

Discussion Paper Series

University of Tokyo
Institute of Social Science
Panel Survey

東京大学社会科学研究所 パネル調査プロジェクト
ディスカッションペーパーシリーズ

私的学校外教育のもたらす高校進学への効果
—傾向スコア解析の応用

The Educational Effect of Private Tutoring on the Progression to High Schools: An Application of Propensity Score Analysis

中澤 渉

(大阪大学大学院人間科学研究科)

Wataru NAKAZAWA

March 2013

No.67

私的校外教育のもたらす高校進学への効果 傾向スコア解析の応用

中澤 渉（大阪大学大学院人間科学研究科）

学校外教育を利用することに真のメリットがあるといえるのかどうかは、多くの人々の関心をよぶといえるだろう。しかし従来の回帰分析をもとにした分析では、学校外教育の利用の有無と相関のある個人の異質性の影響を受けて、正確な効果を推定できない可能性がある。またこのような回帰分析で見出した効果は、サンプル全体の平均効果を述べているだけで、実質的な効果の多様性（異質性）を測定しているわけではない。教育の現場では、特定の教育的処置の効果が誰にでも一様に及ぶと考える方が不自然な場合が多いにもかかわらず、そうした問題は等閑視されてきた。以上の問題を解決するための方法として、傾向スコア・マッチングを本稿で提案する。①家庭教師、②通信添削、③塾・家庭教師・通信添削のうち少なくとも1つの学校外教育、の3点について、その利用経験が大学進学率100%の高校進学に有利に作用したかを検証した。その結果、家庭教師や通信添削は概ね一貫した傾向は見いだせなかったが、学校外教育全体としては、男女とも傾向スコアの低い層、つまり学校外教育をもともと利用しにくい人々が、学校外教育を利用することによって有意に進学校進学率を高めていた（メリットを得ていた）ことが明らかになった。

謝辞 本研究は、科学研究費補助金基盤研究（S）（18103003, 22223005）の助成を受けたものである。東京大学社会科学研究所パネル調査の実施にあたっては、社会科学研究所研究資金、株式会社アウトソーシングからの奨学寄付金を受けた。パネル調査データの使用にあたっては社会科学研究所パネル調査企画委員会の許可を受けた。

1. 問題設定

従来から、日本をはじめ東アジアは受験競争の激しさが指摘されてきた。いわゆる一発型の筆記試験（high-stakes test）を重視する入試が背景にあるため、受験準備教育が多くの人に関心をよんだ。その準備のためには、予めカリキュラムの定められている公教育だけでは不十分とされ、学校を終えてから塾や予備校に通う、通信添削を行う、場合によっては家庭教師をつけるといった学校外教育を利用することが極めて広範囲で観察されてきたのである。

多くの人にとっては、こうした学校外教育を受けることで、どれだけメリットがあるのか（具体的には成績が上がるのか、受験に成功するとかいうような「効果」があるのか）といったことが関心事となるだろう。ではその効果はどうすれば確かめられるのだろうか。こうした学校外教育は、当然のことながら個人の自由選択によって行われるものであり、その受講には金銭的負担を要する。その際、「誰が」学校外教育を受けることを選択するのかが問題になる。常識的には、本人が勉強に熱心だとか、保護者が教育熱心だとか、授業料を支払う経済的余裕があるといったことが、学校外教育を受けるか否かの選択に依存していると考えられるだろう。そこで、学校外教育を受講していた人とそうでない人との間で、成績や進学率を比較し、前者の成績が高い、もしくは進学率が高ければ、学校外教育に効果があったと見なすのは妥当なのだろうか。

当然この答えはノーであろう。日本の学校の例を挙げれば、こうした推論が誤っているのはすぐに想像がつく。日本の高校は入試難易度（いわゆる偏差値）によるランク付けが明瞭に存在している。したがって入口で既に生徒の選別が行われている。そして一般的には、大学の合格者数などでその高校の実績が評価される。では大学の合格者数の違いは、その高校の教育効果の有無の違いと即断できるのだろうか。もちろん部分的にはそうかもしれないが、当然入学時の選別の影響が大きく効いていることは、多くの人認識しているだろう。要するに、ここでは「学力を伸ばす教育」が行われているから進学実績が高いとは限らず、そもそも「学力が高い」者を選んでいるから進学実績が高いに過ぎないのかもしれない¹。学校外教育についても、同じようなことが言える。つまり出てきたアウトプットを単純に比較しただけでは、本当の「効果」を確かめたことにならないのだ。

本稿では、こうした「教育の効果」について考察する。教育社会学では、何らかの教育的処置に実際に効果があるか否かは重要な関心事の一つとなりうるはずである。というのも、一般的には教育によって知識技能を身につけることが可能となり、それが自由な職業選択を可能にし、平等化を進めると考えられているからである。しかし一方で、教育が平

¹ 類例として Lieberson（1985）の chapter2 などを参照。

等化に果たす機能に限界がある（平等化には寄与しない）という見解も多く発表されてきた（Bourdieu and Passeron 1970=1991; Bowles and Gintis 1976; Collins 1979=1984）。しかし教育は平等化にほとんど寄与しないという知見は、教育が果たすとされる機能や役割期待に照らせば、深刻な結論であるに違いない。こうした問題点を克服しようと、アメリカではヘッドスタート計画のような政策が大々的に実施されてきた。日本でも近年、階層差を克服しようとする「効果のある学校」研究の蓄積が見出されるようになってきている（鍋島 2003; 川口 2006; 川口・前馬 2007）。

データの制約もあるのだろうし、また何を効果と見なすかという定義にもかかわってくる部分もあるが、こうした研究には次のような限界があると考えられる。第一に、これらの研究は学校（集団）ごとに階層別のテストの基準点の通過率を算出し、その通過率が全体より高かったり、あるいは不利な階層出身者の通過率が中流階層を上回っていたり、ほとんど差がなかったりするときに、その学校では効果を上げていると定義している。もちろん効果の有無の判定には何らかの操作的な定義が必要であり、それ自体問題ないのだが、ただしここでいう階層の中身や基準点の設定などが恣意的である。第二に、集合レベルで特定の基準を突破するということが一つの基準となること自体、それはそれで構わないかもしれないが、それが一般的に教育の効果といったときにイメージされる内容なのか、という問題がある。極端なケースとして、大半が基準点ギリギリであってもその基準点を超える人が多い学校と、全体として基準点を超える人の数がその学校より少なくても、基準点より大幅に多く得点できる人がたくさん出た学校（そして後者の学校が平均点が高い）があったとき、前者より後者に効果があると判断する人は珍しくないだろう。それ以外に、データを扱う際には、個人レベルの相関と、集合レベルの相関に矛盾が生じるという「生態学的誤謬」の問題が起りうるが、それを考慮しているとはいえない。第三に、インタビューやフィールドワークによって解釈の補強を行っているが、結局計量的データにより導き出された結果と、質的調査により得られた結果の結びつきが恣意的であり、都合のよい解釈を行っているのではないか、という疑いが残る。つまり通過率の高いとされる学校の特徴をいろいろ聞きだし、効果がありそうなものを選んで記述しているだけであり、真の効果があるかの厳密な論証になっていない。彼らの関心が学校という集団レベルにあることからやむを得ない面もあるが、結局学校のなかのどういう処置に意味があるのかよくわからない、ということである。具体的にどういう教育的な取り組みを行っているか効果が上がるのか、といったことが、おそらく実践の現場からも求められる知見だと思われる。実際、質問紙によってそれぞれの学校の特徴を明らかにし、効果のある学校とそうでない学校の特徴の違いを明確にすることで効果の有無を検証しようとしている。これはこれで一定の意味はあると思われるが、そこには具体的な処置によって対処できるものや、学校の雰囲気のようなものなど、様々なものが含まれており、結局漠然とした学校文化によっ

て結果に違いが起こる，といっているに過ぎないようにも思える。そして最後に，これが分析上最も問題になるところだが，一応彼らは階層という変数を統制して学力を測定しているが，学力に差を生ぜしめるもともと個人がもっている原因は，階層に限定されるのだろうか，という問題である。つまり階層と学校以外の（地域や，その学校に通う生徒たちがもともと持っているような）特性がコントロールされておらず，**selectivity** によって説明できることを「効果」とよんでいる可能性が否定できないのである。

