

# **Institute for Economic Studies, Keio University**

## **Keio-IES Discussion Paper Series**

### **Do Superintendent Turnovers Matter for Student Performance?**

**Soichiro Sugita**

**25 March, 2025**

**DP2025-004**

**<https://ies.keio.ac.jp/en/publications/25217/>**

Keio University



Institute for Economic Studies, Keio University  
2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan

[ies-office@adst.keio.ac.jp](mailto:ies-office@adst.keio.ac.jp)

25 March, 2025

Do Superintendent Turnovers Matter for Student Performance?

Soichiro Sugita

Keio-IES DP2025-004

25 March, 2025

JEL Classification: I21, I22, M12

Keywords: Superintendents, Accountability, leadership, Difference in Differences, Education production function

### Abstract

While numerous empirical studies—particularly in the United States—have examined the management skills and leadership of educational leaders, there remains limited evidence in the Japanese context. This study investigates the causal effect of superintendent turnover at the municipal level on student academic achievement, using a large-scale panel dataset from a standardized academic assessment conducted in a prefecture in Japan. The results suggest that changes in superintendents have limited effects on students' academic performance.

Soichiro Sugita

Graduate School of Media and Governance, Keio University

sugita3520@keio.jp

Acknowledgement: I am grateful to my advisor, Professor Makiko Nakamuro (Faculty of Policy Management, Keio University), for her guidance and support. I also thank Mr. Hirotake Ito, currently at CyberAgent, for his insightful feedback throughout the development of this paper. All remaining errors and omissions are my own.

# 教育長交代は重要か？——大規模学力調査データを用いた分析

杉田 壮一朗 \*

## 要旨

アメリカ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を評価する実証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、埼玉県学力・学習状況調査の大規模な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因果的効果を検証した。分析結果より、教育長交代が児童生徒の学力に与える影響は限定的であることが示された。

*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、市町村教育長の交代が児童生徒の学力に与える因果的な効果を大規模な学力調査データを用いて検証したものである（以下、特に断りがない限り教育長とは市町村教育長を指す）。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に位置づけられる。教育の管理職が与える影響は、主に日本国外において教育・人事経済学などの分野で2010年代以降実証分析が蓄積してきた。例えば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。

さらに、本稿は学校や教員が児童・生徒の学力等に与える効果を検証する研究群の延長線上にある。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年議論してきた。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)。投入変数により教員の質を近似するこのようなアプローチに対し、学力などのアウト

\*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。

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）によって教員評価を行う研究群 (Hanushek, 1986, 2006; Chetty et al., 2014; Kraft, 2019; Kane and Staiger, 2008; Rockoff, 2004) が発展した。

特にアメリカを中心とした先行研究では、(1) 上述の教員に対する付加価値モデル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 et al., 2015; Branch et al., 2012)、(2) 校長の交代・異動などをを利用して、児童生徒のアウトカムに与える影響を分析したものがある。児童生徒と校長の情報が紐づけられ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場合が多い。(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検証した文献も、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 et al. (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、教育長・校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。ただし、日本の教育行政制度下における効果は必ずしも明らかでなく研究蓄積も少ない。アメリカの教育行政制度では、教育長・校長に対してテストスコアの結果に対するアカウンタ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。関連する日本の研究としては、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) などがあるものの、教育におけるリーダーの効果を実証的に検証した文献はかなり限られている。

以上の背景に基づき、本稿における研究目的を次のように設定した。本稿では、日本における大規模な児童生徒の個票パネルデータを用いて、日本の教育行政制度のもとで教育長交代（以下、処置）という地域教育行政のリーダーの変更が児童生徒の学力へ与える効果を検証し、教育長の影響を実証的に明らかにする。本稿の分析は、アカウンタビリティが強く問われるアメリカなどのエビデンスが中心であり、日本の教育行政における文脈での検証が手薄であった教育におけるリーダーの効果に関する研究蓄積に貢献する。特に、埼玉県学力・学習状況調査を用いることの意義は、定量的評価に耐えうる十分なデータの質と量にある。

埼玉県のデータは、児童生徒の学力が経年比較が可能なIRT(項目反応理論)を用いて測定されているかつ、児童生徒の属性情報も含んでいる。さらに、さいたま市を除く埼玉県下の全ての市町村が参加する悉皆調査に基づく個票データであり、計量経済学的な効果検証に耐えうるものである。

本稿で行った分析より次のような結果が得られた。Staggered Designを用いた差の差分析の結果より、教育長の交代は国語および数学に対して統計的に有意な効果を持つことは確認されなかった。交代による影響の大きさも小さなものであると考えられる。したがって、教育長交代という地域教育行政のリーダーの変更は、その市町村の学校に通う児童生徒の学力向上には寄与しないと考えられる。

以下、本稿の構成は次の通りである。第3節では、教育長に関する制度上の背景、埼玉県の教育制度や学校・児童生徒の背景情報、データの詳細等を明らかにし、記述統計分析の結果を示す。第4節では実証戦略を提示し、第5節では推定結果を報告する。最後に第6節ではディスカッションと結論を提示し、本稿の研究上の限界や追加的議論をまとめる。

