

# 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－\*

伊藤寛武<sup>†</sup>

田端紳<sup>‡</sup>

March 31, 2019

## Abstract

教員が子どもの能力の成長に与える影響について、多くの研究者や政策担当者がそれぞれの国のデータを用いて検証してきた。その一方で日本の教員を対象に行われた分析は非常に少なく、子どもの能力の成長に教員がどの程度影響を持っているのかはよく分かっていない。本稿では日本のある自治体における個票データを用いて教員付加価値を推定し、教員が子どもの能力の成長に与える影響について記述的に分析した。その際には、学力のみならず学習方略や非認知能力並びに学級での子ども・教員間の関係性を対象にして教員付加価値を求めた。その結果として学力や学習方略に対して個々の教員の違いが与える影響力が取りうる値の範囲は広がった一方で、非認知能力や学級での子ども・教員間の関係性に対して教員の違いは子ども全体の変動のうち少なくとも5%から8%以上を説明していた。更に、教員付加価値間の相関係数から、学力を伸ばすことが得意な教員と非認知能力・学習方略を伸ばすことが得意な教員は異なるという結果を得た。以上の結果から、教員の違いは学力より非認知能力に対して重要である可能性が示唆された。

JEL Classification: I21

キーワード: 非認知能力, 教員効果

---

\*本稿の執筆にあたって、「埼玉県学力・学習状況調査のデータを活用した効果的な指導方法に関する調査研究」として埼玉県学力・学習状況調査データの分析研究業務委託を受けました。埼玉県教育委員会及び戸田市教育委員会から多大な協力を得たことをここに感謝致します。本研究は科学研究費補助金挑戦的研究(開拓)「科学的根拠(エビデンス)に基づく教育政策の推進研究課題」(研究代表者:中室牧子、研究課題番号:18H05314)の支援を受けています。

<sup>†</sup>責任著者。慶應義塾大学政策メディア研究科博士課程、Email: itouhrtk@keio.jp

<sup>‡</sup>慶應義塾大学経済学研究科修士課程、Email: stabata@gs.econ.keio.ac.jp

## I はじめに

教員が子どもの能力の成長にどのように寄与しているのかという問いは、現代において教員付加価値に関する研究として深化している。教員付加価値とは教員が担当した子どものアウトプット（往々にして学力を指す）の平均的な増加分のことであり<sup>1</sup>、教員の質を計測した指標としてしばしば用いられる。多くの研究がこの教員付加価値をどのように推定するか、または推定された教員付加価値が子どもの能力の成長にどのように関係しているかを考察してきた (Kane and Staiger, 2008; Chetty et al., 2014a; Lefgren and Sims, 2012)<sup>2</sup>。それのみならず、実務面でも教員の評価に用いようとする動きもあり、研究・実務両面においては教員付加価値という指標は注目を集めている。

しかし、学力以外の教育成果に対して教員付加価値を分析した研究は必ずしも多くはない。教員付加価値についての研究の多くは学力を対象にしており、非認知能力などを対象に教員付加価値を推定し考察する研究は一部にとどまる (Jackson, 2018)。教育成果について多くの人が最初に思い浮かべるのが学力ではあるものの、子どもにとっての教育成果は学力のみに限られるわけではない。特に、子どもの人的資本蓄積において非認知能力などの学力で表現できない能力の重要性が指摘されて久しい。そして教員が子どもに与える影響もまた、学力に対する影響に限られるわけではない。教員は子どもとのコミュニケーションを通じて非認知能力に対する影響を含めた多様な影響を子どもに与える。そのため、教員付加価値については様々な教育成果を対象とした研究が必要である。

更に、日本を対象にした教員の効果に対する定量的研究は非常に少ない。教育システムや学校を取り巻く状況は国によって大きな違いがあるため、海外での研究成果をどの程度日本においてそのまま適用可能かについては慎重な議論が必要である。そのため、その国独自のデータに基づく研究の蓄積は重要であるにもかかわらず、日本の教員を対象にした研究は乏しいのが現状である。例えば、国際学力調査を用いた教育生産関数の文脈で教員の影響を分析する研究は存在する (Hojo and Oshio, 2012; Hojo, 2012) が、これらは教員の効果を分析することを目的とした研究ではなく教員の特徴を教育生産関数の1つの引数として捉えたものである。また二木 (2017) が日本における教員効果に関する数少ない量的研究であるが、あくまで都道府県レベルでの集計値に対する議論であり教員個人の教育効果を分析するものではない。教員の効果について知ることには政策的な意味も大きいため、日本を対象にして教員付加価値の研究を行う意義は大きい。

本稿では小学校の教員を対象にして教員付加価値の推定を行いその記述的な分析を行う。教員付加価値の推定には日本のある自治体における全ての子どもと教職員を対象に3年に渡って行われた調査のデータを用いた。その調査では、毎年4月に子どもに対する学力調査と質問紙調査及び教員に対する質問紙調査を行っている。特に、子どもに対する質問紙調査の中では非認知能力や学習方略に加えて学級の様子について尋ねており、本稿では学力に限らない多面的な教育成果に対する教員付加価値の推定をすることができた。そして、得られた教員付加価値の分布や相関を見ていくことで、教員付加価値が持っている性質を考察した。

結果として、学力や学習方略に対して個々の教員の違いが与える影響力が取りうる値の範囲は広がった一方で、非認知能力や学級での子ども・教員間の関係性に対して教員の違いは子ども全体の変動のうち少なくとも5%から8%以上を説明していた。更に、学力の教員付加価値と非認知能力の教員付加価値の間の相関は小さく、統計的に有意ではなかった。これは、学力を伸ばすことが得意な教員と非認知能力を伸ばすことが得意な教員が異なることを示唆する。最後に、高い教員付加価値を持つ教員の観察可能な特徴を見つけることはできなかった。すなわち、年齢や性別といった教員の観察可能な特徴と推定された教員付加価値との間に統計的に有意な相関関係はなかった。

これらの結果は教員付加価値研究において次の様な意味を持つ。まず第一に、非認知能力や学習方略に対する教員付加価値のあり方を調べた数少ない研究であるということである。管見の限りセルフコントロールや自己効力感、学習方略に対する教員付加価値を現代的な手法で推定した研究は本研究が初めてである。特に学習方略については、認知能力が伸ばすことができる教員であっても学習方略を伸ばすことが出来る教員とは限らず、そのため子どもの学習のあり方を教員が帰ると言うよりむしろ授業の方法など授業時間における教員のあり方の違いが教員付加価値の違いにもつながっているのではないかという示唆を得た。

第二に、日本における現代的な手法による教員付加価値として管見の限り初出であるということである。その結果、非認知能力に対する教員付加価値について、その大きさの水準については小さい可能性があるものの、日本においてもアメリカにおける研究と同様の結論を得た。ただし、よりバイアスが少ない推定の必要性や子どもの長期的なアウトカムに対する教員の影響など、日本においては教員付加価値研究はさらなる進展を必要とする。

本稿の構成は以下の通りである。第 II 節では先行研究を概観する。第 III 節ではデータについて述べる。第 IV 節では推定結果について述べる。第 V 節では本稿における結論を述べる。

## II 先行研究

教員の子どもの教育成果への影響についての研究は古くから存在し<sup>3</sup>、Hanushek and Rivkin (2006) によればそれらの研究は 3 つの文脈に類型化できる。1 つ目は労働市場一般における教員の立ち位置 (賃金や志望倍率) を用いて教員の分析を行うものである。例えば日本を対象にした研究である二木 (2017) は、労働市場における他職に比した教員の人気度を用いて分析を行なっている。これらの研究では総じて、他の産業と比較して優秀な人材が教員になっている事が子どもの成績を上げる可能性を検証している。

2 つ目は教員の属性が子どもの教育成果に与える影響を分析するものである。ここで言う教員の属性とは、具体的には教員免許の有無や経験年数などを指す。近年のこの種の研究例として Jackson and Bruegmann (2009) や Clotfelter et al. (2010) などを挙げることができる。しかし、これらの方向性での議論では研究によって結果が異なることが多く (Hanushek and Rivkin, 2006)、研究間で一致して「有効である」とされるような教員の属性は見つかっていない。

そして、3 つ目が子どもの教育成果に対する教員付加価値を計算して分析する研究群である。教員付加価値とは子どもの教育成果の成長に対する教員の貢献を指した言葉であり<sup>4</sup>、もとより観察不可能な概念である。そのため、多くの場合は教員ごとの教員付加価値をデータから推定することになる。しかし、この推定をどの様に定式化するかについては必ずしも決まりきった手法があるわけではない。近年は Kane and Staiger (2008) に示された経験ベイズ的な手法を用いて教員付加価値を推定する研究が多く (Chetty et al., 2014a; Lefgren and Sims, 2012, など)、これらの研究は総じて教員付加価値が子どもの教育成果に対して一定の重要な役割を担っていることを報告している。その推定が含意している仮定の妥当性を巡った議論 (Rothstein, 2010; Kinsler, 2012) もなされているが、一方で実験的な状況を用いて教員付加価値が因果的な効果を持つ事を示す研究が存在する (Chetty et al., 2014a)。

しかし、学力以外の教育成果に対する教員付加価値の研究には蓄積が乏しい。教員が子どもに影響を与えるのは学力だけではなく、教室におけるコミュニケーションを通じて教員は様々な影響を子どもに与える。例えば Chetty et al. (2014b) では、良い教員は子どもの大学進学率や将来的な収入、貯蓄率や 10 代での妊娠率に影響を明らかにしていることを示している。しかし、上述の研究を始め多くの教員付加価値を扱った研究はテストスコアに対する付加価値を分析しており、学力以外への教育成果に対して教員が持っている影響力についてはよく分かっていない。数少ない例として、Jackson (2018) は子どもの出席率や停学といった社会行動に対する教員付加価値を推定し分析を行っている。しかしこの結果がどれほど一般的であるかなどはよく分かっておらず、更なる研究が必要である。

更に、非認知能力と教員付加価値に関する議論には教員の能力の多次元性という論点が存在する。すなわち、学力以外の教育成果を対象にして教員付加価値を計算できるということは、ある教員の教員付加価値は学力や非認知能力など教育成果の数だけ存在することになる。そのため、教員の能力は「良い/悪い」で表される 1 次元的な概念ではなく、「学力について良い/悪い」「非認知能力について良い/悪い」と続く多次元的なものになる。そして、この論点は、教員評価などの文脈では一層重要な問題になる。何故ならば、ある教員の付加価値への評価として「良い/悪い」を一概に定めることができなくなるからである。認知能力を伸ばすことに長けた教員が、非認知能力を伸ばすことに長けている保証はない。もしくは同じ認知能力であっても国語を伸ばすことに長けた教員が算数を伸ばすことにも長けてい

るとは限らない。このように、教員の能力の多次元性について言及した論文として、上述の Jackson (2018) では中学 3 年生という比較的高年齢の子どもを対象に、認知能力と非認知能力の教員付加価値の間の相関が弱いことを報告している。また、Goldhaber et al. (2013) では米国ノースカロライナ州のデータを用いて、小学校レベルでの教員の付加価値の科目間相関を算出し、国語科目と数学科目の付加価値相関係数が 0.7 程度であることを報告している。しかし、非認知能力についての研究がもとより少ないため、同様に上述の結果がどれほど一般的であるかなどはよく分かっていない。