教育現場のことを想定すれば，1つの教育的処置や試みがダイレクトに生徒に影響を与え，何らかの効果を及ぼすというようなことはむしろ稀であろう。だから効果の発生するメカニズムがあまりに複雑であることから，こうした「効果のある学校」研究方法にやむを得ない部分もあるのかもしれない。例えば，特定の教育方法の効果の及び方は，生徒の属性や態度，状況によっても変化するだろう。同じ教育方法であっても，教師の属性で効果に違いがあるかもしれない。心理学や疫学では「交絡」とか，社会学では「交互作用」効果などとよばれるものだ。近年計量分析では，こうした複雑な現象を測定，推定しようとする試みが進められている。残念ながら，これらの研究は，そういった計量研究の枠組みを踏まえているとはいえない。

そこで本稿では学校外教育の例を題材に，傾向スコア・マッチングの方法を用いて，上記の方法の問題点を解決する一つの方法を提案したい。

2. 効果の有無の推定について

(1) 教育効果の考え方

教育の効果がある，といったとき，私たちは何を想像するだろうか。ここでは学校外教育によって，大学進学率 100%の高校の志望校に受かるかどうか，ということをめぐる教育の効果を議論することにしよう。

a. 反実仮想的な仮説の設定

自分，もしくは自分の子が，学校外教育（塾，家庭教師や通信添削など）を受けるか否かの選択に迫られているとしよう。そうした学校外教育は，友人がやっているから自分もやる，という側面もあるかもしれないが，まずは成績を伸ばしたい，という目的に焦点を絞って，その目的の達成に十分意味があるかどうかによって受けるかどうかを決断することになるだろう。そのとき，多くの人は頭の中で，自分が学校外教育を受けた場合とそうでない場合に思いを巡らし，結果を想像するはずである。そして一方の決断を下す。この決断の結果どうなるかは，しばらくすればわかるのだが，ただしわかったその結果はあくまで選択した結果のみについてである。つまり学校外教育を受けるという決断を下した場

合は、当然観察される結果は学校外教育を受けた場合の結果であり、学校外教育を受けなかった場合にはどうなっていたかはわからない。これは逆の場合も同様である。もっとも日常的には、これ以上深く掘り下げることなく、出てきた実際の結果を本人は受け止めて次の決断を下す、というのが人生の選択だろう。しかし科学の論証はそうはいかない。

もちろん、選択されなかったケースについては、実際にはその事象は発生していないので、観察のしようがない。だからわからない、というのがある意味では最も科学的に真摯な解答なのかもしれないが、これでは学問上何もしたことになる。そこで科学的な厳密性を担保しつつ、「もし実際のケースと別の選択をしたならば」という反実仮想的（counterfactual）なケースを想定して推定を行うことを考える（Morgan and Winship 2007）。つまり自分や自分の子のコピー人間が存在したと仮定して、そのコピー人間が実際の選択と違う選択をした場合に、どんな結果が生じるかを想像するわけである。実際にはコピー人間は存在しないから、自分や自分の子に似た人の中で、かつ自分や自分の子の選択と異なる選択をしている人を探し出して、その結果を比べれば、近似的に反実仮想的な仮説を検証することが可能になるのではないかと考えるのである。

b. ランダムサンプリングの重要性

しかし反実仮想的な仮説を、上述の手順で行うのは現実には不可能に近い。そこで研究者は、限界を認識しつつも、何らかの形で「効果」を推定できる方法はないかと考える。そこで次善の策として、関心のある処置を受ける人とそうでない人のグループを分けて、処置前後の変化を比較する、という実験的方法が思い浮かぶ。いうまでもなく重要なのは、そのグループ分けである。ただしそのグループの性質が大きく異なっていたら、既に述べたように処置の効果の違いなのか、selectivityの問題なのか、といった問題が区別できない。処置の影響に交互作用がある可能性もある（特定の性質のグループにのみ、効果が強く及ぶということもあるかもしれない）。したがって、当面は一定数のサンプルを確保し、そのサンプルは特定の性格の人に偏らないようにランダムに抽出する。そして一方に処置を行い、もう一方は対照群として何も行わないようにして経過観察を行う。その結果生じた結果の違いの差が有意か否かを検証すれば、事実上一般的な実験科学で行われている検証を踏襲することになる。つまりこの例でいえば、学校外教育を受けさせた群の成績がより伸びたり、あるいは進学率が高くなったりすれば、学校外教育による有意な効果があったと見なせることになる。

c. パネル固定効果推定の意味

効果というのは時間的な推移に伴って発生するものである。つまりどこかで観察の基準になる時点があり、そこから一定のトリートメントを経て、結果がどうなったかという変

化を見極めることになるわけだ。ところが、この方法は現実にはすぐに壁にぶつかるだろう。そもそも学校外教育の選択は個人の自由であり、しかも場合によっては進路選択という人生を左右する結果につながりかねないものだ。それを研究目的とはいえ、一方的な選択を強制することは倫理的に許されないし、そんなことを指示するのは不可能であろう。

またこうした実験的方法を、人間社会に適応できると考えること自体に問題がある。仮に学校外教育を受けられるかどうかをランダムに振り分けることが可能になったとしよう。そして学校外教育を受けるか否か以外の要因が成績や進路決定に影響を与えないように、それ以外の環境は強い統制を受ける必要があるだろう。しかしこの現実社会において、それは明らかに不自然な状態である。こうした実験を行う際には、倫理的にも何らかの説明を事前に被験者は受けることになる。したがって被験者は、何らかの実験の対象となっていることを必然的に自覚する。つまりこうした特殊な環境下に置かれた被験者は、その環境それ自体と、被験者であることを強く意識することで、自らの行為選択や意識に影響を与える可能性がある。それゆえこうした環境下で得られた結果が、現実社会に適用できる保証はない。いわゆるホーソン効果の問題である（星野・岡田 2006）。

だから社会学者は、真に実験的環境をつかって実験を行うのではなく、現実社会から得られるデータを、この実験的環境に近づける、というような方法を考えることになる（Rothman 2002=2004: 63-69）。現実の社会は、様々な異なる要素を保持している個人が存在し、しかも相互に錯綜して生活している。そしてその要素の変化の有無や程度も人それぞれである。もしパネルデータのように、同一個人を追跡するデータが獲得できていたとする。そしてその間で学校外教育を受けるようになった、逆にそれらをやめた、というような学校外教育に関する変化のデータと、観察期間内の成績の変化のデータが得られているとする。学校外教育以外にも、学校生活に関するデータ、生活習慣に関するデータなどが同時に獲得されていると仮定する。固定効果モデルの発想は、説明変数の学校外教育の受講の有無の変化を含む個人内でのそれらの変数の変化と、従属変数の成績の変化との間に、有意な関連性があるかを確認しようとするところからきている。説明変数には学校外教育以外の変数も考慮されているだろうから、もし学校外教育が有意ではなく、他の変数（生活習慣や学校生活に関する変数）が有意になれば、むしろ学校外教育の有無よりも他の変数の変化が成績の変化に影響を与えているのだと推測できる。また観察期間内に変化の起きなかった変数については、考慮する必要はない（実際、固定効果モデルの推定式からは、変化しない変数は排除される）。そもそも特定の変化が観察期間内に起こった（処置を行った）ことが、従属変数に影響を与えたか、というのがここでの関心事だから、変化が起きていないというのは処置を何もしていないということで、その因果効果を確かめ

ようがないからである²。

ところが現実には、こうした条件を満たすようなパネル・データの獲得自体が、現状では困難である。現実によく存在するのは1時点の調査で集められたデータであり、その中に回顧的な情報が含まれている、というものであろう。そこでこうしたデータの活用方法として、傾向スコア・マッチングの考え方が提案されることになる。

d. 傾向スコア・マッチング

本項の a の内容に戻って考えると、質問紙調査では大量の個人データが含まれているのが普通であるから、その大量の個人データから、学校外教育を受けている人とそうでない人とで似た人を選び出して、出てきた成績や進路決定というアウトプットを比較すればよいのではないかと、という発想が自然に出てくるだろう。ところが問題なのは、何をもって似ている個人と判断するかである。

個人を特定する変数は、性、年齢、出身階層（階層についても、複数の指標があり得る）などたくさん存在する。それらを一つ一つ考慮して似ているサンプルを探し出そうとした場合、いくら大量データを集めたとしても、結局その作業は困難に陥るだろう。さらに連続変数の場合、同じ値をもつ別の人を探すのは、むしろ奇跡に近い。つまり似た人を探し出そうとしたときに、変数をいくつも同時に考慮しなければならない上、更に一つの変数についてもマッチングさせる基準が無限に拡大する可能性があるため、マッチング自体が不可能になるという次元の呪い（curse of dimensionality）に陥ってしまうのである（星野 2009: 57, Brand and Xie 2010）。

