

2022年度 統計データ分析コンペティション
審査員奨励賞 [大学生・一般の部]

日本の出生率に関するパネルデータ分析
—男性の育児参加と女性社会進出の重要性—

中根 智樹 (同志社大学経済学部)

日本の出生率に関するパネルデータ分析 —男性の育児参加と女性社会進出の重要性—

中根智樹
同志社大学経済学部経済学科

1 はじめに

1.1 目的

日本の合計特殊出生率は1989年の1.57ショック以降も減少し続け、2005年に底を打ち回復したが、2015年以降また減少を始めている（図1）。出生率の低下要因として挙げられるのは未婚率の上昇、晩婚化、教育コストの上昇、女性の所得上昇や女性の社会進出による出産の機会費用の発生などである。

実際に日本の出生率の実証分析を行った高山他（2000）⁽¹⁾、阿部、原田（2008）⁽²⁾や足立、中里（2017）⁽³⁾では女性の賃金が負の影響を及ぼしていることを実証しており、女性の社会進出による機会費用の発生が出生率の減少要因だとしている。

しかし、近年その傾向は変化していることをDoepke他（2022）⁽⁴⁾は主張しており、収入の大きな国々では女性の社会進出は正の効果を持つようになってきたことを指摘している。Doepke他（2022）は先進国の出生率の新たな決定要因として家族政策、協力的な父、社会規範の変化、そして柔軟な働き方を挙げている。

そこで、本研究では女性の賃金や失業率などの先行研究にある説明変数に加え、女性の社会進出と、協力的な父の指標として男性の育児時間の合計特殊出生率への影響を分析する。



図1 日本の合計特殊出生率の推移（Rを用いて作成）

1.2 先行研究

代表的な理論モデルの先行研究として、出生率の決定要因をモデル化したBecker（1960）⁽⁵⁾やWillis（1973）⁽⁶⁾がある。Becker（1960）は出生率には子供の量と質のトレードオフがあり、子供の質を上げるためにはコストが掛かり、それにより子供の数が減少することをモデル化した。Willis（1973）は女性の社会進出は出産の機会費用を発生させるため、出生率に負の効果を持つことをモデル化して示した。

日本の出生率の実証分析の先行研究として、高山他（2000）、阿部、原田（2008）、足立、中里（2017）、加藤（2017）⁽⁷⁾が挙げられる。高山他（2000）は日本の少子化の理由として、結婚の費用が要因になっているのかを検証した。WLSの結果、男性賃金は正の相関、女性賃金や住宅費、児童手当や初婚年齢が負の効果を持った。高山他（2000）は女性の賃金が機会費用となり、出生率が低下するため、育児休暇期間中の賃金保証が出生率を高めると結論付けた。阿部、原田（2008）は重み付きで操作変数法を使用し、被説明変数に合計特殊出生率をとり、説明変数に所得、女性賃金、地価、通学者数対15～24歳人口、そして保育所の制約を置いた。その結果、すべての変数は有意で負であった。十万人未満の都市に絞ったところも同様の結果が出た。足立、中里（2017）は都道府県データを用いて出生率の決定要因を分析した。その結果、OLSとWLSでは生涯未婚率の上昇と女性賃金の増加が出生率に負の影響をもたらしていることがわかった。一方、女性の社会進出や晩婚化が出生率を低下させるという効果は見られなかった。加藤（2017）はLutz他（2006）⁽⁸⁾が固定効果モデルを用いて出生率減少の要因に人口密度が寄与していることを示したことを参考に、日本で人口密度が合計特殊出生率にどのような影響を持つのかOLSとWLSを用いて実証した。その結果、負の影響を持つことを実証している。近藤（2014）⁽⁹⁾は不況による失業率と出生率・既婚率への影響を都道府県パネルで分析した。その結果、失業率は合計特殊出生率に負の効果を与えることがわかった。しかし、係数の大きさが時系列で見た失業率と合計特殊出生率の変動と比較して小さかった。さらに、データの期間を分割すると結果が安定しないことがわかった。符号が変化する訳では無いが、1985-1990と1995-2000は統計的に有意な相関が見られなかった。

