
人口爆発

- ・人口増加は生活水準に影響を及ぼす

→Malthusが18世紀に「人口論」で人口増加に警鐘を鳴らす

歴史の中でその悲観的な経済成長に対する見方は偽りであると立証されたが、運命論者の間では人気となった 例)Paul EhrlichのPopulation Bomb

家族サイズと生活水準の関係

・家族サイズが大きくなると、子供一人当たり割く資源が少なくなってしまう

→Beckerの「quantity-quality tradeoff」

この考え方は政治にも導入され、中国の一人っ子政策などを生み出した

(発展途上国に住む人々のうち85%が、高出生率が貧困の原因である国に住んでいる)

家族サイズと学歴の影響

- ・家族の子供の数と、子供の学歴の関係を調べたい

→ランダム化比較試験を行うことは現実では不可能(子供をランダムに追加することなどできない)

- ・Mark RosenzweigとKenneth Wolpinは「双子」のランダム性に注目し、小さいサンプルで検証

- ・Master Joshwayらは、イスラエルで大きいサンプルで検証

→イスラエルは人種が多様かつ年齢や性別といったデータが取りやすい

ALS

- ・対象は最低一人以上の兄弟姉妹を持つ、長男または長女
 - ・二人目の子供が一人っ子だったときの平均家族サイズは3.6人
 - ・一方、二人目の子供が双子だったときの平均家族サイズは3.92人
- 双子の出生がなぜ0.32人という微妙な家族サイズの上昇を引き起こすのか？

ALS

- ・多くのイスラエルの家族が3、4人の子供を望む一方で、一部は2人で満足する(7:3)
 - ・二人目が双子でない場合、前者は3人目を望むが後者は望まない
 - ・二人目が双子の場合、どちらもそこで落ち着く
- 双子でない親は0.7人の追加補正を受け、結果的に双子であることの効果は0.32人分となる

ALS

- ・単純な回帰分析の結果、子供が一人増えることによって、長男の被教育年数は/4年減少した
 - ・一方で、双子家庭に生まれていても、非双子家庭と同程度の教育を受けていることが分かった
- 双子であれば家族サイズは変化するはずなのに、教育差は出ていない
- (量と質のトレードオフは見られない)

ALSの問題点

- ・双子が生まれるのは完全にランダムではなく、母親の年齢や人種によっても左右される
- omitted variables biasを引き起こす(本来入るべき変数が入っていない)
- ・そのような変数を入れるには、双子ではサンプル数が少なすぎる

sex-mix instruments

- ・イスラエルや欧米では、子供の性別が異なっていることが望まれる
 - 1人目と2人目が同性だと3人目を欲しがりやすい
- ・異性だった場合の平均家族サイズは3.60、同性だった場合は3.68
 - 0.08の差しかないが、サンプルの対象となる比率が高い(双子が1%に対し、50%)
 - 母親の出産時の年齢や人種などのバイアスを受けない

sex-mix instruments

.

sex-mix instruments

- ・実験の結果、同性の家族も異性の家族も、長男長女の教育年数は2.6年で変わらなかった

→ALSと同様量と質のトレードオフは確認されず

- ・考えられるバイアスとして、同性の兄弟姉妹は、部屋の共有やお下がりなどによって浮いたお金を教育に回せている可能性がある

- ・また、性別に関係なく、宗教上の理由で多くの子供を産む女性は常にsex-mix instrumentsの影響を受けやすく、逆に高学歴で大きな家族を望まない女性は影響を受けづらい

LATEの復習

以下のようにパラメータを設定する

- First stage (ϕ)

$$\phi = E\{D_i|Z_i = 1\} - E\{D_i|Z_i = 0\}$$

- Reduced form (ρ)

$$\rho = E\{Y_i|Z_i = 1\} - E\{Y_i|Z_i = 0\}$$

- LATE (λ)

$$\lambda = \frac{\rho}{\phi} = \frac{E\{Y_i|Z_i = 1\} - E\{Y_i|Z_i = 0\}}{E\{D_i|Z_i = 1\} - E\{D_i|Z_i = 0\}}$$

First Stage= $Z_i \rightarrow D_i = \phi$

Reduced Form= $Z_i \rightarrow Y_i = \rho$

LATE= $D_i \rightarrow Y_i = \lambda$

LATEの復習

$$\rho = \phi\lambda.$$

- ・式より、 ϕ が0の時、 ρ も0になってしまう
 - ・first stageについて証拠がない、reduced-formの統計的に優位な推定値は不安
- 説明変数(この場合は家族サイズ)意外にも、結果(この場合は教育)と繋がってしまっている