ところが、Rosenbaum と Rubin は、いわゆる「強く無視できる割り当て条件³」のもとで、ある処置 t （この場合学校外教育）に割り当てられる理論上の確率 e_t による調整を行いさえすれば（多くの説明変数を同時に考慮するのではなく、 e_t が同じ、もしくは非常に近いもの同士が事実上同じ性質をもつサンプルだと見なしてマッチングさせれば）、そこで考慮されている全共変量を用いて調整を行ったのと同等の偏りの補正が可能になることを証明した（Rosenbaum and Rubin 1984）。この e_t が傾向スコア（propensity score）とよばれるものであり、通常は t の有無を従属変数とするロジスティック回帰分析によって推

² 固定効果モデルの最大のメリットは、変数化できないような個人の変わらない性格も、推定値に影響を与えないという点にある。これは固定効果推定の計算の中ですべて消去されるからである。通常の回帰分析では、この変数化できないような個人の変わらない性格、つまり観察されない異質性が推定値にバイアスをもたらしている可能性が否定できない。このことがパネル分析を発展させる大きなインセンティブになっている（中澤 2012）。

³ Strong ignorability assumption とよばれており、 t の有無の割り当ては観察されている共変量「のみ」によって決まり、それ以外の要素は全くランダムである、という仮定を置くことを意味する。

定される。

(2)効果自体の異質性

以上の議論は、学校外教育という処置と、もともと持っている個人の性質との間に何らかの相関があり（そもそも特定の性格をもつ人が学校外教育を受ける傾向があり）、学校外教育の効果と思っていたものが実際はもともと持っている性格の違いに過ぎなかったというセレクション・バイアスによるものであるという疑念が拭えないため、そのセレクション・バイアスを取り除いて純粋な効果を推定しようとするものである。

しかし教育の現場で、特定の処置（教育方法）がどの生徒にも均等に期待できる、と考えるのは、かなり特殊な例か、前提が甘すぎるといわざるを得ないだろう。現実には、特定の教育的措置をとっても、うまくいく例とそうではない例が混在していると考えるのが自然ではないだろうか。しかしこれを「人それぞれ」と言ってしまえば、元も子もない話になってしまう。科学的には、どういった人に効果があり、また逆にどういった人には効果がないかを確認しなければいけないだろう。

仮説が非常に明確で、特定の共変量に関心があって、その共変量の値によって処置変数の効果に違いが生じるか否かを知りたいのであれば、交互作用効果を推定すればよい。例えば、学校外教育と家庭の所得との間に交互作用効果があると仮定するのであれば、それぞれの主効果と交互作用効果を説明変数に投入し、交互作用効果が有意であれば、学校外教育の影響の仕方（傾き）は収入によって有意な違いがあることを示すことになる。

ただし社会科学的に見て、収入は出身階層を構成する一つの要素であるが、仮説が「階層によって学校外教育の効果は異なる」と考えるのであれば、収入だけを考慮するのは不十分である。親の学歴や職業なども考慮する必要があるだろう。そもそもこれらの変数は、お互い相関があると考えるのが自然である。さらに交互作用項はもともとの主効果とも相関があるので、回帰分析における多重共線性の問題が懸念されることになる（推定は可能であっても、互いの説明変数同士の相関があるため、純粋に見たい回帰係数の推定値が求められているかは怪しいことになる）。また交互作用効果を入れると、それだけ推定モデルの共変量が増加し、推定結果も不安定になりがちになる。

ところが既述したように、傾向スコア e_f は、そういった考慮している共変量のすべてを1つに要約した値とも見なすことができる。つまり傾向スコアの大小は、その措置（この場合は学校外教育）の受けやすさを示す値であり、傾向スコアの大小によって層別にグループ分けして、各層ごとに処置群と非処置群のアウトプットを比較して差が見いだせれば、その層において効果があったと見なすことができるだろう。つまり傾向スコアを活用することで、処置を受けやすい人と受けにくい人の効果の違いが推定できる、ということである。

学校外教育についていえば、冒頭に述べたように、その利用選択を行うか否かには一定の傾向があると考えられる。例えば Kuan (2011) はこうした傾向スコアの層別解析を通して、台湾の塾 (cram school) の学習効果を推定している例である。それによれば、第 9 学年 (中学 3 年) 時点の数学の学力に対する塾の効果は違いがあり、塾に行かない傾向がある子ども (そういった子どもは、もともと数学の点が低く、親が低学歴である傾向がある) ほど学力の伸びがみられ、その一方もともと通塾する傾向のある性質をもつ子ども (数学が得意で、親が高学歴の傾向がある) での学習効果は少なくなるという。

このように、傾向スコア (理論上、そのトリートメントの受けやすさを示すスコアになる) に基づいて個人を n 分位 (通常は 5 分位が多い) に分割し、各分位において、実際にトリートメントを受けた人の群とそうでない人の群との間でアウトプットの平均値の比較 (t 検定) を行う。その結果、平均値に有意な違いがあれば何らかの効果があったと見なせるし、そうでなければトリートメントによる効果は特に見いだせない、という結果になるわけである。

3. データと変数

(1) 利用データ

ここでは主として、社研パネル調査の第 2 波の変数を中心に扱う。この年の調査で、教育に関する情報を多く獲得しているからである。学校外教育については、小中高それぞれについて、塾、家庭教師、通信添削の利用経験を尋ねている。本稿では、家庭教師、通信添削、そして塾・家庭教師・通信添削を区別せずいずれか 1 つでも学校外教育を利用していた場合の 3 つについて推定を行う⁴。なお、ウェーブ 2 ということでサンプル脱落 (attrition) が問題になるが、パネル調査では事前に継続調査に協力できる人に依頼しているため、定着率は 80% を超えており、脱落による深刻なバイアスはないことを確認している⁵。調査対象は 2007 年時点での 20~35 歳の若年層、36~40 歳の壮年層で、その後の追跡調査も含め、必要な変数すべてに回答しているサンプルを用いている。サンプルサイズは男性が 1,213、女性が 1,352 である。

(2) 利用する変数

本稿の処置変数は、中学校時代において、①家庭教師利用経験があれば 1、なければ 0、

⁴ 塾の同様の分析については、中澤 (2013) で示す。

⁵ バイアスの補正を行う上では、そもそもの初年度のサンプルが何らかの傾向をもっていないか問題になるが、それについては三輪 (2008) にあるように、他の同様の全国規模の調査と比較して、社研パネル調査が特に大きな何らかの偏りがあるとは言えない。

②通信添削利用経験があれば 1, なければ 0, ③塾, 家庭教師, 通信添削のいずれか少なくとも 1 つの利用経験があれば 1, いずれも利用したことがなければ 0, というダミー変数をとる。一方, 従属変数であるトリートメントの効果は, 大学進学率の高い「進学校」に進学できたか否かで測ることとする。ここでは自分の出身高校の大学進学率について尋ねる質問への回答を利用し, 「ほぼ全員が進学した」と回答した人を 1, それ以外を 0 とするダミー変数とする。学校外教育を利用することで, こうした進学校に進学できた人の比率が高くなったか否かで効果をみることにする。多くの人にとって, 大学に進学できるか, さらにいえばどの大学に進学できるかが重要な関心事となっているが, それは進学した高校が大きな決定要因となっていることが伝統的な教育社会学のトラッキング研究などからも, あるいは経験則からも明らかであるといえる。つまり学校外教育を受けることで, 実質的に(有名)大学への進学率を高めるような進学校に行けるようになるか否かは, 本人にとって重大な意味をもつと考えられる。こうした分析目的から, 高校受験がそもそも存在しない中高一貫校の出身者は, この分析から除外している。

これら以外の変数(共変量)としては, 出生コーホート(5年単位で分割), 両親の学歴, 出身地域(15歳時点での居住地), きょうだい数, 父職, 15歳時に家にあった本の大凡の冊数, 15歳時の個室(自分の勉強部屋)の有無, 15歳時の勉強机の有無, 15歳時の暮らし向き(5段階)である。換言すれば, これらの変数が, 学校外教育の受講の有無と関連があると考えられる。したがって学校外教育の説明変数に一定のバイアスをもたらしていると考えられるので, 共変量調整が必要になると予想される。