## 2 背景

### 2.1 教育委員会に関する制度上の背景

本項では、教育委員会に関する制度的な背景を簡潔に説明する。市町村教育委員会(以下、市町村教委)は、首長から独立した行政機関であり、地域の教育行政の中心を担う組織である。教育委員会は主に学校教育の振興、地域・社会教育の振興、芸術文化の振興ならびに文化財の保護、スポーツ振興の4領域を担う。本稿と関連する学校教育分野については、具体的に、学校の設置管理、教職員の人事及び研修、児童・生徒の就学及び学校の組織編制、校舎等の施設・設備の整備、教科書その他の教材の取扱いに関する事務の処理などを行う。

また、地方教育行政法(以下、地教行法)第13条の1で「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」と規定されており、市町村教育行政における第一義的な責任者である。教育長の任期は3年で、首長が議会の同意を得た上で直接任命する。また、教育長の再任も認められている。教育長が責任者を務める教育委員会は、常勤の教育長と非常勤の教育委員(原則4名)で構成される合議体と、地域の教育行政に関する事務を行う教育委員会事務局からなり、教育長は合議体ならびに事務局の指揮監督者であるとされる。

本稿が分析対象とする市町村教委は、教員の人事や給与に関する直接的な権限を持たない。市町村の教職員は都道府県教育委員会(以下、県教委)が任命し、政令市を除き県教委がその給与負担も行う。この点は、教育行政制度におけるアカウンタビリティが厳しく問われるアメリカなどの他の国と異なる。例えば、アメリカでは学校区(school district)ごとの教育委員会ならびに各学校が教職員人事・給与負担に関する権限を強く持ち、学力テストの結果等に応じて教育長の説明責任が強く問われる場合が多い。一方で、日本の教育行政制度のもとでは、市町村教委は市町村が設置した学校に対する指導は行うが、教職員の雇用・

配置変更を行ったり教職員給与を負担したりするといった直接的なマネジメントを行うことはできない。さらに、上述のように現行の制度では教育長の任命責任は首長にあり、毎年度開催される総合教育会議において首長と教育委員会との検討によって教育政策の方針が決定されることも踏まえると、教育長は県教委や各首長の影響を強く受け、他市町村の教育長の教育行政方針に影響を受けることは少ないと考えられる。

## 2.2 教育長のリーダーシップに関する教育学・社会学研究

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、日本の教育学や社会学の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では、教育長のリーダーシップは多元的であるとされ、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)。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、社会的包摶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)。2000年代以降のアメリカでは、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)。

ただし生田(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。

## 2.3 埼玉県内の市町村における教育制度

本項では、分析対象となる埼玉県の各市町村における教育制度について言及する。学校教育法施行令第5条では、市町村教委は就学予定者の就学すべき小学校、中学校又は義務教育学校を指定しなければならないと規定されており、児童生徒は、住民登録上の住所で定められた公立小中学校に通うことが原則である。しかし、保護者が何らかの理由で指定校以外への就学(いわゆる「越境通学」)を希望する場合は、事由を添えて就学指定校変更・区域外

就学の申請を行い、各市町村教委が許可することで可能となる。各市町村が定める事由は、身体的な理由、いじめ等への対処、児童生徒の帰宅地が自宅以外（祖父母宅等）の場合などであり、学校選択制のように希望学校を選ぶ仕組みではない。

埼玉県下の一部の市町村（朝霞市、川口市、越谷市等）では、公立学校の学校選択制を実施しているものの盛んに実施されているとは言えない。未実施の市町村も多く、実施している場合であっても、市内の特定のブロック内での自由選択など、完全な自由選択制でないケースがほとんどである。さらに、学校選択制は原則「市町村内の」就学校選択における自由を認めるものであり、教育長交代の効果検証を行う本稿の分析に対し基本的には影響を与えないと考えられる。また、本稿のデータ上では、「さいたま市を除く埼玉県内で市町村間を移動した」と考えられるサンプル（以下、転校サンプル）は観測できる。3.3項では、本稿で用いるデータで市町村間の個人の移動と教育長交代との関連を検証する。

本稿で用いるデータが対象とする埼玉県下の市町村における児童生徒の特徴については、第3節にて記述統計分析とともに補足する。

### 3 データ

#### 3.1 使用データの概要

本稿における分析では、埼玉県学力・学習状況調査（以下、埼玉県学調）の2016年から2019年の児童・生徒個票データを利用した。埼玉県学調は、さいたま市を除く埼玉県内全市町村における公立小中学校の小学校4年生から中学校3年生の全ての児童・生徒を対象に実施されており、本稿では国語と数学のスコアを利用する。

埼玉県学調の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、埼玉県学調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。

#### 3.2 記述統計量

推定に用いるサンプルの記述統計量を表1に示す。推定では処置の効果を捉えるため、全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。

