

(共同研究：経済学部独自アンケート (E-folio) の深化に向けて)

「やる気」は偏差値の差を逆転しうるか

——因果フォレストによるアプローチ——*

荒 木 英 一
井 田 憲 計
矢 根 眞 二

1 はじめに

病に臥せる人が担当医からある治療法を提案された。これを受けるか否かの決断にあたって医師から提供される参考資料は、多くの場合、RCT (Randomised Controlled Trial, ランダム化比較試験) を前提にした ATE (Average Treatment Effect, 平均処置効果) であろう。すなわち、過去にこの治療を受けた人たち (処置群) と受けなかった人たち (対照群) の予後に関する全体の (群間の) 平均差である。しかし実は、これだけでは必ずしも有用な参考資料とはいえない。なぜならば、ひとりひとりの患者は固有の健康状態にあり、ほんとうの処置効果は個別に異なって然るべきだからである。

大学におけるさまざまな教育指導についても、同様の場面を想定しうる。たとえば、勉学や就活指導のために実施された特別なイベントの効果を検証したい場合、ATE の推定だけでは不十分で、個々の学生に固有の諸条件に配慮した HTE (Heterogeneous Treatment Effect, 効果の異質性) を推定してみることが望ましい。

小論では、このような見地に立って、新入生アンケートから抽出される各学生のいわば「やる気」が、その後4年間を通じた学業成績に及ぼす影響について考察したい。すべての大学入学者は出身高校偏差値や出身高校における評定平均値といった「過去」に条件づけられ、それぞれが異なるスタートラインから大学における学習をはじめ。こうした共変量の差異を考慮したうえで、「やる気」がもたらす効果の異質性を析出してみよう。

小論では、因果フォレスト (Causal Forest)¹⁾ と呼ばれる手法を用いて CATE (Conditional Average Treatment Effect, 条件付き平均処置効果) を推定し、それをもとに効果の異質性について考察を進めたい。

* 小論は、桃山学院大学総合研究所共同研究プロジェクト (20 共 275「経済学部独自アンケート (E-folio) の深化に向けて」) の成果である。

1) Athey and Imbens [2016], Wager and Athey [2018] を嚆矢とする。

キーワード：やる気, 因果フォレスト, GPA, 偏差値, 設定平均値

まず、つづく第2節で分析手法の概略を説明し、第3節で我々の分析結果を示す。第4節にて結論を要約する。

2 分析手法について

小論では、因果木・因果フォレストにより、効果の異質性を推定する。本節ではまず、この手法の概略をまとめる²⁾。

与えられた標本のサンプルサイズを N として、 i 番目の個人が「ある処置」を選択した場合を $W_i = 1$ 、選択しなかった場合を $W_i = 0$ とし、選択した場合の潜在的な結果を $Y_i(1)$ 、選択しなかった場合の潜在的な結果を $Y_i(0)$ とする。

この個人について観測された結果変数を Y_i^{obs} とすると

$$Y_i^{obs} = W_i Y_i(1) + (1 - W_i) Y_i(0) \quad for \ i = 1, 2, \dots, N \quad (1)$$

たとえば、この個人が「処置」を選択した場合には $Y_i^{obs} = Y_i(1)$ となり、 $Y_i(0)$ は反実仮想 (counterfactual) となるから、 i 番目の個人に対するこの「処置」の効果 $Y_i(1) - Y_i(0)$ は直接には観測不能である。

そこで、 X_i を「処置」からの影響を受けない共変量 (観測値ベクトル) として、次の $\tau(x)$ を、CATE (Conditional Average Treatment Effect, 条件付き平均処置効果) として定義する。

$$\tau(x) = E[Y_i(1) - Y_i(0) | X_i = x] \quad (2)$$

因果フォレストとは、決定木 (フォレスト) によるパーティション分割に基づいて、この CATE を推定するものである³⁾。

いま、 Π をある決定木、 $\#\Pi$ を Π によって生成されるパーティションの個数、 ℓ_i を i 番目のパーティションとして

$$\Pi = \{\ell_1, \ell_2, \dots, \ell_{\#\Pi}\}$$

さらに、 $\ell(x; \Pi)$ を $x \in \ell$ となるパーティション $\ell \in \Pi$ とすると、このパーティションにおける結果変数の期待値は

$$\mu(x; \Pi) = E[Y_i | X_i \in \ell(x; \Pi)]$$

2) 以下、数式の記法は Athey and Imbens [2016] に準拠している。

3) パーティションとは決定木の各ノードに対応しており、(ラフに言うと) 各パーティション内では共変量 X の性質は類似しているはずだから、パーティション内での処置群 (「処置」を選択した者 $W_i = 1$) と対照群 (選択しなかった者 $W_i = 0$) との平均差を用いて、パーティション間の共変量の違いがもたらす効果の違いを推定しようというものである。