(3) 分析の手順

まず学校外教育の経験に関する傾向スコアを推定する必要がある。これは家庭教師, 通信添削, 学校外教育の経験の有無を従属変数とするロジスティック回帰分析を行い, その推定モデルからそれぞれのサンプルに対して理論上の従属変数が 1 となる確率(予測確率)を割り当てる。このロジスティック回帰分析の予測確率が傾向スコアである。ただし同等の傾向スコアであっても, 「実際に」経験があるサンプルとそうでないサンプルが存在する。その両者をマッチングする作業が必要になるが, これには汎用統計ソフト **Stata** を用いる。ここでは最近傍キャリパー・マッチングを用いる(Guo and Fraser 2010: 146-147)。問題になるのは, 何をもちいて最近傍と見なすかであるが, 本稿では 2 つの基準を当てはめてみることにした。1 つは傾向スコアの標準偏差の 4 分の 1 (0.25 倍) の値までの差をマッチングできる許容範囲とするもの, もう 1 つは同様の値の差を 0.1 までとするものである。なお, マッチングできなかったサンプルは, その後の分析では用いられない。

その後は, マッチング前のサンプルで推定した進学校進学に対する学校外教育の効果のロジスティック回帰分析と, マッチング後のサンプルに行った同じ回帰分析の結果を比較す

る。そして傾向スコアの大きさに基づいて5分位で分割し、層ごとに進学校進学率の差をt検定して、実際に差があるのはどこか（サンプル特性による効果の違い）について検討する。

4. 分析の結果

(1) 処置(学校外教育)を予測するロジスティック回帰分析の結果

まず、最初の手順である傾向スコアの計算のため、従属変数を学校外教育にしたロジスティック回帰分析を推定する。その前に、実際のデータでどの程度学校外教育の利用者がいるのかを確認しておこう（表1）。

	利用者数	非利用者数	利用者での進学率	非利用者での進学率	ピアソンカイ二乗値	p値
家庭教師(男性)	128	1085	14.8%	19.7%	1.757	.185
家庭教師(女性)	153	1199	20.3%	20.6%	.010	.922
通信添削(男性)	143	1070	27.3%	18.1%	6.793	.009
通信添削(女性)	223	1129	29.6%	18.8%	13.344	.000
学校外教育(男性)	904	309	21.7%	12.0%	13.983	.000
学校外教育(女性)	1030	322	23.0%	12.7%	15.861	.000

注)言うまでもなく、利用者と非利用者の間での進学率をt検定した結果も一致する。
その場合、平均値がここでの「進学率」となり、p値がt検定の有意確率と同じ値になる。

表1は、家庭教師、通信添削、学校外教育（塾・家庭教師・通信添削いずれか少なくとも1つ）の利用状況をサンプルから示したものである。そしてそれぞれの学校外教育利用の有無と、進学校進学の有無をクロス表にとって、その結果をカイ二乗検定した結果がピアソンのカイ二乗値とp値(有意確率)である。女性の通信添削利用率が16%を超えるが、家庭教師と通信添削(男性)は10%~12%に過ぎない。ただし塾も含めたいずれか1つの学校外教育、ということになると、利用していない人の方が少なく、男女ともほぼ4人のうち3人(75%前後)が学校外教育を利用していたことになる。家庭教師や通信添削の利用率を考慮すれば、学校外教育の大半は塾が占めていることがここからも理解できる⁶。興味深いのは、家庭教師と通信添削の利用率には大差ないが、そのアウトプット(進学校進学)に及ぼす影響は対照的であることだ。家庭教師については、利用したところで進学校進学率が高くなる傾向は見いだせず、他の学校外教育とは異なる機能をもっているのではないか、と考えられる。

⁶ ちなみに、塾の利用者は男性766人、女性841人で、63%前後である。また進学校進学者は、男性の塾利用者の22.1%、非塾利用者の14.3%、女性の場合は塾利用者の23.1%、非塾利用者の16.4%であり、ともに1%水準で有意な差がある。

	男		女	
	Coef.	S.E.	Coef.	S.E.
1971～75年生(vs 1966～70年生)	-.062	.271	-.264	.249
1976～80年生(vs 1966～70年生)	-.126	.293	.010	.260
1981～86年生(vs 1966～70年生)	.241	.280	-.059	.262
父中学校卒(vs 父高校・専門学校卒)	.181	.316	-.407	.275
父短大・高専・大卒(vs 父高校・専門学校卒)	-.022	.256	-.129	.241
母中学校卒(vs 母高校・専門学校卒)	-.632	.353 +	-.035	.266
母短大・高専・大卒(vs 母高校・専門学校卒)	-.140	.282	.011	.262
15歳時北海道居住(vs 関東)	-1.511	1.031	-.396	.503
15歳時東北地方居住(vs 関東)	.473	.371	-1.050	.458 *
15歳時中部地方居住(vs 関東)	.773	.267 **	.059	.250
15歳時近畿地方居住(vs 関東)	.320	.320	.564	.258 *
15歳時中国地方居住(vs 関東)	-.029	.475	.201	.334
15歳時四国地方居住(vs 関東)	.769	.429 +	-.506	.626
15歳時九州地方居住(vs 関東)	.639	.353 +	-.334	.351
きょうだいなし(vs きょうだい1～2人)	-.147	.419	.633	.307 *
きょうだい3人以上(vs きょうだい1～2人)	-.197	.421	.020	.345
父専門職(vs 父事務職)	-.070	.320	.181	.298
父販売職(vs 父事務職)	.287	.352	.279	.359
父自営・農業(vs 父事務職)	.427	.307	.718	.292 *
父熟練職(vs 父事務職)	-.176	.390	.474	.354
父半・非熟練職(vs 父事務職)	-.333	.369	.239	.319
本の冊数(10冊単位)	.004	.007	.008	.006
15歳時に自分の個室あり(vsなし)	.192	.258	.389	.234 +
15歳時に自分の勉強机あり(vsなし)	.362	.546	.468	.616
暮らし向き(5段階, 大きいほどよい)	.304	.125 *	.410	.108 **
定数項	-3.961	.740 *	-4.411	.758 **
-2 Log Likelihood	774.250		899.808	
N	1213		1352	
C統計量	.661		.670	
+<.10 *<.05 **<.01				

表2は家庭教師の利用経験について推定したモデルの結果である。男女で共通しているのは暮らし向きのよさであるが、特に男性では、地域変数を除いて目立った有意な変数はない。女性については自営や農業層、また個室をもっていた者に利用経験が多いようで、豊かな層において利用されていたという傾向が見て取れる。

次ページの表3は通信添削だが、やはり男性ではあまり有意な変数は存在しない。暮らし向きは10%水準で有意であるが、家庭教師と逆に豊かでない人が利用する傾向が強い。ただしこれも10%水準であるが、母学歴は高い傾向がある。女性は男性よりは傾向が見て取れるが、若いコーホートと、母学歴が高い人でより多い利用がみられるようである。

最後に、家庭教師、通信添削のみではなく、塾も含め、いずれか1つでも学校外教育を経験していたか否かを推定したロジスティック回帰分析の結果を掲載している(表4)。細かな違いはあるが、男女とも親の学歴が低いと利用する傾向が少なくなることがわかる。暮らし向きに関する変数は女性のみ有意で、当然豊かな方が利用する傾向がある。またき

