

Institute for Economic Studies, Keio University

Keio-IES Discussion Paper Series

教育長は重要か？—大規模学力調査データを用いた分析

杉田壮一郎

2024年10月12日

DP2024-021

<https://ies.keio.ac.jp/publications/24298/>

Keio University



Institute for Economic Studies, Keio University
2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan
ies-office@adst.keio.ac.jp
12 October, 2024

教育長は重要か？—大規模学力調査データを用いた分析

杉田壮一郎

IES Keio DP2024-021

2024年10月12日

JEL Classification: I21, I22, M12

キーワード: 教育長、アカウンタビリティ、差の差法

【要旨】

米国を中心に教育の管理職のマネジメント能力やリーダーシップに関する研究が行われてきたが、日本の文脈における実証例は少ない。本研究では、匿名都道府県における学力調査のパネルデータを利用し、市町村教育長の交代というイベントが子どもの学力に与える影響を分析した。分析より、教育長交代が学力に与える影響は市町村ごとに異なり、正の影響と負の影響が0を平均として比較対称な分布をなすこと、国語・数学の各スコアに対する教育長交代イベントの影響は、互いに正の相関関係にあることがわかった。本稿の結果より、日本における教育長のマネジメント・リーダーシップ能力の高さは、子どもの学力の規定要因の1つであることが示唆される。

杉田壮一郎

慶應義塾大学大学院政策・メディア研究科後期博士課程

sugita3520@keio.jp

謝辞：本稿は指導教授である中室牧子教授の承認を得て投稿している。また、中室教授の他、現サイバーエージェントの伊藤寛武氏には本稿の執筆の上で多くのアドバイスをいただいた。記して感謝したい。また、本稿にありうべき主張、誤りの一切の責任は筆者個人に帰する。

教育長は重要か？——大規模学力調査データを用いた分析

杉田 壮一朗*

要旨

米国を中心に教育の管理職のマネジメント能力やリーダーシップに関する研究が行われてきたが、日本の文脈における実証例は少ない。本研究では、匿名都道府県における学力調査のパネルデータを利用し、市町村教育長の交代というイベントが児童・生徒の学力に与える影響を記述的・探索的に分析した。分析より、教育長交代が学力に与える影響は市町村ごとに異なり、正の影響と負の影響が0を平均として比較的対称な分布をなすこと、国語・数学の各スコアに対する教育長交代イベントの影響は、互いに正の相関関係にあることがわかった。本稿の結果より、日本における教育長のマネジメント能力・リーダーシップの高さは、児童・生徒の学力を規定する要因の1つであることが示唆される。

JEL Classification: I21, I22, M12

キーワード — 教育長, アカウンタビリティ, リーダーシップ, 差の差法, 教育生産関数

1 はじめに

本稿は、教育長が児童・生徒の学力に与える影響を匿名都道府県における大規模な学力調査データを用いて検証したものである。本稿は、校長や教育長といった、教育における管理職のマネジメント能力・リーダーシップが与える効果に関する研究群の1つに位置づけられる¹⁾。

米国を中心とした先行研究では、主に校長が児童・生徒の学力等のアウトカムに与える効果が検証されてきた。校長の効果を検証した文献には、(1) 教員に対する付加価値モデル²⁾を校長レベルに応用し、校長一人ひとりの「校長付加価値（アウトカムへの寄与度）」を計測したもの (Coelli and Green, 2012; Grissom, Kalogridis and Loeb, 2015; Branch, Hanushek and Rivkin, 2012)、(2) 校長の交代・異動などを利用して、児童・生徒のアウトカムに与える影響を分析したのものがある。児童・生徒と校長の情報が紐づけられ

*慶應義塾大学大学院政策・メディア研究科後期博士課程 (sugita3520@keio.jp)。本稿は指導教員である中室牧子（慶應義塾大学総合政策学部）教授の承諾を得て投稿している。

たパネルデータが利用可能な場合は(1)、そうでない場合は(2)のアプローチがとられるケースが多い。

(1)に含まれる先行研究では、校長付加価値の分布の下位に位置する校長を、分布の上位に位置する校長に置き換えることで、児童・生徒の学力・非認知能力が向上することを報告する研究が一定数存在する (Grissom et al., 2015; Dhuey and Smith, 2014, 2018)。(2)に関する先行研究では、校長の異動がアウトカムの時系列変動に対する影響を推定する「分散推定」と呼ばれるアプローチなどが代表的であり、校長が児童・生徒のアウトカムに対してプラスの効果を与えることを報告している (Coelli and Green, 2012; Branch et al., 2012)。

教育長が児童・生徒のアウトカムに与える効果を報告する文献も、校長に関する文献に比して少ないものの存在する。数少ない研究の1つであり代表的なものが Lavy, Rachkovski and Boiko (2023) であり、イスラエルの行政データを用いて教育長が児童・生徒の学力に与える効果を検証し、教育長の質が児童・生徒のテストスコアや行動に正の影響を与え、特に教育の質が相対的に低い学校において効果が大きいことを報告している。

これらの先行研究より、校長や教育長といった教育におけるリーダーは、児童・生徒の学力等のアウトカムに対して無視できない影響を与えていると考えられる。日本の教育行政制度においては、教育長 (Superintendent) とは自治体における教育行政の第一義的な責任者であり、教育長は地域の教育行政における中心的人物である³⁾が、その効果は必ずしも明らかでない。米国の教育行政制度では、教育長や校長に対してテストスコアの結果に対するアカウントビリティが強く求められるが、日本の教育長・校長はそのような文脈に置かれていない。そのため、日本の文脈⁴⁾における教育のリーダーの効果を実証的に検証する必要がある。

以上の背景より本稿では、日本における市町村教育長が児童・生徒のアウトカムにどの程度の影響を与えているか、つまりどの程度重要であるかを、大規模学力調査のパネルデータを用いて記述的・探索的に分析した。分析の結果、(a) 教育長交代イベントの効果は0を分布の中心 (平均) として、国語では $-0.08SD$ (標準偏差) から $0.2SD$ の間、数学⁵⁾では $-0.15SD$ から $0.2SD$ の間で単峰型かつ対称的な分布をしていること、(b) 国語に対する教育長交代イベントの影響の大きさと、数学における教育長交代イベントの影響の大きさは正の相関関係にあることがわかった。

