\}\cdot P_A\\
&+\{\mathbb{E}[Y_{0i}|i\in B,X_i=1]-\mathbb{E}[Y_{0i}|i\in B,X_i=0]\}\cdot P_B\\
&+\{\mathbb{E}[Y_{1i}|i\in C,X_i=1]-\mathbb{E}[Y_{1i}|i\in C,X_i=0]\}\cdot P_C\\
&-\mathbb{E}[Y_{1i}-Y_{0i}|i\in B,X_i=0]\cdot P_B\\
=&\{\mathbb{E}[Y_{0i}|X_i=1]-\mathbb{E}[Y_{0i}|X_i=0]\}-P_B\cdot\beta_B,
\end{aligned}
$$

where the first equality uses A1 and the last equality uses A5. The term $\mathbb{E}[Y_{0i}|X_i=1]-\mathbb{E}[Y_{0i}|X_i=0]$ is the difference in the potential outcome in child quality by the policy—that is, a "selection effect." We have $\rho_3=-P_B\cdot\beta_B+\{\mathbb{E}[Y_{0i}|X_i=1]-\mathbb{E}[Y_{0i}|X_i=0]\}$.

A2.5 Second Stage

Consider the "second-stage" regression,

$$Y_i=\gamma_0+\gamma_1 D_i+\gamma_2 D_i\cdot X_i+\gamma_3 X_i+\varepsilon_i,$$

where we use Z_i and $Z_i \cdot X_i$ as the instrumental variables for D_i and $D_i \cdot X_i$. The coefficient γ_1 is the effect of D_i on Y_i when $X_i = 0$,

$$\begin{aligned}
\gamma_1 =& \frac{\mathbb{E}[Y_i|Z_i = 1, X_i = 0] - \mathbb{E}[Y_i|Z_i = 0, X_i = 0]}{\mathbb{E}[D_i|Z_i = 1, X_i = 0] - \mathbb{E}[D_i|Z_i = 0, X_i = 0]} \\
=& \frac{\beta_A \cdot P_A}{P_A} \\
=& \beta_A.
\end{aligned}$$

The coefficient γ_1 identifies the quantity-quality effect for type-A mothers, i.e., $\mathbb{E}[Y_{1i} - Y_{0i}|i \in A]$. The coefficient $\gamma_1 + \gamma_2$ is the effect of D_i on Y_i when $X_i = 1$,

$$\begin{aligned}
\gamma_1 + \gamma_2 =& \frac{\mathbb{E}[Y_i|Z_i = 1, X_i = 1] - \mathbb{E}[Y_i|Z_i = 0, X_i = 1]}{\mathbb{E}[D_i|Z_i = 1, X_i = 1] - \mathbb{E}[D_i|Z_i = 0, X_i = 1]} \\
=& \frac{P_A \cdot \beta_A + P_B \cdot \beta_B}{P_A + P_B} \\
=& \frac{P_A}{P_A + P_B}\beta_A + \frac{P_B}{P_A + P_B}\beta_B.
\end{aligned}$$

We can also show that γ_3 equals negative γ_2 plus a selection effect,

$$\gamma_3 = -\gamma_2 + \{\mathbb{E}[Y_{0i}|X_i = 1] - \mathbb{E}[Y_{0i}|X_i = 0]\}.$$

The sign of γ_2 informs the size of β_A versus β_B:

$$\gamma_2 = \frac{P_B}{P_A + P_B}(\beta_B - \beta_A).$$

If $\gamma_2 > 0$, then $\beta_B > \beta_A$.

A2.6 Continuous Policy

In the previous derivations, we have assumed that the policy exposure is a dummy variable. Relaxing the setting to a continuous policy does not change the model implications.

Consider a continuous policy that ranges from zero to $\bar{X}$, $X_i \in [0, \bar{X}]$. The definitions of type-A and type-C mothers do not change: Type-A mothers always desire and realize two children, and type-C mothers always desire and realize three children. Suppose N_i is realized fertility, and N_i^o is optimal fertility in P1:

$$N_i = N_i^o = 2\Big|_{Z_i=0,} \quad \forall i \in A,$$

$$N_i = N_i^o = 3\Big|_{Z_i=0,} \quad \forall i \in C.$$

The definition for type-B mothers changes a bit. Type-B mothers desire three children, but when the policy is strong enough, a type-B mother will realize two children.

$$\begin{aligned}
\forall i \in & B, \exists \theta_i \in [0,1), \\
N_i = & N_i^o = 3\Big|_{Z_i=0, X_i \leq \theta_i}, \\
N_i = & 2 < 3 = N_i^o\Big|_{Z_i=0, X_i > \theta_i},
\end{aligned}$$

where θ_i is the minimal policy strength that reduces a type-B mother's fertility from three to two. The distribution of type-B mothers is described by the distribution of θ, $F_B(x) \equiv \mathbf{Pr}(\theta < x | i \in B) = \mathbf{Pr}(D_{i0} | i \in B, X_i = x)$.

