

## 第 II 部：調査／分析編

## 第5章. 調査の概要

川口俊明

本章では、SES 代替指標の妥当性・信頼性を確認するために、本研究で実施した調査の概要について述べる。

### 1. 調査対象の選定について

本研究では、(1) 日本全体を対象にするには予算額が小さい、(2) 調査期間が1年弱と非常に短い、さらに(3) SESに関わるきわめてセンシティブな情報を扱う、といった理由から、日本全体を対象にした標本抽出調査は行っていない。代わりに、ごく限られた自治体を対象に、SESに関わるさまざまな情報をできるだけ収集することを優先して調査を実施することにした。

対象となる自治体を選定するにあたっては、西日本の複数の自治体に調査の受け入れを打診し、同意を得られた自治体を対象に調査を実施している。また、最終的なサンプルサイズとして、各学年5000名程度を目安としていたため、基本的に、規模の大きい学校を有する都市部の自治体を中心に選定した。

結果として、本研究の調査対象は日本全体の平均よりややSESの高い地域となっている。当然のことながら、本調査の結果は、日本の特定の地域の実態を表しているに過ぎない。本調査で得られた結果が日本全体に妥当するかどうかは、今後実施されるであろう、全国学力・学習状況調査（保護者に対する調査）の結果を通して、あらためて確認していく必要がある。

なお、調査対象となった自治体の特定を防ぐため、本報告書では、対象自治体の詳細を記載していない。また分析の際に、自治体によって基準が異なる指標（たとえば就学援助）を用いる場合は、その都度、分析対象を適切な範囲に限定するなどの対応を行っているが、その具体的な操作についても十分な説明を行っていない。この点については、ご了承ください。

### 2. 対象学年と対象学級

本研究の調査対象学年は、小学校4年生、小学校6年生、および中学校3年生である。全国学力・学習状況調査の実施学年である小学校6年生、中学校3年生に加えて、小学校4年生を加えた理由は、次の二つである。

一つは、調査の実施時期に関わる問題である。本研究では児童生徒・保護者を対象とした質問紙調査を実施するが、その実施時期は9月か10月になることが想定された。この場合、中学校・高校受験を控える小学校6年生・中学校3年生の回答傾向が、受験の影響を受け、

全国学力・学習状況調査の実施時期（4月）と大きく変わる可能性がある。そこで、受験の影響を受けにくい学年を一つ選定したのである。

もう一つの理由は、調査対象の代表性に関わる問題である。すでに述べたように、今回の調査結果を、日本全体に一般化することはできない。とはいえ、調査対象がどのような地域特性を持つかを記述しておくことは必要であろう。幸い、国際学力調査である TIMSS が 2015 年の 3 月に小学校 4 年生を対象に実施されている<sup>(1)</sup>。そこで、小学校 4 年生の情報があれば、本調査の結果を日本全体と比較する際に、有益な情報が得られると考えたのである。

当然、小学校 4 年生については、全国学力・学習状況調査は実施されていないが、今回の調査対象に、小学校 4 年生を対象とした学力調査／生活実態調査を実施している自治体が含まれていたため、小学校 4 年生のデータを検討する際には、利用可能な範囲でこのデータを利用している。

質問紙調査は、対象となる自治体のすべての学校から、学校の規模に応じて 1 ないし 2 学級を無作為に抽出し、そこに所属するすべての児童生徒とその保護者、および学校長を対象に実施した。ただし、対象学年の在籍者数が 10 名以下の小規模校については、調査対象から除かれている。また、特に課題を抱えた学校（たとえば、前年度の全国学力・学習状況調査で就学援助率が 50%を超えている学校）については、SES 指標の妥当性を検討するという本研究の趣旨と照らして、より重要性が高いという判断から、当該学校の全学級を抽出している。

以上のような抽出の結果、各児童生徒の抽出確率は、自治体・学校ごとに異なっている。そのため、分析結果が対象となる自治体全体にあてはまるように、各児童生徒の抽出確率を計算し、その逆数を **weight** とすることで、分析結果を補正している。

### 3. 使用できるデータ

今回の分析で使用できるデータは、大きく分類すると、次の 3 種類である。

- (1) 全国学力・学習状況調査（学力調査、児童生徒質問紙、学校質問紙）
- (2) 就学援助受給の有無、虫歯の本数、身長・体重、体力テストの情報
- (3) 本研究で配布した児童生徒質問紙、保護者質問紙、学校質問紙

以下、それぞれのデータの概要について説明する。まず、(1) 全国学力・学習状況調査については、文部科学省より平成 28 年度分のデータ提供を受け、分析に使用した。また、先ほど述べたように、小学校 4 年生については、自治体が独自に実施していた学力調査／生活実態調査のデータを利用している。小学 4 年生の生活実態調査については、基本的な設問が全国学力・学習状況調査の児童質問紙に準じて設計されていたため、多くの項目は同じように分析に利用できた。全国学力・学習状況調査で出題されている学力調査の内容や、児

児童生徒質問紙の設問に関する情報は、国立教育政策研究所のウェブサイト<sup>(2)</sup>を参照されたい。

次に、(2) 就学援助受給の有無、虫歯の本数、身長・体重、体力テストのデータについては、各学校が所有する児童生徒の情報を、全国学力・学習状況調査の個人番号に紐付けてもらった後に、教育委員会を通して回収した。なお、体力テストの実施要項は、文部科学省のウェブサイトに掲載されているので、そちらを参照してほしい<sup>(3)</sup>。

最後に、(3) 本研究で配布した児童生徒質問紙、保護者質問紙、学校質問紙について述べる。これらの質問紙は、教育委員会と内容について調整した後、9月下旬に各学校へ郵送した。その上で、学校質問紙については、各学校長に回答を依頼した。児童生徒質問紙と保護者質問紙については、両者を同一の封筒に入れ、あらかじめ共通のIDを印字した。その上で、児童生徒質問紙のみ学校で記入し、保護者質問紙は封筒に入れたまま各家庭に持ち帰って保護者に記入してもらった。その後、記入した質問紙を厳封し、各学校に提出してもらった後で、10月中旬を目処に、学校質問紙・児童生徒質問紙とひとまとめにした上で、各学校からデータ入力を担当する委託業者に郵送してもらうという形をとった。なお、教育委員会との調整の中で、児童生徒質問紙には氏名を記入しないことになったため、代わりに児童生徒の出席番号を記載させることで、全国学力・学習状況調査と接続することになっている。

各データの紐付けは、次のような手順で実施している。まず、児童生徒質問紙・保護者質問紙の紐付けは、あらかじめ印字しておいた共通IDによって行った。次に、全国学力・学習状況調査と児童生徒質問紙を、児童生徒質問紙に記入されている出席番号と、学校から提出された「出席番号と全国学力・学習状況調査のIDを対応づけた表」をもとに接続した。全国学力・学習状況調査のIDと(2) 就学援助受給の有無、虫歯の本数等のデータは、各学校で紐付けられているので、この時点で(1)から(3)のデータがすべて接続できることになる。最後に、出席番号については、児童生徒を特定する手がかりになる可能性を持つ情報であるため、紐付けに利用した後は、データセットから削除している。

なお、データの紐付けに際して、接続ができないデータが100件ほど生じている。これは、児童生徒が出席番号を間違えたり、学校側が児童生徒の出席番号を間違えたりしたためである。特に、特別支援教育の対象となっている児童生徒や、不登校傾向のある児童生徒を挟んだ前後の番号で、こうしたミスが発生している。ただ、ほとんどの場合は、学級の人数が一致しなくなったり、体力テストに付随した性別と児童生徒質問紙で質問している性別が逆転したりといった不自然なデータが生じるため、その都度、教育委員会に問い合わせることで対応した。最終的に各学年3~40件程度の接続できないデータが生じたが、発生率は調査対象全体から見れば1%以下であり、許容範囲内であると判断している。

(1)から(3)の調査の実施時期は、全国学力・学習状況調査が4月末日、体力テストなどが自治体・学校にもよるが5~6月頃、本研究の児童生徒質問紙の実施が9月下旬~10月上旬である。実施時期にズレがあるため、各調査で欠席・転入・転出などの影響を受け、個々のデータに欠損が生じている。また、調査時期が最大で半年以上ずれていることから、たとえば通塾の有無に関して、全国学力・学習状況調査の質問紙の回答と、児童生徒質問紙・保

