

外部資金獲得の効果は？

— 母集団情報を反映した因果効果の試行的分析 —

中 尾 走
(2019年10月3日受理)

Effect of Competitive Fundraising
— The population average treatment effect from
a covariate balancing propensity score and sampling weight —

Ran Nakao

Abstract: The effect of competitive research funding mechanisms on research productivity in various disciplines was clarified. Previous studies used multivariate analyses, such as multiple regression analysis, and we built on the existing research by controlling for confounding effects using a balancing score. Typically, a propensity score is used for balancing scores, but this approach has received criticism about its external validity. Therefore, the matching sample in this study was visualized using a propensity score, which was compared with the population and the sampling data for the distribution of disciplines. The main findings of this study were as follow: first, it became clear that the matching sample analyzed using a propensity score strikingly differed from the population. Second, this problem was solved using a covariate balancing propensity score and sampling weights, which estimated the population's average treatment effects.

Key words: EBPM, Academic Profession, Propensity Score

キーワード：EBPM, 大学教員, 傾向スコア分析

1. はじめに

本稿の目的は、観察データにおける因果推論の手法を応用して、母集団に対する大学教員の外部資金獲得の効果を明らかにすることである。

ここで因果推論の手法として着目するのは傾向スコア分析 (Propensity Score, 以下 PS) である。昨今、根拠に基づく政策立案 (Evidence-based policy making, 以下 EBPM) の必要性が喧伝され、政策立

案におけるエビデンスの質は手法に基づいて序列がつけられた (図1参照)。このような流れが教育政策に大きな影響を及ぼしており、エビデンスの必要性和同時に RCT とメタアナリシスを頂点としたエビデンスの階層ピラミッドが前面に出されている。このような手法の序列の中で、図1に示した通り、社会科学系でよく用いられてきた回帰分析などはエビデンスの質が低い手法として位置づけられている。PS は、観察データにおける因果推論では、より質の高いエビデンスをもたらすとして上位に位置づいている (内閣府2018)。このような影響もあってか、近年では PS を用いた論文も増えてきた (呉・島・西村2019)。中村 (2019) は、このような状況を「猫も杓子も傾向スコア」と揶揄している。

本論文は、課程博士候補論文を構成する論文の一部として、以下の審査委員により審査を受けた。

審査委員：渡邊 聡 (主任指導教員), 藤村正司,
黄 福涛, 村澤昌崇



図1 EBPMで推進される手法の序列

本稿では、このように喧伝されるEBPMの前提として示されている、エビデンスとしての信頼性の高低を判断する分析手法の階層ピラミッドについて、高等教育を素材に応用可能性および信頼性・妥当性を検証する。ここで高等教育を分析対象として選ぶのは、昨今の高等教育政策が「根拠なき大学改革」(羽田2018)として指摘されるとともに、高等教育という対象が政策からの影響を受けやすく(吉田2013)、EBPMによる直接的な影響を大きく受ける可能性があるからである。実際に、過去の教育政策において「エビデンス」という言葉がどのような意味で使われているかをレビューした小野(2015)は、大学改革実行プランで「エビデンス」という言葉が、達成すべき条件のような形で用いられていることを指摘する。また、中教審義務教育特設部会において用いられている「エビデンス」は、政策の検証という意味で用いられており、大学改革実行プランでの「エビデンス」の用法と異なっている。このような「エビデンス」の用いられ方の違いから、文科省が大学に対して介入しやすい対象であることが推察される。

このように、政策レベルで推奨される因果推論ではあるが、教育領域で取得されることの多い観察データにおける因果推論の困難な理由の一つが、交絡要因の統制である。そのため、本稿では交絡の統制が困難な対象の一つとして大学教員の外部資金獲得の効果を上げる。近年、効果的・効率的な資金の活用として、基盤的な研究費配分から競争性の高い研究費配分へと変更されてきた(内閣府2016)。このような変化の中で、大学教員は外部資金を獲得しなければ、研究継続が困難な状況に陥っていることが先行研究より指摘されている(藤村2017)。加えて、これまでの先行研究では競争性の高い外部資金の獲得が研究成果を高めるといった関係には、疑問が残されている(藤村2017, 2018, 浦田2017)。しかし、これらの先行研究では交絡の統制に必ずしも十分な配慮があったわけではない。同時に、図1に示す分析手法の序列における、より上位の手法を使用したものは管見の限りない。