表 1: 記述統計量

項目	観測数	個人数	平均	標準偏差	最小値	最大値
<b>処置・時点に関する変数</b>						
教育長交代後ダミー	875606	354313	0.26	0.44	0.00	1.00
交代イベントあり（処置群ダミー）	875606	354313	0.40	0.49	0.00	1.00
<b>アウトカム</b>						
国語スコア（素点）	875606	354313	0.41	1.52	-7.34	6.17
数学スコア（素点）	875606	354313	0.38	1.32	-5.80	5.80
<b>生徒に関する変数</b>						
女子ダミー	875606	354313	0.49	0.50	0.00	1.00
通塾ダミー	875606	354313	0.67	0.47	0.00	1.00
蔵書ありダミー	875606	354313	0.89	0.31	0.00	1.00
<b>学校に関する変数</b>						
学校規模	875606	354313	350.87	173.36	4.00	961.00
クラスサイズ	875606	354313	32.73	4.79	1.00	42.00

注) 観測数はパネルにおける観測されたデータ点の数、個人数はユニークな個人（ユニット）の数である。連続変数は平均0、標準偏差1に標準化する前の値について統計量を計算している。個人の平均観測年数は2.47年である。

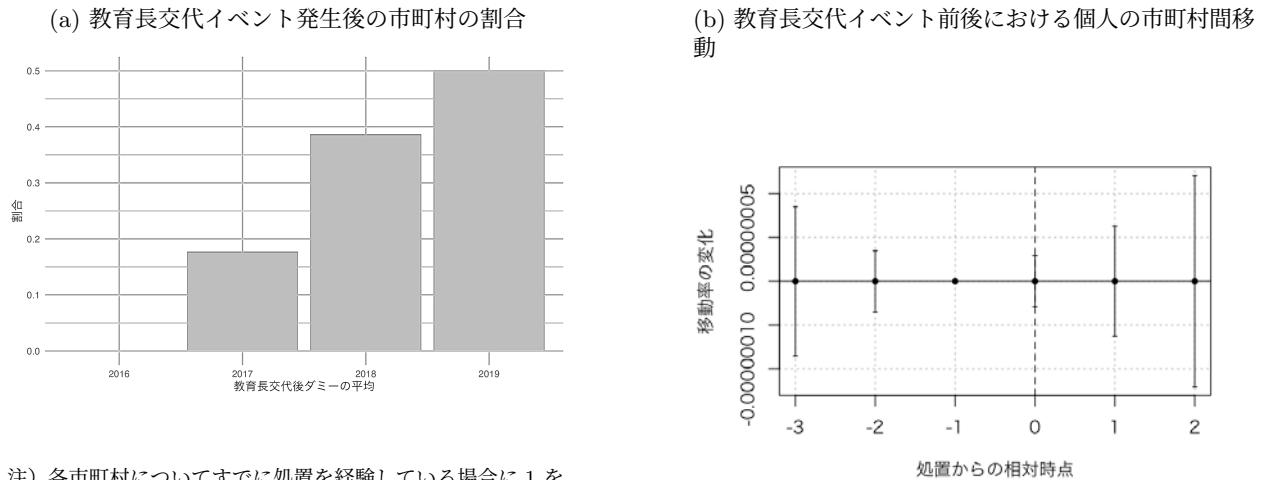
本稿の推定では「処置群のサンプルにもかかわらず処置前に卒業や転出などによって退出したサンプル」および「2016年から2019年の間に市町村間を移動したサンプル」を除外している。後者を除外するのは、処置群の市町村から対照群の市町村に移動することによって、処置時点以降ならば1をとり続けるというパターンが崩れることを防ぐためである。

### 3.3 教育長・市町村に関する記述統計分析

本稿では、観測期間内における埼玉県の各市町村で誰が教育長を務めているのかを埼玉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。埼玉県学調は各年度の5月に実施されるため、その前年度に発生した教育長の交代を学力の測定が行われた年度に紐づけている。本稿の分析で用いたサンプルでは、図1(a)では、年ごとにどれくらいの割合の市町村が処置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(a)を見ると、各年で埼玉県全体の市町村のうち10~15%程度に当たる市町村で処置が発生しており、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。なお、本稿における分析では、処置が観測期間中に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。

さらに、第2.3項で述べたように、パネルデータの期間内での市町村間の移動が教育長交代に合わせて発生していないかを確認する。図1(b)は処置タイミング前後の市町村間移動の様子をイベントスタディプロットで示している。各時点における推定値は0付近の値をとっており、処置による統計的に有意な影響は認められない。したがって、教育長交代タイミングに合わせて市町村間を移動することを支持する証拠は本稿では得られなかった。