護者調査の回答の間に齟齬が生じている箇所も見受けられた。

#### 4. 質問紙調査の回収率

本調査で実施した質問紙調査の回収率は、表1の通りである。まず、全体というのは、本研究の質問紙調査実施時点(9月下旬)での調査対象学級全体の人数である。学力テストは、そのうち学力テストの得点が確認できた児童生徒の数を示している。体力テストは、同じく体力テストのデータが確認できた児童生徒であり、児童生徒質問紙・保護者質問紙は、それぞれ質問紙が返却され、かつ紐付けがなされて分析に使用できる数を示している。最後に、年収に関する設問(問27のキ)は、一般的に回答拒否が生じる確率の高い、保護者質問紙の年収に関わる設問の回答者数を示している。回収率は、それぞれの度数を、全体の人数で割った値である。

表1. 回収率

	小4		小6		中3	
	度数	回収率	度数	回収率	度数	回収率
全体	5577	100.0	5642	100.0	4637	100.0
学力テスト	5450	97.7	5481	97.1	4430	95.5
体力テスト	5424	97.3	5459	96.8	4339	93.6
児童生徒質問紙	5434	97.4	5336	94.6	4317	93.1
保護者質問紙	4495	80.6	4309	76.4	3486	75.2
年収に関する設問(問27のキ)	4305	77.2	4150	73.6	3337	72.0

表1を見ると、学力テスト/体力テスト/児童生徒質問紙のそれぞれで、データが回収できていない児童生徒が、小学校4年生で2.3~2.6%、小学校6年生で2.9~5.4%、中学校3年生で4.5~6.9%存在している。その理由としては、下記4点が考えられる。

まず、1点目として、今回の調査対象全体の人数は、9月時点の数値であり、特に2学期に転入してきた児童生徒が含まれている。転入してきた児童生徒については、4月から6月にかけて実施された学力テスト・体力テストのデータを得ることはできない。

2点目として、学力テスト・体力テスト・児童生徒質問紙のそれぞれで、調査を欠席している児童生徒が存在する。かれらのデータも当然得ることはできない。

3点目は、特別支援教育の対象となっている児童生徒の存在である。今回、特別支援教育の対象となっている児童生徒は、各学校の判断で調査に参加した。そのためかれらのデータは、「そもそも全体から削除されている」「まったく調査に参加していないが全体には含まれ

ている」「一部だけ参加している」「すべて参加している」など、学校ごとにいくつかのパターンが存在している。こうしたデータが含まれたため、回収率が低下している可能性がある。

4点目に、児童生徒質問紙に関しては、主に児童生徒が出席番号を間違えたなどの理由により、データが接続できなかったケースがある。

とはいえ、1, 3, 4点目に関しては、各学年で大きく割合は変わらないはずである。にも関わらず、小学校4年生から中学校3年生にかけて、回収率が低下している。これより、中学校3年生の回収率と小学校4年生の回収率の差である2~3%は、おそらく生徒の欠席によって生じていると考えられる。

全国学力・学習状況調査をはじめとする、各種の学力調査では、欠席した児童生徒は、はじめから分析対象に含まれていないことも多く、分析の対象となることは希である。確かにその割合は高いものではないが、無視してよい値でもない。今回のようにサンプルサイズが5000前後の場合、3%といえど、そこには100人を超える児童生徒が該当しているはずである。学力調査において、欠席した児童生徒に焦点が当たることは少ないが、回収率の分析は、こうした分析の必要性も示唆している。

なお、児童生徒質問紙の度数で保護者質問紙の度数を割ると、今回の保護者質問紙の配布数に対する回収率が計算できる。その数値は、小学4年生が82.7% ( $4495 \div 5434$ )、小学6年生が80.7% ( $4309 \div 5336$ )、中学3年生が80.9% ( $3486 \div 4309$ )である。2013年に実施された全国学力・学習状況調査(きめ細かい調査)の保護者調査では、回収率が約85%であったことからすると、やや低い値になったといえる。

## 5. 対象地域の特性

ここでは、対象地域の特性について、データから得られる範囲で説明する。なお、比較の対象として、2013年度の全国学力・学習状況調査(きめ細かい調査)の保護者調査、およびTIMSS2015のデータを使用する。いずれも、日本全体を対象にした標本抽出調査であり、本調査で得られたデータの特性を理解する上で役に立つであろう。

表2は、本調査と、2013年度の全国学力・学習状況調査の保護者調査で得られた年収の分布について示している。小学6年生、中学3年生のいずれもで、わずかに本調査の方が保護者の年収が高い傾向が見られる。

これは、TIMSS2015と比較した場合も同様の傾向が見られる。TIMSS2015は、2015年の3月に実施された調査であり、日本全体から分析対象となる児童生徒を抽出している。TIMSSは、小学4年生と中学2年生を対象とした学力調査であるため、ここでは本調査の対象である小学4年生と中学3年生のデータと比較してみよう。

表3は、家庭にある本の冊数に対する回答を、本調査とTIMSS2015で比較している。101-200さつ、200さつより多いと回答する子ども割合が、いずれもTIMSSより高い値を示している。

表 2. 世帯年収の比較

	本調査	全国調査	本調査	全国調査
	小 6	小 6	中 3	中 3
200 万円未満	5.4	6.7	6.1	7.5
200～300 万円未満	6.6	8.2	8.0	8.6
300～400 万円未満	10.5	12.6	10.0	11.8
400～500 万円未満	12.3	14.9	11.9	13.3
500～600 万円未満	13.3	14.0	12.4	13.7
600～700 万円未満	12.0	11.9	11.4	12.1
700～800 万円未満	10.6	10.4	10.8	10.2
800～900 万円未満	8.5	6.3	8.9	7.0
900～1000 万円未満	6.8	5.0	7.1	5.5
1000～1200 万円未満	7.4	5.3	7.8	6.0
1200～1500 万円未満	3.1	2.6	3.4	2.8
1500 万円以上	3.4	2.1	2.0	1.4

表 3. 家庭にある本の冊数 (TIMSS2015 との比較)

	本調査 (小 4)	TIMSS (Grade4)	本調査 (中 3)	TIMSS (Grade8)
0-10 冊	10.95	12.20	11.90	11.58
11-25 冊	20.62	28.93	17.97	21.09
26-100 冊	39.04	37.38	36.31	32.13
101-200 冊	16.76	13.19	16.67	17.23
200 冊より多い	12.63	8.30	17.15	17.97

表 2・表 3 の結果から、本調査の対象となった地域は、全国平均よりやや SES の高い地域であるということが出来る。以降の分析結果は、こうした前提を踏まえた上で解釈する必要があると考えられる。

## 6. SES 指標の作成

本報告書の第 6 章以降の分析では、さまざまな方法で作成した SES 指標の妥当性を検討していく。そのために、まずは保護者調査から基準となる SES 指標を比較する必要がある。本研究では、全国学力・学習状況調査 (きめ細かい調査) の分析において SES 指標を作成した手順 (お茶の水女子大学 2014) に従って、SES 指標を作成した。具体的には、以下の通りである<sup>(4)</sup>。

まず、「家庭の年収」「父親の学歴」「母親の学歴」のそれぞれの変数を、以下の手順で数量化する。「家庭の年収」については、各回答の中央値（たとえば、200～300万であれば250万）をとる。「父親の学歴」「母親の学歴」については、教育年数に変換する（たとえば、大卒であれば16年間）。次に、個々の変数をそれぞれ標準化した後、平均値を算出する。なお、その際、欠損値が増えることを防ぐため、二つの変数しか得られていない場合はその平均を、一つの変数しか得られていない場合は、その値をそのまま使うようにしている。このようにして得られた値を、さらにもう一度標準化することで、平均0、標準偏差1のSES指標として分析に利用している。

作成されたSES指標と成績の相関は、表4の通りであり、どの学年・どの教科でも0.3～0.4程度の関連性を示している。なお、ここで小学校6年生・中学校3年生の成績の指標として使用しているのは、国語A・国語Bそれぞれの正答率を合計して2で割った数値である。算数（数学）についても同様である。

表4. SESと成績の相関係数

	小4	小6	中3
国語	0.33	0.31	0.31
算数（数学）	0.34	0.34	0.39

<註>

- (1) TIMSS 2015については、TIMSS & PIRLS 2015 (<https://timss.bc.edu/>) を参照のこと。
- (2) 平成28年度の全国学力・学習状況調査に関しては、国立教育政策研究所のHP (<http://www.nier.go.jp/16chousa/16chousa.htm>) を参照のこと。
- (3) 文科省のウェブサイト【新体力テスト実施要項】を参照。  
([http://www.mext.go.jp/a\\_menu/sports/stamina/03040901.htm](http://www.mext.go.jp/a_menu/sports/stamina/03040901.htm))
- (4) この手順については、垂見（2014）を参照している。

<参考文献>

垂見裕子，2014，「家庭の社会経済的背景（SES）の尺度構成」お茶の水女子大学『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. 13-15.