最後に、日本を対象にした教員の付加価値についての研究は非常に少ないことについて詳述する必要がある。国際学力調査を用いた教育生産関数の推定の文脈で教員の影響を分析する場合は存在する (Hojo and Oshio, 2012; Hojo, 2012)。しかし、これらは教員の効果を分析することを目的とした研究ではない。上述の二木 (2017) は日本におけるある程度の規模のデータセットを用いた数少ない教員効果に関する研究であり、教員の質が子どもの欠席率を減少させようとしている。しかし、あくまで都道府県レベルでの集計値を用いた分析であり、個々の教員の違いを十分に反映できていない可能性がある (集計バイアス)。

### III データ

本稿では 2016 年－2018 年 (調査は 2015 年から行われている) にかけて行われた日本のとある県 S における学力テストのデータを用いる。テストは毎年 4 月に行われ、テストと同時に非認知能力や普段の生活の様子を尋ねる質問紙調査を行った。テストを受けた生徒にはユニークな ID が振られ、経年で同一生徒を特定することが可能な設計になっている。調査の対象になったのは、S 県における政令指定都市を除く全ての市町村の公立小学校及び公立中学校に所属する小学 4 年生から中学 3 年生までの生徒である。すなわち、小学 4 年生から中学 3 年生を対象にした悉皆調査として行われた。およそ生徒数では 30 万人程度、学校数では 1064 の学校 (小学校が 708 校、中学校が 356 校) の生徒が参加した。

上記の調査に加え、S 県の市町村の一つである T 市に勤める小中学校の全ての教員を対象に質問紙調査を 2016 年から 2018 年にかけて行った。これは教員の授業での実践を尋ねるもので、教員一人一人にユニークな ID が振られ経年で同一教員を特定することが可能な設計になっている。ただし日本における公立学校の人事制度においてある教員は必ずしも同じ学校に所属し続けるわけではなく、T 市以外の市町村への転出も多い。また全ての教員が毎年学級を一つ担当するわけでもなく、中には副担任などの形で自分の担当学級を持たない教員もいる。これらの理由から、複数年度で担当クラスを持っている教員は少なくなってしまうことに注意をしなければいけない。本項では教員のデータを用いるために、この T 市のみのデータを取り扱った。

また今回分析に用いたのは小学校での教員データだけで、中学校での教員データは分析に用いなかった。その理由は、小学校と中学校では学級に対する教員の割り当て方として教科担任制か学級担任制かという違いがあるからである。日本の教育システムにおいては、多くの小学校の教員は一つの学級を受け持ち全ての科目を教えることになる<sup>5</sup>。一方で、多くの中学校では教員は専門科目を持ち、複数の学級で授業を行うことになる。その場合、中学生の能力の成長をどの教員の貢献と見做すことができるかは不明瞭である。例えば、子どもの国語の能力が成長していた場合は国語を担当した教員の貢献であると言えるだろう。しかし、非認知能力が成長していた場合は、どの科目の教員の貢献なのかを判断するのは困難である。以上の理由から、本稿では小学校の教員のみを分析の対象とした。まとめると、本稿で分析の対象となるのは、T 市の小学校 (その小学校数は 12 校) に通う 4 年生から 6 年生の子ども及び教員である。

S 県 T 市の子どもが置かれている状況は必ずしも日本の教育状況そのものを表しているわけではないため、外的な妥当性には十分な検討が必要である。S 県は東京近郊に所在する県であり、その中でも T 市は人口 13 万人程度の中規模都市である。また、T 市の公立小学校における 2016 年度の所属教員 1 人当たり児童数は 20.5 人であり<sup>6</sup>、日本の公立小学校平均の 15.5 人と比べると大きい。一方で、T 市に住む人の平均的な課税対象所得は 350 万円<sup>7</sup>であり日本平均の 321 万円と比べ高く、日本の中では比較的所得の高い人が居住する都市で

あると言える。この様に T 市をとりまく教育環境や家庭の状況は必ずしも日本のあらゆる地域と同一ではない。そのため、本校における T 市を対象にした分析が他の市町村を対象にしたときにどの程度適用可能であるかについては十分に注意をする必要がある<sup>8</sup>。

表 1 および表 2 ではデータで用いる変数についての記述統計を示した。表 1 には子どもを単位とする変数について平均値、標準偏差及びサンプルサイズを年度・学年ごとに示した。「標準化の有無」と記された列に「○」がついている行の値は全て、標準化を行った後の値である<sup>9</sup>。また、調査が行われたのは毎年 4 月であることには注意が必要である。ある年の 4 月に行われた調査の結果は、その前年度の状況を反映していると考えられる。例えば、表 1 において 2016 年 5 年生の値は、2015 年 4 年生時の状況を反映した値と考えられる。そのため、教員もその前年の状況に紐付けられる（例えば、2016 年 5 年生の子どものデータに紐付けられる教員は、2015 年 4 年生時点の教員である）。表 2 では得られた教員についての情報をまとめている。表 2 上段では教員の数や年度及び学年ごとに分けて示し、表 2 下段では教員を単位とする変数の記述統計を示した。以降ではデータで用いる変数について記述する。

## 認知能力

認知能力を表す値として、国語と算数の学力の推定値を用いる。すなわち、S 県学力調査では小学生に対して国語と算数のテストを実施しており、その結果から IRT (Item Response Theory) を用いて個人の科目ごとの学力の推定値を算出した。加えて本項では分析にあたって全ての学力の推定値を年度・学年及び科目ごとに平均 0・分散 1 に標準化を行なった。この操作によって標準化された学力は、実質的に学年内での相対的な高低を表すことになる。この IRT によって算出された標準化された学力推定値を以降では認知能力の値として分析の対象とする。

## 非認知能力及び学習方略

また、調査ではテストとは別に質問紙調査によって、子どもの非認知能力や学習方略を尋ねている。本稿ではそのうち、(1) セルフコントロール (Duckworth et al., 2013)、(2) 自己効力感 (Pintrich et al., 1991)、(3) 学習方略 (佐藤, 1998) を教育成果として分析の対象にする。これらの値はそれぞれ複数の質問項目 (表 A.1・表 A.2) の回答 (5 件法) を足しあげることで得られる。更に分析においては学年ごとに平均 0・分散 1 に標準化を行い、学年内での相対的な高低を表す値に直した。

このうちセルフコントロールと自己効力感の 2 つの非認知能力は一部の学年でしか尋ねていないことに注意が必要である。すなわち、セルフコントロールは 2016 年に中 1 年生であった学年と 2017 年に 5 年生であった学年にのみ尋ねており、自己効力感も 2016 年に 5 年生であった学年にのみ尋ねている。そのため非認知能力をアウトカムとして用いた分析では、他の変数をアウトカムとした分析に対して用いることができるサンプルが少なくなる。

## 子ども・教員間の関係性

更に学級風土 classroom climate、特にその中でも学級での子ども・教員間の関係性を対象にして教員付加価値を推定する。学級風土 (もしくは学校風土 school climate) とは子どもの学級生活の質や特徴を表し、主に教育心理学や教育社会学で研究されている概念である (Wang and Degol, 2016; Berkowitz et al., 2017)。その中では学級・学校における規範、価値観及び人間関係などが子どもの学級での生活に働きかけ、学力や行動にまで影響を与えるのではないかと考えられている (Hopson and Lee, 2011)。そして、その学級風土の要素の一つとして、学級における子ども・教員間の関係性は挙げられている (Wang and Degol, 2016)。

本稿では、この学級における子ども・教員間の関係性について分析を行う。調査では前年度の学級における教員及び友人との関係について、4 つの質問 (全て 4 件法) によって子どもに尋ねた (表 A.2)。それらの項目全てを足し合わせて学級での子ども・教員間の関係性を表す変数とした。更に分析においては学年ごとに平均 0 分散 1 に標準化を行った。

## その他の変数

調査において子どもの家庭での状況などを子どもに尋ねて調査している。そのうち、本稿では(1)週あたりの通塾時間(8件法)及び(2)家庭にある本の冊数(5件法)を用いる。これらは共に子どもの家庭に社会経済的地位(Socio-economic Status、以降SESと表記)を表す代理変数として用いる。前者は経済資本を表す変数であり、後者は文化資本を表す変数である。

## 教員

ある教員は必ずしもT市に留まり小学校高学年を対象に教鞭を取り続けるとは限らない。表2上段では教員質問紙への回答を得ることができた教員の数や年度及び学年に分けて示し、同時に年度ごとに調査に新たに参加した教員の数及び調査への参加が認められなくなった教員の数や示した。調査に参加する教員数は学校内の配置転換や人事異動によって変わると考えられる<sup>10</sup>。例えば、T市からの転出もしくは転入によって教員質問紙への調査の可否は変わってくる。そのようにして、ある年に調査に参加した教員のうち凡そ半分は次の年には調査に参加しなくなる。例えば、2017年度において調査に参加した小学校教員数は80人であったが、その内37人が2018年度の調査には参加していない。このように、調査に参加した教員ののべ総数に対して、継続して調査に参加し続ける教員数は少なくなってしまう。本稿では2016年度－2018年度に渡って一度でも調査に参加した教員を対象に分析を行うが、そのうち単年度しか調査に参加していない教員も存在することには注意が必要である。

また小学4年から小学6年の学級を担当した全ての教員が調査に参加したわけではない。調査に参加していない教員の数や示すため、表2上段では調査の対象となった学級の数や同時に示している。例えば、2016年度において調査の対象となった学級は109学級あるのに対し、調査に参加した教員は82人であり全体の約80%である。これらの調査に参加していない教員が存在することには注意が必要である。

教員質問紙では担当した学級の他に、教員の特徴について尋ねている。表2下段にはその記述統計を示した。例えば平均的な年齢は36.2歳であり、全国の公立小学校教員の平均年齢である43.4歳<sup>11</sup>に比べると若いことがわかる。また、教員の出身大学については注意が必要である。すなわち、出身大学(大学院)について回答が得られたのは62人のみであり、データを得られた教員の総数の半分以下であった。これは、2016年に行った調査においてのみ教員に卒業した大学(大学院)を尋ねることが可能であり、2017年度－2018年度に新たに調査に参加した教員については相当する質問を行うことができなかったためである。

年度 学年		2016			2017			2018		
		5	6	7	5	6	7	5	6	7
	標準化の有無									
国語	○	-0.020 (849,1.009)	0.020 (837,1.000)	-0.003 (801,1.001)	0.049 (830,1.002)	0.013 (944,0.983)	0.025 (713,1.002)	0.022 (1212,0.984)	0.016 (1045,0.990)	0.020 (955,0.984)
算数	○	-0.054 (848,0.915)	0.024 (837,0.987)	0.003 (801,0.980)	0.046 (830,1.005)	-0.007 (944,0.989)	0.006 (713,0.998)	0.019 (1212,0.990)	0.012 (1045,0.964)	0.006 (955,0.976)
学習方略	○	-0.010 (785,0.987)	-0.026 (777,0.992)	-0.035 (756,1.010)	-0.028 (767,1.004)	-0.013 (876,0.992)	0.040 (674,0.991)	0.009 (1106,0.999)	-0.003 (968,1.009)	-0.006 (887,0.997)
セルフコントロール	○			-0.028 (795,0.997)	-0.016 (817,1.009)				-0.019 (1030,1.016)	
自己効力感	○	-0.031 (839,0.998)				0.001 (925,0.996)				-0.007 (942,1.003)
子ども・教員間の関係性	○	-0.057 (837,1.015)	-0.045 (834,1.037)	0.070 (796,0.959)	-0.022 (828,1.006)	0.010 (942,1.003)	0.084 (713,0.963)	0.002 (1209,1.000)	0.022 (1045,0.999)	-0.008 (955,1.013)
家庭にある本の冊数		3.018 (837,1.130)	3.220 (833,1.117)	2.955 (796,1.174)	3.048 (827,1.097)	3.116 (941,1.070)	3.179 (711,1.132)	3.145 (1205,1.143)	3.202 (1043,1.099)	3.184 (953,1.142)
通塾時間		2.850 (834,2.227)	2.765 (831,2.160)	2.700 (797,1.886)	3.150 (800,2.431)	2.992 (920,2.245)	3.010 (696,1.982)	3.226 (1187,2.438)	3.179 (1026,2.325)	2.980 (947,1.990)