図 1: 教育長交代イベントに関する記述統計分析

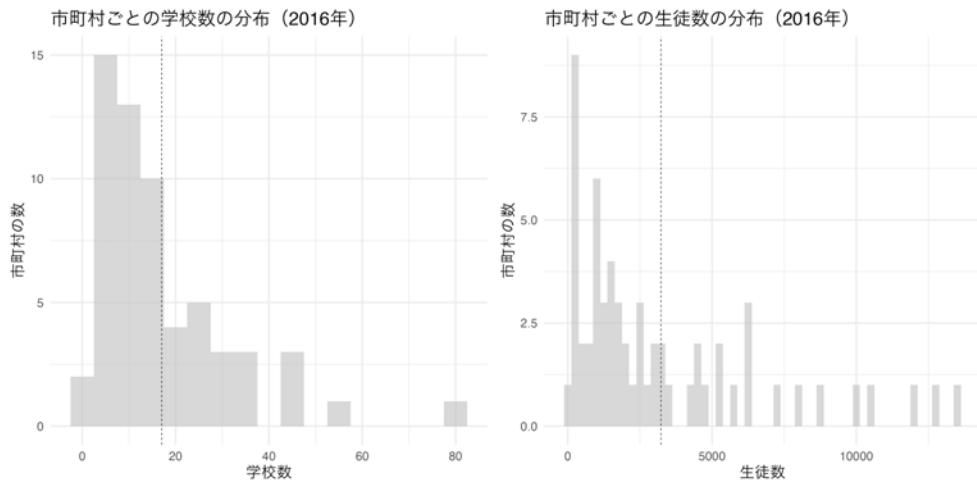


注) 各市町村についてすでに処置を経験している場合に 1 をとるダミー変数（処置後ダミー）を定義し、各年における平均値を算出している。

注) Sun and Abraham (2021)に基づき、市町村間移動ありダミーをアウトカムとした時のイベントスタディを実行した。エラーバーは 95% 信頼区間。使用したサンプルは、市町村間を移動したサンプルを含み、これらを除くと 5.1.1 の推定に使用したサンプルが得られる。

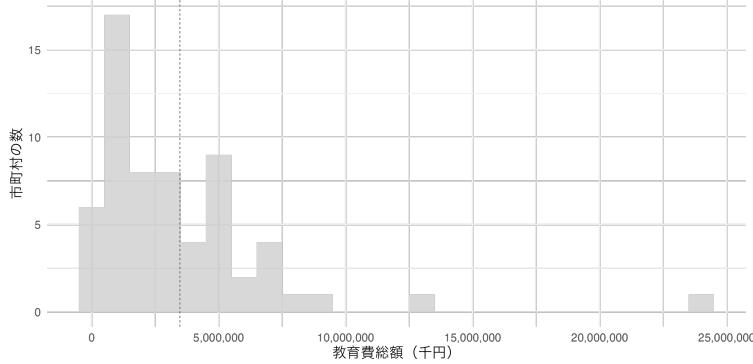
3.2 項の表 1 より、平均して児童生徒の 7 割弱は通塾しており、家庭の蔵書が全くない子どもは 1 割程度である。また、学校数や生徒数の分布を図 2 に、各市町村の教育支出（教育費）の分布を図 3 に示してある。学校数・生徒数・市町村数とともに分布が左側に偏っており、多数の小中規模の市町村と一部の規模の大きな市町村で構成されたサンプルであることがわかる。これらの変数についての処置効果の異質性については 5.1.1 で検証する。

図 2: 市町村ごとの学校数と生徒数の分布（2016 年）



注) 破線は分布の平均を表す。学校数の平均は 17.02、中央値は 12.5、標準偏差は 14.80。生徒数の平均は 3,230、中央値は 1,874、標準偏差は 3,379。推定に使用するサンプルのうち、市町村ごとに学校数・生徒数をカウントした。

図 3: 市町村ごとの教育費分布 (2015 年度)



出所) 平成 28 年度 (平成 27 会計年度) 地方教育費調査。

注) 破線は分布の平均を表す。平均は 3,472,390、中央値は 2,505,184、標準偏差は 3,641,180。

さらに表 2 に観察可能な変数によるバランステストの結果を示す。上述したように、この分析で用いた処置群は「処置を受けた後もデータにとどまる」サンプルのみを用いている。女子ダミーを除き統計的に有意な平均値の差が検出されているものの、各変数の平均値の差に対する検出力は本稿のデータのサンプルサイズが大きいためほぼ 1 に近い値をとっている。したがって、平均値の差の有意性は両群がバランスしないことを必ずしも意味しない。

表 2: 処置前における処置群と対照群の比較

変数	対照群	処置群	p 値	検出力
観測数	528,252	114,367		
女子ダミー	0.49 (0.50)	0.49 (0.50)	0.198	0.05
通塾ダミー	0.68 (0.47)	0.74 (0.44)	<0.001	1.0
蔵書ありダミー	0.89 (0.31)	0.90 (0.30)	<0.001	1.0
学校規模	359.77 (178.95)	306.48 (151.45)	<0.001	1.0
クラスサイズ	32.93 (4.67)	31.85 (5.26)	<0.001	1.0

注) 検出力 (Power) は、効果量  $d = \frac{|\mu_1 - \mu_0|}{s_p}$  に基づいて算出している (各変数の添字は 1 が処置群、0 が対照群を表す)。 $s_p$  はプールされた標準偏差で、 $s_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_0-1)s_0^2}{n_1+n_0-2}}$ 。有意水準  $\alpha = 0.05$  の両側検定における正規分布近似を用い、検出力は以下の式で計算している :  $\text{Power} = \Phi(z - z_{1-\alpha/2}) + [1 - \Phi(z + z_{1-\alpha/2})]$ 。ただし  $z = d\sqrt{\frac{n_1 n_0}{n_1+n_0}}$ 、 $\Phi(\cdot)$  は標準正規分布の累積分布関数。