## 第6章．公開データを利用した学校単位のSES代替指標の作成

土屋 隆裕

### 1. 学校単位のSES指標およびその代替指標

この章では、分析の単位が学校である場合のSES指標およびその代替指標について検討する。指標の値は各校に一つずつ与えられることになる。まず考えられるのは、当該校に在籍する(当該学年の)児童生徒の保護者のSES指標の値を平均し、当該校のSES指標とすることである。また、学校が所在する地域の特性をその代替指標として用いることも考えられる。具体的には学区の住民の学歴や職業などを、一般に公開されているデータ、たとえば国勢調査の小地域集計結果を用いて再集計し、当該校のSESの代替指標として用いるのである。

学区の住民の中には児童生徒の保護者も含まれるが、保護者以外の者も多数含まれる。そのため児童生徒の学力調査結果の分析に用いるという目的に照らせば、SESを表す指標としては学校の地域特性よりも保護者のSES平均を用いる方が妥当と考えられる。しかし、保護者のSES指標は一般に入手が容易ではなく、全国どこでも利用可能な指標とはいえない。一方で学校の地域特性は公開されたデータから誰でも作成可能であり、全国どこでも利用可能な指標である。

そこでこの章では、まず調査によって得られた保護者のSES平均を、公開データから作成した学校の地域特性を用いてどの程度予測可能なのかを検証する。さらに、学力調査結果の学校平均を予測する上で、保護者のSES平均と学校の地域特性のいずれが有用なのかを明らかにするとともに、両者を同時に用いることで予測精度がどの程度向上するのかを確認する。

### 2. 無回答バイアスを補正するためのウェイト算出とSES指標の学校平均

今回の保護者調査の対象には、全国学力・学習状況調査のデータが欠損となっている児童生徒の保護者も含まれている。しかし本章の後半では、SES指標と学力調査結果との関係を見ていく。そのため分析の対象は、学力調査の4科目(国語A, 国語B, 算数/数学A, 算数/数学B)の全てが欠損でない児童生徒の保護者とした。具体的な人数等は表1のとおりである。

表1: 用いるデータ

	小学校	中学校
保護者の対象者数	5,468人 (5,504人)	4,424人 (4,436人)
保護者の有効回答者数	4,216人 (4,237人)	3,411人 (3,414人)
保護者の(重みなし)有効回答率	77.1% (77.0%)	77.1% (77.0%)

カッコ内は学力調査データが一部欠損を含む

また、学校ごとに保護者のSES平均を求めるには、基本的に、在籍する(当該学年の)全ての児童生徒の保護者のSES指標の値が必要である。しかし、今回の保護者調査の有効回答率は小学校・中学校ともに8割弱であり、有効回答があった保護者の単なるSES平均、あるいは標本抽出デザインだけを反映した抽出ウェイトによる加重平均を求めたのでは、無回答バイアスが生じるおそれがある。そこで無回答バイアスを低減するために抽出ウェイトの調整を行うこととした。

## 2.1. CART による重みづけクラス分け

ウェイト調整のために、まず保護者の回答・無回答を基準変数とした CART を用いて、保護者を重みづけクラスに分類した。説明変数としては、学校ごとの有効回答率と、児童生徒の全国学力・学習状況調査結果 (国語A, 国語B, 算数A/数学A, 算数B/数学Bそれぞれの学力層と質問紙調査の85項目への回答), さらに就学援助の有無, 母語が日本語か否か, 日本語指導の必要性の有無を用いた。学校ごとの有効回答率を用いたのは, 回答率が高い学校の保護者に過大な調整用ウェイトを与えないためであり, 学校を回答率が90%以上, 80%以上90%未満, 80%未満の3群に分類した。なお, CART を行うに当たっては抽出ウェイト  $W_i$  を用いた。

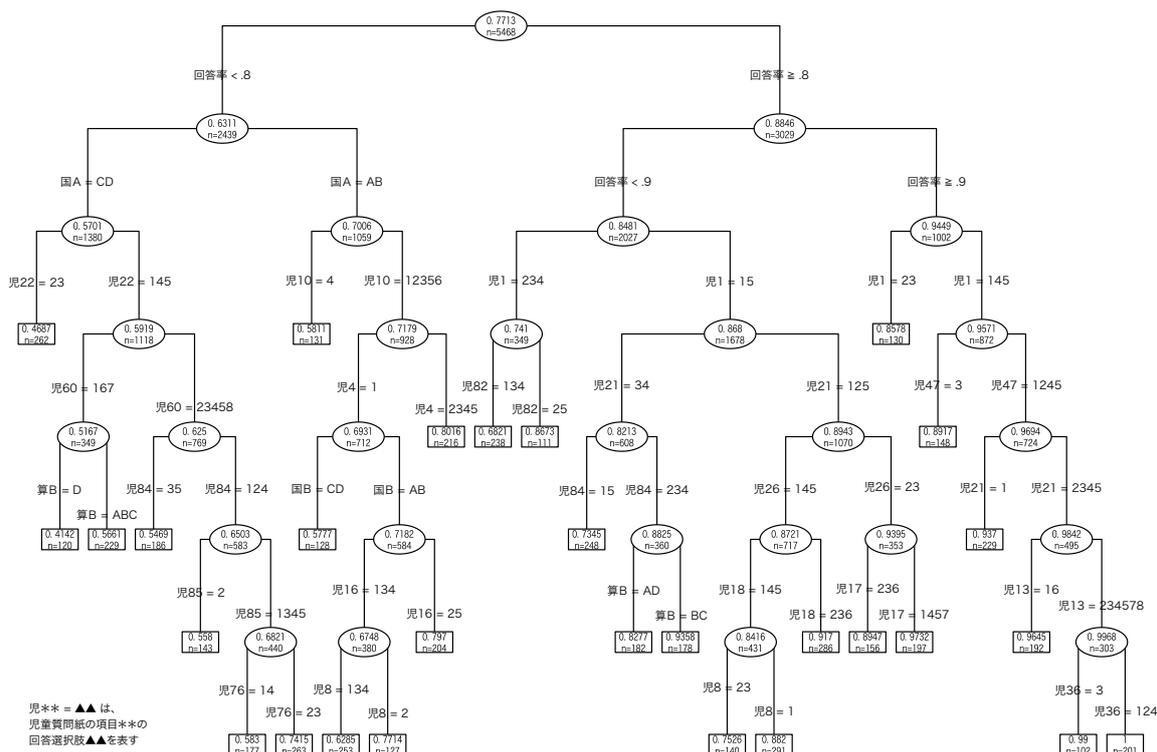


図 1: CART による重みづけクラス分け (小学校)