表 1: 記述統計 (子ども)

注) 本表では、本稿で用いている変数のうち子どもを単位とする変数の記述統計を示している。各々のセルでは、対応する変数について「平均 (サンプルサイズ、標準偏差)」という記述方法を用いている。データ単位は全て、子ども 1 人である。

(a) 調査に参加した教員 (学級) の数							
	4 年	5 年	6 年	ALL	IN	OUT	学級数
年度							
2015	28	26	28	82	82	0	109
2016	26	30	24	80	42	44	120
2017	37	32	34	103	60	37	122

(b) 教員の特徴				
	年齢	教員歴	女性教員	出身大学偏差値
	36.210	12.416	0.509	49.338
教員	(167,10.320)	(166,10.101)	(167,0.501)	(62,7.465)

表 2: 記述統計 (教員)

注) 本表では、本稿で用いている変数のうち教員を単位とする変数の記述統計を示している。上段では、本稿で教員付加価値の推定の対象とした教員の数を示した。小学校では 1 人の教員が 1 つの学級を担当するため、教員数はそのまま学級数を表す。行は年度を表す。1 列目 (「4 年」) から 3 列目 (「6 年」) までは、学年ごとに教員が何人いるかを示し、4 列目 (「ALL」) はその合計を表す。5 列目 (「IN」) は新たに教員質問紙に参加した教員数を表し、6 列目 (「OUT」) は教員質問紙に参加しなくなった教員数を表す。7 列目 (「学級数」) は、調査の対象となった学級数を表す。下段では、教員の特徴を表す変数の記述統計を示した。各々のセルでは、対応する変数について「平均 (サンプルサイズ、標準偏差)」という記述方法を用いている。データ単位は全て、教員 1 人である。

## IV 推定戦略と推定結果

### 1 教員付加価値の推定

教員付加価値を推定するために、ある子どもの教育成果を次のような教育生産関数で表すことができるとする。

$$y_{it} = \alpha y_{it-1} + X_{it}\beta + \mu_j + \theta_c + \epsilon_{it} \quad (1)$$

$y_{it}$  は  $t$  年における子ども  $i$  の教育成果を表す。本稿では教育成果として、学力 (国語、算数) の他に非認知能力 (自己効力感、セルフコントロール) や学習方略及び学級での子ども・教員間の関係性を考える。 $X_{it}$  は観察可能な子ども  $i$  の特徴を表し、本稿では SES の代理変数として家庭にある本の冊数及び通塾時間を用いる。 $\theta_c$  は子ども  $i$  が所属する学級  $c$  の効果を表し、 $\mu_j$  は学級  $c$  を担当する教員  $j$  の付加価値を表し、 $\epsilon_{it}$  は子ども  $i$  の  $t$  年におけるその他の要因全てを表す。 $\mu_j \cdot \theta_c \cdot \epsilon_{it}$  は全て観察不可能であり、互いに独立であることを仮定する。

式 1 は付加価値モデルと広く呼ばれるモデルの一つである (Todd and Wolpin, 2003)。前期の能力を表す  $y_{it-1}$  をコントロールすることで、これまで蓄積された能力もしくは時間変動しない要因の影響を取り除くことを意図している。そのため、分析に用いる変動は子どもの能力の成長である。教員付加価値はこの子どもの能力の成長に対する教員の貢献  $\mu_j$  を評価することになる。

本稿では教員付加価値  $\mu_j$  を Kane and Staiger (2008) の手続きに従って求める。Kane らの手法では式 1 を推定し、その残差を用いて  $\mu_j$  の推定値を計算する。すなわち、式 1 を  $s_{it} = y_{it} - \alpha y_{it-1} - X_{it}\beta = \mu_j + \theta_c + \epsilon_{it}$  と書き直せば、 $s_{it}$  は  $\mu_j$  という求めたい値に  $\theta_c$  と  $\epsilon_{it}$  というノイズが加わった情報を表す変数と見做せる。この時、ある学級  $c$  はある教員  $j$  によって担任されるため、 $s_{it}$  を子ども  $i$  が所属するクラス (クラス  $c$  に所属する子どもの集合



を  $I_c$  と表記する) で平均した値  $s_{jc}$  は

$$\begin{aligned} s_{jc} &= \frac{\sum_{i,i \in I_c} s_{it}}{\|I_c\|} \\ &= \mu_j + \theta_c + \frac{\sum_{i,i \in I_c} \epsilon_{it}}{\|I_c\|} \\ &= \mu_j + v_{jc} \end{aligned}$$

となる。ただし学級  $c$  に所属する子どもの集合を  $I_c$  で表している。この  $s_{jc}$  は教員付加価値  $\mu_j$  についてのノイズ付きの情報と見做せるため、教員  $j$  の教員付加価値の期待値  $\mu_j$  はある重み  $\alpha_{jc}$  を用いて

$$E(\mu_j | s_{jc}) = \sum_{c, c \in I_j} \alpha_{jc} s_{jc}$$

で表される<sup>12</sup>。本稿では以降特に断り無く「推定された教員付加価値」と記述されている場合、上述の手続きで求めた教員付加価値を指す。現代的な教員付加価値研究のほとんどが、このデータから得られる情報にその正確度に応じた重みを与える手法に準拠し教員付加価値を推定しており (Kane and Staiger, 2008; Lefgren and Sims, 2012; Chetty et al., 2014a; Jackson, 2018; Kraft, 2019)、この手法で求めることはそれらの研究と比較しながら本研究を評価することが出来るため有用である。

教員付加価値を求める手法には幾つかの派生形がある。Kane and Staiger (2008) が示した教員付加価値を推定する方法は、ある教員が担当するクラスや生徒数は有限であることから必ずしも一致性などの統計的な性質が保証されているわけではない。また教員付加価値に期待される性質も状況によって異なる。例えば、上述の定式化では教員付加価値は時間を通じて一定であることを前提にしているが、新任の教員が業務経験を通じて徐々にその能力を高めていくことなどを考慮に入れようとするれば教員付加価値は時間を通じて変化しうるものとして定式化する必要がある。そのため、教員付加価値を推定する手法には複数の教育成果で共通する教員の付加価値を評価した Lefgren and Sims (2012) や教員の付加価値の時間による変化を許容する Chetty et al. (2014a) などの幾つかの派生系が存在する。更に、似た手法として式 (1) のうち  $\theta_c + \epsilon_{it}$  を誤差項を含む観察できない要因  $v_{it}$  などとして  $y_{it} = \alpha y_{it-1} + X_{it}\beta + \mu_j + v_{it}$  という式を推定し、そのうち教員の固定効果  $\mu_j$  を教員付加価値として用いる研究も存在する<sup>13</sup>。本稿では比較のため、この固定効果を用いた推定も報告する。

この式 1 において  $\mu_j$  を識別するための仮定は、 $\mu_j$  が  $\theta_c$  や  $\epsilon_{it}$  と独立である事である。すなわち、教師  $j$  が内生的に学校や学級に配置されているはず、そのような関係があれば推定された教員付加価値はその割り当て方に応じたバイアスを含むことになる。このバイアスの方向はその内生的割り当てがどのようにおこなわれているかによって異なる。例えば、経験の浅い教員の授業負担を軽くするために SES の高い子どもを優先的に割り当てているとする。そのような子どもが塾などの外部の教育リソースを活用して能力を伸ばしているといった場合、担当教員の教育効果を高く見積もってしまう可能性がある。本研究は実験的もしくは準実験的な状況を利用した研究ではないため、この仮定が成立しているかどうかは注意深く検討する必要がある。Appendix B 節では、観察可能な教員の特徴と担当している子どもの特徴の間にある相関を見る中でそのような内生的な配置の可能性を検討し、その結果として観察可能な特徴においては内生的な配置の証拠は見つからないことを示している。

また  $\mu_j$  が  $\theta_c$  や  $\epsilon_{it}$  に対して外生であったとしても、個々の  $\mu_j$  を正確に推定するためにはある教師に対して複数の学級のデータが必要である (Jackson et al., 2014)。すなわち、教員が担当するクラス  $c$  固有の効果  $\theta_c$  は学級の生徒間では固有であり、学級内のサンプルを多く集めたところでその影響をとりのぞくことはできない。そして、複数の学級のサンプルを用いることで、そのノイズの影響を軽減することができる。逆に言えばデータにおいて担当学級の数が少ない教員付加価値の推定値  $\hat{\mu}_j$  は  $\theta_c$  の影響を強く受けてしまう<sup>14</sup>。Kane and Staiger (2008) の様な経験ベイズ的手法が用いられるのは、まさにその様なノイズの大きさ

だけ教員の教員付加価値  $\hat{\mu}_j$  を割り引くためであった。本稿では小学校でのデータを取り扱うが、小学校ではある教員は年度ごとに1つの学級を担当するため、中には担当した学級が1つしかないような教員も存在する。そのため、 $\hat{\mu}_j$  の標準偏差  $\sigma_\mu$  を報告する際には(表3)のような教員をデータから除いた場合の結果も報告する。

特にこの問題が重要になってくるのは、年度が変わってもある教員( $j'$ と表記する)は同じクラスを持ち上がって担当し、教員に割り当てられた子どもの構成が時間を通じて一定のケースの場合である(このような教員割り当てを「持ち上がり」とここでは表記する)。「持ち上がり」が教員  $j'$  に発生している時、教員効果  $\mu_{j'}$  とクラス効果  $\theta_c$  は識別することができない。ある教員  $j'$  の付加価値の推定値はある同じクラスの観察不可能な特徴による効果を含むことになり、内生性によるバイアスはより深刻化する。日本の教育システムにおいては、ある教員とあるクラスの割り当てがどの程度持続するかは自治体や学校に依存する。クラス替えが毎年発生し教員にクラスを毎年割り当てし直す学校もあれば、2年から3年程度同じクラスを教員に割り当て続ける学校もある。しかし、本稿におけるデータ内で「持ち上がり」の発生割合を計算すると、T市においては持ち上がりクラスは全体の1%と非常に稀にしか起きていなかった。そのため、クラスの持ち上がりの問題はほとんど発生していないと考えることができる。

更に、推定手法によって教員効果  $\mu_j$  の分散  $\sigma_\mu$  の推定値には異なるバイアスが存在する。教員固定効果を用いた推定では、測定誤差と真の教員効果  $\mu_j$  の識別ができない。そのため、その推定値の標準偏差  $\sigma_{\hat{\mu}_{FE}}$  は真の分散  $\sigma_\mu$  よりも大きくなるというバイアスが存在する。一方 Kane and Staiger (2008) による推定方法で求められた推定値の分散  $\sigma_{\hat{\mu}_{KS}}$  は真の分散  $\sigma_\mu$  よりも小さくなるというバイアスが存在する (Jacob and Lefgren, 2005, Online Appendix C; Kraft, 2019)。すなわち、真の教員付加価値の分散  $\sigma_\mu$  と  $\sigma_{\hat{\mu}_{FE}}$  および  $\sigma_{\hat{\mu}_{KS}}$  の間の関係は次の様にまとめることができる。

$$\sigma_{\hat{\mu}_{KS}} \leq \sigma_\mu \leq \sigma_{\hat{\mu}_{FE}} \quad (2)$$

