

2021年度 統計データ分析コンペティション
特別賞（統計活用）〔大学生・一般の部〕

都道府県別パネルデータを用いた
学力の決定要因分析

桑名 聖人

（早稲田大学政治経済学部政治学科）

豊野 拓巳・吉田 賢汰

（早稲田大学政治経済学部経済学科）

論文の概要

学力テストスコアの都道府県別パネルデータを用いて、個人属性や家庭・社会環境、教育の質等が与える影響についてPOLS（Pooled OLS）、固定効果法、変量効果法による分析を行い、世帯教育費や通塾率、不登校率や離婚率などが学力に影響していることを示した。

論文審査会コメント

先行研究に基づき、厳密な仮説検証型の手順に準拠した手堅い実証研究を行っている。共線性、変量効果モデルの検討も含めた手堅い計量経済実証能力は評価できる。考察の自己批判も適切で一定の学術研究として成立しているが、原因と結果の因果の方向性について考察があるとよい。

都道府県別パネルデータを用いた学力の決定要因分析

桑名聖人^{*1}・豊野拓巳^{*2}・吉田賢汰^{*3}

*1: 早稲田大学政治経済学部政治学科

*2: 早稲田大学政治経済学部経済学科

*3: 早稲田大学政治経済学部経済学科

1. はじめに

わが国における 2021 年度の文教費予算は 4 兆 216 億円になり⁽¹⁾、2019 年度「教育費負担の実態調査」では世帯年収に占める教育費が 3 割を超える⁽²⁾など、現在教育に対する投資額は莫大なものとなっている。教育の目的の一つは、学力の形成にある。主にどのような要因が学力形成に対して大きく影響するのかを知るとは、能率的な教育政策の立案や世帯教育投資の促進などにとって有用である。また、効率的な教育によってもたらされる人的資本蓄積の増加がその国の経済発展を導くといった理論⁽³⁾をはじめとして、学力形成の要因を分析することの社会的意義は大きい。

国内における学力形成要因の分析は、全国レベルの包括的な学力調査である「全国学力学習状況調査」がスタートし、そのデータが蓄積されてきた 2000 年代後半からある程度活発に行われるようになった⁽⁴⁾。その先行研究として、説明変数ごとにさまざまなものが存在する。例えば、大学・大学院卒業者の割合といった地域特性が学力テストの結果と強い正の相関を持つとする研究（土屋, 2014）⁽⁵⁾や、離婚率や不登校率といった指標が近年になって児童の学力と強い相関を見せるようになったという研究（志水, 2014）⁽⁶⁾、中国の事例ではあるが気温が高いと学生のテストスコアが低くなる傾向があるとした研究（Zivin et al., 2020）⁽⁷⁾などが挙げられる。また、世帯所得と児童の学力の間に正の相関を導き学校外教育支出を媒介する効果について示唆した研究（野崎・樋口ほか, 2018）⁽⁸⁾や、社会経済的状况において不利な学校では学級規模の縮小によって学力の有意な向上がみられたとする研究（妹尾・北條, 2016）⁽⁹⁾なども存在する。学級規模、つまりクラスサイズに関する研究は妹尾・篠崎ほか（2013）⁽¹⁰⁾や Akabayashi and Nakamura（2012）⁽¹¹⁾など多数存在する。しかし、例えば野崎・樋口ほか（2018）で行われているのはクロスセクションデータによる OLS 分析のみであり、各個人に特有の効果を制御できていないため説明変数の内生性の問題が残っていると考えられる。また、全く領域分野の異なる説明変数を多数用いて学力を回帰させるような分析はそれほど活発に行われていない。

そこで本研究では、4 時点の都道府県別パネルデータを利用し、経済学における教育の生産関数という理論を用いて学力形成の要因を分析する。13 の説明変数に、被説明変数として小学校 6 年生を対象とした全国学力学習状況調査の国語と算数それぞれのテストスコア偏差値を用いて回帰した。固定効果法による分析の結果、通塾率、不登校率は国語と算数で共通して正に有意であり、世帯教育費や自治体教育費、クラスサイズ（教員一人あたり児童数）は国語のみ正に有意であった。特にクラスサイズについては、負の関係があるとする先行研究が多い中、真逆の結果を示すものとなった。