Assumptions 4 and 5 should be modified as:

Assumption 4′ *A mother's X_i does not change her type.* $\mathbf{Pr}(i \in S | X_i = x) = \mathbf{Pr}(i \in S) = P_S, \forall x \in [0, \bar{X}], \forall S = A, B, C.$

Assumption 5′ *The X_i can be excluded from the average treatment effect of D_i on Y_i for each type of mother.* $\mathbb{E}[Y_{1i} - Y_{0i} | i \in S, X_i = x] = \mathbb{E}[Y_{1i} - Y_{0i} | i \in S], \forall x \in [0,1], \forall S = A, C.$ *Also* $\mathbb{E}[Y_{1i} - Y_{0i} | i \in B, \theta_i \leq x, X_i = x] = \mathbb{E}[Y_{1i} - Y_{0i} | i \in B], \forall x \in [0, \bar{X}].$

The effect of twinning on fertility for mothers under a given X_i is

$$
\begin{aligned}
&\mathbb{E}[D_i|Z_i = 1, X_i = x] - \mathbb{E}[D_i|Z_i = 0, X_i = x] \\
=&\mathbb{E}[D_{1i} - D_{0i}|X_i = x] \\
=&\mathbf{Pr}(D_{0i} = 0|X_i = x) \\
=&\mathbf{Pr}(i \in A|X_i = x) + \mathbf{Pr}(D_{0i}|i \in B, X_i = x) \cdot \mathbf{Pr}(i \in B|X_i = x) \\
=&\mathbf{Pr}(i \in A) + F_B(x) \cdot \mathbf{Pr}(i \in B), \\
=&P_A + F_B(x) \cdot P_B,
\end{aligned}
$$

where the first equality uses A1 and the second-last equality uses A4′. The moderation effect of the policy on the effect of twining on fertility is

$$
\frac{\partial\{\mathbb{E}[D_i|Z_i = 1, X_i = x] - \mathbb{E}[D_i|Z_i = 0, X_i = x]\}}{\partial x} = f_B(x) \cdot P_B,
$$

where $f_B(x)$ is the density function of type-B mothers' θ.

The effect of twinning on child quality is

$$
\begin{aligned}
&\mathbb{E}[Y_i|Z_i = 1, X_i = x] - \mathbb{E}[Y_i|Z_i = 0, X_i = x] \\
=&\mathbb{E}[Y_{1i} - Y_{0i}|i \in A] \cdot \mathbf{Pr}(i \in A) + \mathbb{E}[Y_{1i} - Y_{0i}|i \in B] \cdot \mathbf{Pr}(i \in B) \cdot F_B(x) \\
=& P_A \cdot \beta_A + F_B(x) \cdot P_B \cdot \beta_B,
\end{aligned}
$$

and the moderation effect of the policy is

$$
\frac{\partial\{\mathbb{E}[Y_i|Z_i = 1, X_i = x] - \mathbb{E}[Y_i|Z_i = 0, X_i = x]\}}{\partial x} = f_B(x) \cdot P_B \cdot \beta_B.
$$

Interpretation of the regression parameters changes accordingly:

$$
\begin{aligned}
\alpha_1 &= P_A, \\
\alpha_2 &= F_B(1) \cdot P_B, \\
\rho_1 &= P_A \cdot \beta_A, \\
\rho_2 &= F_B(1) \cdot P_B \cdot \beta_B, \\
\gamma_1 &= \beta_A, \\
\gamma_2 &= \frac{F_B(1) \cdot P_B}{P_A + F_B(1) \cdot P_B} \cdot \beta_B.
\end{aligned}
$$

If $\bar{X} = 1$, interpretation of the parameters in the model of continuous policy is the same as that in the discrete case.

References

Angrist, Joshua D, Guido W Imbens, and Donald B Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91 (434):444.

Becker, Gary S. 1991. *A Treatise on the Family*. Cambridge, Massachusetts: Harvard University Press, enlarged e ed.

Becker, Gary S. and H. Gregg Lewis. 1973. "On the Interaction between the Quantity and Quality of Children." *Journal of Political Economy* 81 (2):S279–S288.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2010. "Small Family, Smart Family? Family Size and the IQ Scores of Young Men." *Journal of Human Resources* 45 (1):33–58.

Ebenstein, Avraham. 2010. "The 'Missing Girls' of China and the Unintended Consequences of the One Child Policy." *Journal of Human Resources* 45 (1):87–115.

Heckman, James J. 2007. "The Economics, Technology, and Neuroscience of Human Capability Formation." *Proceedings of the National Academy of Sciences* 104 (33):13250–13255.

Hull, Peter. 2018. "IsoLATEing: Identifying Counterfactual-Specific Treatment Effects with Cross-Stratum Comparisons." *Woring Paper* .

Imbens, Guido W and Joshua D Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62 (2):467.

Mogstad, Magne and Matthew Wiswall. 2016. "Testing the Quantity-Quality Model of Fertility: Estimation using Unrestricted Family Size Models." *Quantitative Economics* 7:157–192.

National Bureau of Statistics. 2009. *China Compendium of Statistics (1949-2008)*. Beijing: China Statistics Press.