図 1 と図 2 がその結果である。図 1 の小学校では、保護者調査への回答・無回答と最も関係しているのは、もちろん学校ごとの有効回答率であり、次に、回答率が80%未満の学校においては国語Aの学力層である。また図 2 の中学校では、回答率が80%未満の学校においては「(21)家で、自分で計画を立てて勉強をしているか」という質問に対する生徒の回答である。CARTの結果に基づき、保護者調査の対象者を全部で小学校では29、中学校では23の重みづけクラスに分類し、各クラスでは抽出ウェイト  $W_i$  を用いた加重回答率の逆数を調整ウェイト  $F_i$  とした。

$$\text{回答率調整後のウェイト} = W_i F_i \quad (1)$$

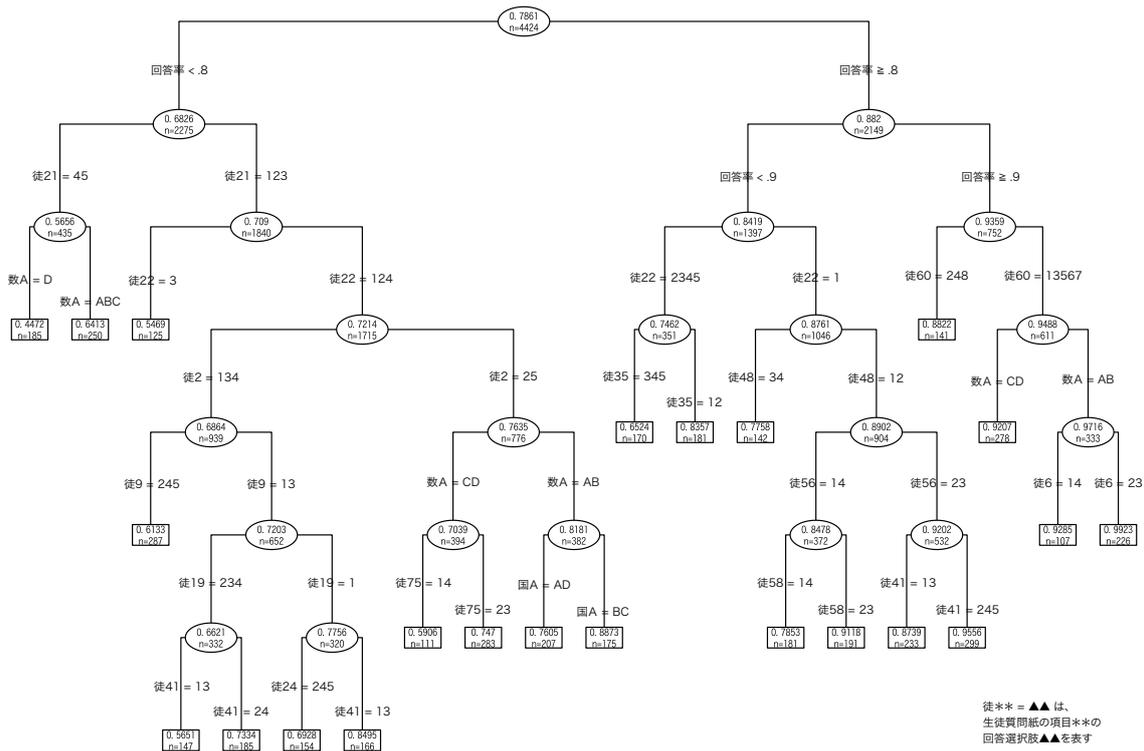


図 2: CART による重みづけクラス分け (中学校)

## 2.2. ウェイトのキャリブレーション

さらに、学校内の有効回答率 3 群と、全国学力・学習状況調査における学力層 (国語 A, 国語 B, 算数 A / 数学 A, 算数 B / 数学 B) の 5 変数の分布をベンチマークとして、回答率調整後のウェイト  $W_i F_i$  のキャリブレーションを行った。キャリブレーションのための距離関数としては線形関数を用いた。他の距離関数を用いても結果はほとんど変わらなかったためである。有効回答した保護者を  $S$  とし、保護者  $i$  の補助変数ベクトルを  $\mathbf{x}_i$ 、ベンチマーク分布を  $\boldsymbol{\tau}_x$  とすると、キャリブレーションのためのウェイトは

$$G_i = 1 + \left( \boldsymbol{\tau}_x - \sum_{i \in S} W_i F_i \mathbf{x}_i \right)' \left( \sum_{i \in S} W_i F_i \mathbf{x}_i \mathbf{x}_i' \right)^{-1} \mathbf{x}_i \quad (2)$$

であり、最終的なウェイトは

$$\text{最終的なウェイト} = W_i F_i G_i \quad (3)$$

となる。その分布は図 3 と図 4 のそれぞれ最上段に示すとおりである。最終的なウェイトは最大値が 10 近くまで広がっているが、その不等加重効果は小学校では 1.22, 中学校では 1.21 であったため、問題ないと判断した。

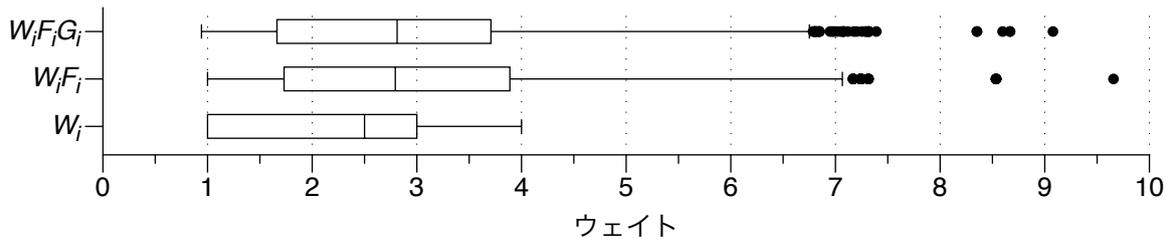


図 3: ウェイトの分布 (小学校)

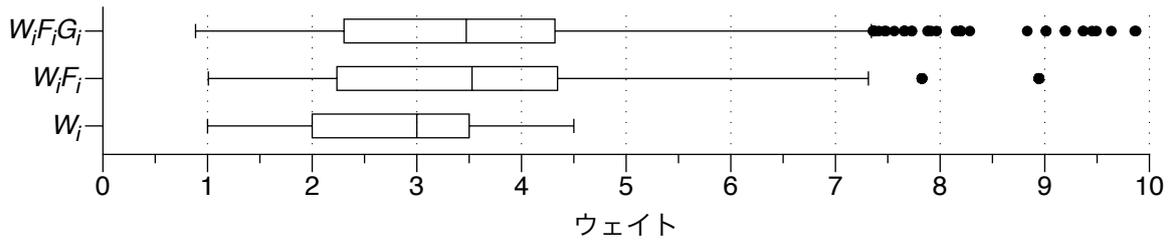


図 4: ウェイトの分布 (中学校)

### 2.3. SES の学校平均

学校  $j$  の SES 平均  $\overline{\text{SES}}_j$  は、学校  $j$  の有効回答保護者を  $S_j$  とし、保護者  $i$  の SES 指標値を  $y_i$  とすると次式により求めた。

$$\overline{\text{SES}}_j = \frac{\sum_{i \in S_j} W_i F_i G_i y_i}{\sum_{i \in S_j} W_i F_i G_i} \quad (4)$$

なお、保護者  $i$  の SES 指標の値  $y_i$  は、第 5 章の方法で求められた値を用いた。

図 5 と図 6 は、ウェイトを調整する前の  $W_i$  を用いた学校ごとの SES 加重平均と、ウェイト調整後の SES 平均とを比較したものである。点の大きさは各学校の有効回答保護者数を表す。多くの学校ではウェイト調整後に SES 平均が小さくなっている。これは、SES 指標が低い保護者は無回答となっており、有効回答した保護者の SES 指標を単に平均したのでは上方バイアスが生じることを意味し、頷ける結果である。また回答率が高い学校では、ウェイト調整前と調整後の SES 平均はあまり異なっておらず、その点で不適切なウェイト調整は行われていないことが分かる。以降の分析では、調整後のウェイトを用いた SES 平均を用いることとする。

### 3. 国勢調査を利用した学校の地域特性

各校の地域特性を表す指標としては、平成 22 年国勢調査 (総務省統計局) の小地域集計を利用することとした。最新の平成 27 年調査を用いなかった理由は、平成 27 年調査は簡易調査であり、教育の状況に関する調査項目が含まれていないためである。小地域集計の結果は、町丁字単位で