## 4 実証戦略

### 4.1 回帰モデル

処置効果を推定するため、まずは次の 2 方向固定効果 (TWFE) 回帰モデルを考える。

$$Y_{it} = \alpha_i + \lambda_t + \delta D_{it} + \epsilon_{it} \quad (1)$$

添字の  $i$  は生徒個人、 $t$  は時点 (年度) を表す。 $\alpha_i$  は個人 (ユニット) の固定効果、 $\lambda_t$  は時

点の固定効果である。 $D$  は個人  $i$  が処置群（処置が観察期間内に生じた市町村）に属するかつ  $t$  が処置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。 $Y$  は標準化されたテストスコア（国語・数学）である。

式 (1) 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、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（時間を通じて不变な個人固有の影響）を制御しているため児童・生徒の性別など時間を通じて不变な属性をモデルに含められない。

第 2 に、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。そのため、本稿では共変量を含めない推定結果を提示する。なお、平行トレンドの成立についての検討は、イベントスタディプロットの確認、事前トレンドに対する検出力の確認などとともに第 5 節にて行う。

また、本稿の設定では処置の発生タイミングが市町村ごとに異なる。処置タイミングがユニットごとに異なる場合、すでに処置を受けたグループを対照群として扱うことにより Negative Weights の問題が生じ、深刻なバイアスをもたらしうることが近年指摘されている。次項では、Negative Weights の問題が生じているかを診断する。

## 4.2 Negative Weights の検証

本項では、Frisch–Waugh–Lovell (FWL) 定理に基づき Jakiela (2021) が提案した方法によって Negative Weights の診断を行う。ここでは、以下の回帰モデルを考える。

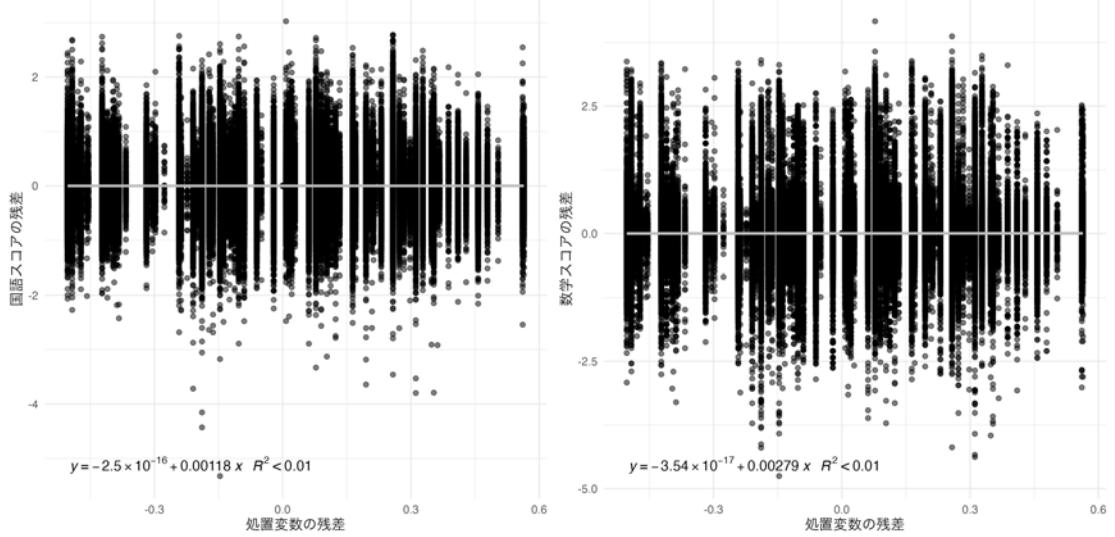
$$\tilde{Y}_{it} = \delta \tilde{D}_{it} + \tilde{\varepsilon}_{it} \quad (2)$$

$\tilde{Y}_{it}$  はアウトカムを、 $\tilde{D}_{it}$  は処置変数を個人固定効果と時点固定効果に回帰して得られる残差である。この回帰式を最小 2 乗法 (OLS) で推定し、傾き  $\delta$  が負であるかを確認する。

図 4 は、 $(\tilde{Y}_{it}, \tilde{D}_{it})$  をプロットした散布図ならびに式 (2) を OLS で推定して得られた回帰直線を示したものである。図から視覚的に分かるように、回帰直線の傾きはほぼ 0 かつ回帰モデルの説明力 (決定係数) は 1% 以下であり、アウトカムの残差と処置変数の残差との間に強い相関関係は確認されない。図中に示している回帰分析の結果は、処置変数の残差 ( $x$ ) の傾きの係数は 1 % 水準で統計的に有意であるものの、符号は正であり、係数の大きさについては国語は 0.0011、数学は 0.0027 であり 0 に近い値をとっている。

以上より、Negative Weights の問題が生じていることを支持する結果は得られなかった。ただし、本稿では Staggered Design に基づく分析を主な分析結果として提示し、比較対象としてナイーブな TWFE の結果も報告する。

図 4: Negative Weights の診断 (Jakiela, 2022)



注) 縦軸はアウトカム（国語・数学のスコア）の残差、横軸は処置変数の残差である。残差は個人固定効果・時点固定効果に回帰して得られたものである。グレーの直線はアウトカムの残差を処置変数の残差で回帰して得られる直線である。