	男		女	
	Coef.	S.E.	Coef.	S.E.
1971～75年生 (vs 1966～70年生)	-.020	.260	.272	.224
1976～80年生 (vs 1966～70年生)	.221	.262	.598	.232 *
1981～86年生 (vs 1966～70年生)	.028	.276	.604	.234 *
父中学校卒 (vs 父高校・専門学校卒)	.305	.307	-.094	.225
父短大・高専・大卒 (vs 父高校・専門学校卒)	.027	.246	-.125	.204
母中学校卒 (vs 母高校・専門学校卒)	.025	.314	-.023	.227
母短大・高専・大卒 (vs 母高校・専門学校卒)	.460	.253 +	.429	.212 *
15歳時北海道居住 (vs 関東)	.382	.393	-.195	.362
15歳時東北地方居住 (vs 関東)	-.079	.364	-.117	.277
15歳時中部地方居住 (vs 関東)	-.124	.265	-.301	.213
15歳時近畿地方居住 (vs 関東)	-.123	.296	-.680	.259 **
15歳時中国地方居住 (vs 関東)	.783	.322 *	-.282	.301
15歳時四国地方居住 (vs 関東)	-.012	.472	-.222	.415
15歳時九州地方居住 (vs 関東)	-.478	.406	-.478	.288 +
きょうだいなし (vs きょうだい1～2人)	.135	.361	.164	.303
きょうだい3人以上 (vs きょうだい1～2人)	.287	.341	-.396	.313
父専門職 (vs 父事務職)	.218	.286	.327	.242
父販売職 (vs 父事務職)	-.182	.351	-.037	.313
父自営・農業 (vs 父事務職)	-.034	.303	.001	.253
父熟練職 (vs 父事務職)	-.439	.365	.120	.295
父半・非熟練職 (vs 父事務職)	-.492	.346	.149	.253
本の冊数 (10冊単位)	.006	.006	-.002	.006
15歳時に自分の個室あり (vs なし)	-.039	.225	-.107	.180
15歳時に自分の勉強机あり (vs なし)	.807	.542	.535	.453
暮らし向き (5段階, 大きいほどよい)	-.217	.124 +	-.154	.099
定数項	-2.308	.693 **	-1.808	.589 **
-2 Log Likelihood	844.945		1175.67	
N	1213		1352	
C統計量	.649		.622	
+<.10 *<.05 **<.01				

ようだい数が多くなると、こうした学校外教育の利用傾向が弱まることがわかる。父職については、男女とも有意な結果をもたらさなかった。

なお、それぞれの表にある c 統計量は、いわゆる「強く無視できる割り当て条件」を満たしているかの条件となり、通常 0.8 以上が望ましいとされる。換言すれば、この値が小さいということは、ここに投入している共変量によって、従属変数の処置変数の有無を完全に区別しきれない可能性が高い（依然、処置の割り当てに有効な変数が残されている）ことを意味する。したがって完全にバイアスが除去しきれない可能性が高いことになるが、星野・岡田（2006）によれば c 統計量が 0.67 程度でも従属変数に強い関連のある共変量が入っていれば、偏りのない調整が可能であるという。本稿では、それよりやや小さな値を示しているものもあるが、ひとまずどのモデルでも同じ共変量の選択に基づいて推定を行った。

	男		女	
	Coef.	S.E.	Coef.	S.E.
1971～75年生(vs 1966～70年生)	.454	.183 *	.366	.172 *
1976～80年生(vs 1966～70年生)	.623	.202 **	.600	.195 **
1981～86年生(vs 1966～70年生)	.656	.212 **	.993	.222 **
父中学校卒(vs 父高校・専門学校卒)	.076	.214	-.409	.178 *
父短大・高専・大卒(vs 父高校・専門学校卒)	.082	.198	.220	.206
母中学校卒(vs 母高校・専門学校卒)	-.613	.208 **	-.478	.177 **
母短大・高専・大卒(vs 母高校・専門学校卒)	.401	.236 +	.182	.235
15歳時北海道居住(vs 関東)	-.883	.311 **	-.503	.335
15歳時東北地方居住(vs 関東)	-1.096	.246 **	-1.268	.246 **
15歳時中部地方居住(vs 関東)	-.357	.197 +	-.719	.201 **
15歳時近畿地方居住(vs 関東)	.331	.245	-.064	.241
15歳時中国地方居住(vs 関東)	-.281	.304	-.674	.278 *
15歳時四国地方居住(vs 関東)	.038	.381	-.609	.372
15歳時九州地方居住(vs 関東)	-.055	.272	-.516	.251 *
きょうだいなし(vs きょうだい1～2人)	-.160	.276	-.075	.292
きょうだい3人以上(vs きょうだい1～2人)	-1.084	.255 **	-.451	.236 +
父専門職(vs 父事務職)	-.140	.245	.024	.242
父販売職(vs 父事務職)	.002	.277	.048	.285
父自営・農業(vs 父事務職)	-.071	.238	.014	.220
父熟練職(vs 父事務職)	-.047	.264	.162	.268
父半・非熟練職(vs 父事務職)	-.257	.245	.272	.233
本の冊数(10冊単位)	.007	.006	-.008	.005
15歳時に自分の個室あり(vsなし)	.005	.173	.247	.161
15歳時に自分の勉強机あり(vsなし)	.366	.287	.565	.327 +
暮らし向き(5段階, 大きいほどよい)	.005	.096	.181	.091 *
定数項	.687	.446	.090	.470
-2 Log Likelihood	1263.735		1341.541	
N	1213		1352	
C統計量	.693		.704	
+<.10 *<.05 **<.01				

(2) マッチング前後の比較

本稿での関心は、学校外教育の変数が、進学校進学に有意な影響をもっているか否かにある。そこで従属変数を進学校進学の有無とし、学校外教育の説明変数のほかに、出生コーホート、親学歴、出身地域、きょうだい数、父職、その他 15 歳時点での生育環境に関する変数によってコントロールする。一般的には、こうした共変量統制によって、純粋な説明変数の効果が導かれると考えられてきたが、そうではないのでは、という疑念から、傾向スコア・マッチングによる共変量統制の必要が叫ばれるようになった、というのがこれまで述べてきたことである。そこで、傾向スコア・マッチングの前後で、説明変数の推定量に変化が起こるかどうかを確認することにしたい。

表 5 は傾向スコア・マッチングを行う前の、表 6 はマッチングを行った後の、ロジスティック回帰分析の係数の結果を示したものである。本稿の関心に即して、学校外教育の変

	家庭教師		通信添削		学校外教育全体		N
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	
男性	-.406	.271	.397	.219 +	.470	.206 *	1213
女性	-.186	.229	.538	.178 **	.425	.198 *	1352

+<.10 *<.05 **<.01

注)いずれも、以下の共変量を同時に投入している。

出生コーホート, 父学歴, 母学歴, 15歳時点での居住地域, きょうだい数, 父職, 15歳時の本の数, 15歳時の自分の勉強机と自分の個室の有無, 15歳時の暮らしぶり

	Caliper 0.25×SD男性		Caliper 0.1 男性		Caliper 0.25×SD女性		Caliper 0.1 女性	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
家庭教師	.140	.388	.150	.388	-.363	.298	-.363	.298
N	251		249		298		298	

通信添削	.537	.313 +	.537	.313 +	.485	.241 *	.485	.241 *
N	277		277		433		433	
学校外教育全体	.583	.236 *	.627	.240 **	.592	.241 *	.453	.251 +
N	614		602		587		585	

+<.10 *<.05 **<.01

注)いずれも、以下の共変量を同時に投入している。

出生コーホート, 父学歴, 母学歴, 15歳時点での居住地域, きょうだい数, 父職, 15歳時の本の数, 15歳時の自分の勉強机と自分の個室の有無, 15歳時の暮らしぶり

数以外の結果はここでは省略している。マッチング前をみると、通信添削と学校外教育のみが有意で、家庭教師は有意ではない（有意ではないが、符号は負となっている）。この結果は表 1 からもある程度推測がつくものである。

表 6 はマッチング後の結果である。マッチング前後で大きく変化したところはないが、有意水準を 5%とすると、女性の学校外教育については、マッチングのキャリパーの数値によっては有意ではなくなる（つまり学校外教育独自の効果はない、ということになる）。なお、マッチングは 1 対 1 で行われるから、進学校進学者数と非進学校進学者数は同じになる。したがって特に非利用者の非進学校進学者はマッチングできないケースが多数あり、それらは分析対象から除外される。つまり極端な性質のサンプルがより除外されやすくなる傾向がある。厳密に言えば、そうして得られた結果と元のマッチング前の結果で適用された標本の性質は異なるので、同列に比較するのが適切なのか、という問題も指摘できる。これについては、最後に考え方を紹介したい。

(3) 効果の異質性

ここまでの分析では、マッチングによる共変量調整を行う前後で大きく結論が変わったわけではない。したがって、マッチングを行う大した意味はない、と考えられるかもしれない。しかし傾向スコア・マッチングを用いて分析する際のメリットが、ここまでは十