Neary, J. P. and K. W. S. Roberts. 1980. "The Theory of Household Behaviour under Rationing." *European Economic Review* 13 (1):25–42.

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

Tobin, James and H. S. Houthakker. 1950. "The Effects of Rationing on Demand Elasticities." *Review of Economic Studies* 18 (3):140–153.

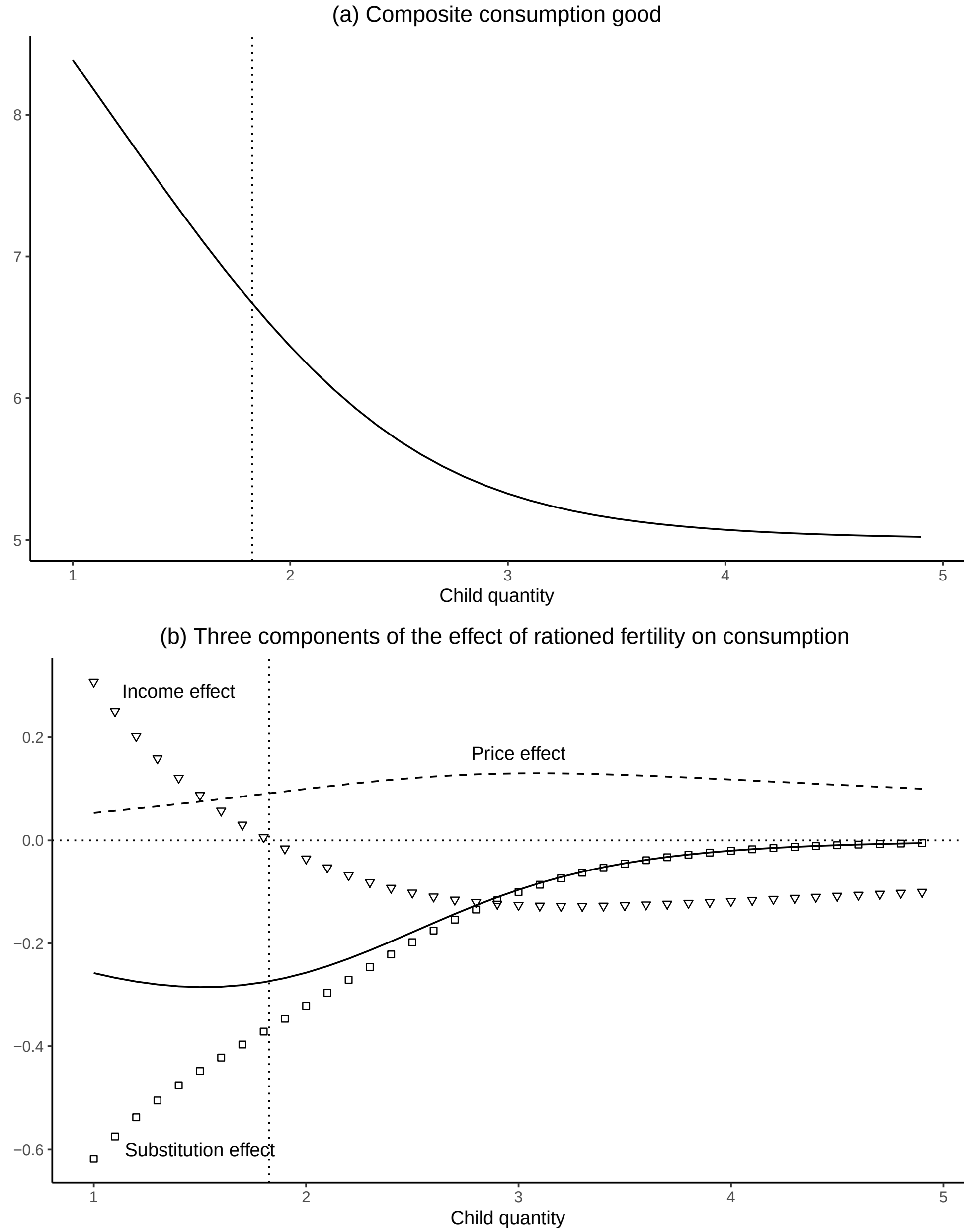

Figure A1: Simulation results for the composite consumption good

Notes: The x-axis is the rationed fertility level. In subfigure (a), the y-axis shows optimal parental consumption s in P2. In subfigure (b), the y-axis shows the semi-elasticity of the effect of rationed fertility on parental consumption ($\frac{\partial s/s}{\partial \bar{n}}$, solid line) and its three components: price effect (dashed line), substitution effect (line in squares), and income effect (line in triangles). Using a semi-elasticity form, the results are insensitive to the unit of measurement of child quality. Parental utility is in the form of nested CES, $U(n,q,s) = U_1^{\theta} s^{1-\theta}$, where $U_1 = (\alpha n^{\rho} + (1-\alpha) q^{\rho})^{\frac{1}{\rho}}$. The budget constraint is $\pi_{nq} nq + \pi_n n + \pi_q q + \pi_s s \leq y$. We set $\theta = 0.5$, $\alpha = 0.5$, $\rho = -3$, $\pi_{nq} = 1$, $\pi_q = 0$, $\pi_n = 0$, $\pi_s = 1$, and $y = 10$. The optimal fertility in P1 when fertility is not rationed, n^o, is 1.83, which is denoted by the vertical dotted line.