サンプル S が与えられたとき、これの不偏推定量は次式で与えられる。

$$\hat{\mu}(x, S; \Pi) = \frac{1}{\#\{i \in S : X_i \in \ell(x; \Pi)\}} \sum_{i \in S; X_i \in \ell(x; \Pi)} Y_i^{obs}$$

そこで

$$\mu(w, x; \Pi) = E[Y_i(W_i) | W_i = w, X_i \in \ell(x; \Pi)]$$

として、CATE を次式で定める。

$$\tau(x; \Pi) = \mu(1, x; \Pi) - \mu(0, x; \Pi)$$

サンプル S から $\mu(w, x; \Pi)$ を推定するには、 $S_w = \{i \in S | W_i = w\}$ として

$$\hat{\mu}(w, x; S, \Pi) = \frac{1}{\#\{i \in S_w : X_i \in \ell(x; \Pi)\}} \sum_{i \in S_w; X_i \in \ell(x; \Pi)} Y_i^{obs} \quad (3)$$

そして、CATE は次により算定される。

$$\hat{\tau}(x; S, \Pi) = \hat{\mu}(1, x; S, \Pi) - \hat{\mu}(0, x; S, \Pi) \quad (4)$$

以下、小論では、この(4)式によって算定される CATE にもとづいて分析を進める⁴⁾。

3 「やる気」の学業成績への影響

3.1 「やる気」の代理変数、共変量、結果変数

周知のごとく「やる気」などというものは観測不能であるが、小論では、かつて本学経済学部において実施されてきた新入生アンケート⁵⁾のなかのひとつの質問項目への自由記述回

4) なお、因果フォレストでは、通常の決定木（フォレスト）とは異なる手順と基準により最適なパーティションが求められる。まず、与えられた標本を S^{tr} , S^{est} , S^{te} （学習用データ、推定用データ、検証用データ）の3つに分割する。そして、訓練データ S^{tr} により学習された決定木 Π の MSE（平均二乗誤差）を、推定用データ S^{est} により評価する。

$$MSE(S^{te}, S^{est}, \Pi) \equiv \frac{1}{\#(S^{te})} \sum_{i \in S^{te}} \{(\tau_i - \hat{\tau}(X_i; S^{est}, \Pi))^2 - \tau_i^2\}$$

この MSE の期待値をとると

$$\begin{aligned} EMSE(\Pi) &= E_{S^{te}, S^{est}}[MSE(S^{te}, S^{est}, \Pi)] \\ &= E_{S^{te}, S^{est}} \left[\frac{1}{\#(S^{te})} \sum_{i \in S^{te}} \{(\tau_i - \hat{\tau}(X_i; S^{est}, \Pi))^2 - \tau_i^2\} \right] \\ &= -E_{(Y_i, X_i), S^{est}}[\tau(X_i; \Pi)^2] + E_{X_i}[V_{S^{est}}[\hat{\tau}(X_i; S^{est}, \Pi)]] \end{aligned}$$

右辺最後の表現で、第1項は CATE の二乗の期待値、第2項は CATE の分散の期待値となる。つまり、この基準は、CATE の絶対値が大きく、CATE の分散が小さい決定木を高く評価することになる。

5) “E-folio” と名付けられた、学部独自の新入生アンケート調査が 2012 年度から 2019 年度まで実施されてきた。

答をもとに、「やる気」の代理変数を設定する⁶⁾。

質問：経済学部でこれから学習したいことや関心のあることについて、自由に記してください。

この質問に対して、学習目標や関心事を具体的に回答した者を「やる気あり」、無回答もしくは回答不備を「やる気なし」として、これを、分析全体を通して「介入変数」 W として設定する。

$$W = \begin{cases} 1 & \text{: やる気あり} \\ 0 & \text{: やる気なし} \end{cases}$$

また「共変量」として、各学生の出身高校偏差値と評定平均値を用いる。

SV : 出身高校偏差値

HS : 出身高校における評定平均値

「結果変数」を各学生の GPA とし、1 回生終了時 GPA と 4 回生終了時 GPA の 2 種類について考えてみる。

$Y1$: 1 回生終了時の GPA

$Y4$: 4 回生終了時の GPA

サンプルとして、本学経済学部への 2012 年度入学生から 2015 年度入学生までの 1270 名に関するデータを用いた⁷⁾。図 1 は、横軸に偏差値をとった全サンプルのヒストグラムである。黒い部分が $W = 1$ （「やる気あり」の学生たち）、白い部分が $W = 0$ （「やる気なし」の学生たち）をあらわしている。