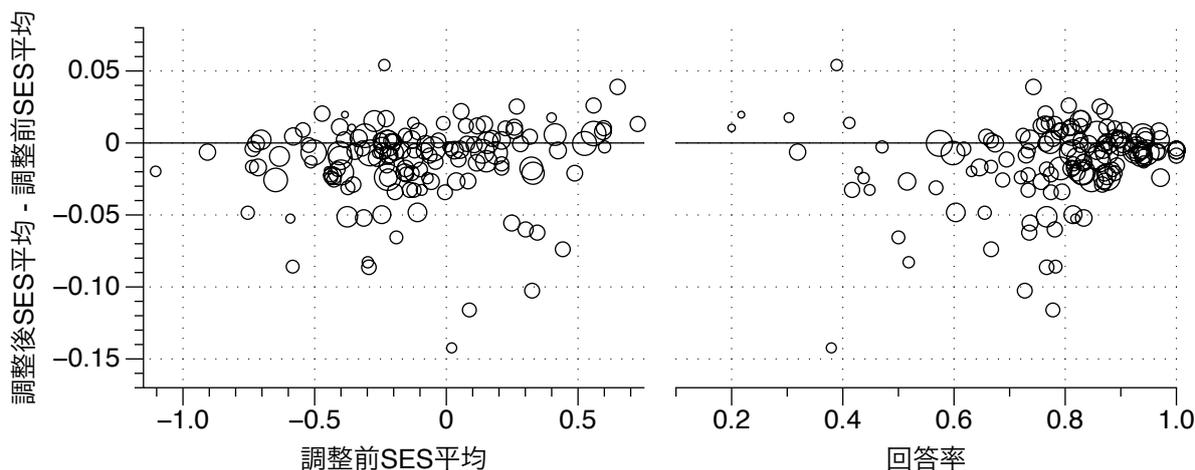


図 5: ウェイト調整前の SES 平均と調整後の SES 平均 (小学校)

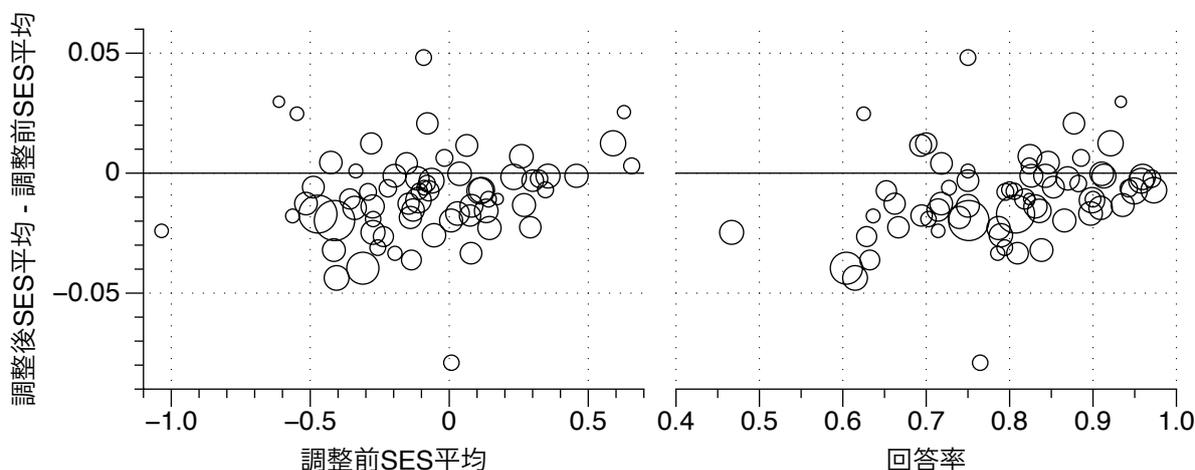


図 6: ウェイト調整前の SES 平均と調整後の SES 平均 (中学校)

公表されているが、学区は必ずしも町丁字単位ではない。そこで、学区をその一部にでも含む町丁字の結果は全て合算し、当該校の指標とすることとした。

算出した指標は表 2 に示す 74 である。指標名の T で始まる番号は小地域集計表の番号を表す。

#### 4. 学校の地域特性を用いた学校の SES 平均の予測

学校の SES 平均を地域特性でどの程度予測可能なのかを調べるため、学校の SES 平均を基準変数とし、表 2 に示す 74 の地域特性の指標全てを説明変数とした回帰分析を行った。ただし、説明変数が多数のため、LASSO による変数選択を行った。また、学校の SES 平均の推定精度を考慮するため、各校の有効回答数を重みとして用いた。その回帰係数は表 3 のとおりである。なお、選択されなかった説明変数は表 3 には示していない。学歴や産業、職業に関わる変数が選択されており、中でも「T14 大学卒業者割合」や「T12 L 学術研究等割合」、「T17 B 専門的等職業割合」

表 2: 学校の地域特性

T3 15歳未満割合	T11 雇用者割合 (男)	T13 1年未満居住者割合
T3 65歳以上割合	T11 自営業割合 (男)	T13 5年未満居住者割合
T4 未婚者割合	T11 雇用者割合 (女)	T13 10年未満居住者割合
T4 未婚者割合 (男)	T11 自営業割合 (女)	T14 大学卒業者割合
T4 未婚者割合 (女)	T12 A農業割合	T14 短大・大学卒業者割合
T5 1人世帯割合	T12 D建設業割合	T15 小中学生割合
T6 核家族世帯割合	T12 E製造業割合	T17 A管理的職業割合
T6 核家族人割合	T12 F電気等割合	T17 B専門的等職業割合
T7 持ち家世帯割合	T12 G情報通信業割合	T17 C事務割合
T7 借家世帯割合	T12 H運輸業等割合	T17 D販売割合
T7 持ち家人割合	T12 I卸売・小売業割合	T17 Eサービス職業割合
T7 借家人割合	T12 J金融・保険業割合	T17 F保安職業割合
T8 一戸建世帯割合	T12 K不動産業等割合	T17 G農林漁業割合
T8 共同住宅世帯割合	T12 L学術研究等割合	T17 H生産工程割合
T8 一戸建人割合	T12 M宿泊業等割合	T17 I輸送等割合
T8 共同住宅人割合	T12 N生活関連サービス業等割合	T17 J建設等割合
T9 延べ面積 29 以下世帯割合	T12 O教育等割合	T17 K運搬等割合
T9 延べ面積 49 以下世帯割合	T12 P医療・福祉割合	T17 L分類不能の職業割合
T9 延べ面積 29 以下人割合	T12 Q複合サービス事業割合	T17 管理等割合
T9 延べ面積 49 以下人割合	T12 Rサービス業等割合	T17 管理他等割合
T10 労働力人口割合	T12 S公務等割合	T20 現住所割合
T10 労働力人口割合 (男)	T12 T分類不能の産業割合	T20 自市区町村内割合
T10 労働力人口割合 (女)	T12 特定サービス割合	T20 県内割合
T11 雇用者割合	T12 特定サービス等割合	T20 県外割合
T11 自営業割合	T13 出生時から居住者割合	

などを説明変数として用いればよいことが分かる。

表 3: 学校の SES 平均を基準変数とした LASSO 回帰係数

	小学校			中学校		
	$s = .10$	$s = .13$	$s = .15$	$s = .06$	$s = .08$	$s = .10$
切片	-0.7451	-0.6614	-0.5982	-0.7981	-0.7290	-0.6614
T12 F 電気等割合	—	—	—	0.0192	—	—
T12 L 学術研究等割合	0.0080	0.0004	—	0.0110	0.0000	—
T14 大学卒業者割合	0.0271	0.0260	0.0231	0.0312	0.0307	0.0274
T17 B 専門的等職業割合	0.0029	0.0002	—	0.0005	—	—
T17 I 輸送等割合	-0.0033	—	—	—	—	—
T17 K 運搬等割合	-0.0002	—	—	—	—	—
$R^2$	0.789	0.785	0.767	0.828	0.794	0.787

図 7 と図 8 は、二つの地域特性「T14 大学卒業者割合」と「T12 L 学術研究等割合」それぞれと、学校の SES 平均とを用いた学校の散布図である。小学校、中学校ともに「T14 大学卒業者割合」では決定係数は 0.8、つまり相関係数は 0.9 に近い。学校の SES 平均は地域特性を用いることでかなりの精度で予測可能であることが分かる。