## 2 教員付加価値の分布

推定された教員付加価値  $\mu_j$  について、その分布を図1に示した<sup>15</sup>。図においては教員付加価値  $\mu_j$  のヒストグラムを棒グラフで示し、カーネル密度推定によって求められた密度関数を線グラフによって示した。これらの図からは、実際に推定された教員付加価値の分布を視覚的に確認することができる。どのような教育成果を対象にしても、教員付加価値は正規分布のような形の分布関数を描いており、極値が複数あるような複雑な形をしているわけではない<sup>16</sup>。

次に、教員の影響の大きさを評価するために、表3に教員付加価値の推定値の標準偏差を示した<sup>17</sup>。(1)列には推定した教員付加価値の標準偏差を教育成果ごとに記している。比較のために以降の列では、他の手法で求めた教員付加価値や元々の記述統計を同時に示している。(2)列には固定効果法によって求めた教員付加価値の標準偏差を示した。(3)列には教員付加価値が推定された人数を記した。(4)列には教育成果の元々の標準偏差を示した。III節で示している通り各教育成果に対しては標準化を行っているため、ここでの標準偏差はほとんど1に近い。(5)列には教育成果のクラス平均の標準偏差を示した。

結果として、認知能力や学習方略に対する教員付加価値の標準偏差が取りうる値の範囲は広いことがわかった<sup>18</sup>。すなわち、(1)列において認知能力(国語、算数)を対象にした教員付加価値の標準偏差はそれぞれ  $0.003 \cdot 0.013$  であり、学習方略を対象にした教員付加価値の標準偏差は  $0.008$  である。これは全体の変動((4)列)やクラス平均の変動((5)列)のうちそれぞれ高々1%・3%程度しか推定された教員付加価値は説明していないことを表す。一方で(2)で示した固定効果によって推定された教員付加価値の標準偏差は、認知能力(国語、算数)を対象にした場合はそれぞれ  $0.140 \cdot 0.141$ 、学習方略を対象にした場合教員付加価値の標準偏差は  $0.194$  である。これは、全体の変動のうち15%から20%を教員付加価値が説明していることを表す。式2で示したように真の教員付加価値の標準偏差は Kane and

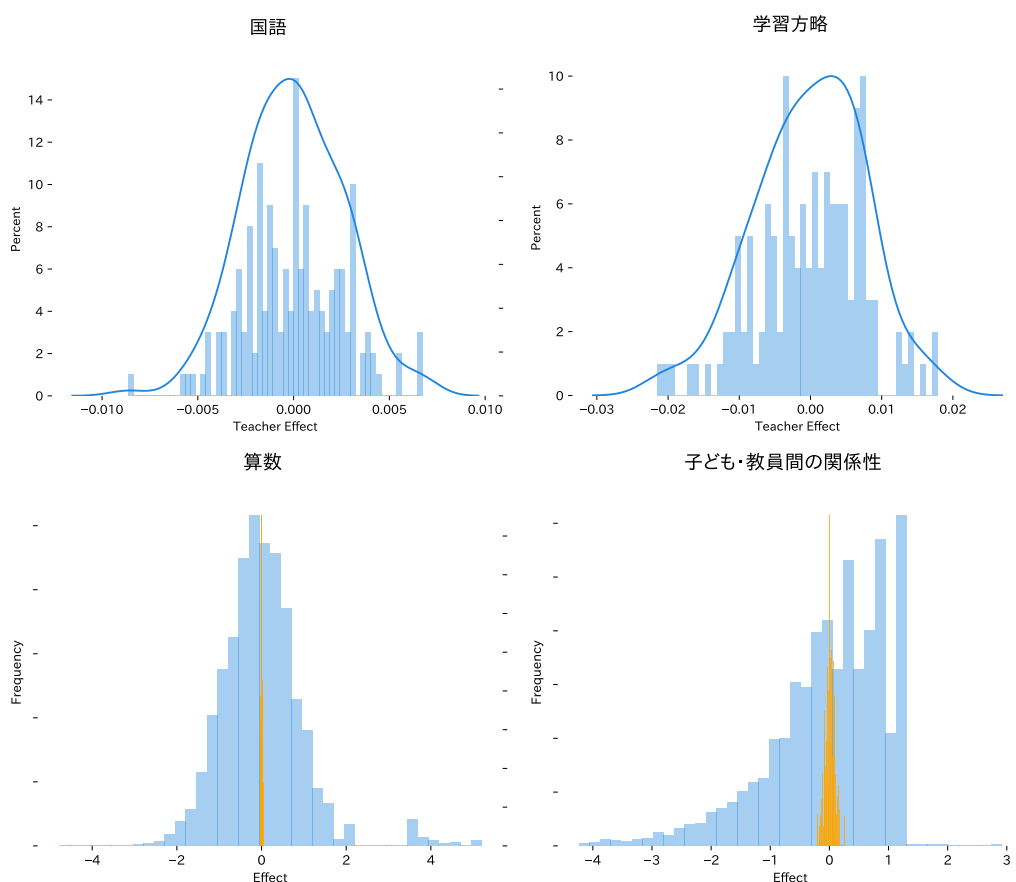


図 1: 推定された教員付加価値の分布

注) 本図では、推定された教員付加価値の分布を示している。上段の図では教員付加価値のヒストグラムを棒グラフで示しており、推定された教員付加価値をカーネル密度推定することによって得られた密度関数を線グラフで示している。下段の図では、推定された教員付加価値の分布及び子どもの教育成果の分布 (ヒストグラム) を示している。オレンジ色で塗られた分布は推定された教員付加価値の分布を表す。青色で塗られた分布は子どもの教育成果の分布を表す。

	(1) 教員付加価値 (標準偏差) KS	(2) 教員付加価値 (標準偏差) FE	(3) 教員の数	(4) $Std(y_{it})$	(5) $Std(\bar{y}_c)$
国語	0.003	0.140	170	0.996	0.330
算数	0.013	0.141	170	0.994	0.343
学習方略	0.008	0.194	140	0.999	0.305
セルフコントロール	0.051	0.165	54	1.000	0.292
自己効力感	0.069	0.159	54	0.997	0.242
子ども・教員間の関係性	0.078	0.323	170	0.999	0.397

表 3: 教員付加価値の標準偏差

注) 本表では求めた教員付加価値の標準偏差を示している。各々の行は教員付加価値を求める対象とした教育成果を表す。(1) 列では本稿で主に用いている教員付加価値の標準偏差を示した。(2) 列では、比較のために教員固定効果推定で求めた教員付加価値 (教員固定効果) の標準偏差の値を示した。(3) 列には教員付加価値が推定された人数を記した。(4) 列では子どもの教育成果の標準偏差を示した。(5) 列ではクラス平均の標準偏差を示した。

Staiger (2008) による方法で求められた標準偏差と教員固定効果による方法で求められた標準偏差の間にあると考えられる。そのため、全体の変動に対して認知能力や学習方略に対する教員付加価値が説明するのは1%から20%の間のいずれかであるといえる。

この「少なくとも1%」という教員付加価値の大きさは、他の研究と比較した時に必ずしも大きいわけではない。Chetty et al. (2014a) では小学校において英語では11%、数学では15%ほどの教員付加価値の大きさを報告している。Kraft (2019) では認知能力に関する各種アウトカムにおいて13%から18%の教員付加価値の標準偏差を報告している。しかし、本研究で得られた分析では「認知能力や学習方略に対する教員付加価値は1%から20%の間にある」というところまでしか言えず、これらの海外における研究よりも教員の役割が小さい可能性もあれば大きい可能性もある。無論もしも教員付加価値の標準偏差の真の値が上限値の20%程度あれば海外の研究と比較しても同水準以上であるが、一方で教員付加価値の標準偏差の真の値が下限値の1%であれば1/10程度と非常に低い水準になる。結果として個々の教員の違いが子どもの学力に与える影響力の大小を評価することは難しいと言える。

Kane and Staiger (2008) の方法で導出した教員付加価値については図1の下段左図から視覚的に確認することができる。図1の下段左図では、算数への教員付加価値のヒストグラム（オレンジ色）と子どもの算数の学力のヒストグラム（青色）を同時に描いている。しかし、オレンジ色で描かれた算数への教員付加価値の分布は非常に小さく、視認することは難しい。この図からも、Kane and Staiger (2008) の方法で推定された場合認知能力に与える個々の教員の影響は小さいことを見て取ることができる。

次に、非認知能力や学級での子ども・教員間の関係性に対する個々の教員の影響力は全体の変動に対して少なくとも5%から8%以上あった。すなわち、(1) 列において示されたKane and Staiger (2008) の方法によって求められた教員付加価値の標準偏差は非認知能力（セルフコントロール・自己効力感）を対象にした時にはそれぞれ0.051・0.069であり、学級での子ども・教員間の関係性を対象にした時には0.078である。これは元々の教育成果の標準偏差が1であることを踏まえると、個々の教員の付加価値の違いは全体の変動に対して5%から8%ほどを説明したことを意味する。そして式2からは、これらの値は真の値に対する下限として捉えることが出来る。

この「少なくとも5%から8%」という教員付加価値の大きさは、他の研究と比較した時にも同水準か少し小さい程度である。Jackson (2018) では中退率で測定した教員付加価値の標準偏差は10%、欠席率で測定した教員付加価値の標準偏差は3%であったとしている。Kraft (2019) では教員付加価値は成長志向やGRITといった非認知能力の変動を15%程度説明するとしている。これらの研究と比較すると今回得られた非認知についての教員付加価値の大きさは少し小さい程度である。無論この値は下限値であるため、場合によっては海外の研究と同水準の大きさである可能性もある。結果として、認知能力に対する教員付加価値と異なり、非認知能力や子ども・教員間の関係性に対しては教員による一定の大きさの影響力を捉えることが出来たと言える。

このような関係性については図1の下段右図から視覚的に確認することができる。図1の下段右図では、学級での子ども・教員間の関係性への教員付加価値（オレンジ色）及び元々の子どもによる回答（青色）を同時にヒストグラムによって描いている。そしてこの図表では、図1の下段左図では算数への教員付加価値とは異なり、オレンジ色で描かれた教員付加価値の分布を視認することができる。

総じてこれらの結果は、小学校において個々の教員は非認知能力や子ども・教員間の関係性において重要な役割を担っていることを示している。子供を担当する教員が変わることは、子どもとのコミュニケーションのあり方の違いなどを通じて、その子どもの非認知能力や教員との関係性を大きく変えうる。この非認知能力の変化は、労働条件などの長期的に成果に結実することを多くの研究が示している。教員付加価値研究では、Jackson (2018) は非認知についての高い教員付加価値は高校卒業や大学入学といった長期的なアウトカムにおいても正に有意な影響を持っていることを示している。またHanushek (2011) は質の高い教員の雇用がもたらす社会的価値を試算し、平均よりも1標準偏差優秀な教員は20人の学級であれば現在価値で40万ドルの社会的価値を生むとしている。これらの研究はアメリカを対象に行われており日本では同様の研究がないことから、これらの議論をそのまま適用すること