なお、本研究は 6 章から成り、1 章では研究動機と先行研究、2 章では分析に用いた変数の根拠と仮説、3 章では分析に用いたモデルと推定方法、4 章では用いたデータの出典と加工方法、5 章では分析結果とその考察、6 章で結果の貢献や政策的含意、および本研究の限界について述べる。

2. 説明変数と仮説

本研究では、教育成果を産出量とする生産関数をもとにして説明変数の選択を行った。小塩（2003）⁽¹²⁾、

また Oshio and Seno (2007)⁽¹³⁾によれば、教育の生産関数は以下の式(1)で表される。

$$Y = f(x_1, x_2, x_3) + \varepsilon \quad (1)$$

Y は教育成果である学力、 ε は誤差項を表す。 x_1 は個人属性や家庭・社会環境などであり、児童のIQや親の学歴・所得、通塾率などが挙げられる。 x_2 は学校で供給される教育の質であり、教員の学歴や経験などがこれに当てはまる。 x_3 はともに教育を受ける児童の特性を表し、クラスメイトの持つ学習能力などが代表的な例である。このうち「ピア・グループ効果」と呼ばれる x_3 の効果については、学級単位・学校単位で観察されるものであるために個票データでない統計から説明変数として取り出すのは難しく、今回は分析の対象外とせざるを得なかった。その代わりに、新たに x_4 を導入し、 x_1 、 x_2 、 x_3 のいずれの領域にも入らない変数の効果を代表させた。よって、本研究では x_1 、 x_2 、 x_4 の3つの領域に基づいて説明変数を採択した。

まず、仮説を述べるにあたり、本研究で使用する説明変数について選択理由を簡潔に説明する。 x_1 については、世帯の教育活動に対する支出と学力との関係を見るために世帯教育費と教養娯楽費を説明変数として使用した。また、先行研究より大学院進学率⁽⁵⁾・大学進学率⁽⁵⁾・不登校率⁽⁶⁾・離婚率⁽⁶⁾を用いた。さらに、野崎・樋口ほか(2018)⁽⁸⁾の研究の分析対象である世帯所得を考慮したうえで、保護者の教育に対する姿勢や地域の所得水準などの家庭・社会環境の影響をみるため、通塾率と住宅地の標準価格を変数に追加した。

x_2 では、学校教育における重要なアクターである教員のやる気と熟練度合いを測るための指標として教員給与と教員経験年数を説明変数として使用した。また、複数の研究で分析対象とされていたクラスサイズと学力の関係について再考するために、クラスサイズを算出して説明変数に加えた。そして、地方自治体の教育活動に対する支出と学力との関係を見るために、自治体教育費を変数として導入した。

x_4 では、気温と学力の負の関係を示した Zivin et al. (2020)⁽⁷⁾の研究から、年平均気温を変数に追加した。

これら13個の説明変数と被説明変数である国語と算数のテストスコアについて、先行研究の分析結果などをもとにして相関の正負について仮説を立て、以下の表1にまとめた。

表1 各説明変数の学力に対する正負の仮説

説明変数	世帯教育費	教養娯楽費	大学院進学率	大学進学率	通塾率	標準価格(住宅地)	不登校率	離婚率	教員給与	教員経験年数	クラスサイズ	自治体教育費	年平均気温
正負	+	+	+	+	+	+	-	-	+	+	-	+	-

3. モデルと推定方法

本研究においては、各都道府県の全国学力・学習状況調査の結果のうち学力調査のテストスコアを被説明変数とするパネルデータ分析を行った。パネルデータモデルでは、「個人に特有で、時間を通じて不変な観測できない外的条件」が存在し、説明変数と誤差項の内生性に対処するために制御する必要がある。田中(2015)⁽¹⁴⁾によると、これは「個別効果」といい、内生性がある場合は固定効果モデルを用いて制御することができる。なお、山本(2015)⁽¹⁵⁾によると、内生性がない場合でも、個別効果の特性によって同じ都道府県の異時点の誤差項で自己相関が生じる可能性があるため、これに対処するために変量効果モデルを利用した推定を行う必要がある。そこで、個別効果が存在するか否かということを確かめるためにF検定を行い、個別効果が存在すれば固定効果モデルあるいは変量効果モデルのいずれかを利用することになる。どちらのモデルを