ただし、本稿の研究デザインならびに分析に使用したデータについては、次の3つの留意点がある。第1に、本稿では教育長一人ひとりに対する付加価値の推定は行っていない。各市町村に対して教育長1名が対応するため、教育長ごとのダミー変数と市町村ダミー変数との間には完全な多重共線性が生じ、市町村固定効果から教育長付加価値を識別することはできない。このような背景から、本稿では教育長個人の付加価値パラメータを推定するのではなく、「教育長が交代する」というイベント (以下、教育長交代イベント) に着目した分析を行っている。具体的にはアウトカムに与える影響を2方向固定効果 (Two-way Fixed Effect: TWFE) ・差の差法 (Difference in Differences: DID) によって推定

し、推定値の分布の図示や記述統計量の計算、推定値の教科間の相関分析などの種々の分析を行い、教育長が児童・生徒の学力へ与える影響を記述的・探索的に考察している。

第2に、本稿では、データ使用上の制約のため教員・校長を対象とした分析は行っていない。本来は教育長だけでなく教員や校長の付加価値も計測できることが理想的であるが、本稿で使用した学力調査データには児童・生徒のデータと紐づけられた教員IDは存在しない。また、各学校名の照合を必要とする校長のデータは児童個人情報保護の観点から執筆時点では使用することができなかった。

第3に、教育長交代イベントが発生するタイミングは市町村により異なるため、本稿におけるTWFE/DIDによる分析はいわゆるStaggered Designに該当する。Staggered Designのもとでは、TWFE/DIDによる推定値にバイアスが生じる場合があり、近年いくつかの修正方法が提案されている(Callaway and Sant'Anna, 2021; Goodman-Bacon, 2021; Sun and Abraham, 2021)。本稿ではCallaway and Sant'Anna (2021)で提案された推定量を用いた推定結果もあわせて考察し、結果の頑健性を確認した。

以下、本稿の構成は次の通りである。第2節では使用データの詳細を明らかにし、記述統計量や教育長に関する基本的なデータを示す。第3節では本稿における推定戦略を提示し、第4節では推定結果を報告する。最後に第5節ではディスカッション最終的な結論を提示し、本稿の研究上の限界や議論、今後の研究の展望をまとめる。

2 データ

2.1 使用データについて

本稿における分析では、ある都道府県において行われた学力調査(以下、X県学力調査)の2016年から2019年の児童・生徒個票データを利用した。X県学力調査は、X県内全市町村の小学校4年生から中学校3年生までの全ての児童・生徒を対象に実施されており、今回は国語と数学のテストスコアを利用する。

X県学力調査データには児童・生徒個人にIDが付与されており、同一個人を経年追跡することが可能なパネルデータとなっている。また、IRT(項目反応理論)に基づいたスコアの算出を行っており異なる年度のテストスコアを比較することが可能である。さらに、X県学力調査データにはテストスコアのデータのほか、通塾の有無や家庭における蔵書数など児童・生徒の家庭のSES(社会経済的地位)に関する質問項目を含んだ児童・生徒質問紙も含まれる。

2.2 記述統計量

本節では、推定に用いるサンプルの記述統計量を表1に示す。推定では教育長交代イベントの効果を捉えるため、すべての学年をプールしたパネルデータを用いる。ここでは、標準化などの処理を行う前の変数に対する記述統計量を示している。後述の分析では、解釈をわかりやすくするため、国語・数学のテストスコア、ならびに学校規模・クラスサイ

	N	平均	標準偏差	最小値	最大値
処置・時点に関する変数					
教育長交代後ダミー	922967	0.24	0.43	0.00	1.00
交代イベントあり（処置群ダミー）	922967	0.43	0.49	0.00	1.00
アウトカム					
国語スコア（素点）	922967	0.45	1.52	-7.34	6.17
数学スコア（素点）	922967	0.41	1.33	-5.80	5.80
生徒に関する変数					
女子ダミー	922967	0.49	0.50	0.00	1.00
通塾ダミー	922967	0.68	0.47	0.00	1.00
蔵書ありダミー	922967	0.89	0.31	0.00	1.00
学校に関する変数					
学校規模	922967	357.35	175.99	4.00	961.00
クラスサイズ	922967	32.79	4.79	1.00	42.00

表 1: 記述統計量

ズなどの連続なコントロール変数は平均 0、標準偏差 1 になるよう標準化した。また、使用するサンプルのうち、欠測のある値を含むレコードは推定に用いていない。

2.3 教育長に関するデータ

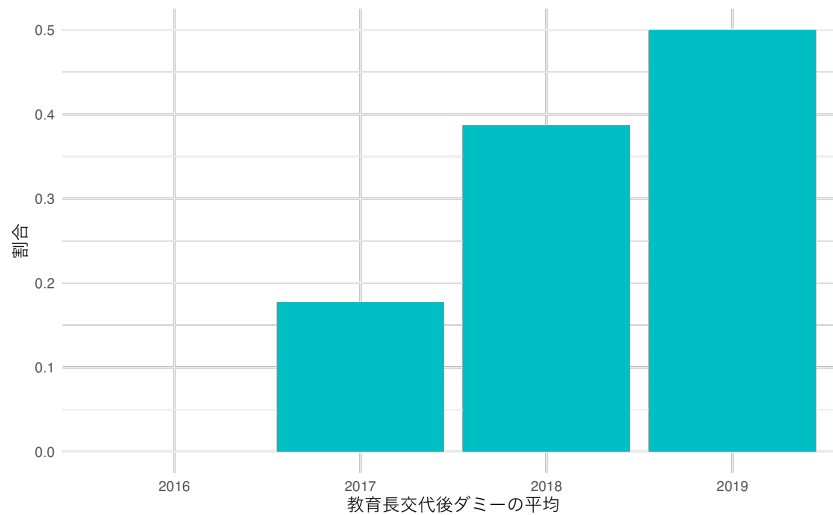
本稿では、観測期間内における X 県の各市町村で誰が教育長を務めているのかを X 県の公表情報から調べ、教育長の交代に関するデータを構築した。本稿の分析で用いたサンプルでは、図 1 では、各年ごとにどれくらいの割合の市町村が教育長交代イベントを経験しているのかを示しており、棒グラフの高さは各時点までに教育長交代を経験した市町村を累積し全体の市町村数で割って算出した割合に等しい。図 1 から、各年で X 県全体の市町村のうち 10~15% 程度に当たる市町村で教育長交代イベントが発生しており、最終的に 2016 年から 2019 年までの間に約半分の市町村が教育長の交代を経験している⁶⁾。