表7 マッチング後の家庭教師による進学校進学率の差の検定						
男性(Caliper 0.25 × SD)	処置あり	処置なし	差	標準誤差	t値	N
サンプル全体	.148	.134	.014	.044	.359	254
第一五分位	.160	.077	.083	.091	.910	51
第二五分位	.000	.120	-.120	.065	-1.846	51
第三五分位	.231	.192	.038	.115	.333	52
第四五分位	.200	.200	.000	.115	.000	50
第五五分位	.160	.080	.080	.093	.859	50
第二分位以下	.078	.980	-.902	.057	-.346	102
第三分位以上	.197	.158	.040	.062	.633	152
男性(Caliper 0.1)	処置あり	処置なし	差	標準誤差	t値	N
サンプル全体	.151	.135	.016	.044	.359	252
第一五分位	.160	.077	.083	.091	.910	51
第二五分位	.000	.120	-.120	.065	-1.846	51
第三五分位	.200	.160	.040	.111	.361	50
第四五分位	.240	.240	.000	.123	.000	50
第五五分位	.160	.080	.080	.093	.859	50
第二分位以下	.078	.098	-.020	.057	-.346	102
第三分位以上	.200	.160	.040	.063	.634	150
女性(Caliper 0.25 × SD)	処置あり	処置なし	差	標準誤差	t値	N
サンプル全体	.296	.206	.090	.041	2.191	446
第一五分位	.200	.089	.111	.074	1.501	90
第二五分位	.273	.133	.139	.085	1.643	89
第三五分位	.289	.341	-.052	.099	-.523	89
第四五分位	.467	.114	.353	.090	3.928	89
第五五分位	.250	.356	-.106	.098	-1.078	89
第二分位以下	.236	.111	.125	.056	2.225	179
第三分位以上	.336	.271	.065	.056	1.156	267
女性(Caliper 0.1)	処置あり	処置なし	差	標準誤差	t値	N
サンプル全体	.204	.263	-.059	.049	-1.219	304
第一五分位	.161	.200	-.039	.100	-.387	61
第二五分位	.133	.290	-.157	.105	-1.500	61
第三五分位	.290	.300	-.010	.119	-.082	61
第四五分位	.300	.161	.139	.108	1.284	61
第五五分位	.133	.367	-.233	.110	-2.131	60
第二分位以下	.148	.246	-.098	.072	-1.366	122
第三分位以上	.242	.275	-.033	.065	-.506	182

分生かされているとは言えない。そこで一步進めて、傾向スコアの大きさに基づいてサンプルを5分割し、傾向スコアの大きさ、つまり学校外教育の受けやすさによって、進学結果に有意な違いが生じているかどうかを確かめる。

表7は家庭教師利用経験の有無で、進学校進学率の差を比較し、t検定した結果である。男性については基本的に5%水準で有意な差はない、ということになるが、Nが小さく、タイプIIエラーが発生しやすくなる可能性は否定できない。有意水準を10%にした場合有意となり、利用がない方が進学率は高いという結果になるが、傾向スコアの大小による明瞭な傾向があるわけではないので、その理由を述べるのは難しい。タイプIIエラーを避け検出力を高めるため、第二分位以下と第三分位以上の2分割にして検定しなおす。結局男

性については、有意な違いはない、という結論となる。女性の場合、もう少し事態は複雑になる。マッチングのキャリパーの基準によっても、若干結果に違いがある。サンプル全体でみても、キャリパーを標準偏差の0.25倍にした場合は有意で、家庭教師利用者の方が進学率は有意に高い。ところがキャリパー0.1の場合、有意ではなくなる（有意でないことに注意が必要だが、率の大小の関係も逆転し、利用していない方が進学率は高くなっている）。こうした結果の大きな揺らぎは、傾向スコアの大きな層、第四分位と第五分位の結果にもみられる。第四分位では利用者の進学率が高く、キャリパーが標準偏差の0.25倍のときのみ有意で、第五分位では逆に非利用者の進学率が高く、キャリパーが0.1のときのみ有意である。こうした入り組んだ結果のためか、第三分位以上でまとめた場合には両者

表8 マッチング後の通信添削による進学校進学率の差の検定						
男性(Caliper 0.25 × SD)	処置あり	処置なし	差	標準誤差	t値	N
サンプル全体	.275	.190	.085	.050	1.689	284
第一五分位	.207	.167	.040	.103	.390	59
第二五分位	.214	.222	-.008	.113	-.070	55
第三五分位	.250	.241	.009	.116	.074	57
第四五分位	.276	.107	.169	.104	1.623	57
第五五分位	.429	.214	.214	.124	1.732	56
第二分位以下	.211	.193	.018	.018	.231	114
第三分位以上	.318	.188	.129	.066	1.951	170
男性(Caliper 0.1)	処置あり	処置なし	差	標準誤差	t値	N
サンプル全体	.275	.190	.085	.050	1.689	284
第一五分位	.207	.167	.040	.103	.390	59
第二五分位	.214	.222	-.008	.113	-.070	55
第三五分位	.250	.241	.009	.116	.074	57
第四五分位	.276	.107	.169	.104	1.623	57
第五五分位	.429	.214	.214	.124	1.732	56
第二分位以下	.211	.193	.018	.018	.231	114
第三分位以上	.318	.188	.129	.066	1.951	170
女性(Caliper 0.25 × SD)	処置あり	処置なし	差	標準誤差	t値	N
サンプル全体	.204	.263	-.059	.049	-1.219	304
第一五分位	.161	.200	-.039	.100	-.387	61
第二五分位	.133	.290	-.157	.105	-1.500	61
第三五分位	.290	.300	-.010	.119	-.082	61
第四五分位	.300	.161	.139	.108	1.284	61
第五五分位	.133	.367	-.233	.110	-2.131	60
第二分位以下	.148	.246	-.098	.072	-1.366	122
第三分位以上	.242	.275	-.033	.065	-.506	182
女性(Caliper 0.1)	処置あり	処置なし	差	標準誤差	t値	N
サンプル全体	.204	.263	-.059	.049	-1.219	304
第一五分位	.161	.200	-.039	.100	-.387	61
第二五分位	.133	.290	-.157	.105	-1.500	61
第三五分位	.290	.300	-.010	.119	-.082	61
第四五分位	.300	.161	.139	.108	1.284	61
第五五分位	.133	.367	-.233	.110	-2.131	60
第二分位以下	.148	.246	-.098	.072	-1.366	122
第三分位以上	.242	.275	-.033	.065	-.506	182

が相殺され、有意ではなくなっている。ただし第二分位以下をまとめた場合、キャリアパーが標準偏差の0.25倍のときに有意で、家庭教師利用グループの進学率が高い。以上が導き出された結果だが、サンプルサイズの小ささや、傾向スコアの大きさによりはっきりした傾向が見出せるわけではなく、すっきりした解釈は難しい。ただ場合によっては、家庭教師を利用しない群で進学率が高くなるという、いわゆる利用によるポジティブな効果とは全く逆の結果が出ることもある、という点に注意しておきたい。

次に表8が通信添削の結果である。こちらは家庭教師よりは解釈しやすい。なお、男女ともキャリアパーの設定が標準偏差の0.25倍でも、0.1でも同じ結果になっている。男性については、基本的に有意な差はないとみてよいが、有意水準を10%にとれば、第三分位以上にまとめたグループで、10%水準で有意に通信添削利用者の進学率が高い。女性については男性ほど明瞭な傾向は見いだせず、むしろ非利用者の進学率が高いような数値が出ているが、全体的に有意な差とは言えない。ただし傾向スコアの最も大きな層（第五分位）のみ有意で、非利用者の進学率が高くなっている。通信添削は、塾や家庭教師と異なり、第三者の大人（教師）の影響が及びにくく、基本的には自分の自助努力、自学自習の姿勢があるかどうかが続の必要条件だと考えられる。つまり家庭での自学自習の態度がもともと身につけているものであれば、それなりに継続し、力を蓄えることが可能になるだろうが、自学自習の習慣がないと、実際にそれを継続すること自体が困難になるだろう。男性の場合は、そうした態度の影響が比較的出ている、もともと学習の習慣のあるような通信添削利用に向けた人ほど、成果を出しやすいものと思われる。しかし女性の場合は一貫した結果が出ていない。サンプルサイズの問題もあると思われるが、やや解釈しにくい。

最後に表9が学校外教育全体についての結果である。サンプル全体については、表6の結果を追認する結果となっている。これを傾向スコアの大きさに基づき細かくみてゆくと、男性については傾向スコアの小さなグループ（第二分位以下）で、いずれのキャリアパー・マッチングでも有意に差が出ており、学校外教育を受けていると進学校進学率が高くなる。つまり学校外教育を受けないような、年齢が上の世代で、母学歴が低く、きょうだい数が3人以上で、（関東に対して）北海道・東北・中部地方出身者といった特性をもつ人ほど、学校外教育を受けたことによる実益（進学校への進学率）が大きくなっている。学校外教育は、もともとそれを利用しない傾向のある人にとって、より大きなメリットをもたらしてきたということである。