利用するかということは、個別効果が説明変数と相関しているか否かによって決まるため、「帰無仮説：個別効果と説明変数は無相関」とする Hausman 検定を行った後に採択した。

まず、パネルデータ分析で使用した以下の回帰モデル(2)について検討する。

$$Y_{it} = \alpha + \sum_{j=1}^{12} \beta_j X_{j,it} + \beta_{13} X_{13,i(t-1)} + F_i + U_{it} \quad (2)$$

Y_{it} は第 i ($i = 1, 2, \dots, 47$) 都道府県の t ($t = 2007, 2010, 2013, 2016$) 年度におけるテストスコアの偏差値、 $X_{j,it}$ ($j = 1, 2, \dots, 12$)、 $X_{13,i(t-1)}$ はその説明変数である。ここで、 $X_{13,i(t-1)}$ は唯一のラグ付き説明変数である離婚率を表す。 α は定数項、 β_j ($j = 1, 2, \dots, 13$) は回帰係数であり、 F_i は各都道府県の個別効果（例えば教育文化）、 U_{it} は誤差項を表す。なお、被説明変数 Y_{it} として用いたのは小学校 6 年生を対象とした国語・算数の調査結果のみであり、教科ごとにそれぞれ Pooled OLS・固定効果法・変量効果法による推定を行い比較対照した。

また、全国学力調査に含まれている中学校 3 年生を対象とした調査結果は被説明変数として分析に使用しなかったが、これは特定の都道府県のテストスコアにバイアスがかかることを防ぐためである。中学校は小学校と比べて私立学校の比率が大きく（小学校約 1.14%、中学校約 7.52%、2016 年）⁽¹⁶⁾、また中学校総数における私立中学校の割合は都道府県ごとによって大きく異なる（最大値：東京約 23.30%、最小値：秋田約 0.86%、分散 0.16%、2016 年）⁽¹⁷⁾。中学受験を経て、都道府県をまたいだ通学が一般化すると、高偏差値の私立中学校が集中する都道府県のテストスコアには上方バイアスがかかることになる。このような効果は今回の分析の対象外であるため、中学 3 年生における調査結果は分析対象から除外した。

4. データ

13 の説明変数の出典と詳細な変換方法について、以下の表 2 に示す。

各種変数の出典元である統計データのうち、教員給与と教員経験年数のために必要となる学校教員統計調査は 3 年ごとに調査が行われるため、SSDSE-2021B にデータのある 2007 年～2018 年のうち 2007 年、2010 年、2013 年、2016 年の 4 時点を分析対象期間とした。同様に、残りの変数を裏付ける統計データについては、この 2 つの説明変数の調査時点に合わせて 4 時点のもののみ抜き出して分析に使用した。

被説明変数であるテストスコアの偏差値は、国立教育政策研究所の教育課程研究センターが実施した全国学力調査の都道府県別データをもとに算出した。2007 年、2010 年、2013 年、2016 年の平均正答率のデータからそれぞれ年ごとの平均と標準偏差を求め、各都道府県の偏差値を計算した。なお、2010 年のデータに関しては平均正答率の値が記載されていなかったため、95%信頼区間の平均値を使用した。また、本研究で用いたテストスコアはほぼ完全なパネルデータであったが、2016 年の熊本県のデータについてのみ熊本地震の影響で学力調査の実施が見送られたために欠損値となっており、全体としてはアンバランスパネルとなっている。さらに、全国学力調査の対象は公立小学校のみであり、得られたテストスコアのデータは全国の私立・国立小学校に通う児童のものを反映していなかったため、説明変数は可能な限り公立小学校に関するデータのみを用いるよう努めた。しかし、不登校児童数のデータだけは公立のみのものが存在しなかったため、公立小学校の児童数で除することによって公立小学校の不登校率を近似した。