図 1 からわかるように、教育長交代イベント（処置）のタイミングは市町村によって異なる。したがって、本稿における TWFE / DID による分析は、いわゆる Staggered Design に該当する。

3 推定戦略

3.1 教育長交代イベントの有無による市町村間比較

使用するサンプルのうち、「観測期間内で教育長交代イベントを経験した市町村（処置群）」と「観測期間内で教育長交代イベントがなかった市町村（対照群）」の 2 つのグループの間に、統計的にどのような差が見られるかを記述的に分析する。具体的な分析枠組みは以下の 3.1.1 ならびに 3.1.2 に記載する。



注) 各市町村についてすでに処置を経験している場合に1をとるダミー変数（処置後ダミー）を定義し、各年における平均値を算出している。

図 1: 教育長交代イベント発生後の市町村の割合

3.1.1 2群間での共変量・市町村属性の比較

処置群の市町村と対照群の市町村との間で、児童・生徒や学校の共変量、市町村ごとの属性がどの程度異なるかを、バランステスト（平均値の差）の形式によって比較する。これにより、教育長交代イベントが生じた（生じなかった）市町村の特徴を考察する。

3.1.2 処置群ダミーによる回帰

処置群ダミー変数 Z を「(各市町村で) 観測期間内に教育長交代イベントがあった」場合に1をとる変数と定義する。

$$Y_{icst} = \beta Z_c + \mathbf{x}_{ist}^T \boldsymbol{\gamma} + \mu_s + \lambda_t + \epsilon_{icst} \quad (1)$$

\mathbf{x} は児童・生徒、学校に関する共変量のベクトルであり、通塾ダミー、蔵書ダミー、女子生徒ダミー、学校規模、学級規模が含まれる、添字の i は児童・生徒個人、 c は市町村、 s は学校、 t は時点(年度)を表す。 μ_s は学校固定効果、 λ_t は時点の固定効果である。 Y は項目反応理論 (IRT) に基づき算出されたテストスコア (国語・数学) を、科目ごとに年・学年単位で平均0、標準偏差1となるよう標準化したアウトカムである。つまり、アウトカムは、同じコホート内での相対的な位置を示す指標となる⁷⁾。

本稿では、式 (1) に基づき、「共変量のみモデルに含める (固定効果なし)」、「共変量・固定効果ともモデルに含める」の2つの特定化で回帰分析を行う。これにより、教育長交代イベントという処置の有無によって児童・生徒のアウトカムに統計的に有意な差が生じているかを検証する。

3.2 TWFE / DID 推定

教育長交代イベントの効果を推定するため、次のような2方向固定効果 (Two-way Fixed Effect: TWFE) 回帰を考える。

$$Y_{it} = \alpha_i + \lambda_t + \delta D_{it} + \epsilon_{it} \quad (2)$$

式 (2) における TWFE 推定は本質的には差の差推定 (Difference in Differences: DID) と同一であるため、以下では便宜上、 δ の推定値 (または推定量) を TWFE / DID 推定値 (推定量) と表記する場合がある。

添字の i は生徒個人、 t は時点 (年度) を表す。 α_i は個人 (ユニット) の固定効果、 λ_t は時点の固定効果である。 D は個人 i が処置群 (教育長交代イベントが観察期間内に生じた市町村) に属するか t が教育長交代イベントが生じた後の時点である場合に 1 をとるダミー変数である。 Y は 3.1.2 節で説明したように、標準化されたテストスコア (国語・数学) である。

式 (2) の定式化に基づく推定では、共変量を含んでいない。これにはいくつかの理由がある。第 1 に、TWFE / DID 推定における技術的な問題がある。具体的には、個人固定効果を制御しているため児童・生徒の性別や時間を通じて不変な属性をモデルに含められない (識別できない) こと、時間を通じて変化する共変量を入れた場合は処置変数が共変量の影響を受ける場合があるため注意すべきであること、共変量を含んだ場合に条件付き平行トレンドの仮定における定式化の誤りに対して頑健とは限らないことが挙げられる⁸⁾。本稿では記述的・探索的に教育長の影響を考察することが主眼にあるため、個人ならびに時間の固定効果のみを含んだ TWFE / DID をベースラインの推定とし、学校固定効果を含めた場合の特定化とあわせて報告する。

第 1 節ならびに第 2 節で述べたように、教育長交代イベントの発生タイミングは市町村によって異なるため、本稿におけるデザインはいわゆる Staggered Design に該当する。本稿では、Callaway and Sant'Anna (2021) で提案された手法を用いた分析結果もあわせて、報告する。また、平行トレンドの成立については 4.3 項にてイベントスタディプロットとともに言及する。

4 推定結果

4.1 教育長交代イベントの有無による市町村間比較の結果

4.1.1 2 群間での各変数のバランス

表 2 は、教育長交代イベント (処置) の有無によって、各変数の標本平均に差があるかどうかを示したものである。表より、教育長交代イベントが生じた市町村とそうでない市町村との間では、各変数に大きな差が生じていないものと考えられる。

バランステストの形式で、平均値の差の検定における p 値を念の為付記している。0.1% 水準で統計的に有意な結果が多いものの、本稿で用いたデータのサンプルサイズは 90 万以上であるため検出力が高いことを考慮に入れると、表 2 の結果は実質的に有意な差を示したのではないという解釈に至った。