そこで本稿では、PSを用いて外部資金獲得の効果の再分析を行う。さらに、PSの問題点として指摘される架空集団への因果効果であるという批判に対して、CBPSとrakeによるサンプリングウェイトを用いて母集団に対する因果効果の推定の試行的分析を行う。

2. データと方法

2.1. データと変数

本稿では、「大学教授職の変容に関する国際調査」の二次データ(以下、AP2017)を用いた。調査期間は2017年10月からであり、大学教員8807人に対して質問紙による郵送調査を行い、2018年1月末までに回収を行った。回収率は2128人(24.2%)であった。

分析に用いた変数の全ては、表1にまとめた通りである¹⁾。今回の分析手法で重要な処置変数である外部資金獲得の有無は、過去3年間の研究費の総額に占める割合を①所属大学、②公的研究支援機関、③政府機関、④企業、⑤私的・非営利機関、⑥外国政府機関、⑦その他の7つに分けて聞いている。その中で、①所属大学以外を全て外部資金として、外部資金獲得の有無というダミー変数を作成した。処置変数に基づき、処置を受けた群(T=1)を処置群、処置を受けなかった群(T=0)を対照群とする。また、研究成果としての結果変数には、村澤(2008)、村澤・立石(2018)、中尾・村澤(2019)でも用いられている指標である①論文数(査読なし)、②査読論文数を、加えて昨今の大学ランキングで研究力の指標としても用いられることのある③SCI・SSCI論文数を設定した。分析結果については、平均処置効果(Average treatment effect, 以下 ATE)を比較した。

表1 基礎統計量

	n	mean	s.d.
外部資金の獲得	1990	0.727	0.446
性別	1990	1.18	0.388
年齢	1990	49.2	10.5
設置者	1990	1.98	0.943
職階	1990	2.16	1.27
専門分野	1990	3.45	1.39
博士取得	1990	1.81	0.39
論文数	1710	9	23.3
査読論文数	1815	7.23	18
SCI・SSCI論文数	1735	2.81	9.33
所属大学からの研究	1990	177	345

ることは因果効果の推定においてバイアスが残されたままとする。

3. 分析

3.1. PS 分析

まず、処置群と対照群で統制をしない場合の2群の比較を行なった。その結果を表2に示す。この結果に基づく、外部資金を獲得すれば論文数、査読論文数、SCI・SSCI論文数はそれぞれ7本、6本、2～3本増えることになる。しかし、この結果は、純粋な外部資金獲得の効果ではない。なぜなら、外部資金を獲得しやすい専門分野や大学の設置者などの交絡要因が結果変数にも影響を与えており、元々外部資金を獲得しやすい大学教員が外部資金を獲得したことによる効果がこの2群の差には含まれている可能性があるからである⁴⁾。そのため、観察データによる因果推論では前述の通りPSを用いた共変量の調整を行うことで、擬似的なランダム化を行い、外部資金を獲得したことの効果をみる必要がある。

表2 処置群と対照群の基礎統計量と検定の結果

	Group	n	Mean	Median	Mean difference	Mann-Whitney U
論文数	1	1395	10.770	6	7.250	<.001***
	0	417	3.520	2		
査読論文数	1	1475	8.710	5	6.080	<.001***
	0	450	2.633	1		
SCI論文数	1	1418	3.490	0	2.820	<.001***
	0	424	0.665	0		

p.value: .000***, .001***, .05**

表3、4には、PS分析の分析結果を示した。PS分析は、二段階の統計的な操作を行う。まず、一段階目に処置変数を従属変数としたモデルを構築し、その分析結果からPSが同じような個体間で結果変数を比較することで、因果効果を推定している。

表3 PS推定のロジスティック回帰分析

	Coef	se	p.value
切片	2.198	0.337	***
個人属性			
女性 (v.s.男性)	-0.053	0.143	***
年齢(基準: 25歳)	-0.057	0.008	***
准教授 (v.s.教授)	-0.502	0.173	***
講師 (v.s.教授)	-1.285	0.234	***
助教	-1.129	0.220	***
助手	-1.782	0.329	***
専門分野			
社会科学 (v.s.人文科学)	0.048	0.195	***
理学 (v.s.人文科学)	0.341	0.225	***
工・農学 (v.s.人文科学)	0.784	0.211	***
医・歯・薬学 (v.s.人文科学)	0.143	0.205	***
その他 (v.s.人文科学)	0.229	0.288	***
博士号取得 (v.s.未取得)	1.036	0.140	***
設置者			
私立 (v.s.国立)	-0.530	0.193	***
公立 (v.s.国立)	-1.043	0.129	***
研究費(単位: 100万円)	0.002	0.000	***
AIC		1970.647	
BIC		2060.181	
Log Likelihood		-969.323	
Deviance		1938.647	
Num. obs.		1990	