通塾率については、全国学習状況調査の児童質問紙に記載のあるアンケート結果を利用して導出した。全ての時点の質問紙に「学習塾で勉強をしているか」と問う質問項目があり、「学習塾に通っていない」と回答した児童の割合を 1 から減じることによって通塾率を定義した。また、離婚率については 2006 年、2009 年、2012 年、2015 年のデータを用いる、つまりラグ付きの変数とした。なぜなら、離婚が学力に影響を及ぼすすれば少なくとも離婚の事実が学力の測定に先行する必要があるため、例年 4 月に行われる全国学力調査に先行した離婚率は 1 年前のデータから算出せざるを得ないからである。そして、気温の変数に関しては SSDSE-

2021B に含まれる年平均気温のデータを用いることで、全国学力調査が実施される 4 月の気温を近似した。

表 2 説明変数のデータの出典と変換方法

領域	変数名	単位	変数の説明	出典
x_1 (個人属性あるいは家庭・社会環境)	世帯教育費	円	二人以上の世帯における教育費の支出額	SSDSE-2021B
	教養娯楽費	円	二人以上の世帯における教養娯楽費の支出額	SSDSE-2021B
	大学院進学率	%	大学卒業者のうち進学者数×100/大学卒業者数	SSDSE-2021B
	大学進学率	%	高等学校卒業者のうち進学者数×100/高等学校卒業者数	SSDSE-2021B
	通塾率	%	塾に通っている児童の割合	全国学力・学習状況調査
	標準価格 (住宅地)	円	住宅地の標準価格における平均価格	SSDSE-2021B
	不登校率	人 (千人対)	不登校児童数×1000/小学校児童数	児童生徒の問題行動・不登校等生徒指導上の諸課題に関する調査、学校基本調査
	離婚率	人 (千人対)	1 年 前 の 離 婚 件 数 ×2×1000/1 年前の総人口	SSDSE-2020B
x_2 (学校で供給される教育の質)	教員給与	千円	教員の平均給料月額	学校教員統計調査
	教員経験年数	年	教員の平均勤務年数	学校教員統計調査
	クラスサイズ	人	小学校児童数 (公立) / 小学校教員数 (公立)	学校基本調査
	自治体教育費	円	自治体支出小学校教育費総額/小学校児童数 (公立)	地方教育費調査、学校基本調査
x_4 (その他の変数)	年平均気温	℃	都道府県庁所在地等の年平均気温	SSDSE-2021B

表 3 記述統計量

変数	観測数	平均	標準偏差	最小値	最大値	VIF
国語偏差値	187	49.9997	10.0271	18.3200	85.3900	-
算数偏差値	187	50.0000	10.0265	15.7600	89.95	-
世帯教育費	188	11764.9500	3328.7900	5689.0000	26910.0000	1.9500
教養娯楽費	188	29785.0300	5035.3230	17320.0000	47739.0000	2.3500
大学院進学率	188	14.1289	4.5793	5.8048	28.9118	1.4000
大学進学率	188	50.6113	6.8797	36.0852	66.1925	3.8400
通塾率	187	43.1177	8.5243	20.3000	59.4000	3.5400
標準価格 (住宅地)	188	151617.6000	231136.4000	26100.0000	1699300.0000	3.8700
不登校率	188	3.6105	1.0234	1.7813	6.9055	1.4300
離婚率	188	3.6831	0.4621	2.6818	5.3324	2.1700
教員給与	188	353.2165	20.4770	300.9000	391.8000	3.3700
教員経験年数	188	20.3484	2.1661	13.2000	24.8000	4.1600
クラスサイズ	188	15.1593	2.0780	10.9400	20.3600	6.9800
自治体教育費	188	990146.4000	144831.9000	744808.0000	1453748.0000	6.0700
年平均気温	188	15.8697	2.3698	9.2000	24.1000	1.7100

各変数の記述統計量は表 3 にまとめてある。VIF（分散拡大因子）が 10 より大きい説明変数は多重共線性が疑われるが、本研究で選択した説明変数の VIF は最大で 6.98 だったので、多重共線性の問題は重篤でないといえる。

5. 結果と考察

被説明変数が国語の偏差値の場合、算数の偏差値の場合それぞれについて、POLS（Pooled OLS）、固定効果法、変量効果法で分析した結果は表 4 の通りである。なお、固定効果法に基づいて解釈を行うが、比較対象のために、POLS や変量効果法の結果を示している。