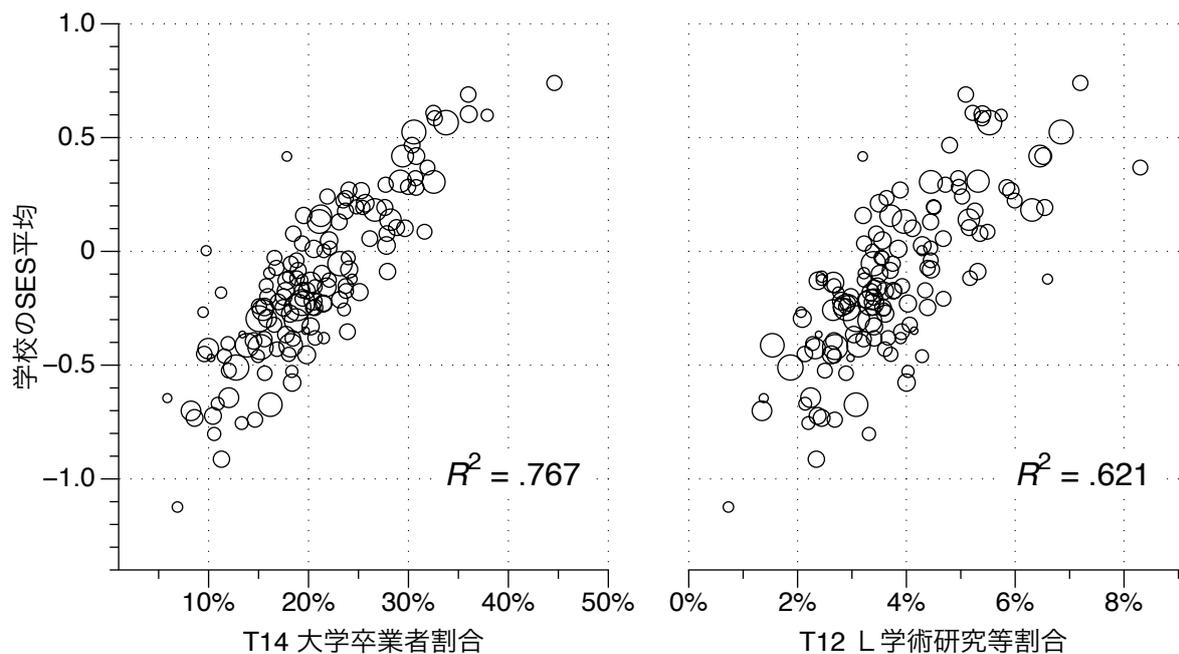


図 7: 学校の地域特性と SES 平均 (小学校)

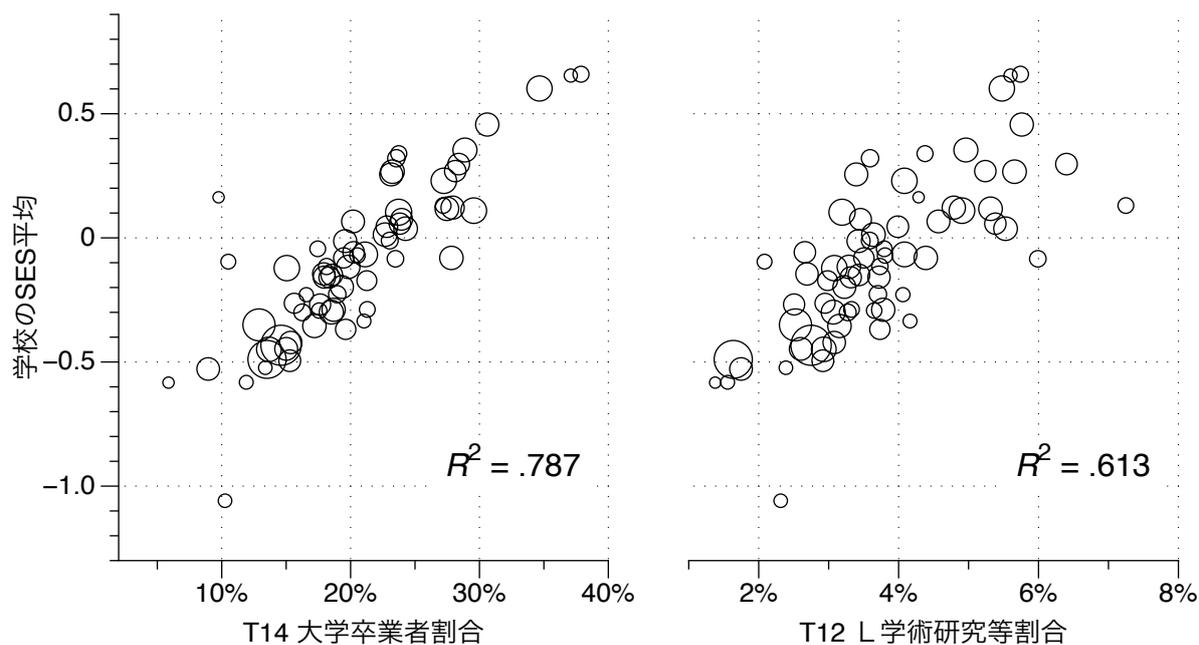


図 8: 学校の地域特性と SES 平均 (中学校)

### 5. 学校単位の学力調査結果と SES 平均, 地域特性

次に、学力調査結果の学校平均を予測する上で、学校の SES 平均や地域特性がどの程度有用なのかを検討する。地域特性としては「T14 大学卒業割合」のみを取り上げることとする。他の

地域特性と組み合わせれば、予測精度は向上すると期待されるが、どのように組み合わせればよいのかは必ずしも明らかではないからである。まず、表4と表5は、国語A、国語B、算数A／数学A、算数B／数学Bの正答数の学校平均を基準変数とし、学校のSES平均あるいは地域特性のいずれか、さらには両者を説明変数としたときの回帰分析結果である。ただし、学力調査の解答者数を重みとして用いた。

表 4: 学校の正答数平均を基準変数とした回帰分析結果 (小学校)

	国語A			国語B			算数A			算数B		
切片	9.321	10.990	8.759	4.627	5.921	4.299	10.659	12.522	10.139	4.770	6.364	4.328
SES 平均	0.550	1.768		0.321	1.266		0.510	1.869		0.434	1.597	
T14 大卒	0.073		0.097	0.056		0.070	0.081		0.103	0.069		0.088
$R^2$	0.595	0.509	0.582	0.653	0.545	0.643	0.594	0.500	0.585	0.669	0.562	0.658

表 5: 学校の正答数平均を基準変数とした回帰分析結果 (中学校)

	国語A			国語B			数学A			数学B		
切片	24.323	25.207	21.797	5.770	6.184	4.791	19.469	22.726	15.637	5.769	6.788	4.078
SES 平均	2.434	3.132		0.943	1.271		3.693	6.265		1.630	2.435	
T14 大卒	0.039		0.151	0.018		0.062	0.144		0.313	0.045		0.120
$R^2$	0.736	0.728	0.654	0.720	0.709	0.647	0.765	0.737	0.717	0.756	0.738	0.695

小学校では、学校のSES平均のみを用いると学校の正答数平均の分散の5割以上を説明でき、地域特性のみを用いると6割前後を説明できることが分かる。つまり、学校の正答数平均を予測する上では、学校のSES平均よりも地域特性の方が有効ということになる。図9は、学校のSES平均と正答数平均、あるいは地域特性と正答数平均を用いた学校の散布図である。点の大きさは学力調査の解答者数を表す。この図9を見ても、学校のSES平均よりも地域特性の方が、学校の正答数平均との相関が高いことが分かる。

表4には、説明変数として学校のSES平均と地域特性の両方を同時に用いたときの回帰分析結果も示した。その決定係数 $R^2$ は、説明変数として地域特性だけを用いたときの $R^2$ とほとんど変わらず、学校の正答数平均を予測する上で、地域特性に加えてSES平均を追加する効果はほとんどないといえる。

中学校では、学校のSES平均のみを用いると学校の正答数平均の分散の7割以上を説明でき、地域特性のみを用いると6割から7割程度を説明できる。つまり、小学校とは逆に、学校の正答数平均を予測する上では、地域特性よりも学校のSES平均の方が有効である。図10には中学校の散布図を示した。

次に、児童質問紙のうち「(61) 国語の勉強は好きだ」や「(71) 算数の勉強が好きだ」に「当てはまる」あるいは「どちらかといえば、当てはまる」と回答した児童の割合と、学校のSES平均、あるいは地域特性との関係を示したのが図11である。学校のSES平均と地域特性のいずれとも、相関関係はほとんど見られない。中学校について示した図12でも結果は同様であり、国語が好きな生徒や数学が好きな生徒の割合と学校のSES平均あるいは地域特性との相関は高くない。