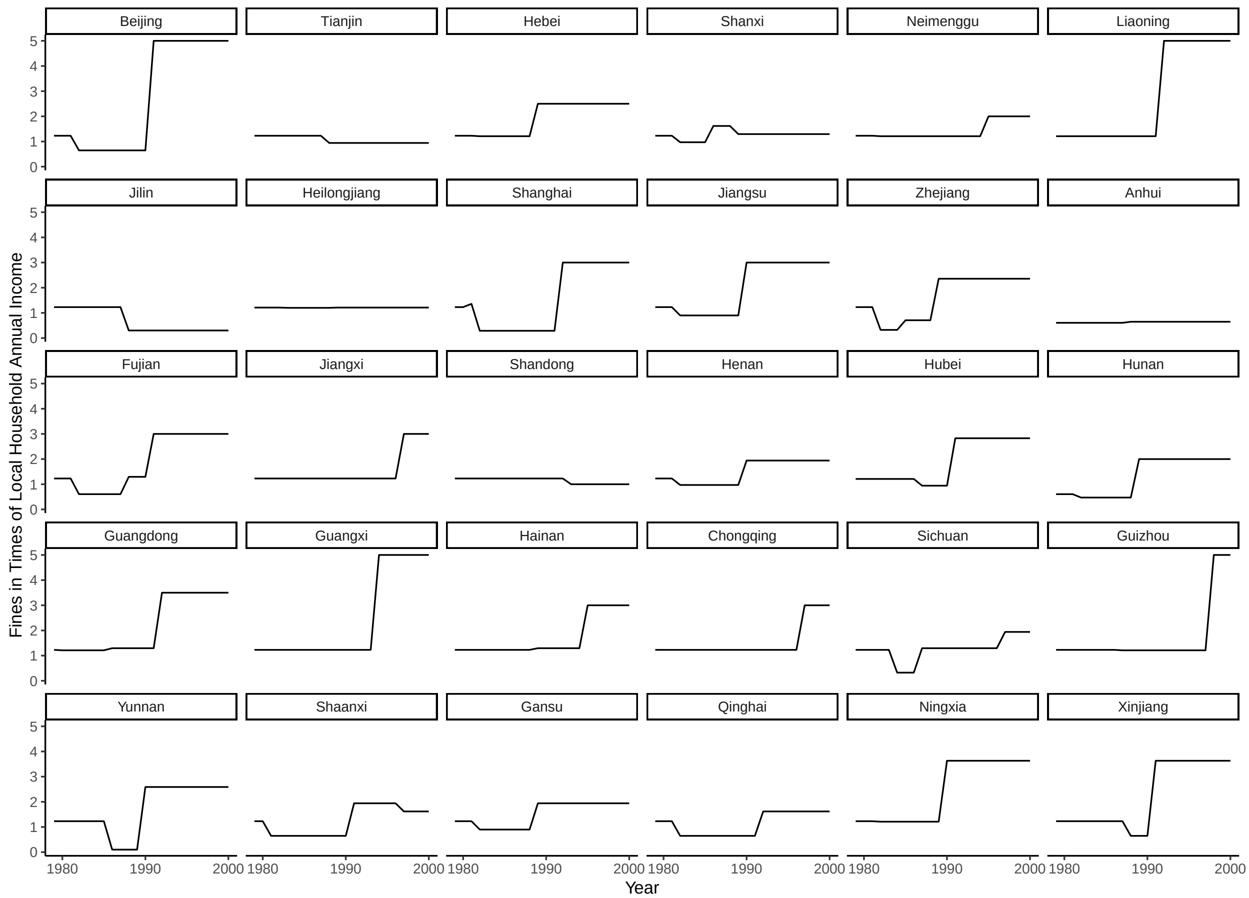

Figure A2: Fines by Province

Notes: This figure shows fines for unauthorized births from 1979 to 2000 in Chinese provinces, excluding Tibet. Fines are measured by multiples of local household annual income. We obtain fines data from Ebenstein (2010).