これは女性についても同じである。ただし女性の場合、男性と異なって、学校外教育を利用する傾向が強いほど進学校進学率も高くなるというような線型的な関係になっていない。特に第二分位において極端な差が出ており、これが全体に影響を与えている可能性もある。こうした結果になるのは、女性の場合、男性ほど成績と進学校（大学進学）が直結しにくい（成績がいいから進学校に進学するという選択が必ずしも自然ではない）世代が

表9 マッチング後の学校外教育による進学校進学率の差の検定						
男性(Caliper 0.25 × SD)	処置あり	処置なし	差	標準誤差	t値	N
サンプル全体	.208	.121	.088	.030	2.955	614
第一五分位	.200	.074	.126	.060	2.093	123
第二五分位	.134	.070	.064	.054	1.182	123
第三五分位	.188	.074	.114	.062	1.833	123
第四五分位	.232	.208	.024	.076	.307	123
第五五分位	.267	.191	.075	.080	.945	122
第二分位以下	.168	.072	.096	.041	2.375	246
第三分位以上	.230	.161	.069	.042	1.662	368
男性(Caliper 0.1)	処置あり	処置なし	差	標準誤差	t値	N
サンプル全体	.209	.123	.086	.030	2.862	602
第一五分位	.179	.077	.102	.030	1.698	121
第二五分位	.125	.063	.063	.529	1.181	120
第三五分位	.190	.086	.104	.063	1.654	121
第四五分位	.241	.194	.048	.076	.631	120
第五五分位	.294	.212	.083	.081	1.020	120
第二分位以下	.152	.070	.082	.040	2.056	241
第三分位以上	.243	.163	.081	.042	1.899	361
女性(Caliper 0.25 × SD)	処置あり	処置なし	差	標準誤差	t値	N
サンプル全体	.221	.130	.091	.031	2.985	614
第一五分位	.102	.027	.075	.042	1.766	123
第二五分位	.237	.031	.206	.058	3.571	124
第三五分位	.131	.197	-.066	.067	-.974	122
第四五分位	.227	.246	-.019	.077	.237	123
第五五分位	.361	.200	.161	.083	1.933	122
第二分位以下	.176	.029	.147	.036	4.062	247
第三分位以上	.246	.214	.032	.044	.721	367
女性(Caliper 0.1)	処置あり	処置なし	差	標準誤差	t値	N
サンプル全体	.184	.131	.053	.029	1.781	610
第一五分位	.071	.030	.042	.039	1.062	123
第二五分位	.246	.016	.230	.057	4.054	121
第三五分位	.164	.180	-.016	.069	-.238	122
第四五分位	.231	.228	.003	.077	.035	122
第五五分位	.197	.232	-.035	.075	-.469	122
第二分位以下	.159	.023	.136	.035	3.888	244
第三分位以上	.198	.213	-.015	.042	-.348	366

サンプルに含まれていることが考えられる。女性の進路の制約が厳しかったため、大学進学率が高い高校より、就職率の高い職業系の高校が好まれたかもしれない。あるいは入試難易度が高くて、女子高の場合は進学先に短大が多く、「ほぼ全員が大学に進学する」という本稿の「進学校」の定義に該当しなかった高校も多数含まれている可能性も考えられる。いずれにしても、表4から、女性は学校外教育の利用の有無が、相対的に男性より家庭環境に左右されやすいといえるだろう。その中で、傾向スコアの低い層、つまり学校外教育を受けない傾向のある人ほど、実質的なメリットを受けやすいこと、逆にもともと学校外教育を受けやすい傾向のある人では、学校外教育による実質的な効果は見出しにくいことが明らかになったといえる。

5. まとめと課題

以上をまとめると、家庭教師と通信添削については、個別には有意と見なせる部分もないわけではなかったが、これらを利用することによる実質的なメリットはほとんどないといえるだろう。これらは利用率も高くなく、利用者がもともと持っている特性が大きな意味をもっていると考えられる。家庭教師は、表 1 をみても、進学校への進学意欲が高いものが利用する手段というより、親の学歴は高いが、本人の学習態度が必ずしも身につけていないような場合に利用するケースが多いのではないかと推察される。したがってここでは進学校進学という高い目標ではなく、まず学習習慣を身につけて成績を少しでも上げることが当座の目標とされる可能性が高いのではないだろうか。その結果、本稿のように、アウトプットを進学校進学とした場合には、その成果が見えにくくなるのだと思われる。一方通信添削は、既に述べたように、本人の学習習慣がないとそれ自体継続の困難な学習法だといえる。表 1 では有意な差が出ていたが、基本的にはもともと勉強ができるなどのセレクションの効果が現れていると見た方が自然だろう。

一方で塾を主体にした学校外教育については、傾向スコアの小さな群で有意に進学校進学率を高めること（メリットがあること）が明らかになった。学校外教育によってもともとあった差を縮め、進学校進学の可能性を広げるとい実質的な効果がある、と見なせるということだ。逆に言えば、もともと学校外教育を受ける可能性の高い層においては、学校外教育の有無にかかわらず、結果には大きな違いは生じない、ということの意味する。

ただしここまでの分析で、傾向スコア・マッチングの問題点も明らかになったと思われる。本稿ではキャリパー・マッチングという方法を採用したが、キャリパーの大きさによって結果に若干の違いが生じている。マッチングの方法は他にもあるため、どの方法を採用するかによって結論が変わる可能性がある（Guo and Fraser 2010: 183-184）。また層別解析についても、5 分位に分けるといことは単なる習慣の問題で、こういったわけ方に統計的な合理性があるわけではない（星野 2009）。傾向スコアの層の分割方法によっても、当然結論に違いが出る可能性がある。したがって採用方法や基準が恣意的に決まりやすく、その結果も不安定であるという問題がある。

そして傾向スコア・マッチングにより調整が可能になるのは、既に述べたように「強く無視できる割り当て条件」が成立したときであり、特に傾向スコアを推定する際のロジスティック回帰分析の共変量選択が問題になる。しかし結局共変量の調整は、存在する変数によってしか行うことはできない。例えば学校外教育の選択には、学校外教育を自分であえて選択するほど勉強熱心とか、向学心があるとか、目標がはっきりしていたなどの性格のようなものが影響していると思われるが、それらは（回顧調査でもあるので）変数に含まれていない。これらは観察できない個人の異質性とよばれるものだが、強く無視できる

割り当て条件が成立するというのは、基本的に観察できない個人の異質性が学校外教育の選択と全く関係がない、ということと同義であり、そうした前提を呑んだうえでの分析結果の解釈ということになる。しかしこうした前提が本当に正しいといえるかは議論の余地があるだろう。パネル調査であれば、固定効果推定によって観察できない個人の異質性の影響を完全に除くことが可能である。しかし傾向スコア・マッチングで補正しているのは、傾向スコアを推定する際に投入された共変量のみである、という限界があることは意識しておく必要がある。

さらに重要なのは、SUTVA (Stable Unit Treatment Value Assumption) の原則である(星野・繁樹 2004; 石田 2012 など)。これは従属変数にもたらす処置の影響は、他の変数の割り当て状況に依存しない、ということの意味する。これだけだとわかりにくいので、少し具体的に示すことにしよう。ある職業訓練プログラムがあって、それが実際に就職のチャンスを広げるかどうかを確かめたいとする。同様に傾向スコア・マッチングを行って、有意に効果があった、という結論が導けたとき、ではその職業訓練プログラムをどんどん広めるべきなのだろうか。しかしよく考えるとこの結論は無理がある。就職のチャンスは、そもそも無限にあるわけではない。職業訓練プログラムは特定の仕事の技能を身につけるものであるが、そうした特定の技能の需要は、産業構造の変化や景気変動によっても大きく左右されるものだ。また有限の就職のチャンスを上回るような数の職業訓練プログラムを受けた人々を供給すれば、結局効果があるどころか、職業訓練を受けたのに仕事につけない人を増やすことになりかねない。つまりこの傾向スコア・マッチングの結果は、分析対象とした標本に基づいて計算された結果に過ぎず、その前提を大きく変えるような社会環境の変化が起きた場合には適用できない。本稿に照らしていえば、学校外教育によってテストの点数が上昇することを効果と定義づけたとすれば、絶対的な学力が身についたことで点数が上昇するのだから、おそらく他者の動向と学習効果は無関係なので、SUTVA の原則は満たされているといえる。しかし学校外教育によって偏差値が上昇するか、という問いになると、話は別である。偏差値は集団内での位置を意味するから、絶対的な学力が上がっても、サンプル全員の絶対的な学力も上昇していれば、偏差値は上がらない(効果はない)ということになるからである。本稿の進学率も同じである。ただし現実の、特に政策的な問いを立てたときに、SUTVA の原則を満たすような場面は極めて少ないといえる。つまり現実の場面に適用しようとするときには、必ずしも SUTVA の原則を満たしていないこと、あくまでその分析結果は分析対象となっているサンプルの特性や、何を「効果」と見なしているかの定義によって結論が異なる可能性があるということ念頭に置く必要がある。