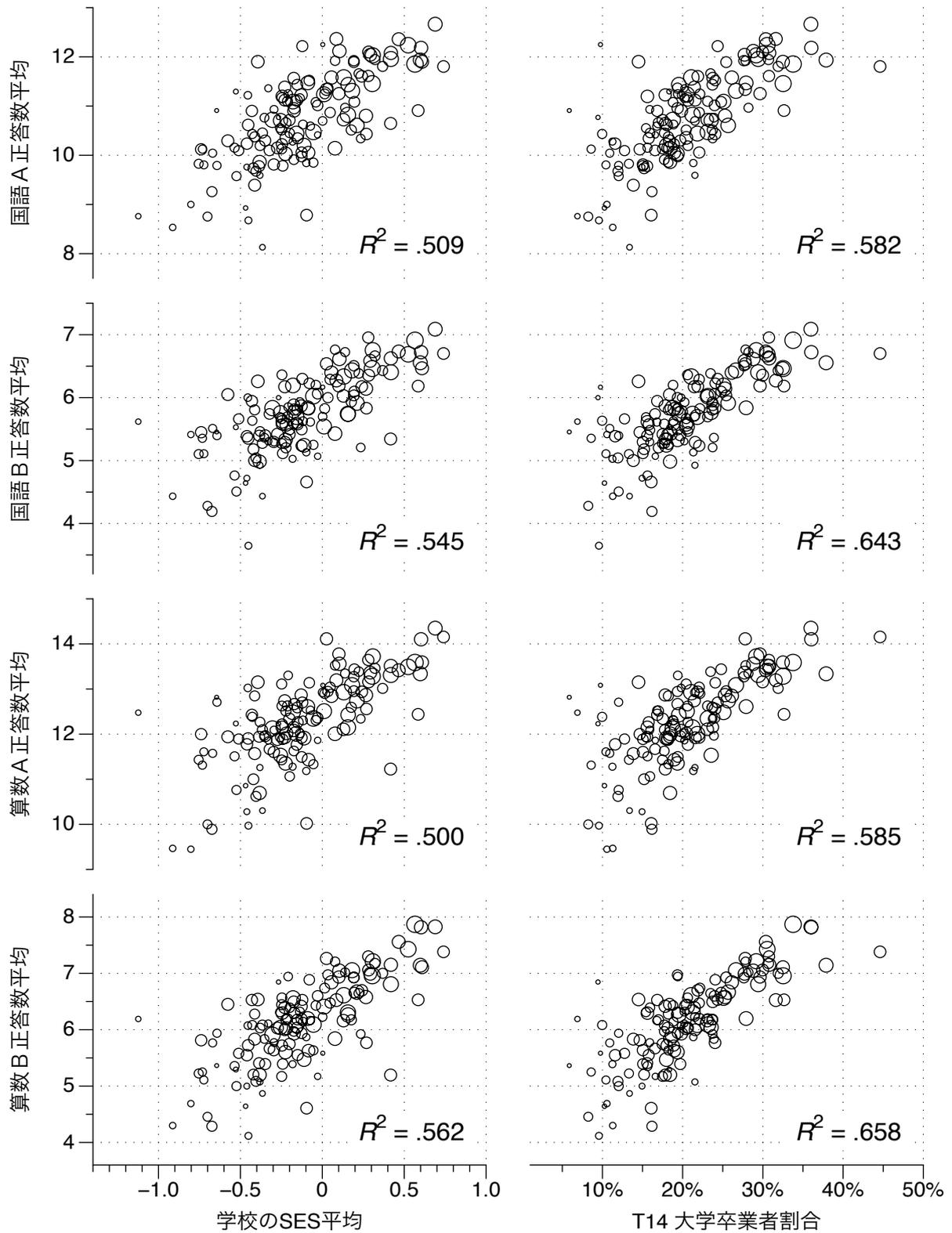


図 9: 学校の正答数平均と SES 平均・学校の地域特性 (小学校)

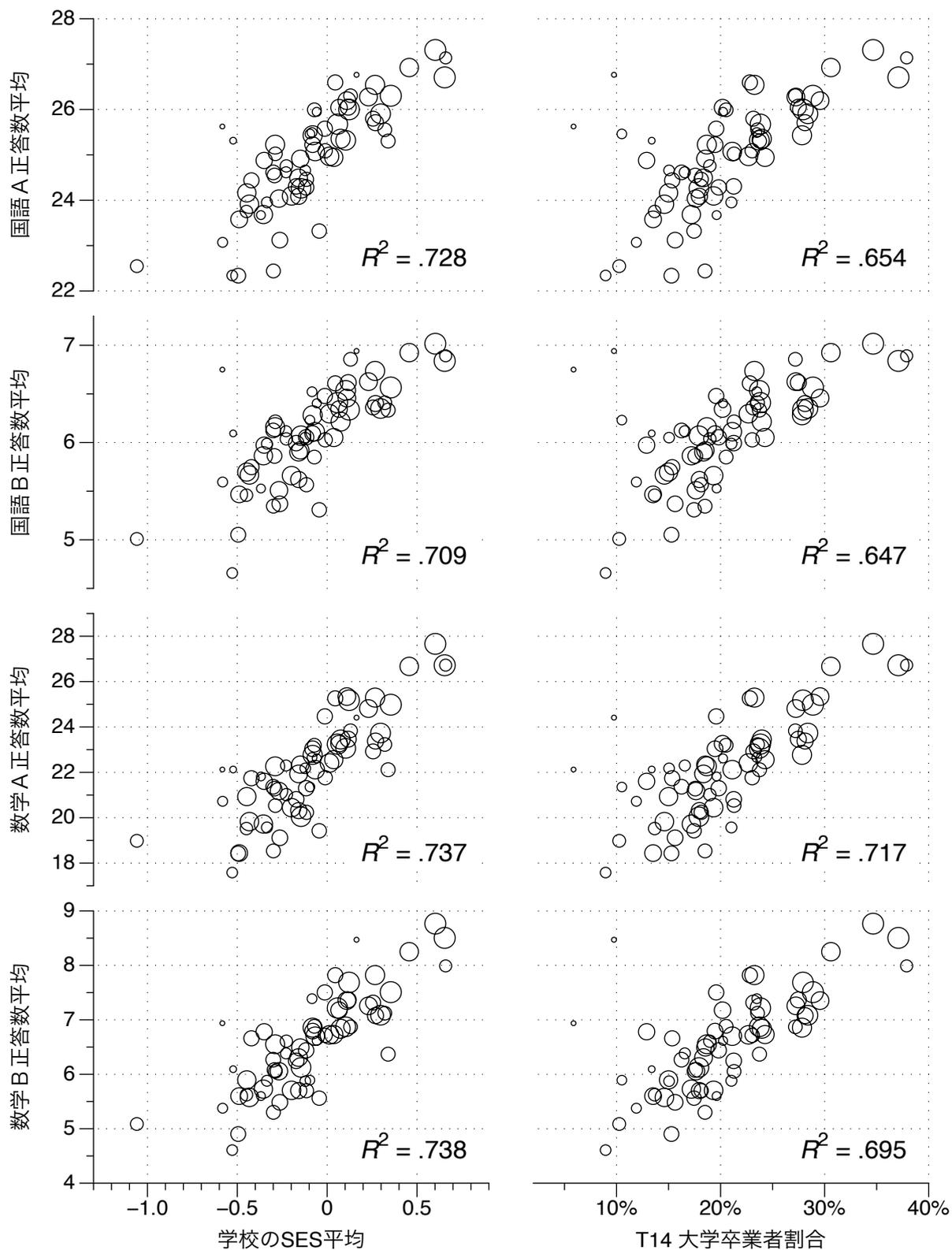


図 10: 学校の正答数平均と SES 平均・学校の地域特性 (中学校)