*** p < 0.001, ** p < 0.01, * p < 0.05

この一段階目の分析が表3、二段階目の分析が表4である。一段階目のPSを推定する分析では、交絡変数として考えられる7つの変数を投入し、ロジスティック回帰分析を行なった。この分析結果から、共変量が二つの群で適切に割り当てられているかの一つの基準であるC統計量という数値を算出した。C統計量が0.7805(信頼区間: 0.7581-0.803)であり、基準よりも悪い値ではなかったため⁵⁾、本稿では2つの群で共変量が調整されたと判断した。その後、PSに基づいて擬似的なランダム化集団を作るがその手法についてはいくつかの方法が紹介されている。その中で、得られたデータの情報を削ることなく分析可能なIPW推定を用いた(星野2009)。

表4 外部資金獲得の効果 (PS)

	E(Y ₁)	E(Y ₀)	E(Y ₁ -Y ₀)
論文数	8.632	3.609	5.023
査読論文数	7.311	3.130	4.182
SCI・SSCI論文数	2.708	1.099	1.609

3.2. 擬似的なランダム化集団とは

このようにPS分析は交絡要因を統制することで、処置変数の純粋な効果を導き出せる一方、共変量を調整した擬似的なランダム化を施した集団の中身についてはほとんど議論されることがない。そこで、本稿では擬似的なランダム化集団がどのような集団になっているのかを明らかにするためにIPW推定後の集団を専門分野別に集計した。その結果が図3の通りである。

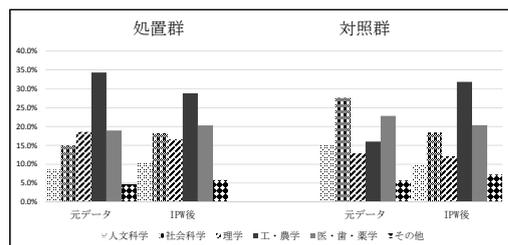


図3 元データとIPW後の集団の比較

処置群と対照群で分け、それら群ごとの専門分野別の割合を計算し、元データと擬似的なランダム集団で比較した。図3を見ていくと、処置群においては理学、工・農学の割合が少し減少しているのに対し、社会科学、医・歯・薬学の割合が微増していることが分かる。大幅に変化している専門分野は処置群においては存在しないと見てよいだろう。対照群においては工・農

学の割合が顕著に増加しており、2倍ぐらいの割合になっている。つまり、工・農学は対照群の期待値（表4のE(Y0)）で元データの期待値よりも2倍の影響力になることが分かる。さらに大学教員の母集団分布は学校教員統計調査より取得可能であるため、その専門分野別の分布と比較したものが図4である。縦軸は、それぞれの集団における専門分野の割合を計算したものをパーセントで表示している。

PSを用いたIPW後の擬似的なランダム化集団においては、中村・南(2017)の指摘通り、稀な事象に過剰に重みがつく(King and Nielsen 2018, Imai and Ratkovic 2012)。

これらの先行研究の指摘を踏まえると、対照群において工・農学の割合が顕著に増加しているのは、外部資金を獲得していない工・農学の教員が今回用いたデータ上は稀であり、そのために過大推定された可能性が考えられる。実際に、元データでは外部資金を獲得していない工・農学の教員は610人中94人(15.4%)と他の専門分野と比べても少ない。