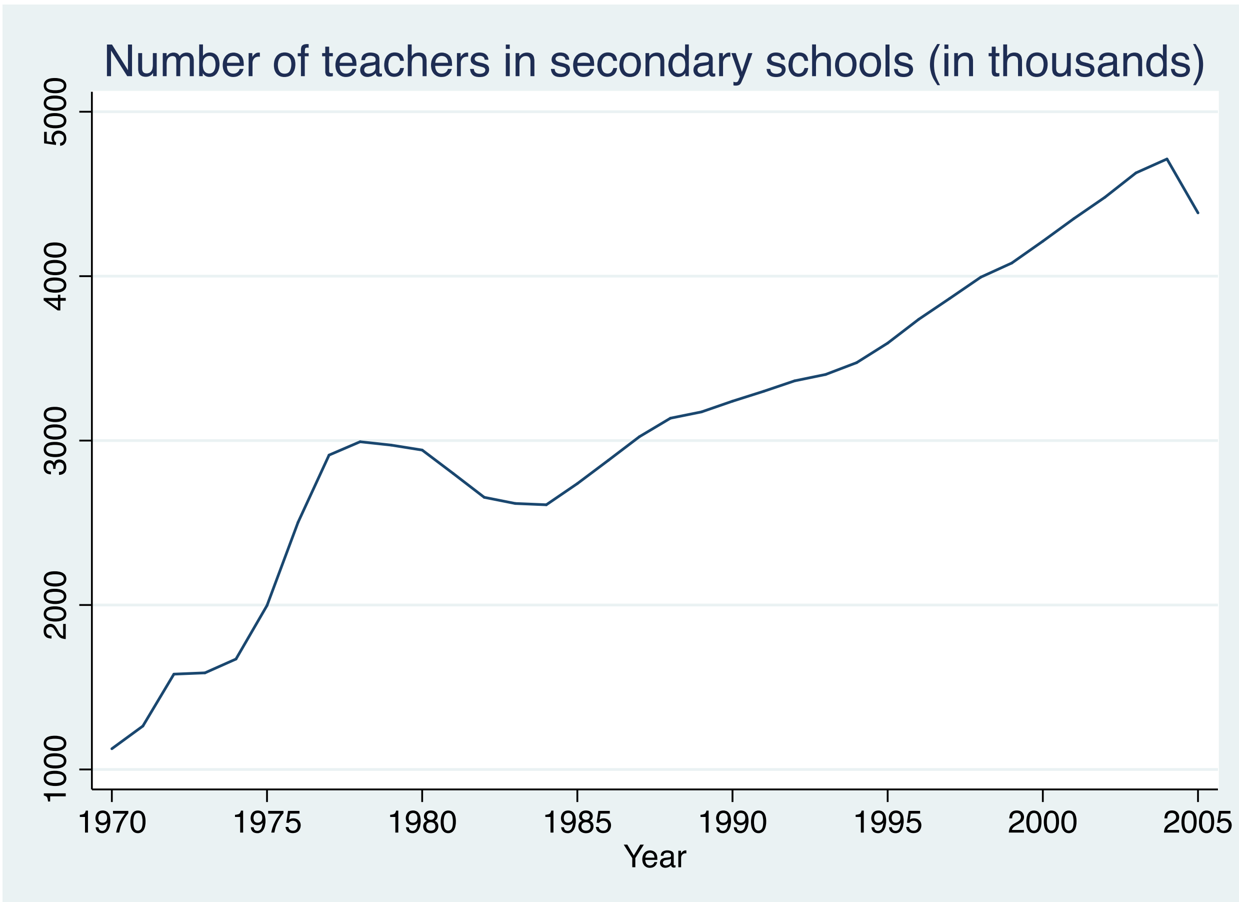

Figure A3: Trend in the number of teachers in secondary schools
Notes: This figure shows the number of teachers in secondary schools from 1970 to 2005 in China. The data source is the National Bureau of Statistics (2009) “China Compendium of Statistics (1949-2008).”

Table A1: Literature review on the child quantity-quality tradeoff

Author(s) and year	Data	Sources of identification	Outcomes measures	Results
Descriptive				
Hanushek, E. A. (1992).	Gary Income Maintenance Experiment between 1971 and 1975.	OLS controlling for confounding factors (family characteristics and school inputs)	scholastic performance (test scores)	Achievement falls systematically with increased family size.
Twins, controversy				
Rosenzweig, M. R., & Wolpin, K. I. (1980).	twins data from farm households in India: the Additional Rural Incomes Survey, collected in three rounds between 1969 and 1971 by the National Council of Applied Economic Research in New Delhi	Twining	Eduction (age-standardized index of school attainment) and Consumption (consumer durables)	Support quantity-quality trade-off (first evidence to confirm the hypothesis that exogenous increases in fertility decrease child quality)
Black, S. E., Devereux, P. J., & Salvanes, K. G. (2005)	Matched administrative files that cover the entire population of Norway who were aged 16-74 during the 1986 to 2000 (First and second births)	Twining	Education (main), earnings, employment, and teenage childbearing	A negative correlation between family size and children's education, but when includeing indicators for birth order or use twin births as an instrument, family size effects become negligible. In addition, higher birth order has a significant and large negative effect on children's education.
Angrist, J. D., Lavy, V., & Schlosser, A. (2010).	The main sources of data used here are the 20% public-use microdata samples from the 1995 and 1983 Israeli censuses. Individuals aged 18-60. The final sample restriction retains only first- and second-born	Sex composition & Twining	Schooling, labor market outcomes, marriage, and feritility	No evidence of a quantity-quality trade-off.
Brinch, C. N., Mogstad, M., & Wiswall, M. (2017).	Administrative registers from Statistics Norway covering the entire resident population of Norway who were between ages 16 and 74 at some point during the period 1986-2000.Family and demographic files are merged from Norwegian educational establishments.(First-born children)	Sex composition & Twining	Education	Family size effects vary in magnitude and even sign and that families act as if they possess some knowledge of the idiosyncratic effects in the fertility decision
Mogstad, M., & Wiswall, M. (2016).	The entire resident population of Norway who were between 16 and 74 years of age at some point during the period 1986-2000	Twining	Education	Estimation using a unrestricted specification for the quality-quantity relationship yields substantial family size effects.
Rosenzweig, M. R., & Zhang, J. (2009).	the Chinese Child Twins Survey (CCTS): households with child twins aged between 7 and 18 in Kunming; households with non-twin children in the same age group were also used for comparison	Twining	Eduction (schooling progress, the expected college enrolment, grades in school) and Health (weight and height)	Modest child quantity-quality effects