3.2 推定された因果木 (Causal Tree)

図 2 は、結果変数を $Y1$ (1 回生終了時の GPA) として推定された因果木である⁸⁾。各パーティション (ノード) に記された 2 つの数値のうち、上の数値が CATE (前節の(4)式によ

6) たとえば、中村 [2021] のように、学生に簡単な作文をさせて動詞の数を数えてみると、動詞の個数の減り具合と「やる気」の減退とが同期するといった興味深い分析もある。が、小論では単純な代理変数を用いることにしたい。逆に、こうした単純な代理変数でも有効となりうることを示す意図もある。

7) 休学や中退除籍等により、有効なサンプルサイズは入学者総数をいくぶん下回っている。なお、このアンケートは経済学部 1 回生春学期の少人数必修科目「入門演習」のなかで実施されたもので、例年の回答率は 90% を超えていた。無回答のなかにはこの必修科目への欠席者も含まれるが、大半の無回答はこの少人数必修科目の授業中での白紙回答である。

8) 因果木の推定には、Athey, Imbens, Kong and Ramachandra [2016] によるパッケージ causal Tree を用いた。

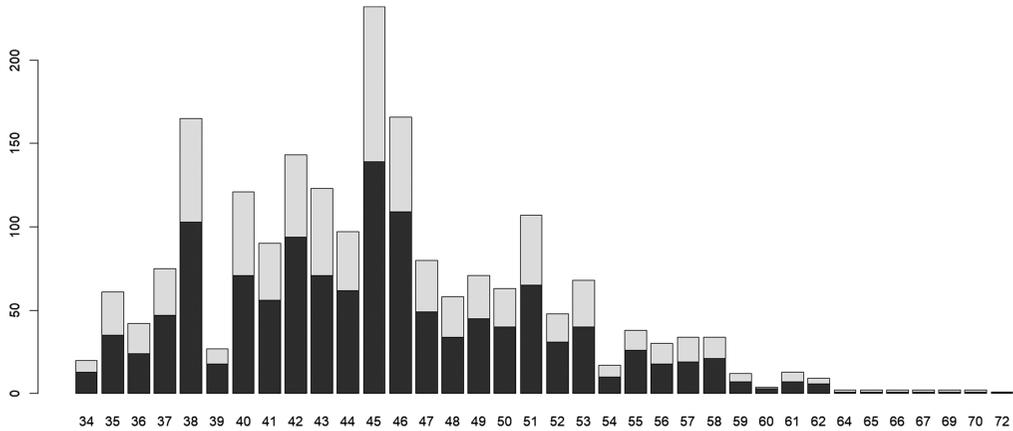


図1 偏差値別のサンプルの分布

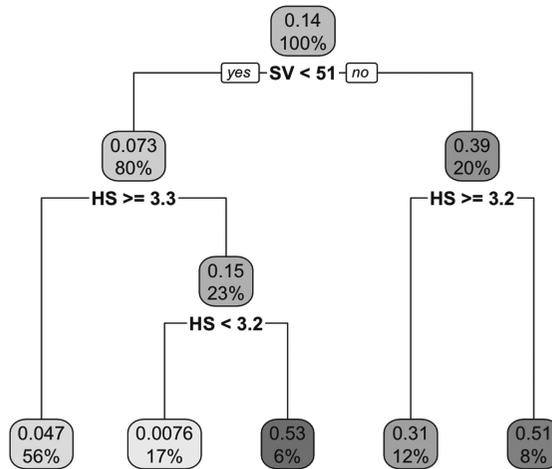


図2 1回生終了時 GPA (Y1) への CATE 推定値 (N = 1270)

り算定されたもの)を、下の数値はそのパーティションに分類されるサンプルの個数 (%)をあらわしている。

たとえば、最下部右端のパーティションには全体 (1270 名) のうち 8% のサンプルが含まれるが、これらのサンプルは、 $51 \leq SV \rightarrow HS < 3.2$ という 2 個の条件を満たしている⁹⁾。