むしろ筆者が重要だと考えるのは、傾向スコア・マッチングのもつ因果効果とは何か、という考え方と、効果の異質性の検討が可能になるという点にある。効果があるといった

ときに、何が効果なのか、ということの中身は、あまり問われてこなかったように思われる。特に社会科学で既存のデータを用いる場合には、科学的な実験的方法を採用することがほぼ不可能であったため、因果効果をみるときの手続きが深く追究されてこなかった。しかし傾向スコア・マッチングの方法は、疑似的にはあるが、既にあるデータを活用して、できるだけ自然科学の実験的方法に近づけてトリートメントによる効果の有無を確認しようとするものであることが理解できる。既述したように、マッチングの欠点は、マッチングできなかったサンプルがその後の分析から除外されてしまう点にあり、それを防ぐために傾向スコアの逆数でウェイトをつける方法も提案されている。バイアスを除去するという目的は同じだが、しかしウェイトイングの方法は、傾向スコアを用いるもとの問題意識を意識させにくくしてしまう懸念がある。マッチングという方法を採用することで、実験的環境を構築しているのだということを分析者が常に意識しておくことが重要なのである。また、説明変数と統制変数の区別が曖昧なまま、思いつく有意になりそうな変数を一度に投入する回帰分析を実行し、有意になった変数を「効果」と見なすのではなく、何がトリートメントなのか、という問題意識を明確にしないと、傾向スコア・マッチングの方法は使えない。それゆえ傾向スコア・マッチングを実行しようとするとき、クリアな仮説の設定や、理論と仮説の接合について分析者は十分意識する必要に迫られるだろうが、そのことは好ましいことでもある。また現実世界では、トリートメントの効果が一律に及ぶと考える方が不自然なことが多く、その中身を詳細に検討する際には、こうした分析が威力を発揮するだろう。SUTVAの問題があったとしても、効果の及びやすい層と及びにくい層がはっきり区別できるということは、政策的、実践的にも大きな意味をもっているのではないかと考えられる。

引用文献

Bourdieu, Pierre, and Jean-Claude Passeron, 1970, *La Reproduction: Éléments pour une théorie du système d'enseignement*, Paris: Éditions de Minuit. (=1990, 宮島喬訳『再生産：教育・社会・文化』藤原書店)

Bowles, Samuel, and Herbert Gintis, 1976, *Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life*, New York: Basic Books. (=2008, 宇沢弘文訳『アメリカ資本主義と学校教育：教育改革と経済制度の矛盾』I・II, 岩波書店)

Brand, Jennie E. and Yu Xie, 2010, "Who Benefits Most from College? Evidence for Negative Selection in Heterogeneous Economic Returns to Higher Education," *American Sociological Review*, Vol.75, No.2, pp.273-302.

Collins, Randall, 1979, *The Credential Society: An Historical Sociology of Education and*

- Stratification*, Academic Press. (=1984, 大野雅敏・波平勇夫共訳『資格社会—教育と階層の歴史社会学』有信堂高文社)
- Guo, Shenyang, and Mark W. Fraser, 2010, *Propensity Score Analysis: Statistical Methods and Applications*, Sage.
- 星野崇宏・繁榊算男, 2004, 「傾向スコア解析法による因果効果の推定と調査データの調整について」『行動計量学』第31巻第1号, pp.43-61.
- 星野崇宏・岡田謙介, 2006, 「傾向スコアを用いた共変量調整による因果効果の推定と臨床医学・疫学・薬学・公衆衛生分野での応用について」『保健医療科学』第55巻第3号, pp.230-243.
- 星野崇宏, 2009, 『調査観察データの統計科学—因果推論・選択バイアス・データ融合』岩波書店。
- 石田浩, 2012, 「社会科学における因果推論の可能性」『理論と方法』Vol.27, No.1, pp.1-18.
- 川口俊明, 2006, 「『効果のある学校』を探す」『大阪大学教育学年報』11: 167-177.
- 川口俊明・前馬優策, 2007, 「学力格差を縮小する学校：『効果のある学校』の経年分析に向けて」『教育社会学研究』80: 187-205.
- Kuan, Ping-Yin, 2011, “Effects of Cram Schooling on Mathematics Performance: Evidence from Junior High Students in Taiwan,” *Comparative Education Review*, Vol.55, No.3, pp.342-368.
- Lieberson, Stanley, 1985, *Making It Count: The Improvement of Social Research and Theory*, Berkeley: University of California Press.
- 三輪哲, 2008, 「働き方とライフスタイルの変化に関する全国調査2007における標本特性と欠票についての基礎分析」『東京大学社会科学研究所パネル調査プロジェクト ディスカッションペーパーシリーズ』No.10
- Morgan, Stephen L. and Christopher Winship, 2007, *Counterfactuals and Causal Inference: Methods and Principles for Social Research*, Cambridge University Press.
- 鍋島祥郎, 2003, 『効果のある学校—学力不平等を乗り越える教育』解放出版社
- 中澤渉, 2012, 「なぜパネル・データを分析するのが必要なのか—パネル・データ分析の特性の紹介」『理論と方法』27(1): 23-40.
- , 2013, 「通塾が進路選択に及ぼす因果効果の異質性—傾向スコア・マッチングの応用」『教育社会学研究』92 (掲載決定・ページ本稿執筆時未定)
- Rosenbaum, Paul R. and Donald B. Rubin, 1984, “Reducing Bias in Observational Studies Using Subclassification on the Propensity Score,” *Journal of the American Statistical Association*, Vol.79, No.387, pp.516-24.
- Rothman, Kenneth J., 2002, *Epidemiology: An Introduction*, Oxford: Oxford University

Press. (=2004, 矢野栄二・橋本英樹監訳『ロスマンの疫学－科学的思考への誘い』
篠原出版新社)

志水宏吉, 2004, 「低学力克服への戦略－『効果のある学校』論の視点から」 苅谷剛彦・
志水宏吉編『学力の社会学－調査が示す学力の変化と学習の課題』岩波書店, 217-235。

東京大学社会科学研究所パネル調査プロジェクトについて

労働市場の構造変動、急激な少子高齢化、グローバル化の進展などにもない、日本社会における就業、結婚、家族、教育、意識、ライフスタイルのあり方は大きく変化を遂げようとしている。これからの日本社会がどのような方向に進むのかを考える上で、現在生じている変化がどのような原因によるものなのか、あるいはどこが変化してどこが変化していないのかを明確にすることはきわめて重要である。

本プロジェクトは、こうした問題をパネル調査の手法を用いることによって、実証的に解明することを研究課題とするものである。このため社会科学研究所では、若年パネル調査、壮年パネル調査、高卒パネル調査の3つのパネル調査を実施している。

本プロジェクトの推進にあたり、以下の資金提供を受けた。記して感謝したい。

文部科学省・独立行政法人日本学術振興会科学研究費補助金
基盤研究 S：2006 年度～2009 年度、2010 年度～2014 年度

厚生労働科学研究費補助金
政策科学推進研究：2004 年度～2006 年度

奨学寄付金
株式会社アウトソーシング（代表取締役社長・土井春彦、本社・静岡市）：2006 年度～2008 年度

東京大学社会科学研究所パネル調査プロジェクト ディスカッションペーパーシリーズについて

東京大学社会科学研究所パネル調査プロジェクトディスカッションペーパーシリーズは、東京大学社会科学研究所におけるパネル調査プロジェクト関連の研究成果を、速報性を重視し暫定的にまとめたものである。



東京大学社会科学研究所 パネル調査プロジェクト
<http://ssjda.iss.u-tokyo.ac.jp/panel/>