Continued...

Author(s) and year	Data	Sources of identification	Outcomes measures	Results
Twin instrument: negative QQ				
Black, S. E., Devereux, P. J., & Salvanes, K. G. (2010).	Birth records for all Norwegians born in the period 1967 to 1998 obtained from the Medical Birth Registry of Norway (First and second births)	Sex composition & Twining	IQ	IV estimates using sex composition as an instrument show no significant negative effect of family size; however, IV estimates using twins imply that family size has a negative effect on IQ scores. Our results suggest that the effect of family size depends on effects of expected increases in family size. However, unexpected shocks to family size resulting from twin births have negative effects on the IQ scores of existing children.
Li, H., Zhang, J. and Zhu, Y. (2008)	1% sample of the 1990 Chinese Population Census (children who were between 6 and 17 years old)	Twining	Education	Negative correlation between family size and child outcome, even after we control for the birth order effect
Caceres-Delpiano, J. (2006).	1980 Census Five-Percent Public Use Micro Sample (PUMS) (First-born children)	Twining	Education (attending private school, grade retention), mothers' labor force participation, parents divorce,	A larger family generated by twins in a later birth reduces the likelihood that older children attend private school, reduces the mother's labor force participation, and increases the likelihood that parents divorce. The impact of family size on a measure of child outcome, such as grade retention, is less clear.
Bhalotra, S., & Clarke, D. (2016).	(1) The US: National Health Interview Surveys (NHIS) for 2004-2014); (2) 68 countries from 1972 to 2012, available from the Demographic and Health Surveys (DHS)	1. Twining & adjust for available maternal health related characteristics; 2. An exogeneous shock to maternal health: natural experiments in the United States and Nigeria to estimate the bounds of the IV estimates following Conley et al. (2012).	Education & subjectively assessed binary indicator of child birth (NHIS only)	1. Mother's health and health-related behaviours and exposures are systematically positively associated with the probability of a twin birth. 2. Studies using twins to isolate exogenous variation in fertility will tend to under-estimate the impact of fertility on parental investments in children, and on women's labour supply. 3. Additional unexpected births do have quantitatively important effects on their siblings' educational outcomes.
Fletcher, J. M., & Kim, J. (2019).	the National Longitudinal Study of Adolescent to Adult Health (Add Health). Begin with students in grades 7 through 12 in 1994-1995 and follow up with a series of in-home interviews of students approximately one year and six years, and 13 years later (first- and second-born children).	Twining	Personality traits (Big five)	Having additional siblings lead to deficiency in personality traits and reduction in labor market earnings.

Continued...

Author(s) and year	Data	Sources of identification	Outcomes measures	Results
Policies and programs on the use of contraceptives or abortion: undesired increase or desired reduction, negative QQ				
Dumas, C., & Lefranc, A. (2019)	Philippine Censuses. Children aged 11 to 16 years in the 1995 and 2007 censuses.	A policy shock: in the late 1990s, the mayor of Manila enacted a municipal ban on modern contraceptives, not not in other cities.	Education	Increased family size led to a sizable decrease in educational attainment.
Joshi, S., & Schultz, T. P. (2013).	Census data from 1974 and 1982 together with the Matlab Health and Socioeconomic Survey (MHSS) of 1996	An experimental maternal and child health and family planning program that was established in Matlab, Bangladesh, in 1977.	fertility, child mortality, women's health and preventive health inputs	Village data from 1974, 1982, and 1996 suggest that program villages experienced a decline in fertility of about 17%. Household data from 1996 confirm that this decline in "surviving fertility" persisted for nearly two decades. Women in program villages also experienced other benefits: increased birth spacing, lower child mortality, improved health status, and greater use of preventive health inputs.
Pantano, J. (2016).	the Uniform Crime Reports (UCR), 33 cohorts: the youngest cohort (born in 1988) is 15 years old in the last year of the sample (2003); the oldest cohort (born in 1956) is 24 years old in the earliest year of the sample (1980).	Differential timing of contraceptive liberalization (policy change regarding removing the restricted access to the contraceptive pill for single women in their late teens)	Crime rate (two decades later)	Greater inflexibility to avoid unwanted pregnancies is likely to reduce crime two decades down the road, when undesired children born to these women would had reached their maximum criminal potential.
Ponczek, V., & Souza, A. P. (2012).	the 1991 Brazilian Census micro database: (1) first-born children in families with two or more births and (2) children in families with three or more births	Twining	Child labor (labor-force participation and household chores) and Eduction (school attendance, school progression, and literacy)	1.Positive relation between family size and labor force participation for boys and girls and to household chores for young women; 2. Negative relation between family size and educational outcomes for boys and girls; 3. Negative impacts of family size on human capital formation for young female adults
Rosenzweig, M. R., & Schultz, T. P. (1987).	The Malaysian Family Life Survey (MFLS)	IV (parental schooling, husband's income, and the local community program variables)	Eduction (school attainment) and Health (birthweight)	Imperfect fertility control leads to fecundity which significantly reduces both the average schooling attainment and birthweight of children in Malaysia, and this effect is more pronounced for couples with wives having lower levels of schooling