	交代なし (対照群)	交代あり (処置群)	p 値
国語スコア	0.46 (1.52)	0.43 (1.53)	<0.001
数学スコア	0.42 (1.33)	0.39 (1.33)	<0.001
女子ダミー	0.49 (0.50)	0.49 (0.50)	0.015
通塾ダミー	0.68 (0.47)	0.68 (0.47)	0.008
蔵書ダミー	0.89 (0.31)	0.89 (0.31)	<0.001
学校規模	359.49 (178.89)	354.46 (171.96)	<0.001
クラスサイズ	32.93 (4.68)	32.59 (4.93)	<0.001
N	530483	392484	

注) N はサンプルサイズ。表の各セルの値は平均値 (標準偏差) の形で表示している。

表 2: 教育長の交代の有無による比較

4.1.2 処置群ダミーによる回帰分析の結果

表 3 に、処置群ダミーを用いた回帰分析の結果を示す。処置群 (教育長交代イベントあり) ダミーの係数の推定値は、列 (1) から列 (4) のすべてにおいて、帰無仮説 $H_0: \beta = 0$ に対して統計的に有意でない。したがって、観測期間内における教育長交代イベントの有無によって差が生じるかどうかはこの分析結果からは明らかでない。

教育長交代イベントの有無、つまり DID における処置群と対照群⁹⁾ の単純な差に基づく分析では、学力への影響に差が認められなかった。4.2 項では、教育長交代イベントの差の差に着目した分析を行い、市町村ごとの影響の大きさを考察する。

4.2 TWFE / DID 推定

4.2.1 TWFE / DID の推定結果

表 4 に TWFE / DID の推定結果を示す。国語スコアに対する推定結果である列 (1) および列 (2)、数学スコアに対する推定結果である列 (3) および (4) では、処置 (教育長交代イベント後) ダミーと処置群ダミーの交差項は統計的に有意ではなく、帰無仮説 $H_0: \delta = 0$ は棄却されない。

教育長が交代した場合、「相対的に優秀な教育長が着任する」ケースと「相対的に優秀でない教育長が着任する」ケースがありうる。そのため、正の影響と負の影響が相殺し合い、教育長交代というイベントの X 県全体での効果は 0 に近づくという解釈は直観的にも整合性がある。

	国語		数学	
	(1)	(2)	(3)	(4)
処置群 (交代あり) ダミー	-0.020 (0.021)	-0.025 (0.046)	-0.026 (0.020)	-0.041 (0.055)
女子ダミー	0.283*** (0.005)	0.283*** (0.005)	-0.030*** (0.005)	-0.030*** (0.005)
通塾ダミー	0.041*** (0.006)	0.025*** (0.005)	0.172*** (0.009)	0.167*** (0.008)
蔵書ダミー	0.514*** (0.011)	0.484*** (0.009)	0.431*** (0.010)	0.402*** (0.009)
クラスサイズ (標準化済み)	0.010** (0.004)	0.017*** (0.004)	0.005 (0.004)	0.012*** (0.004)
Observations	922 967	922 967	922 967	922 967
FE: Year		X		X
FE: School		X		X

注) 係数の推定値について***は 0.1%, ** は 1%, *は 5% で有意であることを示す。() 内の値は各推定値に対する標準誤差を示す。標準誤差は、市町村を単位としたクラスター構造に頑健な標準誤差である。アウトカムは標準化されているため、表内における数値の単位は標準偏差 (SD) となる。

表 3: 教育長交代イベントの有無 (処置群ダミー) による回帰分析の推定結果

4.2.2 市町村ごとの影響の分布

各市町村における教育長交代イベントの影響の差を調べるため、処置後ダミーと処置群ダミーの交互作用項に市町村ごとのダミー変数を掛け合わせた交互作用項を用いた推定を行った。得られた推定値の分布を示したのが図 2 である¹⁰⁾。パネル (a) が国語、パネル (b) が数学のスコアに対する推定値の分布である。どちらのパネルも分布は 0 付近に値が集中し、単峰型であることがわかる。

表 5 は得られた推定値の記述統計量を示す。この表と図 2 を合わせると、教育長交代イベントの市町村ごとの影響の大きさは、分布の中心 (平均) が 0 で、国語が約 -0.08SD から約 0.2SD まで、数学が約 -0.15SD から約 0.2SD までの間に分布することがわかる。言い換えれば、教育長の交代が学力にポジティブな影響を与える場合とネガティブな影響を与える場合のどちらも認められる。市町村ごとで TWFE / DID における交互作用項の推定値が正と負のどちらにも分布することは、表 4 における、市町村全体の推定結果において 0 と有意に異なることと整合的である。

4.2.3 市町村ごとの影響の教科間の相関関係

図 3 は、縦軸に数学のスコア、横軸に国語のスコアに対する 4.2.2 で得られた推定値の大きさをとった散布図である。図より、国語のスコアに対する推定値の大きさと数学に対する推定値の大きさは、正の相関関係にあることが視覚的にわかる。実際に、国語と数学それぞれの推定値の相関係数を計算すると 0.68 であり、正の相関関係が認められる。

上記の結果より教育長交代イベントが国語と数学どちらかのスコアに正 (負) の影響を

	国語		数学	
	(1)	(2)	(3)	(4)
教育長交代後×処置群	0.003 (0.011)	0.002 (0.009)	0.004 (0.017)	0.001 (0.008)
Observations	922 967	922 967	922 967	922 967
FE: Student	X	X	X	X
FE: Year	X	X	X	X
FE: School		X		X

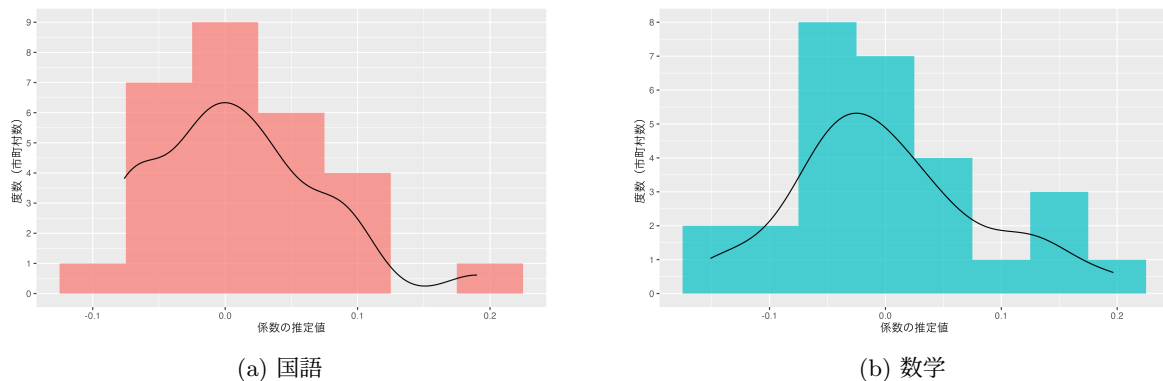
注) 係数の推定値について***は 0.1%, ** は 1%, *は 5%で有意であることを示す。() 内の値は各推定値に対する標準誤差を示す。標準誤差は、市町村を単位としたクラスター構造に頑健な標準誤差である。アウトカムは標準化されているため、表内における数値の単位は標準偏差 (SD) となる。