これらの先行研究の課題は、近藤（2014）を除いてすべてクロスセクション分析であることで、都道府県の個別効果や時間効果を考慮していない。つまり、都道府県の持つ特性や、時間を通じた社会の構造変化などの効果を除去しないまま分析しているため、欠落変数のバイアスが発生していると考えられる。例として、沖縄県の合計特殊出生率は常に高く、2021年の合計特殊出生率は全国平均が1.30のなか、沖縄県は1.80である。この場合、他の都道府県とは違う沖縄県特有の個別効果が存在すると考えるのが自然である。このような都道府県の持つ特性である個別効果を除去せずに分析を行うと、この個別効果に結果が歪められる。多くの場合、このような個別効果は観測ができず、データとして入手できないため欠落変数のバイアスを起こす。そのため、本研究では欠落変数のバイアスに対処するために都道府県レベルのパネルデータ分析を行う。

2 分析方法

本研究では都道府県別パネルデータを用いて合計特殊出生率の要因分析を行う。パネルデータ分析を用いるメリットは2点ある。1点目は標本サイズを大きくできることである。特に都道府県別データなどを用いて分析するとき問題になる自由度の小ささを解決できる。2点目は観測できない、時間を通じて一定な変数の欠落変数のバイアスを回避できることにある。

本研究ではパネルデータ分析の基本的なモデルである固定効果モデル（Fixed Effect Model）と変量効果モデル（Random Effect Model）を用いて分析を行う。固定効果モデルとは各クロスセクションの個体が異なる切片を持つ回帰式を推定する手法で、それに対して変量効果モデルは各切片が平均的な切片の周りを中心とする確率分布から生み出される回帰式を推定する手法である（Studenmund, 2011）⁽¹⁰⁾。固定効果モデルでは本来観測できない個別の効果除去することができる。本研究では都道府県を個体として分析しているため、固定効果モデルでは都道府県の持つ特性などの個別効果を取り除くことができる。

さらに、固定効果モデルを三種類に分ける。観測はできないが、時間を通じて一定な変数を固定効果と呼び、個体間では一定であるが、時間を通じて変化する変数を時間効果という（西山他, 2019）⁽¹¹⁾。そこで、固定効果のみ除去するモデル（Fixed）、時間効果のみを除去するモデル（Time）、そして固定効果と時間効果両方を除去するモデル（Two-way）を考える。

固定効果と時間効果の入ったTwo-way固定効果モデルを解説する。まず以下の(1)式を検討する。

$$Y_{it} = \alpha_i + \lambda_t + \beta_1 X_{1it} + \dots + \beta_k X_{kit} + u_{it} \quad (1)$$

ここで、 α_i が固定効果を表し、 λ_t が時間効果を表す。ここで、 \dot{Y}_{it} を(2)式のように定義する。

$$\dot{Y}_{it} = Y_{it} - \bar{Y}_i - \bar{Y}_t + \bar{Y} \quad (2)$$

ここで、 \bar{Y}_i は*i*期での Y_{it} の平均で(3)式、 \bar{Y}_t はすべての観測値から計算した Y_{it} の平均である(4)式。

$$\bar{Y}_i = \sum_{j=1}^N \frac{Y_{jt}}{N} \quad (3)$$

$$\bar{Y} = \sum_{j=1}^N \sum_{s=1}^T \frac{Y_{js}}{NT} \quad (4)$$

同様に、

$$\dot{X}_{1it} = X_{1it} - \bar{X}_{1i} - \bar{X}_{1t} + \bar{X}_1, \dots, \dot{X}_{kit} = X_{kit} - \bar{X}_{ik} - \bar{X}_{kt} + \bar{X}_k \quad (5)$$