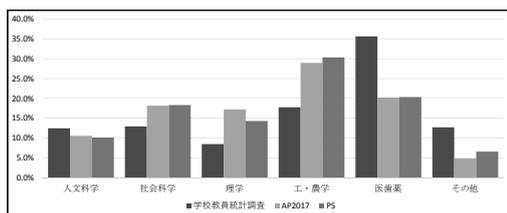


図4 母集団分布と元データとIPW後の比較

3.3. 母集団情報を反映したCBPSの応用

本節では、CBPSを用いてIPW推定を行なった場合とCBPSとサンプリングウェイトを用いてIPW推定を行なった場合の両者の分析結果を提示する。サンプリングウェイトは、平成28年度学校教員統計調査をもとに専門分野別の割合を算出し、rakeによる重み付けを行なった⁶⁾。学校教員統計調査自体は、三年に一度しか発刊されていないため、AP2017の調査時期と最も近いものとして平成28年度の学校教員統計調査を選択した。CBPSのみでIPW推定を行なった後の擬似的なランダム化集団とCBPSと母集団情報を反映させた擬似的なランダム化集団の両方を専門分野別にまとめたものが図5、6である。

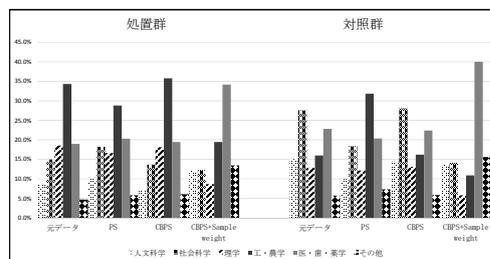


図5 元データと擬似的なランダム化集団の比較

図5は、処置群と対照群それぞれにおける各専門分野の割合を、図6は集団全体に占めるそれぞれの専門分野の割合を示している。

図5をみると、CBPSのみの結果は、処置群ではPSで共変量を調整した場合と大きな変化はない。しかし、対照群においては大きく異なる。PSで調整した場合においては、工・農学が約2倍に増加していたのに対し、CBPSでは元データの割合とほとんど変わらない。つまり、PSで共変量を調整した場合に比べて、対照群においては元データにより近い状態で共変量を調整したと言える。

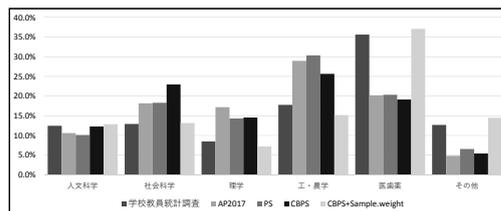


図6 母集団と擬似的なランダム化集団の比較

図6は、学校教員統計調査とそれぞれのバランススコアを用いた調整後の専門分野別の比較である。この結果をみると、PSとCBPSの両方のバランススコアを用いた場合においても元データの医歯薬が専門の教員の回答率が低く、偏ったデータであるためその影響を受けていることが分かる。加えて、社会科学の教員は回答率が高く、PSとCBPSによる共変量を調整後は、母集団の割合からより乖離している。しかし、CBPSとサンプリングウェイトを用いた場合は、学校教員統計調査の分布により近い分布となっていることが分かる。特に、PSとCBPSのみで乖離の大きかった医歯薬とその他の専門分野において学校教員統計調査の割合とほとんど変化のない割合になっている。これらの分析結果からCBPSとサンプリングウェイトを同時に用いることで、母集団と類似のランダム化集団を作り上げることが可能となることが分かった。

た。しかし、母集団情報を反映させることで、元々のバランススコアを用いる利点であった処置群と対照群の共変量の調整が疎かになっている可能性も否定できない。そのため、PSと同じようにC統計量を用いて処置群と対照群の共変量が調整されているかを確認した。CBPSのみを用いた調整では、C統計量が0.9375（信頼区間：0.9275-0.9475）であり、CBPSとサンプリングウェイトを同時に用いた場合では、0.8222（信頼区間：0.8037-0.8408）であった。母集団情報を反映させることによって、CBPSのみで共変量を調整するよりも、C統計量の値は下がるが、C統計量は0.7~0.9の範囲であれば2つの群間で共変量が調整されていることを示唆しており⁷⁾、PSを用いた共変量調整よりは、C統計量の値が上回っていることも加味すると、共変量の調整がうまくいっていると判断してもよいだろう。

最後に、本節で2通りの共変量調整を行なった擬似的なランダム化集団におけるATEの分析結果を表5に示す。表5と表4を比較すると、CBPSのみとCBPSとサンプリングウェイトの場合の両方で、対照群の期待値が小さくなっていることが分かる。これは、比較的論文の数が多くなる傾向のある工・農学の教員をPSでは2倍程度過大に推定していたのに対し、その点が改善されたのではないだろうか。もちろん、その影響のみによって本当に対照群の期待値が小さくなったのかについては厳密には不明であり、他の要因も考えられる。