5.1 被説明変数が国語偏差値の場合

まず、個別効果の有無を判定するために、F 検定を行ったところ、F 値は 7.82、 p 値は少なくとも小数第 4 位までは 0 であるため、1%水準で個別効果が存在した。また、固定効果法と変量効果法のどちらを採択すべきかを判断するために Hausman 検定を実施したところ、カイ二乗値は 36.86 であり、 p 値は少なくとも小数第 4 位までは 0 であるため、「説明変数と都道府県の個別効果は無相関である」という帰無仮説が 1%水準で棄却され、説明変数と個別効果の相関が確認された。したがって、以下では固定効果法による推定結果のみ解釈を行う。

世帯教育費の係数推定値は約 0.0004 で、5%水準で統計的に有意であった。これは、世帯の平均的な教育費が 1 万円増えると、国語の偏差値は約 4 増えることを意味している。次に、通塾率の係数推定値は約 0.8489 で、5%水準で統計的に有意であった。これは、小学生の通塾率が 1%増えると、国語の偏差値は約 0.85 増加することを意味している。また、不登校率の係数推定値は約 2.5092 で、5%水準で統計的に有意であった。これは、児童 1000 人あたりの不登校者数が 1 人増えると、国語の偏差値は約 2.5 増加する傾向にあることを意味している。さらに、クラスサイズの係数推定値は約 2.6857 で、10%水準で統計的に有意であった。これは、教員一人あたり児童数が 1 人増えると、国語の偏差値は約 2.7 増加する傾向にあることを意味する。加えて、自治体教育費の係数推定値は約 0.00002 で、10%水準で統計的に有意であった。これは、児童一人あたりの自治体教育費が 10 万円増えると、その都道府県の平均的な国語の偏差値は約 2 増加することを意味している。

その他の説明変数の統計的有意性は確認されなかった。したがって、その他の説明変数は、国語の学力を有意に説明する要因にはならない。

POLS の結果と比較すると、POLS において正で有意になっている大学進学率、標準価格（住宅地）や定数項、負で有意になっている離婚率は、固定効果法では有意になっていない。また、通塾率や不登校率は、POLS では負で有意になっているものの、固定効果法では正で有意になっているという違いが見られる。これは、単純な POLS による推定で発生していた内生性バイアスが、固定効果法によってある程度克服された結果もたらされた違いであると考えられる。さらに変量効果法の推定結果と比較すると、世帯教育費が正で有意になっているという点が似通っているが、固定効果法の場合ほどその他の変数の統計的有意性は確認されない。

5.2 被説明変数が算数偏差値の場合

F 検定を行ったところ、F 値は 7.50、 p 値は少なくとも小数第 4 位までは 0 であるため、1%水準で個別効果が存在した。Hausman 検定の結果、カイ二乗値は 45.35 であり、 p 値は少なくとも小数第 4 位までは 0 であるため、1%水準で固定効果法が選ばれた。したがって、以下では固定効果法による推定結果のみ解釈を行う。