また、このパーティションでは $CATE = 0.51$ である。つまり、このパーティションに含まれるサンプルについては、 $W = 1$ (「やる気あり」) の GPA は、 $W = 0$ (「やる気なし」) の GPA を 0.51 上回るといふ潜在的な因果効果をあらわしている。

9) 木の頂点から出発して、まず $SV < 51$ という条件に “yes” なら左へ枝分かれし “no” なら右へ枝分かれする。次に $3.2 \leq HS$ という条件に “yes” なら左へ枝分かれし “no” なら右へ枝分かれする。頂点 → 右 → 右と枝分かれすると、最下部右端のパーティションに至る。

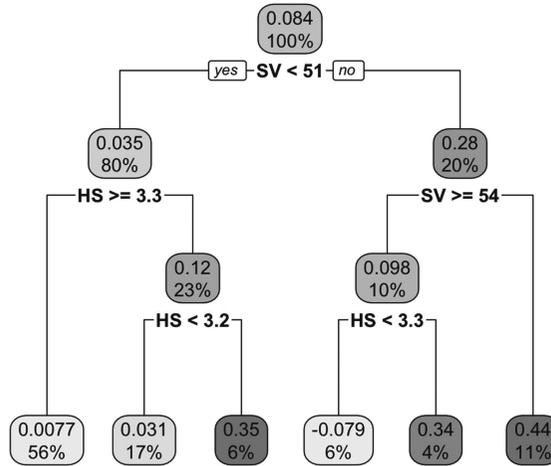


図3 4回生終了時 GPA (Y4) への CATE 推定値 (N = 1270)

また、図3は、結果変数を Y4 (4回生終了時の GPA) として推定された因果木であり、同様の読み方ができる。

図4は、2つの因果木(図2と図3)それぞれについて、パーティション分割と各パーティションにおける CATE の推定値を、より見やすくまとめたものである。

一箇所をのぞいて、いずれのパーティションにおいても CATE はプラスと推定されており、ほとんどの学生にとって、「やる気」を持つことは GPA の増加につながっていると言える¹⁰⁾。

なお、頂点のパーティション(ルートノード)には全サンプル(100%)が含まれるので、この CATE 推定値(Y1 については 0.14, Y4 については 0.084)は、ATE に他ならない。

3.3 「やる気」は偏差値の差を逆転しうるか

上では、「やる気」が GPA 値を引き上げる効果が確認できた。それでは、この GPA の増加によって、各学生のランキング(記録上の成績順位)¹¹⁾はどう変化するのだろうか。これ

10) 一箇所とは、偏差値が 54 以上かつ評定平均値が 3.3 未満のパーティションの Y4 への CATE である。(ラフに言うと)偏差値が比較的に高い高校の出身者で、かつ、高校時代の評定平均値がさほど芳しくないタイプの学生たちのパーティションであるが、たとえば、入学時には「やる気」を有していたものの、4年のうちにその意欲が薄れて消えてしまったという解釈も可能かもしれない。また、全サンプルの半数以上(56%)は、偏差値が 51 未満かつ評定平均値が 3.3 以上のパーティションに分類されているが、このパーティションにおける CATE の値は非常に小さい。これは、56%の学生にとっては「やる気」を持つことは(残念ながら)GPA の向上にさほど大きくは寄与しないことを意味している。

11) 順位は、2012年から2015年までの入学生のデータ(有効サンプル1270名)をすべてまとめて、1番から1270番までの降順の順位付けを施した。同点は同順位としている(たとえば[8, 4, 4, 1, 9]の5個のサンプルならば、順位は[4, 2, 2, 1, 5])。

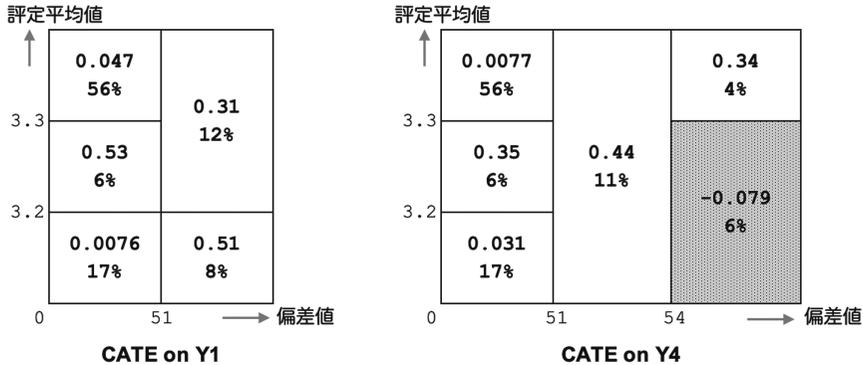


図4 因果木によるパーティション分割と、Y1, Y4 への CATE 推定値

を見るために、以下のような指標を結果変数に設定して、ふたたび、CATE を推定してみよう。

$R1 = rank(SV) - rank(Y1)$: 順位の変動 (偏差値による順位 - 1 回生終了時 GPA による順位)