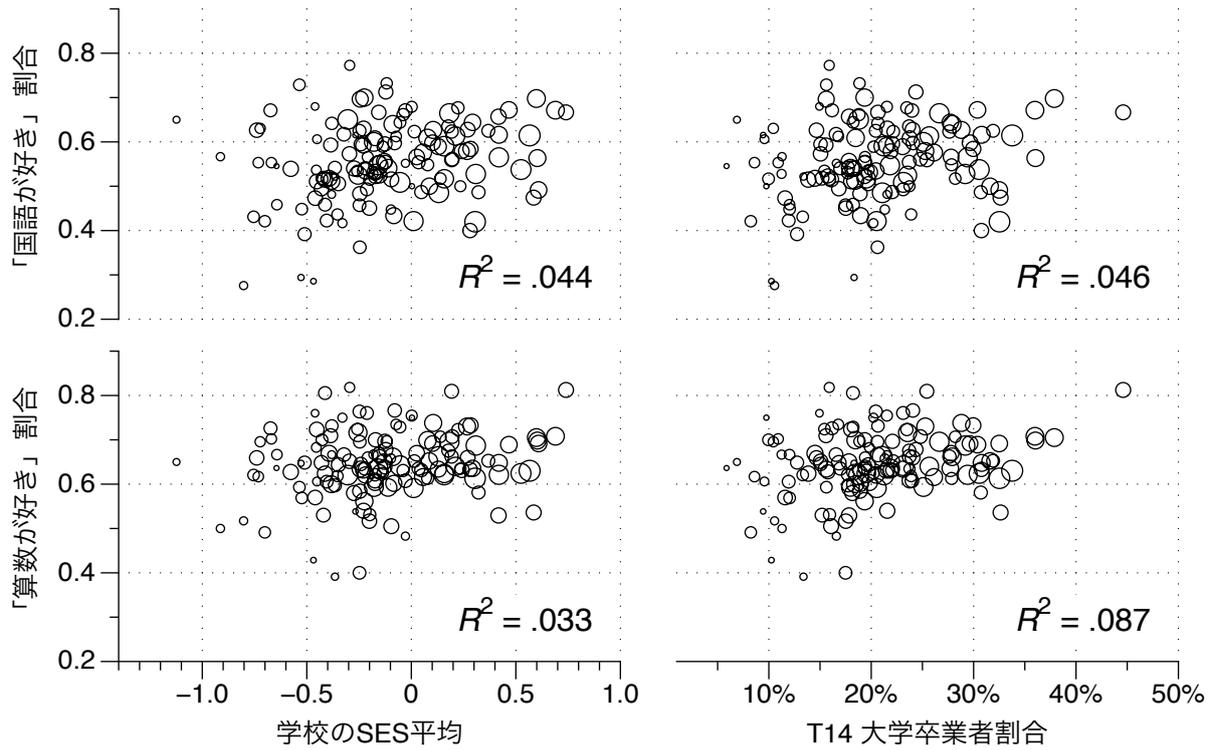


図 11: 国語・算数が好きな児童の割合と SES 平均・学校の地域特性 (小学校)

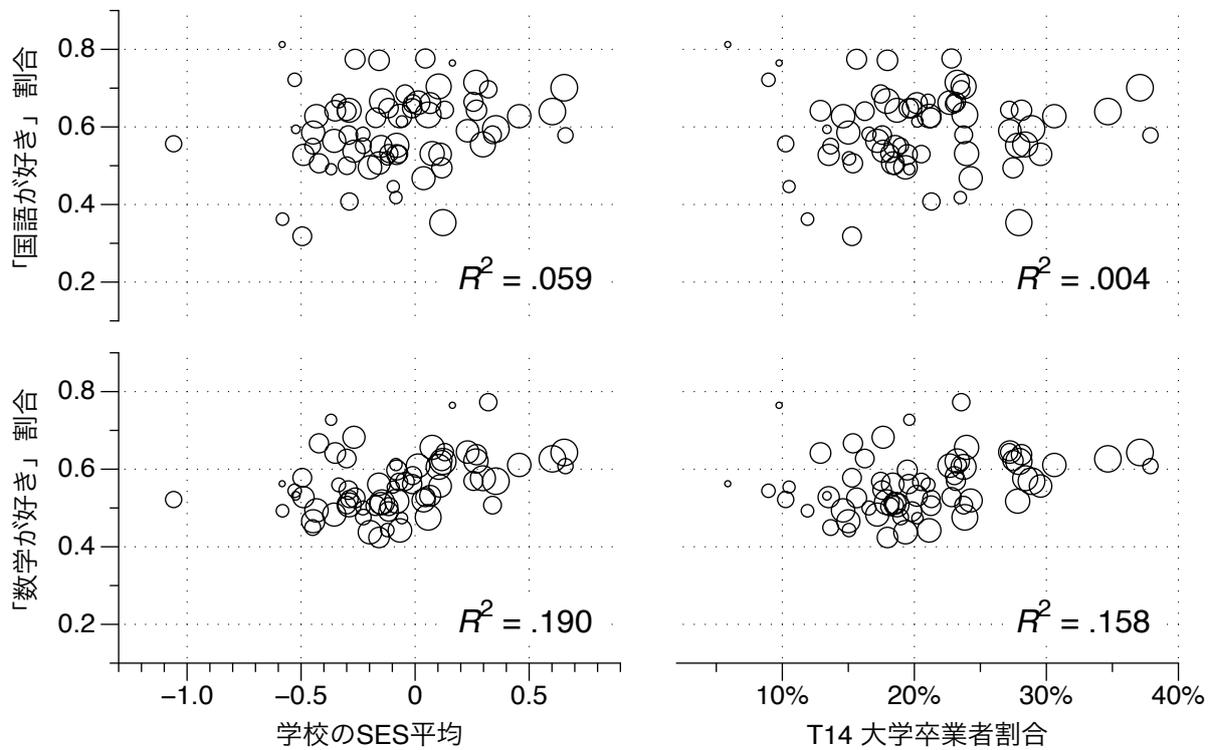


図 12: 国語・数学が好きな生徒の割合と SES 平均・学校の地域特性 (中学校)

## 6. まとめ

学校単位での SES 指標、あるいはその代替指標として、この章では当該校に在籍する児童生徒の保護者の SES 平均と、公開されたデータから算出した学校の地域特性の二種類を検討した。その結果、地域特性として住民における大学卒業者の割合を用いると、保護者の SES 平均と地域特性との間の相関係数は 0.9 に近く、後者は前者の十分な代替となり得ることが示された。また学校の正答数平均は、学校の地域特性を用いることで小学校では 6 割前後、中学校では 6 割から 7 割程度を予測可能であり、さらに小学校では保護者の SES 平均を用いるよりも予測の精度は高かった。ただし、国語や算数／数学が好きな児童・生徒の割合と、保護者の SES 平均あるいは地域特性との関係はあまり見られなかった。

学力調査の結果を予測する上で、小学校では地域特性の方が保護者の SES 平均よりも有効であった理由の一つとしては、有効回答率が全体で 8 割を下回っており、ウェイトの調整を行ったとは言え、学校の SES 指標として偏りがあったということが考えられる。つまり保護者から 100% の有効回答を得られれば、地域特性を用いるよりも正答数の学校平均の予測精度は高くなるかもしれない。しかし、現実の保護者調査では 100% の有効回答を得ることは不可能な場合が多く、そもそも保護者調査の実施自体が困難な場合も少なくない。調査実施に当たっては、学校や教育委員会などに新たに負担を求めることにもなる。また、仮に回答率が高かったとしても、小学校は中学校よりも 1 校当たりの学年在籍数が少なく、学校平均は不安定になりやすい。一方で地域特性は、学校の在籍数よりも大規模な、公開された住民データから算出されるため数値としても安定している。そして、学力調査の学校平均を予測する上で保護者の SES 平均と遜色ない。今後学校単位で学力調査データの分析をする際には、保護者に関するデータの有無に関わらず、本章で検討したような地域特性を SES の代替指標の一つとして考慮し、統制変数として用いることは必須といえるであろう。また児童・生徒を単位とした分析であっても、在籍校が置かれた環境を考慮するために地域特性を積極的に用いるべきであろう。

ただし、土屋 (2014) で示されるように、地域特性と学校の正答数平均との関係は、都道府県によってその様相がだいぶ異なる。都市部を多く含む都道府県では地域特性と学力調査結果との相関は高いが、他の県では必ずしも相関は高くない。今回の保護者調査の対象地域には都市部が多く含まれていたために本稿で示した結果となったが、他の地域でも同じことがいえるのかどうかはさらに検討が必要である。

### <参考文献>

- 土屋 隆裕, 2014, 「学校の地域特性について」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.