Continued...

Author(s) and year	Data	Sources of identification	Outcomes measures	Results
Ananat, E.O., Gruber, J., Levine, P.B. and Staiger, D. (2009)	The 2000 decennial Census of the United States; Vital Statistics data from the natality detail files for each year between 1965 and 1979; the Alan Guttmacher Institute (AGI) beginning in 1973; data reported by the Centers for Disease Control for the period 1970 through 1972 (Centers for Disease Control, 1971, 1972, 1974)	IVs (the cost of acquiring an abortion): 1. whether the cohort was born in one of the five early repeal states (California, New York, Washington, Alaska, Hawaii); 2. the average straight-line distance to the nearest county in which abortion is legal; 3. two measures of the latent social cost of abortion (state political attitudes compiled from 1960s state-level voter surveys and illegal abortion rates by state before 1970)	1. whether the individual lives in a household that is below the poverty line, is receiving welfare, or is a single parent; 2. whether the individual dropped out of high school, or did not graduate from college; 3. whether the individual is incarcerated;15 and whether the individual is not employed	Long-run selection effects through abortion: when abortion costs are lowered, cohort outcomes improve. In particular, lower-cost abortion increased likelihood of college graduation, lower rates of welfare use, and lower odds of being a single parent.
Pop-Eleches, C. (2006)	15 percent sample of the Romanian 1992 census.	A policy shock: in 1966 dictator declared abortion and family planning illegal. Birth rates doubled in 1967 because formerly abortion had been the primary method of birth control.	Educational (school achievements) and labor (skill specialization) outcomes	Children born after the abortion ban attained more years of schooling and greater labor market success. The reason is that urban, educated women were more likely to have abortions prior to the policy change, and the relative number of children born to this type of woman increased after the ban.But after controlling for composition using observable background variables, children born after the ban on abortions had worse educational and labor market achievements as adults.
Shifts of desired fertility level				
Conley, D., & Glauber, R. (2006).	1990 Census 5 percent PUMS (the first-born or second-born boy).	Sex composition	Boys' probabilities of private school attendance and grade retention	Negative effects on the second-born boys and no effects on the first-born boys
Dang, H. A. H., & Halsey Rogers, F. (2016).	Three rounds (2002, 2006, and 2008) of the Vietnam Household Living Standards Surveys (VHLSSs).	Instrument for family size using the commune distance to the nearest family planning center	Eduation investment (especially private tutoring)	Rural families do indeed invest more in the education of school-age children who have smaller numbers of siblings
Kugler, A. D., & Kumar, S. (2017).	Indian District Level Household Survey (DLHS), collected in 2007-2008, and the first round of the NFHS (1992-1993)	Gender of the first child	education (attainment and school enrolment)	children from larger families have lower educational attainment and are less likely to be enrolled in school, with larger effects for rural, poorer, and low-caste families as well as for families with illiterate mothers
Positive QQ at parity one				
Cameron, L., Erkal, N., Gangadharan, L., & Meng, X. (2013).	Economics experiments conducted with 421 individuals born just before and just after the One Child Policy's introduction in 1979.	One child policy	Behavioral and personality traits	The single child is significantly less trusting, less trustworthy, more risk-averse, less competitive, more pessimistic, and less conscientious