表 4 分析結果

	国語の偏差値			算数の偏差値		
	POLS	固定効果	変量効果	POLS	固定効果	変量効果
世帯教育費	0.0002 (0.0002)	0.0004** (0.0002)	0.0004** (0.0002)	-0.0000 (0.0002)	0.0003 (0.0002)	0.0002 (0.0002)
教養娯楽費	-0.0001 (0.0002)	0.0000 (0.0002)	0.0000 (0.0002)	-0.0002 (0.0002)	0.0002 (0.0002)	0.0000 (0.0002)
大学院進学率	0.1725 (0.1546)	-0.1252 (0.4834)	0.1771 (0.2302)	0.2122 (0.1489)	-0.1662 (0.5468)	0.1801 (0.2193)
大学進学率	0.6964*** (0.1617)	-0.1449 (0.4062)	0.2542 (0.2252)	0.8614*** (0.1584)	-0.0590 (0.4221)	0.4531** (0.2176)
通塾率	-0.5829*** (0.1346)	0.8489** (0.3540)	-0.2165 (0.1760)	-0.5789*** (0.1497)	0.9414** (0.4541)	-0.2718 (0.1859)
標準価格(住宅地)	0.0000* (0.0000)	0.0001 (0.0001)	0.0000 (0.0000)	0.0000*** (0.0000)	0.0000 (0.0001)	0.0000 (0.0000)
不登校率	-1.4800** (0.7254)	2.5092** (1.0309)	-0.2676 (0.9821)	-1.4776* (0.7948)	3.5927** (1.4068)	0.0950 (1.2580)
離婚率	-5.5584*** (1.8998)	1.4550 (4.1691)	-3.7402 (3.0154)	-5.0180** (1.9885)	5.9339 (4.0715)	-2.3209 (3.3560)
教員給与	0.0415 (0.0537)	0.0274 (0.0516)	0.0120 (0.0489)	0.0739 (0.0507)	0.0463 (0.0517)	0.0179 (0.0455)
教員経験年数	0.2593 (0.5799)	-0.7680 (0.4967)	-0.0264 (0.4509)	0.0385 (0.6103)	-0.6720 (0.5808)	0.0289 (0.4787)
クラスサイズ	-0.3382 (0.6946)	2.6857* (1.5956)	0.5001 (0.8081)	-0.5960 (0.7569)	1.3738 (1.6334)	-0.1011 (0.6948)
自治体教育費	0.000002 (0.000010)	0.000020* (0.000010)	0.000013 (0.000010)	-0.000003 (0.000010)	0.000015 (0.000010)	0.000007 (0.000010)
年平均気温	-0.0419 (0.3797)	0.9650 (2.2390)	-0.6853 (0.6724)	0.3062 (0.3919)	0.8912 (2.5021)	-0.2788 (0.7068)
定数項	45.7490* (24.2484)	— (—)	40.0704 (30.2022)	36.5991 (24.6044)	— (—)	31.7884 (29.2665)
観測数	187	187	187	187	187	187
決定係数	0.4049	0.0730	0.3180	0.3677	0.0500	0.2752

注：括弧内は、POLS の場合は不均一分散に対して頑健な標準誤差、固定効果法と変量効果法の場合はクラスター構造に頑健な標準誤差。***, **, *はそれぞれ 1%、5%、10%水準で統計的に有意であることを示す。1%水準で有意な場合はグレー、5%水準で有意な場合は青、10%水準で有意な場合は赤で塗りつぶしている。決定係数は、固定効果法と変量効果法の場合は、全体 (overall) での決定係数を示している。

通塾率の係数推定値は約 0.9414 で、5%水準で統計的に有意であった。これは、通塾率が 1%増えると、算数の偏差値は約 0.94 増加することを意味している。また、不登校率の係数推定値は約 3.5927 で、5%水準で統計的に有意であった。これは、児童 1000 人あたりの不登校者数が 1 人増えると、その都道府県の算数の偏差値は約 3.6 増加することを意味している。

その他の説明変数の統計的有意性は確認されなかった。しかし、 p 値が 0.1 よりも大きいものの 0.1 に近かった説明変数に、世帯教育費と自治体教育費があった (p 値はそれぞれ約 0.135、約 0.113 であった)。経済的有意性の観点からは、これらの教育費の係数推定値の大きさは国語の偏差値の場合とほぼ同じで、世帯教育費 (自治体教育費) が 1 万円 (10 万円) 増えると、算数の偏差値は約 3 (1) 上昇する傾向にあるという結果であり、国語の偏差値の場合と似通った結果が得られた。

POLS の結果と比較すると、POLS において正で有意になっている大学進学率や標準価格 (住宅地)、負で有意になっている離婚率は、固定効果法では有意になっていない。また、通塾率や不登校率は、POLS では負で有意になっているものの、固定効果法では正で有意になっているという違いが見られるが、この結果は、国語の偏差値の場合と同様である。さらに、変量効果法の推定結果と比較すると、変量効果法では大学進学率のみが正で有意になっているが、固定効果法では大学進学率は有意になっていないという違いが見られた。