ただし、

$$\bar{X}_1 = \sum_{j=1}^N \sum_{s=1}^T \frac{X_{1js}}{NT}, \dots, \bar{X}_k = \sum_{j=1}^N \sum_{s=1}^T \frac{X_{kjs}}{NT} \quad (6)$$

$$\dot{u}_{it} = u_{it} - \bar{u}_i - \bar{u}_t + \bar{u} \quad (7)$$

ただし、

$$\bar{u} = \sum_{j=1}^N \sum_{s=1}^T \frac{u_{js}}{NT} \quad (8)$$

と定義する。このように変数を変換すると、以下の式 (9) が得られ、固定効果 α_i と時間効果 λ_t が除去されているモデルができる。

$$\dot{Y}_{it} = \beta_1 \dot{X}_{1it} + \dots + \beta_k \dot{X}_{kit} + \dot{u}_{it} \quad (9)$$

3 分析対象とデータ

3.1 分析対象

本研究の分析対象は 2001 年、2006 年、2011 年、2016 年、2021 年の 47 都道府県データである。合計特殊出生率は市区町村別のデータも存在するが、男性や女性の賃金の影響を考慮するため、都道府県別データとなった。欠損値はなかったため、サンプル数は 47×5 の $N=235$ である。

3.2 データの出所

本研究のデータセットは教育用標準データセットから SSDSE-県別推移と SSDSE-社会生活を用いており、その他には国勢調査、人口動態調査、人口推計、労働力調査、賃金構造基本統計調査、家計調査、そして社会生活基本調査のデータを用いている。各説明変数の出典は表 1 のとおりである。

表 1 説明変数一覧

変数名	出典
合計特殊出生率	人口動態調査、SSDSE 県別推移
女性未婚率	国勢調査、人口推計
女性就業率	国勢調査、人口推計
女性賃金	賃金構造基本統計調査
失業率	労働力調査
男性賃金	賃金構造基本統計調査
男性育児時間	社会生活基本調査、SSDSE 社会生活
教育費	家計調査

3.3 説明変数と定義

本研究の被説明変数は合計特殊出生率である。厚生労働省 (2013) ⁽¹²⁾ によると、合計特殊出生率とは、15 から 49 歳までの女性の年齢別出生率を合計したものである。説明変数は女性の就業率、女性の未婚率、女性の月当たり賃金、男性の月当たり賃金、失業率、男性の週当たり育児時間、そして月当たり教育費支出である。

女性就業率、女性未婚率、女性の月当たり賃金と男性の月当たり賃金はデータを加工して取得した。それぞれの定義は以下の通りである。

女性未婚率は 15~49 歳の女性人口のうちの未婚女性の割合を示している。同様に、女性就業率は 15~49 歳の女性人口のうちの女性就業者数の割合を示している。この 2 つの変数は出生率の定義に従い、出生可能年齢である 15~49 歳を対象にした。

月当たりの女性賃金と男性賃金は以下の計算式で求めている。

$$\text{決まって支給する現金給与額} + \frac{\text{年間賞与その他特別給与額}}{12}$$

3.4 記述統計量と相関係数表

各説明変数の記述統計量は表2、相関係数表は表3の通りになっている。VIFはすべて10以下なので、多重共線性は深刻ではない。

表2 記述統計量

	変数名	平均	標準偏差	最小値	最大値
A	合計特殊出生率	1.43	0.14	1.00	1.95
B	女性未婚率	41.75	2.21	35.92	49.65
C	女性就業率	61.87	5.03	49.23	72.54
D	女性賃金(万円)	27.88	3.03	21.46	39.21
E	失業率	3.63	1.18	1.40	8.40
F	男性賃金(万円)	41.28	4.80	30.16	56.96
G	男性育児時間(分)	4.49	1.72	1.20	10.00
H	教育費(万円)	1.17	0.35	0.48	2.71