## 5 推定結果

### 5.1 Staggered Design による推定結果

#### 5.1.1 Sun and Abraham (2021) に基づく推定

本稿では、Sun and Abraham(2021、以下SA2021)に基づく R の `fixest` パッケージ・`sunab()` を用いた差の差分析、ならびにイベントスタディプロットによる可視化を行った。表 3 に Staggered Design の推定結果、図 5 にそのイベントスタディプロットを示す。

国語については、処置ダミー  $D_{it}$  の係数は 5% 水準で統計的に有意でなく、帰無仮説  $H_0 : \delta = 0$  は棄却されない。係数の符号の向きは一貫しておらず、効果の大きさは全時点の平均が 0.002sd で、-0.043sd(標準偏差) から 0.01sd の間の値となっている。数学についても同様に、処置ダミー  $D_{it}$  の係数は 5% 水準で統計的に有意でない。係数の符号の向きは一貫しておらず、効果の大きさは全時点の平均が 0.004sd で、-0.059sd から 0.015sd の間の値となっている。図 5 のイベントスタディプロットからも、各教科・各時点の推定値は統計的に有意に 0 と異ならず、効果の大きさが限定的であることがわかる。

図 6 は、5.1.1 で行った推定に対してさらに 2015 年度の市町村ごとの学校数・生徒数・教育費それぞれについて処置効果の異質性があるかを分析したイベントスタディプロットである。全体としては、統計的に有意な異質性は観察されなかった。ただし統計的有意性はないが、教育費（教育支出）が中央値より上位の市町村はで国語・数学ともに処置後の推定値の符号が正で一貫している。

表 3: Staggered DID による推定結果 (SA2021)

国語の推定結果

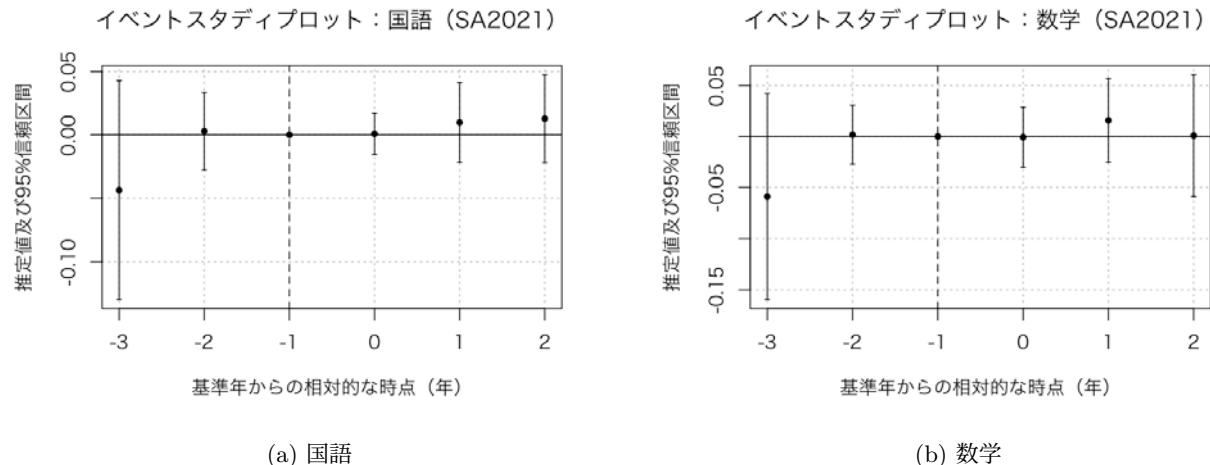
相対時点	推定値	<i>p</i> 値
-3 時点	-0.043	0.318
-2 時点	0.003	0.851
0 時点	0.000	0.954
+1 時点	0.008	0.614
+2 時点	0.012	0.484
<b>全時点平均</b>	<b>0.002</b>	<b>0.900</b>

数学の推定結果

相対時点	推定値	<i>p</i> 値
-3 時点	-0.059	0.247
-2 時点	0.002	0.907
0 時点	-0.001	0.959
+1 時点	0.015	0.458
+2 時点	0.001	0.958
<b>全時点平均</b>	<b>0.004</b>	<b>0.860</b>