5.3 分析結果の考察

国語の偏差値の場合と算数の偏差値の場合とで、共通して正で有意になった変数に、通塾率と不登校率があった。まず、通塾率が高いほど学力が高くなる傾向にあることは、塾通いをする児童ほど、学校での勉強に先立って学習を進められたり、学校での勉強を補う問題演習を重ねたりする傾向にあり、それゆえに学校の勉強の定着度合いが向上することが理由であると考えられる。これは仮説で予想していた通りの結果である。また、不登校率が高いほど都道府県の平均的な学力が高くなる傾向にあることは、仮説とは反した結果である。ところが、不登校児童は不登校ゆえに学力テストを受けない傾向にあることを考えれば、不登校児童はそもそも学習に遅れが生じていて学力が低い可能性があるため、不登校児童が多い都道府県ほど不登校児童を含まない平均的な学力水準が高い傾向にある、という解釈も可能となる。

他には、世帯教育費や自治体教育費は、国語の偏差値の場合に、仮説通り正で有意になった。これは、児童の国語の学力を向上させるために、家庭において保護者が児童に教育投資することや、政府が教育費を増加させることへの実証的根拠を与えるものである。一方で、算数の偏差値の場合には、世帯教育費や自治体教育費は有意にこそならなかったが、 p 値はそれぞれ約 0.135、約 0.113 で 0.1 近くであったため、児童の算数の学力向上のためにも、教育に費用を投じることが有効である可能性がある。

最後に、クラスサイズは、国語の偏差値の場合にのみ、仮説と異なり正で有意となった。つまり、教員一人が担当する児童数が増えるほど、国語の偏差値が向上することを意味しているが、この結果は、多くの先行研究で示されていた少人数学級の有効性と真逆の帰結を示すものである。これは、国語の場合、クラスの人数が多いほど周りの児童の多様な意見に触れることができ、それが豊かな感性を育むことにつながり、結果的に国語に必要な能力を向上させているからかもしれない。一方で、算数の偏差値の場合には、クラスサイズは有意にならなかったため、クラスサイズは算数の学習のうえでは重要な要素とはいえない。

6. おわりに

本研究では、都道府県レベルのパネルデータを用いることで、国語と算数それぞれの公立小学校児童の学力の決定要因に関して、①個人属性あるいは家庭・社会環境、②学校で供給される教育の質、③その他の3つの領域に分けて固定効果法による分析を行った。その結果、国語と算数ともに、①における通塾率と不登校率が正で有意に学力に影響を与えていることがわかった。また、国語の場合には、①に属する世帯教育費と、②に属するクラスサイズおよび自治体教育費が、有意に学力に正の影響を及ぼすことがわかった。その他の変数は、学力に有意な影響を及ぼしていなかった。

本研究の分析は、以下の2点で貢献がある。1つ目に、先行研究では学力に対して効果があるといわれていたクラスサイズの縮小であるが、国語の場合には、逆にクラスサイズの拡大が学力を向上させることがわかり、既存研究とは異なる結果が得られた。2つ目に、国語と算数の2つの教科に分けて分析を行い、両者で異なる点を比較検討したことが挙げられる。その結果、クラスサイズや教育関連費用の面では異なる結果が得られた。

本研究の分析結果は、以下の2点において政策的含意を持つ。1つ目に、自治体教育費は国語の学力に正の影響を持つことから、小学生の国語の学力向上のために、自治体には教育予算を増やす政策が求められるといえる。一方で、算数の場合には自治体教育費は有意ではなかったものの、有意に近いといえるため、算数に関しても同じような政策提言が可能である。2つ目に、国語の場合においてのみ、クラスサイズが学力に正の影響を持つことから、国語の学力向上のために、学校現場において教師一人あたり児童数を増やすことに意味があるといえる。