表3 相関係数表

	A	B	C	D	E	F	G	H	VIF
A	1								
B	-0.29	1							1.69
C	0.41	-0.33	1						3.65
D	-0.52	0.36	-0.43	1					6.75
E	-0.16	0.10	-0.57	-0.23	1				2.88
F	-0.58	0.11	-0.56	0.84	0.04	1			7.08
G	0.11	0.35	0.10	0.24	-0.38	-0.09	1		1.81
H	-0.47	0.03	-0.46	0.44	0.16	0.57	-0.18	1	1.60

4 分析結果

パネルデータ分析の結果は表4のとおりである。各モデルのF統計量はすべて1%有意であった。次に、採択するモデルを検討するためにF検定とハウスマン検定を行う。

4.1 モデル選択

評価するモデルを選択するためにF検定とハウスマン検定を行った。結果は表5のとおりである。

まず3つの固定効果モデルかPooled OLSのどちらを採択するか判断するために、それぞれにF検定を行った。その結果、FixedはF=10.02、TimeはF=19.391、Two-wayはF=38.053ですべて1%有意であった。これは都道府県が個別効果と時間効果を持つことを意味しており、Two-wayの固定効果モデルが適していることがわかる。

次に3つの固定効果モデルか変量効果モデルのどちらかを採択するか判断するために、それぞれにハウスマン検定を行った。その結果、Fixedは $\chi^2 = 85.555$ 、Timeは $\chi^2 = 103.13$ 、そしてTwo-wayは $\chi^2 = 147.85$ で、すべて1%有意であった。これは都道府県の個別効果と時間効果が説明変数と相関がないことを示唆している。よって、変量効果モデルはバイアスを持つ推定量であるため、固定効果モデルを採択する方が良い。結果としてTwo-way固定効果モデルを採択する。

4.2 Two-way 固定効果モデルの解釈と比較

次に、Two-way固定効果モデルの結果について考察する。まず、女性未婚率の係数は約-0.0202で、1%有意であった。これは、女性未婚率が1%上昇するたびに合計特殊出生率が0.0202減少することを意味している。次に、女性就業率の係数は約0.0143で、1%有意であった。つまり、女性就業率が1%増える

と、合計特殊出生率は0.0143増加する。男性育児時間の係数は約0.0050で、10%有意であった。これは、男性の育児時間が1分増加すると出生率は0.0050増加することを意味している。最後に、教育費の係数は0.0301で、5%有意であった。つまり、教育費支出が1万円増加すると合計特殊出生率は0.0301増加する。女性賃金、失業率、男性賃金、そして人口密度は有意でなかった。よって、合計特殊出生率に影響を与えない。

ここで、Pooled OLS、Two-way 固定効果モデル、そして変量効果モデルのいくつかの説明変数に注目する。女性未婚率の符号はすべてのモデルで負に有意であるため、個別効果や時間効果の影響は小さいと考えられる。

女性就業率と男性育児時間はPooled OLSと変量効果モデルでは有意でないが、Two-way 固定効果モデルでは正に有意である。つまり、女性就業率と男性育児時間は強い個別効果と時間効果の影響を持ち、その影響でPooled OLSと変量効果モデルでは正しく評価できていなかったと言える。この結果は日本の既存研究とは反対の結果であるが、Doepke 他 (2022) の結果と整合的であり、女性の社会進出と男性の育児参加は合計特殊出生率に正の効果を持つことを示している。さらに注目すべきなのはそれぞれの係数の大きさである。現在60%近くである女性就業率が10%増加すると合計特殊出生率は約0.143増加し、平均4分程度の男性の育児時間が10分増加すると合計特殊出生率は0.050増加する。