表 4: TWFE / DID の推定結果

	N	平均	標準偏差	最小値	最大値
国語の推定値	28	0.010	0.064	-0.076	0.190
数学の推定値	28	0.002	0.083	-0.151	0.197

注) 国語、数学のそれぞれのスコアに対する「処置後ダミー×処置群ダミー×市町村ダミーの交互作用項の推定値」の記述統計量である。アウトカムは標準化されているため、表内における数値の単位は標準偏差 (SD) となる。

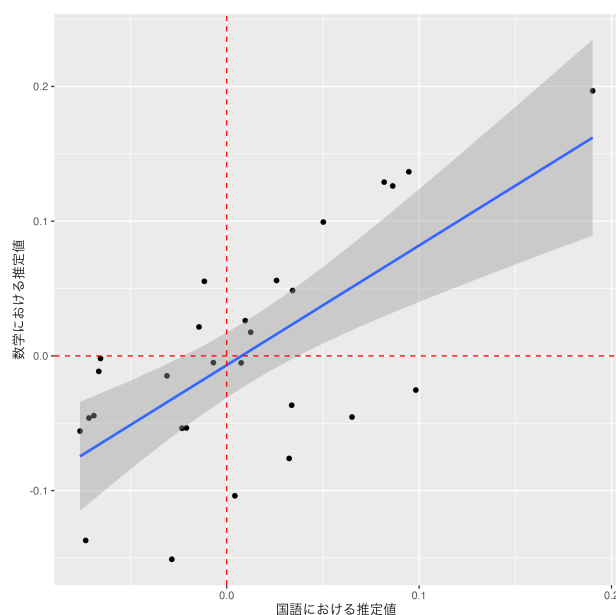
表 5: 交互作用項推定値の記述統計量



注) 国語、数学のそれぞれのスコアに対する「処置後ダミー×処置群ダミー×市町村ダミーの交互作用項の推定値」の推定値に対するヒストグラムである。アウトカムは標準化されているため、表内における数値の単位は標準偏差 (SD) となる。黒実線は密度関数のカーネルプロットである。

図 2: 交互作用項 (処置後ダミー×処置群ダミー×市町村ダミー) 推定値のヒストグラム

与えている場合、もう一方の科目についても正（負）の影響を与えていると解釈される。本結果は国語の学力を向上させる（低下させる）ような教育長に交代した場合は、その教育長は数学の学力も向上させる（低下させる）傾向にあることを含意する。



注) 縦軸、横軸共に単位は標準偏差。縦軸は数学、横軸は国語スコアに対する処置後ダミー×処置群ダミー×市町村ダミーの交互作用項の推定値であり、両者の相関係数は 0.68。網掛け部分は 95%信頼区間。

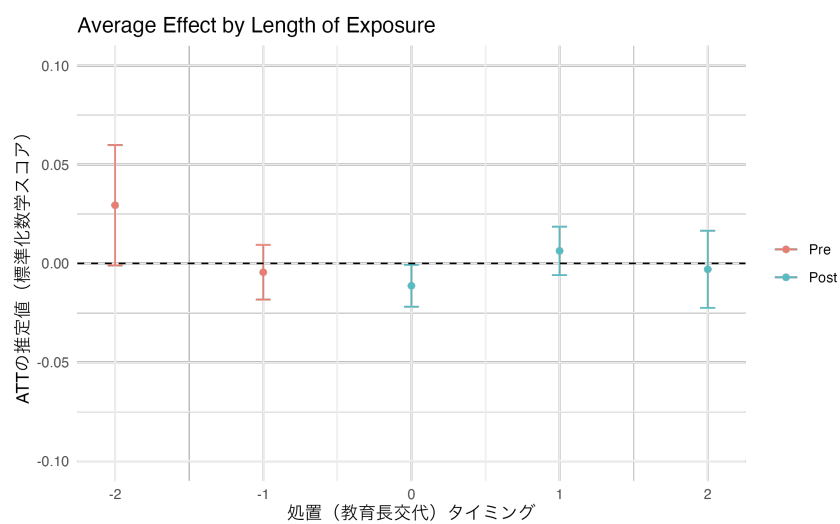
図 3: 交互作用項（処置後ダミー×処置群ダミー×市町村ダミー）推定値の散布図

4.3 イベントスタディプロットの図示

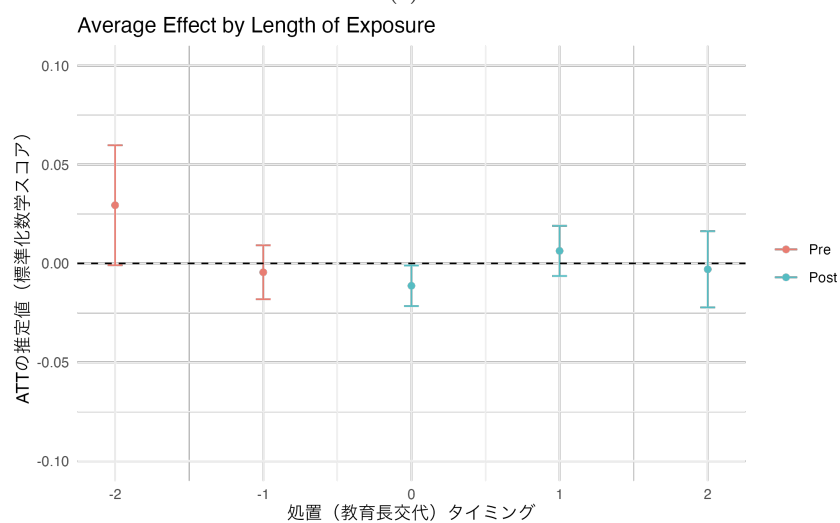
第 1 節および 3.2 項では、本稿における分析のデザインがいわゆる Staggered Design に該当することを述べた。Staggered Design においてナイーブな TWFE / DID 推定を行った場合、バイアスが生じることが近年の計量経済学の研究で指摘されている。そのため、Staggered Design における TWFE / DID 推定を行う場合は、ナイーブな方法での推定結果に加え、近年の研究で指摘されているいくつかの修正推定量に基づく分析結果を示すことが望ましいと考えられる。