しかし、本研究の分析には以下の3つの限界がある。1つ目に、都道府県レベルのデータしか用いることができなかった点が挙げられる。都道府県レベルのデータは集計データであるため、個人の情報が平均点などの一つの数値にまとめられてしまい、多様な個人属性を考慮できない。この問題はしばしば生態学的誤謬として知られ、集計データだけに基くと、個人の行動に関して誤った判断をしてしまう危険がある（浅野・矢内，2018）⁽¹⁸⁾。したがって、より正確な分析を行うためにも、都道府県レベルの集計データではなく、小学校教育の個票データを用いる必要がある。2つ目に、内生性の問題への配慮が不十分であった点が挙げられる。今回の分析では、例えば通塾率や不登校率は学力の決定要因となる一方で、逆に通塾率や不登校率が学力から影響を受けている、逆の因果性が存在している可能性が高い。こうした逆の因果性があると、係数推定値は単なる相関関係を示しているにすぎず、本来の因果関係を反映していない可能性がある。そこで、逆の因果性に対処するために、追加的な操作変数を用いた操作変数法を使用することが案として考えられる（山本，2015）。3つ目に、算出方法に疑問があるような説明変数を使用してしまったことが挙げられる。不登校率とクラスサイズについては、先行研究と異なる結果が得られたが、双方がデータの制約のために十分に適切な代表の仕方を遵守できなかったことは否定できない。特にクラスサイズの算出の際には、担任を務めていない教員も含んだ全ての教員数で除してしまっていたため、平均が15.16と過小推定される原因になったと考えられる。再びクラスサイズを変数に用いる際には、やはりミクロな個票データを用いて代表させることが必要となる。

参考文献

- (1) 文部科学省 HP：“令和3年度予算のポイント” 文部科学省（最終アクセス；2021年8月31日）.
https://www.mext.go.jp/content/20201218-mxt_kaikesou01-000010167_1.pdf
- (2) 日本政策金融公庫 HP：“教育費負担の実態調査結果” 株式会社日本政策金融公庫（最終アクセス；2021年8月31日）.
https://www.jfc.go.jp/n/findings/pdf/kyouikuhi_chousa_k_h30.pdf
- (3) 川口大司：“労働経済学—理論と実証をつなぐ”、有斐閣（2017）
- (4) 川口俊明：“日本の学力研究の動向 Trends of Academic Achievement Research in Japan” 福岡教育大学紀要、第68号、第4分冊、1-11（2019）
- (5) 土屋隆裕「学校の地域特性」、国立大学法人お茶の水女子大学編：“平成25年度全国学力・学習状況調査（きめ細かい調査）の結果を活用した学力に影響を与える要因分析に関する調査研究”，pp.156-206（2014）
- (6) 志水宏吉：“「つながり格差」が学力格差を生む”、亜紀書房（2014）
- (7) Joshua Graff Zivin, Yingquan Song, Qu Tang, Peng Zhang “Temperature and high-stakes cognitive performance: Evidence from the national college entrance examination in China” *Journal of Environmental Economics and Management*, vol.104, pp.15-17（2020）
- (8) 野崎華世、樋口美雄、中室牧子、妹尾渉：“親の所得・家庭環境と子どもの学力の関係：国際比較を考慮に入れて”、NIER Discussion Paper Series No.008、pp.19-24（2018）
- (9) 妹尾渉、北條雅一：“学級規模の縮小は中学生の学力を向上させるのか-全国学力・学習状況調査（きめ細かい調査）の結果を活用した実証分析-”、国立教育政策研究所紀要、第145集、pp5-8（2016）
- (10) 妹尾渉、篠崎武久、北條雅一：“単学級サンプルを利用した学級規模効果の推定”、国立教育政策研究所紀要、第142集、pp.167-172（2013）
- (11) Hideo Akabayashi, Ryosuke Nakamura：“Can Small Class Policy Close the Gap? An Empirical Analysis of Class Size Effects in Japan” *Tokyo Center for Economic Research Paper No.E-51*、pp.27-28（2012）
- (12) 小塩隆士：“教育を経済学で考える”、日本評論社（2003）
- (13) Takashi Oshio, Wataru Seno “The Economics of Education in Japan” *The Japanese Economy*, vol.35, pp.57-58（2007）
- (14) 田中隆一：“計量経済学の第一歩”、有斐閣（2015）
- (15) 山本勲：“実証分析のための計量経済学”、中央経済社（2015）
- (16) 文部科学省：“学校基本調査 平成28年度小学校・中学校 都道府県別学校数” 文部科学省（2016）
- (17) 文部科学省：“学校基本調査 平成28年度中学校 都道府県別学校数” 文部科学省（2016）
- (18) 浅野正彦、矢内勇生：“Rによる計量政治学”、オーム社（2018）