2SLS(2段階最小二乗法)

$$Y_i = \alpha_0 + \rho Z_i + e_{0i}.$$

KIPPとの違いは、一つ以上の因果関係に影響を及ぼす要素があること

→双子と子供の性別の二つの回帰を組み合わせて統計的正確性を上げる

2SLS(2段階最小二乗法)

通常の回帰では誤差項と独立変数の間に相関がないと仮定

※両者に相関がある場合は？

→誤差項と相関のない操作変数を利用すれば良い

2段階最小二乗法=(2回に分けて最小二乗法を行う)

2SLS(2段階最小二乗法)

$$D_i = \alpha_1 + \phi Z_i + e_{1i},$$

普通に回帰をして、 $\alpha_1, \phi, Z_i, e_{1i}$ を求める

$$\hat{D}_i = \alpha_1 + \phi Z_i.$$

\hat{D}_i 得られた値を使って、誤差項と相関の \hat{D}_i を作る

2SLS(2段階最小二乗法)

$$Y_i = \alpha_2 + \lambda_{2SLS} \hat{D}_i + e_{2i}.$$

得られた \hat{D}_i を代入して、もう一度回帰

→ Y_i は誤差項と相関のない操作変数に基づくため、最適となる

2SLS(2段階最小二乗法)

母親の年齢を変数 A_i として導入しても全く同じ形でできる。

$$\text{Reduced form: } Y_i = \alpha_0 + \rho Z_i + \gamma_0 A_i + e_{0i}$$

$$\text{First stage: } D_i = \alpha_1 + \phi Z_i + \gamma_1 A_i + e_{1i}.$$

2SLS(2段階最小二乗法)

誤差項と相関のない操作変数を作成し、それをもとに得られた計算値を用いて二度目の回帰

$$\hat{D}_i = \alpha_1 + \phi Z_i + \gamma_1 A_i.$$

$$Y_i = \alpha_2 + \lambda_{2SLS} \hat{D}_i + \gamma_2 A_i + e_{2i},$$

2SLS(2段階最小二乗法)

2SLSを用いるといくらでも説明変数をぶち込める？

一人目の子供と二人目の子供が同性であった場合、 $W_i=1$ 、そうでなければ $W_i=0$

一人目の子供が男性だったとき、 $B_i=1$ 、女性だったとき $B_i=0$

(男性の方が女性よりもわずかに高い確率で生まれてくるため)

2SLS(2段階最小二乗法)

$$D_i = \alpha_1 + \phi_t Z_i + \phi_s W_i + \gamma_1 A_i + \delta_1 B_i + e_{1i}.$$

$$Y_i = \alpha_2 + \lambda_{2SLS} \hat{D}_i + \gamma_2 A_i + \delta_2 B_i + e_{2i},$$

こうすることで、様々なバイアスを生み出す要因をコントロールできる

Quantity-quality first stages

・双子効果については、高齢出産や民族などのバイアスを取り除いた方は、0.437で0.320より大きくなっている
(列2)

・セックスミックス効果については、コントロールしても値がほぼ変わらず、バイアスの影響が少ないことを示す

・一人目が男の子であることによって、子供数は減少しているが、統計的に優位ではない

(アジア圏では特に男の子が好まれる)

TABLE 3.4
Quantity-quality first stages

	Twins instruments		Same-sex instruments		Twins and same-sex instruments
	(1)	(2)	(3)	(4)	
Second-born twins	.320 (.052)	.437 (.050)			.449 (.050)
Same-sex sibships			.079 (.012)	.073 (.010)	.076 (.010)
Male		-.018 (.010)		-.020 (.010)	-.020 (.010)
Controls	No	Yes	No	Yes	Yes

2SLSによるQuantity-quality trade-off

・OLSによる推定値はマイナスであり、子供の人数と教育に負の相関関係があるように見える

・2SLSによる推定値は大体がプラスの値をとっているが、標準誤差の2倍には達さず、統計的に有意とは言えない

・双子と子供の性別どちらも加味したものは標準誤差は減少しているが、有意ではない

→(結論)OLSによって得た負の相関関係は、選択バイアスによるもの

TABLE 3.5
OLS and 2SLS estimates of the quantity-quality trade-off

Dependent variable	OLS estimates (1)	2SLS estimates		
		Twins instruments (2)	Same-sex instruments (3)	Twins and same-sex instruments (4)
Years of schooling	-.145 (.005)	.174 (.166)	.318 (.210)	.237 (.128)
High school graduate	-.029 (.001)	.030 (.028)	.001 (.033)	.017 (.021)
Some college (for age ≥ 24)	-.023 (.001)	.017 (.052)	.078 (.054)	.048 (.037)
College graduate (for age ≥ 24)	-.015 (.001)	-.021 (.045)	.125 (.053)	.052 (.032)