$R4 = rank(SV) - rank(Y4)$: 順位の変動 (偏差値による順位 - 4 回生終了時 GPA による順位)

図5は、横軸に偏差値 SV をとり、縦軸に $R4$ をとって、有効サンプル 1270 個の点をプロットしたものである。図の左 (低偏差値) から右 (高偏差値) にかけて右下がりの帯状となることが確認できる。図中の太い実線は「やる気あり」の学生 ($W = 1$) だけを対象にした平滑化曲線、太い点線は「やる気なし」の学生 ($W = 0$) だけを対象にした平滑化曲線である¹²⁾。ほとんどの偏差値水準において、実線の位置が点線の位置を上回っていることが見てとれる。これは、「やる気あり」の学生たちのほうが順位の上昇が大きかったことを含意している。具体的にどの程度の上昇が見られたのかを因果フォレストにより確認してみよう。

図6に、1 回生終了時での順位変動 ($R1$) と 4 回生終了時での順位変動 ($R4$) のそれぞれを結果変数とした因果木の推定結果をまとめて示す。

$R1$ (左表) と $R4$ (右表) のいずれについても、偏差値 (SV) 51 を境にして、 $51 \leq SV$ の場合には「やる気」の潜在因果効果は、評定平均値 (HS) に関わりなく、プラスとなる。 $SV < 51$ の場合には煩雑なパーティション分割となるが、3つのパーティションをのぞいて¹³⁾、「やる気」の潜在因果効果はプラスとなっている。

「やる気」による偏差値差の逆転は可能であるといえるだろう。

12) 帯域幅 2.0 のガウスクアーネルを用いた。なお、図5では、「やる気あり」の学生 ($W = 1$) の点は◇で、「やる気なし」の学生 ($W = 0$) の点は+で表示している。

13) $SV < 51$ かつ $HS < 3.2$ のパーティションは偏差値・評定平均値ともに最も低いグループであり、「やる気」が空回りしているという解釈もできるのかもしれない。

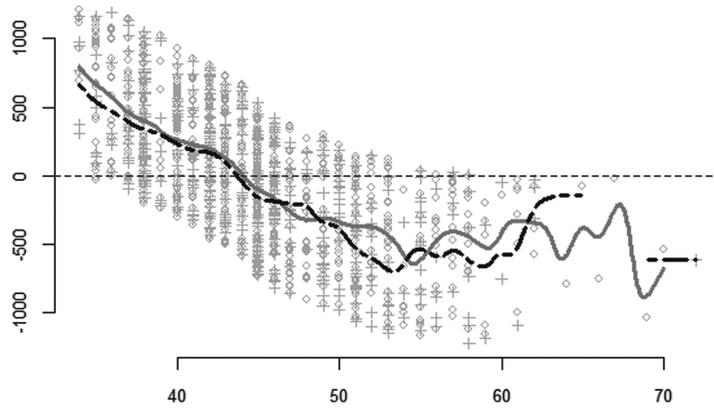


図5 R4：偏差値による順位 - 4 回生終了時 GPA による順位

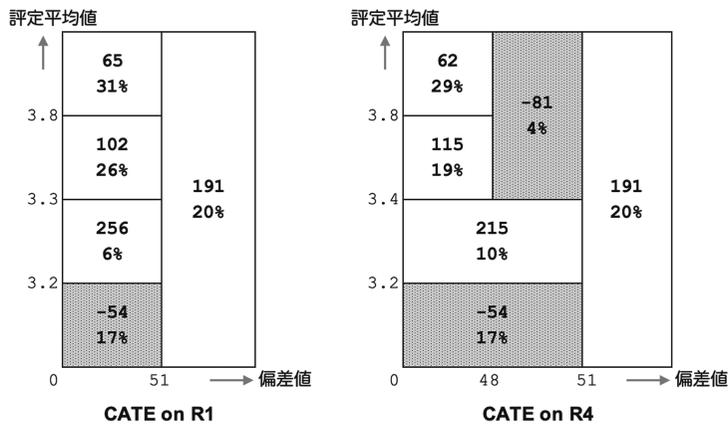


図6 因果木によるパーティション分割と, R1, R4 への CATE 推定値

4 結びにかえて

小論では、経済学部新入生アンケート (E-folio) の回答結果から「やる気」の代理変数を抽出し、因果フォレストによって、「やる気」の CATE (条件付き平均処置効果) を析出した。代理変数設定の妥当性や因果木推定プロセスの精緻化といった課題は残るものの、「やる気」は学業成績 (GPA) を向上させ、入学前の偏差値順位を逆転させうることを示し得たと考える。