が可能かどうかについてはさらなる議論が必要ではあるものの、教員の違いは非認知能力の改善を通じて子どもの長期的な達成や社会的便益において重要な役割をもちうるだろう。一方で認知能力に対しては、推定結果が指し示す教員による違いの大きさの推定値の範囲は広く、外国の研究との比較などからその大小について議論することは難しかった。そのためその長期的影響においても、国による違いの影響などを除いてもなお、認知能力についての教員付加価値の長期的影響を研究した Chetty et al. (2014b) や上述の Hanushek (2011) が示すような高い効果がある可能性もあれば、それらの研究ほどの効果が無い可能性もある。

### 3 教員付加価値間の相関

次に教員付加価値間の関係性について考えるため、表4に教員付加価値の項目間相関を示した。表の行と列は教育成果の種類を表し、各々のセルは対応する教員付加価値間の相関を表す。例えば、[1] 行 (1) 列には、国語の教員付加価値と算数の教員付加価値の間の相関が示されている。

		(1) 算数	(2) 学習方略	(3) セルフコントロール	(4) 自己効力感	(5) 子ども・教員間の関係性
[1]	国語	0.488*** (0.067)	-0.017 (0.085)	0.213 (0.136)	0.087 (0.138)	0.093 (0.077)
[2]	算数		-0.143* (0.084)	0.101 (0.138)	0.211 (0.136)	0.237*** (0.075)
[3]	学習方略			0.345** (0.130)	0.510*** (0.119)	0.248*** (0.082)
[4]	セルフコントロール				0.254 (0.306)	0.363*** (0.129)
[5]	自己効力感					0.344** (0.130)

表 4: 教員付加価値の項目間相関係数

注) 本表では教員付加価値の項目間の相関を示している。各々のセルは相関係数を表し、括弧の中は標準誤差を表す。各々の行及び列は教員付加価値を求める対象とした教育成果を表す。「\*\*\*」・「\*\*」・「\*」はそれぞれ 1%・5%・10%の水準で統計的に有意であることを表す。

結果として第1に、子どもの国語の学力を向上させた教員は、算数の学力をも向上させた傾向があった。具体的には、国語と算数の教員付加価値の間の相関 ((1) 列 [1] 行) は統計的に有意であり、相関係数も 0.488 と正である。すなわち、国語を教える際に高いパフォーマンスを発揮している教員は算数でも高いパフォーマンスを発揮している。高いパフォーマンスを発揮している教員はある特定の科目を教えることが得意というよりむしろ、どのような科目であっても知識を伝達することを得意としているのかもしれない。この結果は英語と国語への教員付加価値の間に 0.7 の高い相関があることを示した Goldhaber et al. (2013) とも整合的な結果である。

第2に、子どもの非認知能力を向上させた教員は、学習方略や学級での子ども・教員間の関係性をも向上させた傾向がある。すなわち、学習方略と非認知能力 (セルフコントロール・自己効力感) 及び学級での子ども・教員間の関係性の教員付加価値はその間に統計的に有意な相関があり<sup>19</sup>、相関係数も正である。この結果は、子どもとの友好的コミュニケーションを通じて学級での子ども・教員間の関係性がよくなり、結果として子どもたちの非認知能力や学習方略が向上したと解釈することができるだろう。

第三に、学力 (国語・算数) の教員付加価値と学習方略・非認知能力の教員付加価値の間の相関は小さく統計的な有意性もなかった。例えば、算数を対象とした教員付加価値とセルフコントロールを対象にした教員付加価値の間の相関係数は 0.101 ([2] 行 (3) 列) と小さく統計的に有意ではない。学力を伸ばすという意味で「良い」教員と学習習慣や非認知能力を伸ばすという意味で「良い」教員は必ずしも同一ではない。この結果は、学力に対する教員付加価値と授業への出席率や停学などの生徒の行動に対する教員付加価値の間には弱い相関しかないとした Jackson (2018) や Kraft (2019) とも整合的である。

またこの学習方略の教員付加価値の性質からは、教員はどのようにして子どもの学力を伸ばすかについて示唆を得ることができる。学習方略をあげることが出来た教員とは、子どもの「学習のやり方」をより効率的にすることが出来た教員である。しかし、そのように学習を効率的にしたからといって、子どもの学力そのものを上昇させているわけではない。ここから分かるのは、子どもの学習のあり方を教員が変えると言うよりむしろ、授業の方法を工夫することなど授業時間における教育のあり方の違いが学力にたいする教員付加価値の違いにもつながっている可能性である。この授業における教授方法の違いが子どもの学力に強い影響を与えているのではないかという論点は Bietenbeck (2014) や Hidalgo-Cabrillana and Lopez-Mayan (2018) といった研究が最近議論を行っている。教員は学習のありかたを教えるのではなく、子どもがより理解できてより授業内容を習得できるような授業を提供することで、子どもの学力を上昇させていると表4は解釈することができる。

まとめると、これらの結果が示唆するのは、教員の効果は多次元的であるということである。確かに、国語及び算数といった学力を成長させることができる教員は比較的共通である。同様に非認知能力・学習方略を伸ばすことができる教員もまた比較的似通っている。しかし学力を伸ばすことができる教員と非認知能力・学習方略を伸ばすことができる教員は必ずしも同一ではない。すなわち、一概に「良い」教員というものが存在するわけではなく、人的資本の種類に応じて良い教員が存在するのである。そのため、例えば、学力についての教員付加価値を用いて教員を評価することは、非認知能力という重要な人的資本に資する能力を評価できない可能性がある<sup>20</sup>。

一方で、学習方略への教員付加価値についての解釈には議論の余地がある。第1に学習方略と学力は、教員が異なることによる子どもの能力の成長への影響力という点で類似している。すなわち学力と学習方略は共に、推定された教員付加価値の標準偏差がとりうる値が広く、場合によっては非常にその影響力が非常に小さい可能性がある(表3)。第2に、学習方略を伸ばすことができた教員は非認知能力を伸ばすことができた教員と共通する傾向がある。すなわち、学習方略についての教員付加価値は学力の教員付加価値との相関は高く有意であった(表4)。更にその一方で、学習方略についての教員付加価値と学力の教員付加価値との間にある相関は小さかった。このように教員付加価値のどのような性質に着目するかによって、学習方略は学力とも非認知能力とも近い性質を持っている。教員付加価値と学習方略の関係については管見の限り本稿が初出であり、以上の関係性についての議論には更なる慎重な検討が必要である。

それでは具体的にどのような教員が、「良い」教員なのだろうか。多くの研究が観察可能な教員の特徴が必ずしも子どもの教育成果とは結びつかないことを報告しており(Hanushek and Rivkin, 2006)、研究間で一貫して効果的だとされる具体的な要因は見つかっていない。しかし、それらの研究はほとんど学力を対象にしたものであり、ここまで考えてきた非認知能力や学級での子ども・教員間の関係性などの学力以外の教育成果と教員の観察可能な特徴についての研究はほとんど存在しない。更に、そもそも日本の教員を対象にした場合、II節でも述べた通りほとんど研究が存在しないため海外での研究での結論を当てはめることが可能なかどうかは分からない。

そのため、教員質問紙で尋ねている教員の特徴と推定した教員付加価値の間の相関について分析した。また、高い教員付加価値を持った教員の観察可能な特徴を調べた。その結果は付録の表E.5に示した。しかし、年齢や教員歴・性別や学力について調べた結果、頑健な関係性を得ることはできなかった。教員の能力と観察可能な教員の特徴と結びつけることは未だに困難があると考えられる。

## V 結論

本稿では子どもの教育成果に対する個々の教員の効果について考えた。教員効果を教育成果に対する付加価値として計測した上で、得られた推定値の性質を記述的に調べた。

まず教員が子どもの教育成果に与えている影響の大きさは、教育成果の種類に応じて異なることが分かった。学力や学習習慣に対する教員の影響力は小さい可能性もあれば海外の研究と同水準に大きい可能性もあった。一方で、非認知能力や学級での子ども・教員間の関

係性に対する教員の影響は比較的大きかった。そのため、「教員は子どもに大きな影響を与える/与えない」といった言説について評価する時には、子どもの人的資本のどの部分を対象にしているかを考える必要がある。子どもの学力の向上を期待して能力の高い教員を雇用することは必ずしも効果的な施策ではない可能性があるが(一方で効果的である可能性もある)、非認知能力や学級への雰囲気への効果を期待して能力の高い教員を雇用することは高い効果を持ちうる。

次に教員付加価値間の相関を調べた結果、学力を伸ばすことが得意な教員と学習方略や非認知能力を伸ばすことが得意な教員は異なるのではないかという示唆を得た。そして、この結果を踏まえた時に、教員付加価値を用いて教員の評価をすることについては慎重に検討するべきであると考えられる。すなわち、教員を評価する時一概に「良い」か「悪い」の2つに分けることは難しく、学力において良いパフォーマンスを発揮した先生であっても、非認知能力では必ずしもそのようではない。

本稿の最大の限界はサンプルサイズの問題である。特に教員の数は少なく、解析の結果がどれほど頑健であるかという点に課題が残る。ただし、本稿では市町村に在籍する教員全ての情報を使っているため、もとより教員のパフォーマンス測定で文脈で教員固定効果を考える時にはより大きい行政区分単位で考えなければならない可能性が高い。

また本稿で得られた教員の付加価値は、教員が子どもの教育成果に対して持つ因果的な効果とも異なることに注意が必要である。Chetty et al. (2014a) らの議論のように、過去のデータから得られた教員付加価値が子どもの将来の教育成果に対して因果的な効果を持つことを示した研究は存在する。そして、本稿ではそれらの議論を踏まえて推定された教員付加価値をそのまま分析に用いている。しかし、日本においても教員付加価値が因果的な効果を持ちうるかどうかは必ずしも自明ではなく今後の検証を必要とする。

子どもの教育成果に対する教員の効果についての研究は日本のみならず全世界で未だ決まった結果の方向性が示されていない研究分野である。本稿で行った議論も多く存在する研究課題の一部を扱ったものであり、決して絶対的な知見であるとは言えないが、それでも本研究は日本における数少ない教員付加価値の研究として重要であると言える。

## 註

<sup>1</sup> Jackson et al. (2014) では “a label given to systematic variation in output across students assigned to the same teacher” と定義されている。

<sup>2</sup> 元々、子どもは学校という場所を通して多くの時間を教員と共に過ごすため、教員が子どもに対して強い影響力を持ちうることを多くの研究が指摘してきた (Hanushek, 1986, 2006; Ammermüller et al., 2005; Hojo and Oshio, 2012)。

<sup>3</sup> 教育経済学の分野で行われた初期の研究については、Hanushek and Rivkin (2006) にまとめられている。

<sup>4</sup> Jackson et al. (2014) は近年の教員の付加価値に関する研究動向をまとめている。

<sup>5</sup> ただし習熟度別学級の実施などにより、一部科目において複数の学級を担当するような状況は存在する。

<sup>6</sup> 平成28年度学校基本調査より算出。なお本統計で言う教員とは学校に所属する教員全てを指し、その中には養護教諭など学級を受け持たない教員も含まれる。そのため、ここで示した値は平均的な学級規模とは異なることに注意が必要である。

<sup>7</sup> 2013年度総務省『市町村税課税状況等の調』より、納税義務者一人当たり課税対象所得を算出した。

<sup>8</sup> 一方で、学習指導要領などを通じて公立小学校では標準化されたカリキュラムが提供されており、T市と日本の他の地域の教育環境には共通する要素も多い。