表5 外部資金獲得の効果

CBPS	E(Y ₁)	E(Y ₀)	E(Y ₁ -Y ₀)
論文数	8.71	1.926	6.784
査読論文数	7.377	1.421	5.956
SCI・SSCI論文数	2.732	0.312	2.421
CBPS + sample weight			
論文数	11.071	2.045	9.026
査読論文数	9.552	1.450	8.102
SCI・SSCI論文数	3.621	0.319	3.302

4. まとめと考察

4.1. 知見の整理と含意

本稿では、大学教員の外部資金獲得の効果を高等教育分野ではあまり用いられてこなかったPSを用いて明らかにした(3-1)。次に、PSを用いた因果推論に対する「あくまでも架空集団に対する因果効果である」という批判に対して、架空集団の構成をみるために、専門分野に着目してその分布を元データと母集団と比較した(3-2)。最後に、上記の批判に対する改善策と

してCBPSとrakeによるサンプリングウェイトを用いることで、母集団の専門分野の分布を反映させた架空集団を作り上げた(3-3)。図1に即して、PSを用いてエビデンスを作り上げ、そのエビデンスに基づいて政策立案を行なったとしても、そのエビデンスが何に対する効果であるのか不明であるという時点で、想定する政策の効果が得られるとは限らない。3-1で分析した通り、大学教員が外部資金を獲得したらSCI・SSCI論文数が2本弱増加するかもしれない。しかし、この因果効果は図4に示す大学教員の専門分野構成になっていなければ期待通りの成果は得られないこととなる。処置群と対照群の共変量を調整することで、擬似的なランダム化集団を作り上げたとしても、母集団として想定される日本の大学教員の専門分野の構成と大きく乖離していることが明らかになった。これはつまり、正しい因果効果が推定できたとしてもその因果効果が対象から期待通り得られるとは限らず、安易な適用には慎重を要するべきであろう。近年は大学改革が連続し、根拠なき高等教育政策とも表現されるが、図1に基づく手法でエビデンスを作りあげたとしても政策が良くなるかの保証はなく、そもそもエビデンスがないことが政策の失敗に繋がっているかどうか不明である。さらに、過度なエビデンス重視には、再現性など科学の脆弱性を突かれることで科学の信頼性を損なうことにもなりかねない(小林2018)。

最後に、今回の分析結果から、大学教員の研究活動にとって外部資金の有効性が証明されたかのような分析結果となっているが、その解釈には注意も必要である。

確かに、一貫して外部資金を獲得している教員の方が論文数は多い。これは昨今、基盤的研究費が少なくなっている中で外部資金を獲得しなければ論文が書けない状況を反映しているのであろう。しかし、これは必ずしも外部資金の有効性を示すものではない。それは、外部資金の有効性とは、比較の視点が入ることで初めて検証可能となるからである。図1で推奨されている手法は、ある処置の効果を明らかにすることに特化した手法であり、それらの処置の効果を比較する手法ではない。なぜなら、それぞれの処置によって異なるランダム化集団を作り上げてしまうからである。異なる集団への処置の効果の比較になるため、優劣をつけることが不可能である。この点についても、母集団情報を反映させることで、ある程度類似の集団に対する効果の比較が可能となり、処置効果の比較といった視点からみても、母集団情報を反映させた本稿の試みは、一つの代替策となり得るのではないだろうか。

4.2. 今後の課題

本稿の課題は以下の2点である。1点目が処置変数を2値としたことによる情報の損失である。PSなどのバランシングスコアを用いた因果推論では、ほとんどの場合で2値変数を処置変数として分析する(Fong, Hazlett and Imai 2018)。しかし、処置変数は2値であるとは限らず、外部資金獲得額などの連続量の処置も考えられる。そのような場合に対して有効な一般化傾向スコア(Generalized Propensity Score, Hirano and Imbens 2004, 星野2009)を用いた分析は今後の課題である。2点目が、一つの変数でしか母集団情報の反映が出来ていない点である。中尾・村澤(2018)では、専門分野以外にも性別、学位、職階、設置者、旧帝大ダミーなど多様な母集団情報を反映させている。しかし、本稿の分析では分析ソフトの限界上、専門分野に絞った母集団情報の反映となっている。そのため、専門分野の変数においては母集団情報を反映できているが、その他の変数は考慮されていない。表4と表5の因果効果の違いについて、母集団情報に基づいた専門分野の調整といった観点から考察を行なったが、厳密には、その真の影響については不明である。