次に、変量効果モデルでは女性賃金が合計特殊出生率に正の効果を持つ結果が出ているが、Two-way 固定効果モデルとPooled OLSでは効果を持たない。これは足立・中里 (2007)、阿部・原田 (2008) の負の効果を持つという結果と反対である。失業率もPooled OLSと変量効果モデルでは負に有意だが、Two-way 固定効果モデルでは有意でない。これも近藤 (2014) の失業率が負の効果を持つという結果に反するものである。同様に男性賃金もPooled OLSと変量効果モデルでは負に有意だが、Two-way 固定効果モデルでは有意でない。教育費はPooled OLSでは負に有意だが、Two-way 固定効果モデルでは正に有意である。結果として、ほとんどの変数が強い個別効果と時間効果を持ち、Pooled OLSでは欠落変数のバイアスがかかった結果となることがわかる。

表 4 分析結果

被説明変数	合計特殊出生率					
	手法	Pooled OLS	Fixed	Time	Two-way	Random
切片		3.4703860*** (0.6431919)				2.4666830*** (0.3905296)
女性未婚率		-0.0165819*** (0.0059704)	-0.0197951*** (0.0042183)	-0.0101712 (0.0061588)	-0.0202728*** (0.0041479)	-0.0160130*** (0.0026131)
女性就業率		-0.0066996 (0.0057732)	0.0129268*** (0.0041299)	-0.0063122 (0.0051362)	0.0142514*** (0.0036332)	0.0029293 (0.0042308)
女性賃金		-0.0070558 (0.0056370)	0.0107713*** (0.0039581)	-0.0262562*** (0.0067967)	-0.0037174 (0.0031514)	0.0083468** (0.0037271)
失業率		-0.0319921** (0.0134045)	-0.0002225 (0.0048645)	-0.0378597 (0.0247290)	0.0024836 (0.0065815)	-0.0114624** (0.0047350)
男性賃金		-0.0135211** (0.0066130)	-0.0148873*** (0.0042959)	-0.0036432 (0.0052383)	0.0025452 (0.0035697)	-0.0174129*** (0.0046630)
男性育児時間		-0.0067661 (0.0102478)	-0.0025587 (0.0044909)	0.0220795* (0.0123471)	0.0050458* (0.0028634)	-0.0016348 (0.0039535)
教育費		-0.0790956*** (0.0245740)	0.0215985 (0.0179244)	-0.0714665*** (0.0224858)	0.0301160** (0.0123272)	-0.0117228 (0.0160235)
サンプルサイズ		235	235	235	235	235
自由度調整済み決定係数		0.44	-0.03	0.51	0.08	0.24
F統計量		27.07***	6.55***	36.99***	11.01***	82.06***

注1: ***, **, *はそれぞれ1%、5%、10%有意であることを示す。

注2: 括弧内の標準誤差はクラスター構造に頑健な標準誤差である。

表 5 モデル選択結果

F検定	F値	p値	ハウスマン検定	χ 二乗	p値
Fixed vs Pooled OLS	10.02	2.20E-16	Fixed vs Random	85.555	1.01E-15
Timed vs Pooled OLS	19.391	1.05E-13	Time vs Random	103.13	2.20E-16
Two-way vs Pooled OLS	38.053	2.20E-16	Two-way vs Random	147.85	2.20E-16

5 結論

本研究では、都道府県レベルのパネルデータを用いて固定効果モデルと変量効果モデルで合計特殊出生率の決定要因を分析した。その結果、Two-way 固定効果モデルが採択され、個別効果と時間効果を除去すると、女性未婚率が負の効果を持ち、女性就業率、男性の育児時間、教育費支出が正の効果を持つことがわかった。その他の変数は合計特殊出生率に有意な効果を及ぼしていないことがわかった。

本研究の貢献は二点ある。一点目は日本の既存研究と異なり、新たに日本でも女性の社会進出が正の効果を持つことを実証したことである。先行研究やその理論モデルでは女性賃金などが機会費用となり、出生率に負の効果を与えるとされていたが、本研究ではその反対の結果が示唆された。これは Becker (1960) や Willis (1973) などの機会費用で出生率を説明する理論モデルでは現在の日本の出生率を説明できず、Doepke 他 (2022) の主張する、収入の大きな国々では女性の社会進出は正の効果を持つようになってきたことを実証した。