本節では、Callaway and Sant’Anna(2021) で提唱された、処置のグループ・時間ごとの処置群平均処置効果 (ATT) を推定し、それらを集計した結果に基づくイベントスタディプロットを示す。

図 4 に、Callaway and Sant’Anna(2021) で提唱された方法に基づき作成したイベントスタディプロットを示してある。縦軸は推定値の大きさを、横軸は処置タイミングを 0 とした相対的な時点を示している。図より、各時点の推定値は統計的に有意に 0 と異なる (11)。事前期間 (図における Pre の点) における推定値も統計的に有意ではなく、事前トレンドによる TWFE / DID 推定への影響は深刻でないと推察される。



(a) 国語



(b) 数学

注) 各時点における点推定値は、Callaway and Sant'Anna (2021) で提案された方法に基づきグループ・時間ごとに推定された ATT を、処置タイミングからの相対的な時点ごとに集計した値。エラーバーは 95%信頼区間。

図 4: TWFE / DID 推定値のイベントスタディプロット

5 ディスカッション・結論

本稿では、市町村教育長が児童・生徒の学力に与える影響がどの程度であるかを、匿名都道府県における大規模な学力調査データを用いて記述的・探索的に検証した。

本稿における分析の結果は次のようにまとめられる。(a) 学力スコア、児童・生徒個人の属性、学校規模やクラスサイズなどに教育長交代イベント（処置）の有無による大きな差は認められない。(b) 市町村ごとの教育長交代イベントの影響は、中心（平均）が0かつ単峰型で、国語が約-0.08SDから約0.2SDまで、数学が約-0.15SDから約0.2SDまでの間に分布する。つまり、交代する教育長の違いが学力に正の影響・負の影響のどちらも与えうる。(c) 教育長交代が国語のテストスコアに与える影響と、教育長交代が数学のテストスコアに与える影響は正の相関関係にあり、国語スコアにポジティブな影響を与える教育長は数学スコアにもポジティブな影響を与える傾向がある。したがって、教育長は児童・生徒の学力を左右する規定要因の1つとして重要度が高いものと考えられる。

しかし、上述の分析では教育長交代というイベントと学力スコアへの影響を記述的・探索的な統計分析によって考察したのみであり、教育長がなぜ児童・生徒の学力等のアウトカムに対する影響を持つのかに関するメカニズムを検討する必要がある。さらに本稿には、研究上の限界もある。

以下、本節ではメカニズムに対する先行研究からの考察、ならびに本稿の研究上の限界点を述べる。

5.1 教育長が児童・生徒の学力へ与える影響のメカニズム

5.1.1 経済学における実証研究

第1節で述べたように、公教育における管理職のマネジメント・リーダーシップに関する先行研究のうち、教育長が児童・生徒のアウトカムに与える影響を考察したものは限られている。Lavy et al. (2023) はイスラエルの教育行政における教育長の効果を検証した数少ない経済学の先行研究である。Lavy et al. (2023) では、学校における業務手順を明確にし学校において優先すべき事項を適切に設定できること、ならびに学校内のいじめ・暴力などの問題行動に対処し学校環境を整えることが教育長の質の高さと関連していることを報告している。

5.1.2 教育学分野における理論研究

教育長や校長といった教育における管理職がもたらす効果を理論的に説明した研究は、経済学分野ではほとんど存在しない。一方で、日本の教育学分野では教育長のリーダーシップに関する研究が一定数存在する。

教育学における研究によれば、教育長のリーダーシップは多元的であり、特に日本の教育長については、変革的リーダーシップ、教育的リーダーシップ、社会正義リーダーシップ、政治的リーダーシップの4次元で構成される(山下・諏訪, 2021)¹²⁾。2000年代以降の

米国では、教育長のアカウンタビリティが明確に問われることから教育長のリーダーシップが学力に直接的な効果を与えることが想定されてきたが、日本の教育行政制度のもとでは、教育長の権限は限定的であり教育長が直接学力に効果を与えることは想定しづらいと考えられている(生田 2021)。

生田(2021)は、日本における教育長のリーダーシップが学校を介して児童・生徒の学力に影響を与える経路を検討した。分析の結果、教育長は学校の雰囲気に影響を与え、間接的に児童・生徒の学力に影響を与えることが示唆された。具体的には、変革的および教育的リーダーシップが各学校長の校内でのリーダーシップの発揮を促し、社会正義リーダーシップが教員の教育実践力の向上を促すことで学力に影響を与えることを共分散構造分析に基づき報告している。

以上より、教育長が児童・生徒の学力に与える影響は学校を通じた間接的なものであり、教育長が各学校における校長のリーダーシップの発揮・校内の教育環境の管理を促すことや、教員の教育実践力を高めていくことが学力と関連していると考えられる。

5.2 技術上の限界点

本稿の技術上の限界点は次の2点にある。第1に、本稿では付加価値モデルによる教育長付加価値の推定は行っていない。日本の教育行政制度では、各市町村につき1名の市町村教育長が対応する。そのため、教育長一人ひとりの影響は市町村の固定効果と識別することができない¹³⁾。このような背景から、教育長個人ごとのパラメータを識別・推定するのではなく、教育長の交代というイベントを利用した分析を行っている。