<sup>9</sup> ただし、全ての子どもを対象に標準化を行った後に、前年度に紐づく教員がいないような子どもは取り除き計算したため厳密に平均0分散1になっているわけではない。

<sup>10</sup> 無論、あえて調査に参加をしないという教員も存在する。

<sup>11</sup> 平成28年度学校教員統計調査より。

<sup>12</sup> 重み  $\alpha_{jc}$  は  $\mu_j$  に関する情報  $s_{jc}$  の正確度を表すパラメータである。具体的な計算方法は Kane and Staiger (2008) を参照せよ。

<sup>13</sup> Kane らの手法が教員固定効果の計算と大きく異なる点として、 $s_{it}$  の信頼度で教員付加価値を評価しているという点を挙げることができる。例えば5人しかいないクラスを担当する教員の固定効果と30人いるクラスを担当する教員の固定効果ではその信頼度に大きく差があり、Kane らの手法ではその点を考慮に入れて計算をしているとすることができる。

<sup>14</sup> 本稿では推定された個々の  $\mu_j$  そのものを直接分析の対象にはしなかったが、例えば教員付加価値を教員評価に用いることを考えた時にこの問題は強い制約となる。

- <sup>15</sup> ただし、ここでは国語についての教員付加価値および学習方略についての教員付加価値のみ示した。全ての被説明変数についてのプロットは図 C.1 に示した。
- <sup>16</sup> また得られた教員付加価値は推定の手法に対しても頑健であった。本稿では詳細な議論を展開しないが、本稿で得られた教員付加価値は、教員固定効果として求めた教員付加価値の方法で求めた教員付加価値との間とも高い相関関係があった。そのため、教員付加価値推定手法の選択は本稿における議論の結果には強い影響を及ぼすことはないと考えられる。
- <sup>17</sup> またこれらの値が頑健であることを示すために、Appendix D では推定の対象とする教員を変えた場合における推定値を示した。
- <sup>18</sup> この「教員付加価値の標準偏差」の大きさは、教員の影響力そのものを意味しないことには注意する必要がある。ここで議論しているのは個々の教員の教員付加価値の違いであり、すなわち教員の違いの影響力である。
- <sup>19</sup> セルフコントロールと自己効力感の教員付加価値の間には有意な相関はないが、これはサンプルサイズが極端に少ないからであると考えられる
- <sup>20</sup> ただし両者の関係性はトレードオフというわけではないと言う点には注意が必要である。

## References

- Ammermüller, Andreas, Hans Heijke, and Ludger Wößmann (2005) “Schooling quality in Eastern Europe: Educational production during transition”, *Economics of Education Review*, **24** (5), pp. 579–599.
- Baumeister, Roy F., Kathleen D. Vohs, and Dianne M. Tice (2007) “The strength model of self-control”, *Current Directions in Psychological Science*.
- Behncke, Stefanie (2012) “How Do Shocks to Non-Cognitive Skills Affect Test Scores?”, *Annals of Economics and Statistics*.
- Berkowitz, Ruth, Hadass Moore, Ron Avi Astor, and Rami Benbenishty (2017) “A research synthesis of the associations between socioeconomic background, inequality, school climate, and academic achievement”, *Review of Educational Research*, **87** (2), pp. 425–469.
- Bietenbeck, Jan (2014) “Teaching practices and cognitive skills”, *Labour Economics*, **30**, pp. 143–153.
- Chetty, Raj, John N Friedman, and Jonah E Rockoff (2014a) “Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates”, *American Economic Review*, **104** (9), pp. 2593–2632.
- , ———, and ——— (2014b) “Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood”, *American economic review*, **104** (9), pp. 2633–79.
- Clotfelter, Charles T, Helen F Ladd, and Jacob L Vigdor (2010) “Teacher credentials and student achievement in high school a cross-subject analysis with student fixed effects”, *Journal of Human Resources*, **45** (3), pp. 655–681.
- Duckworth, Angela L, Eli Tsukayama, and Teri A Kirby (2013) “Is it really self-control? Examining the predictive power of the delay of gratification task”, *Personality and Social Psychology Bulletin*, **39** (7), pp. 843–855.
- Goldhaber, Dan, James Cowan, and Joe Walch (2013) “Is a good elementary teacher always good? Assessing teacher performance estimates across subjects”, *Economics of Education Review*, **36**, pp. 216–228.
- Hanushek, Eric A (1986) “The economics of schooling: Production and efficiency in public schools”, *Journal of economic literature*, **24** (3), pp. 1141–1177.



- (2006) “School resources”, *Handbook of the Economics of Education*, **2**, pp. 865–908.
- (2011) “The economic value of higher teacher quality”, *Economics of Education review*, **30** (3), pp. 466–479.
- and Steven G Rivkin (2006) “Teacher quality”, *Handbook of the Economics of Education*, **2**, pp. 1051–1078.
- Harris, Douglas N and Tim R Sass (2011) “Teacher training, teacher quality and student achievement”, *Journal of public economics*, **95** (7-8), pp. 798–812.
- Heckman, James J. and Tim Kautz (2012) “Hard evidence on soft skills”, *Labour Economics*.
- Hidalgo-Cabrillana, Ana and Cristina Lopez-Mayan (2018) “Teaching styles and achievement: Student and teacher perspectives”, *Economics of Education Review*, **67**, pp. 184–206.
- Hojo, Masakazu (2012) “Determinants of academic performance in Japan: an economic perspective”, *Japanese Economy*, **39** (3), pp. 3–29.
- and Takashi Oshio (2012) “What Factors Determine Student Performance in East Asia? New Evidence from the 2007 Trends in International Mathematics and Science Study”, *Asian Economic Journal*, **26** (4), pp. 333–357.
- Hopson, Laura M and Eunju Lee (2011) “Mitigating the effect of family poverty on academic and behavioral outcomes: The role of school climate in middle and high school”, *Children and Youth Services Review*, **33** (11), pp. 2221–2229.
- Jackson, C Kirabo (2018) “What do test scores miss? The importance of teacher effects on non-test score outcomes”, *Journal of Political Economy*, **126** (5), pp. 2072–2107.
- Jackson, C. Kirabo and Elias Bruegmann (2009) “Teaching Students and Teaching Each Other: The Importance of Peer Learning for Teachers”, *American Economic Journal: Applied Economics*, **1** (4), pp. 85–108, October.
- Jackson, C Kirabo, Jonah E Rockoff, and Douglas O Staiger (2014) “Teacher effects and teacher-related policies”, *Annu. Rev. Econ.*, **6** (1), pp. 801–825.
- Jacob, Brian A and Lars Lefgren (2005) “Principals as agents: Subjective performance measurement in education”, Technical report, National Bureau of Economic Research.
- Kane, Thomas J and Douglas O Staiger (2008) “Estimating teacher impacts on student achievement: An experimental evaluation”, Technical report, National Bureau of Economic Research.
- Kinsler, Josh (2012) “Beyond levels and growth estimating teacher value-added and its persistence”, *Journal of Human Resources*, **47** (3), pp. 722–753.
- Kraft, Matthew A (2019) “Teacher effects on complex cognitive skills and social-emotional competencies”, *Journal of Human Resources*, **54** (1), pp. 1–36.
- Lefgren, Lars and David Sims (2012) “Using subject test scores efficiently to predict teacher value-added”, *Educational Evaluation and Policy Analysis*, **34** (1), pp. 109–121.

- Multon, Karen D., Steven D. Brown, and Robert W. Lent (1991) “Relation of Self-Efficacy Beliefs to Academic Outcomes: A Meta-Analytic Investigation”, *Journal of Counseling Psychology*.
- Pintrich, Paul R et al. (1991) “A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ).”.
- Pintrich, Paul R. and Elisabeth V. de Groot (1990) “Motivational and self-regulated learning components of classroom academic performance.”, *Journal of Educational Psychology*.
- Richardson, Michelle, Charles Abraham, and Rod Bond (2012) “Psychological correlates of university students’ academic performance: A systematic review and meta-analysis”, *Psychological Bulletin*.
- Rothstein, Jesse (2010) “Teacher quality in educational production: Tracking, decay, and student achievement”, *The Quarterly Journal of Economics*, **125** (1), pp. 175–214.
- Todd, Petra E and Kenneth I Wolpin (2003) “On the specification and estimation of the production function for cognitive achievement”, *The Economic Journal*, **113** (485), pp. F3–F33.
- Tsukayama, Eli, Angela Lee Duckworth, and Betty Kim (2013) “Domain-specific impulsivity in school-age children”, *Developmental Science*, **16** (6), pp. 879–893.
- Wang, Ming-Te and Jessica L Degol (2016) “School climate: A review of the construct, measurement, and impact on student outcomes”, *Educational Psychology Review*, **28** (2), pp. 315–352.
- 佐藤純 (1998) 「学習方略の使用と達成目標及び原因帰属との関係」,『筑波大学心理学研究』, 第 20 巻, 115–124 頁.
- 成田 健一・下仲 順子・中里 克治・河合 千恵子・佐藤 眞一・長田由紀子 (1995) 「特性的自己効力感尺度の検討:生涯発達の利用の可能性を探る」,『教育心理学研究』.
- 二木美苗 (2017) 「子どもの学習に対する教員の質の効果: 都道府県パネルデータによる実証分析」,『日本経済研究』, 第 74 号, 56–83 頁.
- 妹尾渉・松繁寿和・梅崎修 (2003) 「公務員および教員の男女間賃金格差—大卒者アンケート調査から」,『大阪大学経済学』, 第 53 巻, 第 2 号, 96-108 頁, sep 月.
- 李嬋娟 (2014) 「非認知能力が労働市場の成果に与える影響について」,『日本労働研究雑誌』.

## Appendix A 非認知能力および学習方略の定義と組成の詳細

「非認知能力」という言葉はある能力を具体的に指し示すものではなく、認知能力の尺度では捉えられないような性格的特徴などの個人的な属性を広く指す概念として用いられている (Heckman and Kautz, 2012)。特にその「能力」という用語法は、それらの属性のうち訓練などによって学習可能である部分に着目していることを表している。この非認知能力という言葉が指し示す広範な概念のうち、本研究では「セルフコントロール」および「自己効力感」の2つの能力を対象に分析を行った。またその他に本研究では「学習方略」という心理尺度を分析の対象にしている。本節ではそれらの概念の定義や概要、加えてどのように測定したかについて記述を行う。

### 1 セルフコントロール

セルフコントロールとは、「理想、価値観、道徳、社会的期待などの基準に合わせて自身自身の反応を変化させ、長期的な目標の追求をサポートする能力」と定義される (Baumeister et al., 2007)。このセルフコントロールの能力が高いことは、自らの行動の意識的な制御を通じて、学業や勉強時間 (Tsukayama et al., 2013)、また賃金などの労働条件などを予測することがこれまで示されてきている (李嬋娟, 2014)。

本研究では Tsukayama et al. (2013) で示されたセルフコントロール尺度を日本語に翻訳したものを用いる。この尺度は学齢期の子どもを対象に開発されたセルフコントロール尺度であり、比較的簡易的にセルフコントロールを測定することができるという特徴をもっている。Tsukayama et al. (2013) では、この尺度で測定されたセルフコントロールの値によって、GPA や勉強時間・TV ゲームに費やす時間などを予測できることが示されている。尺度を構成する具体的な質問項目については表 A.1 に示した。各々の質問は5件法によって回答され、尺度を計算する際には点数が高いほどセルフコントロールが高いことを指し示すように変換を行った (例えば「授業で必要なものを忘れた」という項目に対して1(ほとんどない)と回答された場合は、指し示す値を5とした)。最終的にこれらの質問項目が指す値を合計して、「セルフコントロール」の値として用いた。