二点目は男性の育児参加も正の効果を持つことを発見したことである。筆者の知る限り日本の出生率の決定要因を実証分析した先行研究において、協力的な父の指標になるものを入れて分析したものはない。そのため、男性の育児時間を用いて協力的な父の出生率への影響を明らかにしたことは、日本の出生率や少子高齢化に関する研究に貢献するものである。

本研究の政策的意義として、女性の社会進出と男性の育児時間が合計特殊出生率に正の効果を持つことから、現在の「男性は仕事、女性は家庭」という分業方式から「男性も家庭、女性も仕事」へのさらなるパラダイムシフトが出生率の改善につながることを示した。つまり、女性は社会進出を、男性は育児参加をさらに進める政策を取ることが出生率の改善につながると考えられる。

本研究の限界は二点ある。一点目は都道府県別のデータでパネルデータ分析を行ったことである。パネルデータ分析には個別効果と時間効果を除去できるメリットがあるが、合計特殊出生率へ影響する個別効果は市区町村レベルではまた異なる結果が得られると考えられる。例として、岡山県の奈義町は2014年の合計特殊出生率が2.81を記録していたが、県全体では1.49であった。このことから、市区町村レベルのパネルデータ分析にすることにより、より興味深い分析が可能になると考えられる。二点目は多重共線性の問題が発生するため、時間を通じて一定である変数をモデルに含めることができないことである。例として、Lutz (2006) や加藤 (2017) によって変数として取り上げられている人口密度は、時間を通じてほぼ一定であるため、パネルデータ分析では変数に含めることができない。

参考文献

- (1) 高山憲之他(2000)「結婚・育児の経済コストと出生力 一少子化の経済学的要因に関する一考察」『人口問題研究』56-4、pp.1-18.
- (2) 阿部一知・原田泰(2008)「子育て支援策の出生率に与える影響:市区町村データの分析」『会計検査研究』38.
- (3) 足立泰美・中里透(2017)「出生率の決定要因—都道府県別データによる分析」『日本経済研究』75、pp.63-91.
- (4) Doepke, Matthias., et al. (2022), “THE ECONOMICS OF FERTILITY: A NEW ERA”, NBER working paper series, 29948.
- (5) Becker, S. Gary. (1960), “An Economic Analysis of Fertility”, Demographic and Economic Change in Developed Countries, pp.209-240.
- (6) Willis, J. Robert. (1973), “A New Approach to the Economic Theory of Fertility Behavior”, Journal of Political Economy, Vol.81-2, pp.14-64.
- (7) Lutz, Wolfgang., Maria Rita Testa., Dustin J. Penn. (2006), “Population Density is a Key Factor in Declining Human Fertility”, Population Environment, Vol.28, pp.69-81.
- (8) 加藤久和(2017)「市区町村別にみた出生率格差とその要因に関する分析」『フィナンシャル・レビュー』131、pp.6-23.
- (9) 近藤絢子(2014)「不況と少子化 : 失業率と出生率・既婚率の都道府県パネル分析」『経済志林』81、pp.109-125.
- (10) Studenmund, A.H.(2011), “USING ECONOMETRICS: A PRACTICAL GUIDE”/高橋青天監訳、牧大樹他訳(2018)『計量経済学の使い方: 下[応用編]』ミネルヴァ書房.
- (11) 西山慶彦他(2019)『計量経済学』有斐閣.
- (12) 厚生労働省“合計特殊出生率について”、厚生労働省(最終アクセス:2022年8月30日)
<https://www.mhlw.go.jp/toukei/saikin/hw/jinkou/geppo/nengai11/sankou01.html>