第2に、教育長個人に関する属性や教育長がそれぞれの市町村で独自に行った施策内容などに関する情報は本稿で用いたデータに含まれない。したがって、どのような教育長が児童・生徒の学力に大きな影響を与えるかについての具体的な分析は十分に行えておらず、本稿では「教育長は重要か?」、つまり教育長が児童・生徒の学力の左右する規定要因であるかを考察するにとどまっている。この点は本稿の最大の限界である。

注

- ¹⁾Bloom, Lemos, Sadun and Van Reenen (2015) に代表されるように、教育経済学に限らず人事経済学等における管理職研究とも本稿は関連する。Bloom et al. (2015) は8カ国の学校を対象とした大規模なサーベイから、自律的な(人事・財務等の裁量がある)公立学校ほど校長のマネジメント能力が高いことを示した。
- ²⁾本稿は、学校や教員が児童・生徒の学力等のアウトカムに与える効果を検証する研究群の延長線上にもある。コールマン報告(Coleman, 1968)の公表以来、学校や教員が持つ効果は長らく議論されてきた。教員の経験年数や学歴、教員免許の有無といった「投入変数」のうち、一貫して児童・生徒の学力に有意な影響を与えるものは確認されていない(Hanushek and Rivkin, 2006)。こうした投入変数により教員の質を近似するアプローチに対し、学力などのアウトカムに対する寄与度を直接計測する付加価値モデル(Value-Added Model)と呼ばれる方法によって教員評価を行う研究群(Hanushek, 1986, 2006; Chetty, Friedman and Rockoff, 2014; Kraft, 2019; Kane and Staiger, 2008; Rockoff, 2004)が発展した。
- ³⁾地方教育行政法第13条の1において、「教育長は、教育委員会の会務を総理し、教育委員会を代表する。」と規定されている。
- ⁴⁾関連する日本の研究は非常に数が限られてはいるものの、本稿と同じデータを用いて教員付加価値の記述的な解析を行った伊藤・田端(2022)や、全国学力・学習状況調査の繰り返しクロスセクションデータを用いて校長固定効果の解析・分散推定を行った赤林・佐野(2019)などがある。
- ⁵⁾本稿では、小学生における算数も便宜上「数学」と表記する。
- ⁶⁾なお、本稿における分析では、教育長交代イベントが観測期間中に2回以上生じた市町村は分析から除外している。そのような市町村は2市町村のみであり、それらのサンプルを含めた場合の分析結果も、本稿で報告している結果とほとんど変わらないことを確認している。

- 7) 標準化の操作については、本稿と同じ X 県学力調査データを用いてパネルデータ分析を行った Ito, Nakamuro and Yamaguchi (2020) と同様の処理となっている。
- 8) 実際、Callaway and Sant’Anna (2021) では、共変量を時間変化しないものに限定する対応がとられている。
- 9) 言い換えれば、本稿における対照群とは「観測期間内に教育長交代イベントが発生しなかった自治体」である。Staggered Design での TWFE / DID に関する研究の文脈では、never treated と呼ばれる。
- 10) 市町村ダミーと TWFE / DID 推定における交互作用項（処置群ダミー×処置後ダミー）との交互作用項を用いると、観測期間内に処置を受けなかった市町村（never treated なグループ）は多重共線性が生じるため、この分析では除かれることになる。ここでの推定値は「観測期間内に教育長交代イベントが生じた」市町村内での相対的な推定値の差を示すものと考えられる。
- 11) 本稿では R の did パッケージを用いてイベントスタディプロットを行った。did は Callaway and Sant’Anna (2021) に基づき論文の著者らが作成したパッケージであり、イベントスタディにおいて複数時点のパラメータを同時に推定・検定することで生じる多重検定問題を考慮した信頼区間をデフォルトで出力する。ここでは上記の信頼区間に基づく結果を報告している。
- 12) 山下・諏訪 (2021) によれば、変革的リーダーシップとは「現状維持志向の対置概念であり、教育政策の改革を推進し、効果的な実践を志向する」こと、教育的リーダーシップとは「児童生徒の学力向上及び学力格差の是正を志向する」こと、社会正義リーダーシップとは「社会的に排除されやすい人々に目を向け、社会的包摂に向けて支援することを志向する」こと、政治的リーダーシップとは、「教育長としての説明責任の遂行、地域の利害関係調整、資源配分過程における透明性を志向する」ことを指す（それぞれ原文ママ）。
- 13) 教育長が観測期間内に交代したとしても、同じ市町村内の教育長ダミー変数を足し上げると市町村ダミー変数に一致し、完全な多重共線性が生じる。

参考文献

- Bloom, N., R. Lemos, R. Sadun, and J. V. Reenen (2015) “Does Management Matter in schools?”, *The Economic Journal*, **125** (584), pp. 647–674.
- Branch, G. F., E. A. Hanushek, and S. G. Rivkin (2012) “Estimating the Effect of Leaders on Public Sector Productivity: The Case of School Principals”, Working Paper 17803, National Bureau of Economic Research.
- Callaway, B. and P. H. C. Sant’Anna (2021) “Difference-in-Differences with multiple time periods”, *Journal of Econometrics*, **225** (2), pp. 200–230.
- Chetty, R., J. N. Friedman, and J. E. Rockoff (2014) “Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates”, *American Economic Review*, **104** (9), pp. 2593–2632.
- Coelli, M. and D. A. Green (2012) “Leadership Effects: School Principals and Student Outcomes”, *Economics of Education Review*, **31** (1), pp. 92–109.
- Coleman, J. S. (1968) “Equality of Educational Opportunity”, *Integrated education*, **6** (5), pp. 19–28.
- Dhuey, E. and J. Smith (2014) “How Important Are School Principals in the Production of Student Achievement?”, *Canadian Journal of Economics / Revue canadienne d’Economie*, **47** (2), pp. 634–663.
- and J. S. Smith (2018) “How School Principals Influence Student Learning”, *Empirical Economics*, **54** (2), pp. 851–882.
- Goodman-Bacon, A. (2021) “Difference-in-differences with Variation in Treatment Timing”, *Journal of Econometrics*, **225** (2), pp. 254–277.