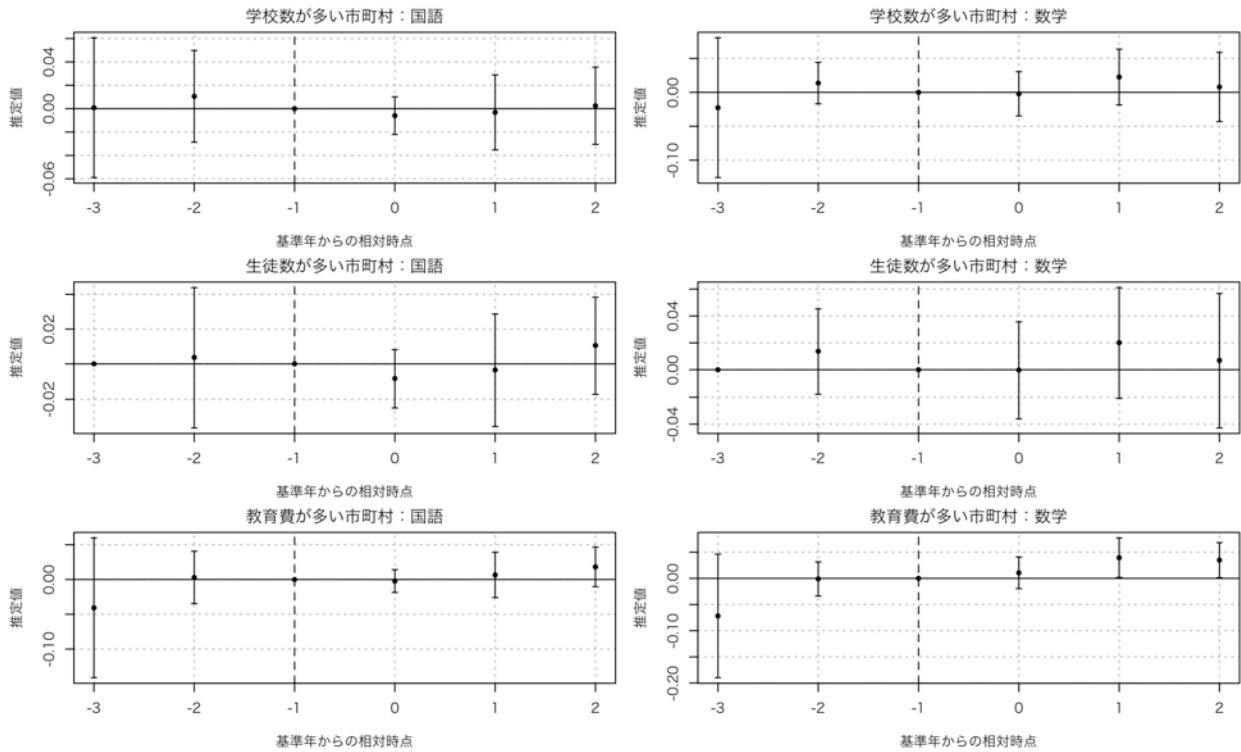
注) アウトカムは標準化されているため、表内の数値の単位は標準偏差 (sd) である。

図 5: Staggered Design による推定結果 (イベントスタディプロット、SA2021)



注) 各時点における点推定値は、Sun and Abraham (2021) で提案された方法に基づく推定値。エラーバーは 95%信頼区間。

図 6: 処置効果の異質性



注) Sun and Abraham (2021) によって得られた推定値を報告している。それぞれ、学校数・生徒数・教育費の中央値より上位に位置する市町村のサンプルに対する異質性を分析している。

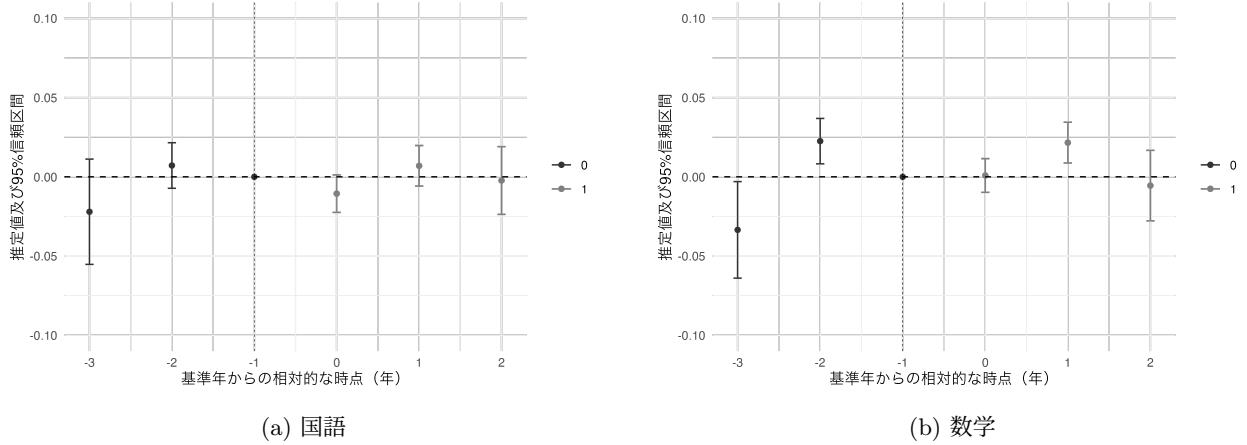
## 5.2 頑健性チェック

### 5.2.1 Callaway and Sant'Anna (2021) に基づく推定

図 7に Callaway and Sant'Anna (2021、以下 CS2021) の方法による推定結果をイベントスタディプロットの形式で示してある。国語については、SA2021 における結果と同じく、統計的に有意な推定値はどの時点においても認められず、処置後における推定値の正負は一貫していない。推定値の大きさは-0.05sd から 0.02sd 程度である。数学については、-2 時点と+2 時点において統計的に有意な正の推定値が得られている。しかしそれぞれ 0.025sd 程度の大きさであり、十分に大きな効果であるとは言えない。ただし、事前時点で統計的に有意な差が見られることについては、5.2.3にて議論する。