Continued…

Author(s) and year	Data	Sources of identification	Outcomes measures	Results
Qian, N. (2009).	Match the 1% sample of the 1990 Population Census with the 1989 China Health and Nutritional Survey (CHNS) at the county level. The sample consists of first-born children in cohorts born during 1962-1981. The reference group in the regression analysis is comprised of individuals born during 1962-1972. The treatment group comprises children who are nine to fourteen years old in 1990.	1.5-child policy & first child gender	Education (school enrollment of first-born child)	No evidence of a quantity-quality trade-off
Unclassified				
Bagger, J., Birchenall, J. A., Urzua, S., & Holloway, R. (2019).	Danish administrative panel dataset for the period 1980-2006 with annual observations on all individuals aged 15-70.	Structural estimation exploring both within- and between-family variation in education	Education	Significant birth order and family size effects in individuals' years of education thereby confirming the presence of a quantity-quality trade off
Klemp, M., & Weisdorf, J. (2018).	Individuals from a sample of 26 English parishes, originally recorded in English church books for the period 1541–871, later transcribed by the Cambridge Group for the History of Population and Social Structure as documented in Wrigley et al. (1997).	Fecundity measured by the time span from a couple's marriage to their first birth	socio-economic achievements	children of parents with lower fecundity were more likely to become literate and employed in skilled and high-income professions
Silles, M. A. (2010).	The intial 1958 survey from the British National Child Development Study (NCDS), known as the Preinatal Mortality Survey (PMS), and three subsequent sweeps taken in 1965 (age 7), 1969 (age 11) and 1974 (age 16).	IVs related to relate to parental reproductive capacity: (a) the number of siblings of the individual's mother, and (b) the length of the interval between marriage and the birth of her first child.	Eduction (scores of math tests) and Behavioral characteristics (measures of aggression)	Sibling size is shown to have an adverse causal effect on test scores and behavioral development: 1) given family size, first-borns ultimately obtain higher test scores than middle-born or last-born children; 2) first-borns and last-borns tend to be better behaved at school than middle-borns, though last-borns have no test score advantage.
Marteleto, L.J. and de Souza, L.R. (2012)	1977-2007 Pesquisa Nacional por Amostra de Domicilio (PNAD), a nationally representative household survey collected annually by the Brazilian Census Bureau (12- to 18-year-old, first- and second- born children)	Twining	Education	The effect of family size on education is not uniform throughout a period of significant social, economic, and demographic change. Rather, the causal effect of family size on adolescents' schooling resembles a gradient that ranges from positive to no effect, trending to negative.
Bougma, M., LeGrand, T. K., & Kobiane, J. F. (2015).	Three complementary sources: the Ouagadougou Health and Demographic Surveillance System (Ouaga HDSS), the Demtrend survey, and the Ouaga HDSS Baseline Health Survey (BHS)	subfecundity (both secondary sterility or infertility and pregnancy problems) as the instrumental variable	Education (1) parents' cumulative investment, (2) parents' schooling strategies concerning their children at each level of education, (3) educational system organization, and (4) educational delays	A net negative effect of sibship size on the level of schooling achieved by children, one that grows stronger as they progress through the educational system.

Continued...

Author(s) and year	Data	Sources of identification	Outcomes measures	Results
Ananat, E.O. and Hungerman, D.M. (2012)	Women born in the US between 1943 and 1965 and observed in the 1980 Census 5-percent individual public use microdata (IPUMS)	policy change regarding removing the restricted access to the contraceptive pill for single women in their late teens (Difference-in-Difference-in-Difference test: within states that had policy changes, some states increased access only for some young women)	fertility, family characteristics (e.g., childern's birthweight), and the prevalence of abortion (both short-term and long-term effects)	1. access to the pill led to a short-term decline in fertility among these women; 2. there are temporary reductions in abortion rates; 3. the pill had no long-term effects on total childbearing; 4. the short-term decline in fertility led to immediate declines in the average birth weight and economic circumstances of children born to young women, but in the long-term early access to the pill increased the likelihood that a child had a college-educated, married mother.

Table A2: Summary statistics of the Chinese Child Twins Survey

	Non-twin		Twin	
	Mean (1)	S.D. (2)	Mean (3)	S.D. (4)
Basic characteristics				
Paternal age	37.20	4.72	40.20	4.74
Maternal age	35.28	4.41	37.99	4.35
Paternal schooling years	8.42	2.65	8.02	2.43
Maternal schooling years	7.36	2.50	6.83	2.39
Family income (¥/year)	10,704.67	8,845.79	11,032.37	9,888.22
Paternal consumption				
Cigarette expense (¥/month)	58.82	65.51	47.49	54.93
Alcohol expense (¥/month)	11.69	15.23	13.15	23.42
Clothing expense (¥/six months)	124.20	177.34	87.77	168.39
Maternal consumption				
Cosmetic expense (¥/six months)	18.12	39.14	10.10	32.93
Clothing expense (¥/six months)	118.55	146.37	82.32	109.33
Paternal labor supply				
Employment (dummy)	0.83	0.37	0.82	0.38
Days worked last month	25.88	4.32	26.38	4.64
Private business (dummy)	0.21	0.41	0.33	0.47
Migration (dummy)	0.10	0.30	0.12	0.32
Maternal labor supply				
Employment (dummy)	0.78	0.42	0.78	0.41
Days worked last month	25.76	4.96	26.53	4.57
Private business (dummy)	0.20	0.40	0.35	0.48
Migration (dummy)	0.02	0.14	0.04	0.20
Observations	364		278	

Notes: ¥ stands for Chinese yuan. "Private business" is an indicator variable that equals one if the father or mother has a private business. "Migration" is an indicator variable that equals one if the father or mother has left home for more than 30 days in the last 180 days.