HTE (効果の異質性) に配慮した因果推論の手法は、教学データの分析にとっても有用である。いわゆる「教学 IR」、教学施策の効果検証作業等において、効果の異質性はもっと注目されるべきであろう。

なお、言うまでもないが、小論の立場は偏差値至上主義をとるものではない。GPA 至上

主義でもないが、これらが無価値とするものでももちろんなく、これらを「人材価値」(山崎 [2016]) を構成する基本要素のひとつとする立場に近い。これが社会や企業にどう評価されているのか、たとえば大学における学業成績(とくに中堅私立大学文系学部における学業成績)と就職活動のいわば成功度との関連等を調査・分析する作業を、本プロジェクトにおける次の課題にしたいと考えている。

参考文献

- 荒木英一「成績・卒業を左右する要因について」 桃山学院大学経済経営論集 62 巻 3 号 pp.115–127 2021 年 1 月
- 荒木英一・井田憲計・矢根眞二「経済学部統一テストの分析—3 年間の軌跡から見えてくるもの—」 桃山学院大学総合研究所紀要 48 巻 3 号 pp.13–28 2023 年 3 月
- 石原卓典・依田高典「因果性と異質性の経済学②: Causal Forest」 京都大学大学院経済学研究科ディスカッションペーパーシリーズ J-20-004 2020 年 8 月
- 依田高典『データサイエンスの経済学』 岩波書店 2023 年 10 月
- 中村勝之『学生の「やる気」の見分け方 経済学者が教える教育論』 幻冬舎 2021 年 9 月
- 中村知繁『ランダムフォレストによる因果推論と最近の展開』 フォレストワークショップ 2023 (招待講演) 2023 年 2 月 https://speakerdeck.com/tomoshige_n (2024 年 3 月 11 日閲覧)
- 山田育矢ほか『大規模言語モデル入門』 技術評論社 2023 年 7 月
- 山崎元『偏差値「10」の差を逆転する』 星海社 2016 年 12 月
- Susan Athey and Guido Imbens. “Recursive Partitioning for Heterogeneous Causal Effects.” *Proceedings of the National Academy of Sciences* vol.113 no.27 pp.7353–7360 July 2016
- Stefan Wager and Susan Athey. “Estimation and Inference of Heterogeneous Treatment Effects using Random Forests.” *Journal of the American Statistical Society* vol.113 no.523 pp.1228–1242 June 2018
- Léo Grinsztajn, Edouard Oyallon and Gaël Varoquaux. “Why do tree-based models still outperform deep learning on tabular data?” 36th Conference on Neural Information Processing Systems (NeurIPS 2022), *arXiv preprint, arXiv: 2207. 08815*
- Susan Athey, Guido Imbens, Yanyang Kong and Vikas Ramachandra. “An Introduction to Recursive Partitioning for Heterogeneous Causal Effects Estimation Using causalTree package.” September 2016, <https://github.com/susanathey/causalTree/blob/master/doc/briefintro.pdf> (accessed 2024–03–11)

(2024 年 3 月 22 日受理)

Can ‘Motivation’ Reverse Academic Rankings?

ARAKI Eiichi

IDA Norikazu

YANE Shinji

In this paper, we constructed a proxy variable for student ‘motivation’ based on the results of a questionnaire for freshmen at our Faculty of Economics, and estimated its conditional average treatment effect (CATE) using an analytical method called Causal Forest. Although issues such as examining the validity of the proxy variable and refining the causal tree estimation process remain, we believe that we have shown that ‘motivation’ can improve academic performance (GPA) and reverse students’ pre-enrolment rankings of standard scores.

Causal forest and other causal inference methods that take account of heterogeneity of effects are valuable and crucial for the analysis of teaching and learning data. Heterogeneity of effects should receive more attention when examining the effectiveness of measures for educational guidance or job search support.