なお、本稿で行った以下 CS2021 に基づく分析は Not yet treated(処置群だが未処置) をベンチマークとしており、推定の際にバランスドパネルとなるよう 5.1.1で用いたサンプル (アンバランスドパネル) から `did` パッケージによって切り出されたものを推定に用いている。

図 7: Staggered Design による推定結果（イベントスタディプロット、CS2021）



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く推定値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%信頼区間であり、`did` パッケージによって多重検定問題を考慮して計算されたものである。

### 5.2.2 ナイーブな TWFE との比較

4.2項で行ったJakiela (2021) による診断では、処置タイミングの異なる TWFE における Negative Weights の存在は支持されなかったが、頑健性確認のためナイーブな TWFE の推定値も確認する。表 4 にナイーブな TWFE の推定結果を示す。結果より、帰無仮説  $H_0 : \delta = 0$  に対して国語・数学とともに 5% 水準で統計的に有意ではなかった。推定値の大きさは、-0.003sd から 0.001sd であり限定的である。Staggered Design における結果と比較しても、推定値は大きく変わらず Negative Weights の影響は大きなものではないと考えられる。

表 4: ナイーブな TWFE の推定結果

	国語		数学	
	(1)	(2)	(3)	(4)
教育長交代後×処置群	0.001 (0.011)	0.000 (0.008)	0.003 (0.017)	-0.003 (0.008)
Observations	875 606	875 606	875 606	875 606
FE: Student	✓	✓	✓	✓
FE: Year	✓	✓	✓	✓
FE: School		✓		✓

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。() 内の値は標準誤差。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、表内の数値の単位は標準偏差 (SD) となる。

### 5.2.3 平行トレンド仮定の検証

本項では、5.1.1で得られた推定結果をもとに、Staggered Design に基づく差の差分析における平行トレンド仮定が成立しているかを検討する。ここでは、事前トレンドにおける検出

力の検討 (Roth, 2022) を行う。平行トレンド仮定の検討において、事前トレンドが有意でないだけでは不十分であり、検出力の確認や感度分析等をすべきであることが近年指摘されている (Roth, 2022; Rambachan and Roth, 2023)。

ここでは、有意水準  $\alpha = 0.05$ 、検出力  $1 - \beta = 0.8$  という設定のもと、処置前の時点 (-3 および -2 時点) それぞれについて最小検出可能効果 (MDE) を  $MDE = (Z_{1-\alpha/2} + Z_{1-\beta}) \times se$  で計算した。 $Z_{\gamma}$  は標準正規分布における上側  $100 \times \gamma\%$  点を、 $se$  は標準誤差を表す。国語については -3 時点で 0.121、-2 時点で 0.043 であった。数学については -3 時点で 0.141、-2 時点で 0.041 であった。実際の推定値よりも MDE の方が大きいことは事前トレンドを検出することができなかった可能性があることを意味する。実際、5.1.1における数学の推定結果では、事前トレンドがないという帰無仮説は棄却されている。

## 6 ディスカッション・結論

本稿では、地域教育行政のリーダーである教育長の交代が、その市町村の学校に通う児童生徒の学力に与える影響を検証した。Staggered Design を用いた差の差分析の結果より、教育長の交代が国語および数学のスコアに対し、統計的に有意かつ十分に大きな効果を持つことは確認されなかった。したがって、教育長交代という地域教育行政のリーダーの変化が児童生徒の学力に与える影響は限定的であると考えられる。この結果は、教育学・社会学分野における研究や、市町村の教員人事・給与権に対する権限が低いという制度内容とも整合的である。

ただし、本稿には以下に述べるように技術上の限界点・留意点がいくつかある。第 1 に、本稿では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 1 名の教育長が対応するため、個人ごとの教育長付加価値は市町村固定効果と識別できない。このような背景から、本稿では教育長付加価値の識別・推定ではなく教育長交代に着目した分析を行った。本稿はあくまで「教育長交代というイベントが与える効果が支持されない」という主旨の結果であり、個々の教育長が持つ管理能力や行政手腕の高さを否定するものではないことには留意が必要である。

第 2 に、教育長個人に関する属性や独自に行った施策内容などに関する情報は本稿の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、教育長交代イベントに対する因果的な効果の検証にとどまっている。

第 3 に、5.2.3 項で検討したように、平行トレンド仮定が成立しているとは必ずしも主張できず、事前トレンドを検出できなかった可能性がある。そのため、本稿の結果は平行トレンド仮定について上記の留保のもとで解釈されたい。

## 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’Economique*, **47** (2), pp. 634–663.
- and J. S. Smith (2018) “How School Principals Influence Student Learning”, *Empirical Economics*, **54** (2), pp. 851–882.
- 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.
- Jakiela, P. (2021) “Simple Diagnostics for Two-Way Fixed Effects”, arXiv:2103.13229.

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.

Rambachan, A. and J. Roth (2023) “A more credible approach to parallel trends”, *Review of Economic Studies*, **90** (5), pp. 2555–2591.

Rockoff, J. E. (2004) “The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data”, *American Economic Review*, **94** (2), pp. 247–252.

Roth, J. (2022) “Pretest with caution: Event-study estimates after testing for parallel trends”, *American Economic Review: Insights*, **4** (3), pp. 305–322.

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.