### 2 自己効力感

自己効力感とは「個人がある状況において必要な行動を効果的に遂行できる可能性の認知」と定義される (成田 健一他, 1995)。すなわち、自己効力感の低い人は特定のタスクを達成することができないと考え、そのタスクを避けようとするのに対し、自己効力感の高い人はタスクが困難であっても達成に向けて動こうとすると考えられる。自己効力感について心理学的研究は、高い自己効力感の高い学業成績 (Multon et al., 1991; Richardson et al., 2012; Behncke, 2012) と関係あることを示している。

本研究では Pintrich and de Groot (1990) で示された尺度のうち Self-Efficacy for Learning and Performance 下位尺度を日本語に翻訳したものを用いる。この尺度は子どもの学習動機づけを構成する尺度として開発され、子どもの学業成績と正の相関を持っていることが示されている。尺度を構成する具体的な質問項目については表 A.1 に示した。各々の質問は5件法によって回答された。最終的にこれらの質問項目が指す値を合計して、「自己効力感」の値として用いた。

### 3 学習方略

学習方略とは学習のやり方を表すメタ的な認知能力として主に教育心理学の分野で研究がなされてきている (佐藤, 1998; Pintrich et al., 1991)。すなわち、学習方略が高い子どもは、学習のやり方を工夫する中でその学習の成果を高めることができると考えられる。Pintrich and de Groot (1990) といった心理学的研究が、高い学習方略が高い学業成績につながることをこれまで示している。

本研究では佐藤 (1998) で示された学習方略尺度を用いる。当該尺度は Pintrich and de Groot (1990) や Pintrich et al. (1991) をもとに、その下位尺度を整理しながら日本の

小学生・中学生を対象に開発した心理尺度である。尺度を構成する質問項目については表 A.2 に示した。各々の質問は 5 件法によって回答され、尺度を計算する際には点数が高いほど学習方略が高いことを指し示すように変換を行っている (例えば「授業の内容が難しいときは、やらずにあきらめるか、簡単なところだけ勉強する」という項目に対して 1(ほとんどない) と回答された場合は、指し示す値を 5 とした)。最終的にこれらの質問項目が指す値を合計して、「学習方略」の値として用いた。

項目	逆転項目	質問
セルフコントロール	✓	授業で必要なものを忘れた
	✓	他の子たちが話をしているときに、その子たちのじゃまをした
	✓	何か乱暴なことを言った
	✓	机・ロッカー・部屋が散らかっていたので、必要なものを見つけることができなかった
	✓	家や学校で頭にきて人やものにあたった
	✓	先生が、自分に対して言っていたことを思い出すことができなかった
	✓	きちんと話を聞かないといけないときにぼんやりしていた
	✓	イライラしているときに、先生や家の人（兄弟姉妹を除きます）に口答えをした
自己効力感		授業ではよい評価をもらえるだろうと信じている
		教科書の中で一番難しい問題も理解できると思う
		授業で教えてもらった基本的なことは理解できたと思う
		先生が出した一番難しい問題も理解できると思う
		学校の宿題や試験でよい成績をとることができると思う
		学校でよい成績をとることができるだろうと思う
		授業で教えてもらったことは使いこなせると思う
		授業の難しさ、先生のこと、自分の実力のことなどを考えれば、自分はこの授業でよくやっているほうだと思う

表 A.1: 質問リスト:非認知能力

注) 本表では変数「セルフコントロール」「自己効力感」及び「勤勉性」を構成する質問をまとめた。列「逆転項目」にチェックマークがついている場合、その質問は測定の向きが他の質問と逆になっている。また全ての質問に対して、回答は5件法で得られた。

項目	逆転項目	質問
学習方略すべて		勉強のやり方が、自分に合っているかどうかを考えながら勉強する
		勉強でわからないところがあったら、勉強のやり方をいろいろ変えてみる
		勉強しているときに、やった内容を覚えているかどうかを確かめる
		勉強する前に、これから何を勉強しなければならないかについて考える
		勉強するときは、最初に計画を立ててから始める
		勉強をしているときに、やっていることが正しくできているかどうかを確かめる
		勉強するときは、自分で決めた計画に沿って行う
		勉強しているとき、たまに止まって、一度やったところを見直す
		勉強するときは、参考書や事典などがすぐ使えるように準備しておく
		勉強する前に、勉強に必要な本などを用意してから勉強するようにしている
		勉強していて大切だと思ったところは、言われなくてもノートにまとめる
		勉強で大切なところは、繰り返して書くなどして覚える
		勉強でわからないところがあったら、友達にその答えをきく
		勉強でわからないところがあったら、友達に勉強のやり方をきく
		勉強のできる友達と、同じやり方で勉強する
		勉強するときは、最後に友達と答え合わせをするようにする
		勉強するときは、内容を頭に思い浮かべながら考える
		勉強をするときは、内容を自分の知っている言葉で理解するようにする
		勉強していてわからないところがあったら、先生にきく
		新しいことを勉強するとき、今までに勉強したことと関係があるかどうかを考えながら勉強する
	✓	学校の勉強をしているとき、とてもめんどくさくつまらないと思うことがよくあるので、やろうとしていたことを終える前にやめてしまう
		今やっていることが気に入らなかったとしても、学校の勉強でよい成績をとるために一生懸命がんばる
	✓	授業の内容が難しいときは、やらずにあきらめるか、簡単などろだけ勉強する
		問題が退屈でつまらないときでも、それが終わるまでなんとかやり続けられるように努力する
子ども・教員間の関係性		学級での生活は楽しかったですか
		学級は落ち着いて学習する雰囲気がありましたか
		学校の友達は自分のよいところを認めてくれましたか
		学校の先生たちは自分の悩みの相談にのってくれましたか

表 A.2: 質問リスト:学習方略、学級での子ども・教員間の関係性

注) 本表では変数「学習方略」及び「学級での子ども・教員間の関係性」を構成する質問をまとめた。列「逆転項目」にチェックマークがついている場合、その質問は測定の向きが他の質問と逆になっている。また「学習方略」を表す質問に対して回答は5件法で得られ、「学級での子ども・教員間の関係性」を表す質問に対しては回答は4件法で得られた。

## Appendix B 教員と子どもの観察可能な特徴間の相関

節 IV.1 で記したように、教員付加価値を推定するためには教員が内生的に配置されてはならない。その仮定の妥当性を考えるために、表 B.3 では教員と子どもの観察可能な特徴の間に相関が無いかを調べた。具体的には子どもの特徴を (学校固定効果・年度・学年の影響は取り除いた上で) 担当した教員の特徴に回帰し、その係数の推定値を示した。

示した推定値は全て統計的に有意ではなかった。もちろん、観察されない子どもや教員の特徴を基に教員が内生的に配置されている可能性もありその検証は今後の研究において重要であるものの、子どもの特徴に合わせて教員を内生的に配置している傾向を見つけることはできなかった。

		前年度学力 (国語)	前年度学力 (算数)	本の冊数	通塾	女性	就学援助
年齢	推定値	0.000 (0.001)	0.002 (0.001)	0.000 (0.001)	-0.001 (0.001)	0.000 (0.001)	0.000 (0.001)
	サンプルサイズ	8051	8048	8011	7904	8051	4904
経験年数	推定値	-0.003 (0.003)	0.001 (0.003)	-0.002 (0.003)	-0.002 (0.001)	0.000 (0.001)	0.001 (0.001)
	サンプルサイズ	6491	6488	6460	6381	6491	3811
女性教員	推定値	-0.003 (0.023)	-0.006 (0.023)	0.023 (0.026)	0.002 (0.011)	-0.001 (0.012)	0.000 (0.010)
	サンプルサイズ	8088	8085	8048	7940	8088	4904
出身学校の偏差値	推定値	0.000 (0.003)	-0.002 (0.003)	0.003 (0.003)	-0.001 (0.001)	0.000 (0.001)	0.000 (0.001)
	サンプルサイズ	3614	3613	3590	3553	3614	2812

表 B.3: 教員と子どもの観察可能な特徴間の相関

注) 本表では子供の観察可能な属性 (前年度学力 (国語・算数)、本の冊数、通塾、女性、就学援助) を教員の観察可能な特徴 (年齢、経験年数、女性教員、出身学校の偏差値) に回帰した結果を表している。推定では学校固定効果、学年、年度の影響を取り除いて推定した。各々のセルでは推定値及び標準誤差・有意水準を示しており、「推定値 有意水準 (標準誤差)」という表記法を用いている。「\*\*\*」・「\*\*」・「\*」はそれぞれ 1%・5%・10% の水準で統計的に有意であることを表す。

## Appendix C 教員付加価値の分布 (詳細)

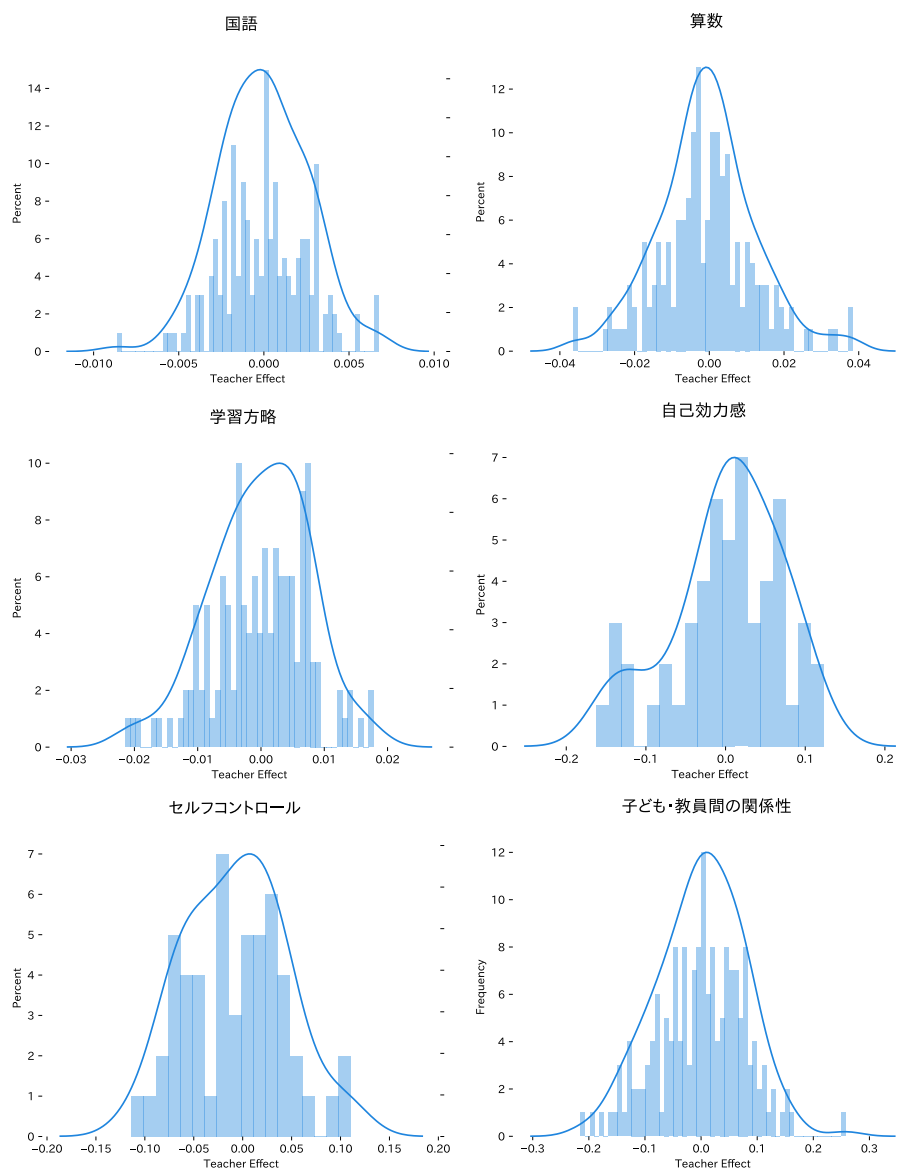


図 C.1: 推定された教員付加価値の分布 (詳細)