- Grissom, J. A., D. Kalogrides, and S. Loeb (2015) “Using Student Test Scores to Measure Principal Performance”, *Educational Evaluation and Policy Analysis*, **37** (1), pp. 3–28, March, Publisher: American Educational Research Association.
- Hanushek, E. A. (1986) “The Economics of Schooling: Production and Efficiency in Public Schools”, *Journal of Economic Literature*, **24** (3), pp. 1141–77.
- (2006) “School Resources”, in Hanushek, E. and F. Welch eds. *Handbook of the Economics of Education*, **2**: Elsevier, 1st edition, Chap. 14, pp. 865–908.
- and S. Rivkin (2006) “Teacher Quality”, in Hanushek, E. A. and F. Welch eds. *Handbook of the Economics of Education*, **2**: Elsevier, 1st edition, Chap. 18, pp. 1051–1078.
- Ito, H., M. Nakamuro, and S. Yamaguchi (2020) “Effects of Class-size Reduction on Cognitive and Non-cognitive Skills”, *Japan and The World Economy*, **53**, p. 100977.
- Kane, T. J. and D. O. Staiger (2008) “Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation”, Working Paper 14607, National Bureau of Economic Research.
- Kraft, M. A. (2019) “Teacher Effects on Complex Cognitive Skills and Social-Emotional Competencies”, *Journal of Human Resources*, **54** (1), pp. 1–36.
- Lavy, V., G. Rachkovski, and A. B. Boiko (2023) “Effects and Mechanisms of CEO Quality in Public Education”, *Economic Journal*, **133** (655), pp. 2738–2774.
- Rockoff, J. E. (2004) “The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data”, *American Economic Review*, **94** (2), pp. 247–252.
- Sun, L. and S. Abraham (2021) “Estimating dynamic treatment effects in event studies with heterogeneous treatment effects”, *Journal of Econometrics*, **225** (2), pp. 175–199.
- 赤林英夫・佐野晋平 (2019) 「校長が在籍児童生徒のアウトカムに与える効果の統計的計測」『フィナンシャル・レビュー』141(6), pp.5–23.
- 生田淳一 (2021) 「教育長のリーダーシップが学校を通して学力に与える間接的効果の発見と経路の解明」露口健司・藤原文雄編著『子供の学力とウェルビーイングを高める教育長のリーダーシップ——校長、教職員、地域住民を巻き込む分散型リーダーシップの効果』学事出版, pp.50–58.
- 伊藤寛武・田端紳 (2022) 「教員付加価値から見た教員の役割について：日本の小学生を例にして」『日本労働研究雑誌』64(2・3) pp.48–69.
- 山下絢・諏訪英広 (2021) 「教育長のリーダーシップの構造」露口健司・藤原文雄編著『子供の学力とウェルビーイングを高める教育長のリーダーシップ——校長、教職員、地域住民を巻き込む分散型リーダーシップの効果』学事出版, pp.33–49.

付録

市町村の影響を考慮した教育生産関数の推定

本稿では、市町村の影響を考慮した上での教育生産関数の推定を行う。具体的には、次の教育生産関数を推定する。

$$y_{isct} = \alpha y_{isp(t-1)} + \mathbf{x}_{it}^T \boldsymbol{\gamma} + \mu_s + \eta_c + \lambda_t + \epsilon_{ispt} \quad (3)$$

\mathbf{x} は児童・生徒・学校に関する共変量のベクトルである。 \mathbf{x} には通塾ダミー、蔵書なしダミー、女子ダミー、学校規模、学級規模が含まれる。 μ は学校固定効果、 η は市町村固定効果、 λ は時点（年）固定効果を表し、 ϵ は誤差項である。

式 (3) は、一般に Value-Added-Model (付加価値モデル) と呼ばれる定式化の 1 つである (Todd and Wolpin, 2003)。ラグ付きアウトカム y_{t-1} を説明変数に含めることにより、それまでに蓄積された児童・生徒の能力や時間に対して変動しない要因が与える影響を取り除くことができる。第 4 節では、式 (3) に基づき、学校固定効果をモデルに含めた場合と含めなかった場合の 2 通りの特定化について結果を報告する。市町村の影響を考慮した教育生産関数の推定結果を付録として記載する。本稿の主旨は教育長交代イベントに着目した教育長の影響の考察にあるため、市町村ごとの教育長生産関数推定の結果を本文には記載していない。

表 6 より、1 期前 (前学年) の時点における国語や数学のテストスコアは統計的に有意な正の相関を示し、前学年におけるテストスコアが高いほど現学年におけるテストスコアが高いことを表している。また、女子生徒であることは数学のテストスコアと負の関連性を示している。学校規模や学級規模といった学校の属性とテストスコアとの関連は非常に小さな値となっている。

	国語		数学	
	(1)	(2)	(3)	(4)
国語 ($t-1$ 時点)	0.681*** (0.002)	0.662*** (0.002)		
数学 ($t-1$ 時点)			0.766*** (0.002)	0.751*** (0.002)
女子ダミー	0.089*** (0.002)	0.096*** (0.002)	-0.018*** (0.002)	-0.019*** (0.002)
通塾ダミー	0.058*** (0.004)	0.045*** (0.003)	0.153*** (0.005)	0.143*** (0.004)
蔵書ありダミー	0.166*** (0.005)	0.185*** (0.005)	0.122*** (0.004)	0.132*** (0.004)
クラスサイズ	-0.001 (0.003)	0.014*** (0.005)	-0.009*** (0.003)	-0.001 (0.003)
学校規模	0.087*** (0.006)	-0.116*** (0.024)	0.064*** (0.006)	0.040** (0.019)
サンプルサイズ	937 849	937 849	937 849	937 849
FE: City	X	X	X	X
FE: Year	X	X	X	X
FE: School		X		X

注) 係数の推定値について***は 0.1%, ** は 1%, *は 5%で有意であることを示す。() 内の値は各推定値に対する標準誤差を示す。標準誤差は、市町村を単位としたクラスター構造に頑健な標準誤差である。アウトカムならびに説明変数のうち連続変数であるものは標準化されているため、表内における数値の単位は標準偏差 (SD) となる。

表 6: 市町村固有の影響を制御した教育生産関数の推定結果