昨今のEBPMブームでは、PSを含め、この手法を使えば頑健なエビデンスであるかのような主張が図1のようなピラミッドなどに暗に含まれている。しかし、本稿の結果は、PSなどのバランシングスコアを用いたとしても、バランシングスコアごとに結果が大きく異なることが示された。これは、PSを使えば頑健なエビデンスであるというような思考停止状態に陥っている、明らかにならなかった事実であろう。同時に、頑健なエビデンスとして推奨されている手法も検討の余地があることを指摘した。母集団情報を反映できたからといって本稿の結果が真の因果効果であるとは限らない。この点について、真の因果効果が既知であるデータを用いたシミュレーションでの検証を要するが、別稿に譲ることとした。

【注】

- 1) いずれの分析にもR ver.3.5.1を用いた。3-1では、ライブラリ Matching を3-2では、ライブラリ plotrix を、3-3ではライブラリ CBPS を用いた。
- 2) GMM 推定は、一般化積率法もしくは一般化モーメント法と訳される。一般的なPSは、 β の推定の際に、ロジスティック回帰分析を用いると最尤法を用いる。しかし、(2)式よりパラメータの次元よりも方程式の数が多くなる可能性があり、CBPSではGMM推定が用いられる(Imai and Ratkovic

2014)。

- 3) このような推定量の詳細は、Nevo(2003)を参照のこと。
- 4) 例えば、代表的な外部資金の一つである科学研究費補助金は、設置者ごとに比較すると、公立・私立に比べて国立に所属する大学教員の方がキャリアを通じて採択率が高い(藤村2017)。
- 5) C統計量は、0.6~0.9程度と低すぎても高すぎても良くないとされる(康永・笹渕・道端・山名2018)。Katz 2010=2016, 呉・島・西村(2019)ではC統計量は最大値1で、大きければ大きい方がよいとのことであるが、この点に関しては懐疑的にならなければならない。C統計量がほとんど1に近いのであれば1段階目の分析で用いた変数を統制変数とした回帰分析を行なったときのモデルは、ほとんど完璧なモデルであり処置変数の係数は因果効果と不偏推定量となる。そのため、PSを用いるメリットは少ない。
- 6) 大学教員データを対象としたrakeによる重み付け法は中尾・村澤(2018)参照。
- 7) 実際には、C統計量の数値だけで機械的に判断できるわけではなく、処置群と対照群で共変量の分布に差がなくなったかどうかをみるのが本質的には重要である(星野・岡田2006)。本稿では、処置群と対照群の共変量の分布について確認した上で、C統計量の数値も加味して総合的に判断を下したが、紙幅の都合上、省略している点もある。

【謝辞】

本稿は、JSPS 科研費 JP15H03487(基盤研究(B) 代表者 有本章)で収集したデータの提供を受け、使用することをご快諾いただいた。また、本研究推進に際し、以下の資金提供を受けた。JSPS 科研費 JP16H03780・JP18K18651(基盤研究(B)・挑戦的研究(萌芽) 代表者 村澤昌崇)。ここに記して感謝申し上げる次第である。

【参考文献】

- 青木周平・木村めぐみ, 2016, 「日本の国立大学の論文生産性分析」『ファイナンシャル・レビュー』第128号.
- Deaton, Angus., Cartwright, Nancy., 2018, "Understanding and misunderstanding randomized controlled trials", *Social Science & Medicine*, **210**, pp.2-21.
- Drake C., 1993, "Effects of Misspecification of the