注) 本図では、推定された教員付加価値の分布を示している。棒グラフで示しているのは推定された教員付加価値のヒストグラムである。線グラフで示しているのは、推定された教員付加価値をカーネル密度推定することによって得られた密度関数である。



## Appendix D 教員付加価値推定の頑健性

節 IV で得られた教員付加価値は頑健な推定値なのだろうか。ここでは教員付加価値を分析する教員の集団を変えながら推定値の頑健性を検討する。ここでは (1) 複数クラスを受け持った教員のみにデータを絞った場合、および (2) 教員歴が 3 年以上ある教員のみにデータを絞った場合のそれぞれについて教員付加価値の推定値の標準偏差がどのような値をとるかを示す。前者は教員付加価値の推定においてクラス固有の効果の影響を一切軽減することができていない教員 (節 IV.1) を除いた場合を検討している。後者は、教員の質が教員歴と共に変わっていく可能性を考慮している。教員の経験年数が子どもの能力に与える影響を分析する研究では、Harris and Sass (2011) の様に最初の数年の経験が教員の質にとって重要であるという知見が存在する。もしそうであれば、教員として勤務を始めて数年の教員の質は低くなりその結果、推定された教員付加価値の標準偏差が高くなってしまう。そのような可能性を検討するために、教員歴が短い教員をデータから除いた上で推定した結果を報告する。

得られた推定値の標準偏差を表 D.4 に示した。(1) 列は本文と示した値と全く同一である。(2) 列では 1 クラスしか担当していない教員を除いて推定された教員付加価値の標準偏差を示した。(3) 列では教員歴が 2 年以下の教員を除いて推定された教員付加価値の標準偏差を示した。各々の行は教員付加価値を求める対象とした教育成果を表す。

結果として、推定された教員付加価値の標準偏差は推定の対象とする教員によらずほとんど同じような値をとった。例えば国語については考えると、本文で報告した標準偏差は 0.003 である ((1) 列) のに対して、複数学級を担当した教員のみに考えた場合の標準偏差は 0.005 であり ((2) 列)、経験年数 3 年以上の教員のみに考えた場合の標準偏差は 0.003 である ((3) 列)。これらの値の違いは本文における議論に修正を迫るような差ではない。この傾向は他のアウトカムにおいても同様である。

	(1) 教員付加価値 (標準偏差) KS	(2) 教員付加価値 (標準偏差) KS 複数クラス	(3) 教員付加価値 (標準偏差) KS 経験年数 3 年以上
国語	0.003	0.005	0.003
算数	0.013	0.017	0.012
学習方略	0.008	0.007	0.007
セルフコントロール	0.051	0.057	0.055
自己効力感	0.069	0.068	0.067
子ども・教員間の関係性	0.078	0.097	0.086

表 D.4: 教員付加価値の標準偏差 2

注) 本表では求めた教員付加価値の標準偏差を示している。各々の行は教員付加価値を求める対象とした教育成果を表す。(1) 列では本稿で主に用いている教員付加価値の標準偏差を示した。(2) 列では 1 クラスしか担当していない教員を除いて推定された教員付加価値の標準偏差を示した。(3) 列では教員歴が 2 年以下の教員を除いて推定された教員付加価値の標準偏差を示した。

## Appendix E 教員の属性と教員付加価値

下記のような定式化で分析を行う。

$$\hat{\mu}_j = \mathbf{X}_{j,2018}\beta + \epsilon_j$$

$\hat{\mu}_j$  は推定された教員付加価値である。 $\mathbf{X}_{j,2018}$  は観察可能な教員の特徴のベクトルを表し、本稿では教員の年齢、教員歴、性別、出身大学の難易度を考える。ただし出身大学の難易度

を考えるとサンプルサイズが非常に小さくなるため、出身大学の難易度をサンプルに加えた場合と加えていない場合でそれぞれ検討した。また、年齢などの年によって変動する値については、全て 2018 年度時点での値を用いた。加えて、非線形性を考慮するために年齢や教員歴については、その二乗項を推定式に加えて推定を行った。推定された  $\hat{\beta}$  が本節で関心を寄せる値である。

推定結果を表 E.5 に示した。以下ではその結果について詳述していく。

#### 年齢、教員歴

どのような教員が良い教員か考えた時に頻繁に挙げられる点として、年齢や教員歴といった仕事の経験の多寡を挙げることができる。教員という仕事は高度な専門職でありその業務経験が長ければ長いほど教員としてのスキルは向上する可能性がある。Hanushek and Rivkin (2006) では教員の年齢が子どもの教育成果に与える影響を分析した研究をサーベイし、研究によってその推定値は様々であり結論を未だに出せるような状況ではないとしている。

被説明変数 (教員付加価値)	(1)	(2)	(3)	(4)	(5)	(6)	(7) セルフ コントロール	(8) セルフ コントロール	(9) 自己効力感	(10) 自己効力感	(11) 子ども・教員間の 関係性	(12) 子ども・教員間の 関係性
年齢	0.000 (0.001)	0.001 (0.002)	0.000 (0.004)	0.003 (0.007)	-0.001 (0.003)	0.015** (0.006)	0.056 (0.043)	-0.256 (0.293)	0.031 (0.047)	-0.017 (0.106)	0.031 (0.047)	-0.017 (0.106)
年齢 (2 乗)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000** (0.000)	-0.001 (0.001)	0.003 (0.004)	0.000 (0.001)	0.000 (0.001)	0.000 (0.001)	0.000 (0.001)
教員歴	0.000 (0.000)	-0.001 (0.001)	0.001 (0.002)	0.002 (0.004)	0.000 (0.001)	-0.005* (0.003)	-0.004 (0.013)	0.180 (0.124)	0.016 (0.016)	0.024 (0.049)	0.016 (0.016)	0.024 (0.049)
教員歴 (2 乗)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000* (0.000)	0.001 (0.001)	-0.006 (0.005)	-0.001 (0.001)	-0.001 (0.002)	-0.001 (0.001)	-0.001 (0.002)
女性教員	0.000 (0.000)	0.000 (0.001)	0.004 (0.002)	0.004 (0.004)	-0.001 (0.001)	-0.007** (0.003)	-0.023 (0.018)	-0.097 (0.096)	0.034 (0.022)	0.050 (0.053)	0.034 (0.022)	0.050 (0.053)
出身大学偏差値		0.000 (0.000)		0.000 (0.000)		-0.001** (0.000)		0.005 (0.003)		0.002 (0.004)		0.002 (0.004)
サンプルサイズ	133	48	133	48	113	30	44	10	47	18	47	18

表 E.5: 教員の特徴と教員付加価値の相関

注) 本表では年齢、教員歴、性別及び出身大学偏差値が教員付加価値に与える影響を示している。各々のセルでは推定値及び標準誤差・有意水準を示しており、「推定値 有意水準 (標準誤差)」という表記法を用いている。「\*\*\*」・「\*\*」・「\*」はそれぞれ 1%・5%・10%の水準で統計的に有意であることを表す。

結果として、教員の年齢と教員付加価値の間に有意な関係は発見されなかった。表 E.5 において年齢及びその二乗項に係る係数推定量は、どの教員付加価値を被説明変数としても統計的に有意ではない。ただし学習方略を被説明変数とし出身大学偏差値を説明変数に加えた推定 ((6) 列) では、年齢に係る係数の推定値は 0.015 で 5% の水準で有意、その 2 乗項に係る係数の推定値は 0.000 で 5% の水準で有意であった。しかしこれらの結果は、出身大学偏差値を説明変数に加えていない推定 ((5) 列) の結果とは統計的な有意水準やその効果の正負が異なり頑健な結果とは言い難い。

同様に、教員歴と教員付加価値の間に有意な関係は発見されなかった。すなわち、表 E.5 において教員歴及びその二乗項に係る係数の推定値は、教員付加価値を被説明変数としても統計的に有意ではないことがほとんどであった。ただし学習方略を被説明変数とし出身大学偏差値を説明変数に加えた推定 ((6) 列) では、教員歴に係る係数の推定値は -0.005 で 5% の水準で有意、その 2 乗項に係る係数の推定値は 0.000 で 10% の水準で有意であった。しかしこれらの結果は、出身大学偏差値を説明変数に加えていない推定 ((5) 列) の結果とは統計的な有意水準やその効果の正負が異なり頑健な結果とは言い難い。

## 性別

教員の性別も良い教員の特徴を考える時には頻繁に論点に挙げられる。例えば、労働市場における男女差は教員の能力にも影響しているかもしれない。日本の労働市場の中では教員という職業は性別による賃金差が比較的小さい (妹尾渉他, 2003)。そのため、相対的に労働市場で能力の高い女性が教員になっている可能性がある。

表 E.5 の「女性教員」行には教員の性別と教員付加価値の間にある関係性を示している。結果として、教員が女性であるか否かと教員付加価値との間の関係のその推定値の分散は大きい。ただし学習方略を被説明変数とし出身大学偏差値を説明変数に加えた推定 ((6) 列) では、「女性教員」に係る係数の推定値は -0.007 で 5% の水準で有意であった。出身大学偏差値を説明変数に加えるか否かで結果が変わっており、その頑健性には注意が必要である。

## 出身大学の偏差値

高い学力・認知能力を持つ教員は高いパフォーマンスを発揮するのではないかという仮説は頻繁に検証されてきた。例えば二木 (2017) など教員の採用時の労働市場の状況が子どもの教育成果に与える影響を調べる研究は、暗に能力の高い労働者ならば教員としても能力が高い可能性があるという仮説を検証していると言えよう。Hanushek and Rivkin (2006) ではそのサーベイを行い、最終的には研究によってその推定値は様々であり結論を未だに出せるような状況ではないとしている。

本稿で用いるデータでは、2016 年度に T 市に所属した教員に対する質問紙調査の中で卒業大学及び卒業学科について尋ねている。それらの卒業大学の情報を 2018 年度に河合塾が算出した大学受験偏差値に紐付けることで、教員の学力を卒業大学の偏差値という形で取得した。無論入試難易度偏差値は時間によって変わるため、このひも付けは必ずしも適切ではないことに十分注意する必要がある。

表 E.5 の「出身大学偏差値」行には教員の卒業大学偏差値と教員付加価値の間にある関係性を示している。結果として、卒業大学偏差値が教員付加価値に与える影響は概ね正である一方でその推定値の分散は大きく統計的に有意ではなかった。ただし学習方略を被説明変数とした推定 ((6) 列) では、その推定値は -0.001 と負で統計的に 10% で有意である。教員の認知能力が高いと担当した子どもの学習方略が下がってしまうことを示す結果であるが、そのサンプルサイズは 30 と非常に少ないこともあり、更なる研究と慎重な議論が必要である。