- Propensity Score on Estimators of Treatment Effect. *Biometrics*, **49**, pp.1231-1236.
- Fong, C., Hazlett, C., & Imai, K., 2018, "Covariate balancing propensity score for a continuous treatment: Application to the efficacy of political advertisements", *The Annals of Applied Statistics*, **12**(1), pp.156-177.
- 藤村正司, 2017, 「基盤研究費削減が外部資金獲得と研究継続性に及ぼす効果」『ディスカッションペーパーシリーズ』No.6, 広島大学高等教育研究開発センター.
- 藤村正司, 2018, 「なぜ研究生産性が失速したのか?」『大学論集』, **50**, pp.1-16.
- 呉書雅・島一則・西村君平, 2019, 「日本学生支援機構貸与型奨学金の受給が生活時間に与える影響」『高等教育研究』 **22**, pp.207-229.
- Hainmueller, J., 2012, "Entropy balancing for causal effects: multivariate reweighting method to produce balanced samples in observational studies", *Political Analysis*, **20**, pp.25-46.
- 羽田貴史, 2018, 「混乱にみち、根柢なき最近の高等教育政策」『教育学術新聞』.
- 林隆之・土屋俊, 2016, 「学問分野による卓越性指標の多様性」石川真由美編『世界大学ランキングと知の序列化』京都大学学術出版会, pp.327-345.
- Hirano, K & Imbens, G. W., 2004, "The Propensity Score with Continuous Treatment", Gelman, A. & Meng, X-L., *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives: An Essential Journey with Donald Rubin's Statistical Family*, John Wiley & Sons, Ltd., pp.73-84.
- 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』 **55**(3), 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.
- Imai Kosuke and Ratkovic Marc., 2014, "Covariate balancing propensity score", *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **76**(1), pp.243-263.
- 片瀬一男, 2007, 『社会統計学』日本放送出版協会.
- Katz, M. H., 2010, "Evaluating clinical and public health intervention: A Practical Guide to Study Design and Statistics, New York, Cambridge University Press. (=2016, 木原雅子・木原正博訳『医学的介入の研究デザインと統計－ランダム化/非ランダム化研究から傾向スコア・操作変数法まで』メディカル・サイエンス・インターナショナル).
- King G., Nielsen R., 2018, "Why Propensity Scores Should Not Be Used for Matching". Working Paper.
- 小林信一, 2018, 「ポスト真実 (Post-Truth) 時代の科学と政治：科学の危機、証拠に基づく政策立案、日本の動向」『研究 技術 計画』 **33**(1), pp.39-59.
- Leite, W., 2017, "Practical Propensity Score Methods Using R", Sage Publications, Inc.
- 村澤昌崇, 2008, 「評価」有本章編著『変貌する日本の大学教授職』玉川大学出版部, pp.280-294.
- 村澤昌崇・立石慎治, 2017, 「計量分析の新展開：過去10年間の経験を振り返って」『高等教育研究』 **20**, pp.135-156.
- 内閣府, 2018, 『平成30年度 内閣府本府 EBPM 取組方針』.
- 中村知繁・南美穂子, 2017, 「Covariate Balancing Propensity Score を用いた、スクイズ作戦の有効性の解析」『統計数理』, **65**(2), pp.217-234.
- 中村知繁, 2019, 「統計的因果推論とデータ解析」(https://speakerdeck.com/tomoshige_n/causal-inference-and-data-analysis : 2019年8月10日最終アクセス)
- 中室牧子, 2015, 『学力の経済学』ディスカヴァー・トゥエンティワン.
- 中室牧子・津川友介, 2017, 『原因と結果の経済学』ダイヤモンド社.
- 中尾走・村澤昌崇, 2018, 「大学教員の生産性分析：再考－新たな計量分析の試み：欠損値補完・サンプリングバイアスの補正・"ゼロ"の意味の解釈－」広島大学高等教育研究開発センター編『ディスカッションペーパーシリーズ』No.9, pp.1-21.
- Nevo, A., 2003, "Using weights to adjust for sample selection when auxiliary information is available" *Journal of Business & Economics Statistics*, **21**(1), pp.43-52.
- 小野方資, 2015, 「教育政策形成における「エビデンス」と政治」『教育学研究』 **82**(2), pp.241-252.
- Rosenbaum, P. R., & Rubin, D. B., 1983, "The central role of the propensity score in observational studies for causal effects", *Biometrika*, **70**(1), pp.41-55.
- 筒井淳也, 2016, 「因果推論の限界についての社会的検討」『危機に対峙する思考』梓出版社, pp.15-34.
- 浦田広朗, 2017, 「改革期における大学教員の研究生産性規定要因」『ディスカッションペーパーシリーズ』

- No.3, 広島大学高等教育研究開発センター . pp.645-655.
- Wyss, R., Ellis, A.R., Brookhart, M.A., Girman, C.J., Func, M.J., LoCasale, R., Sturmer, T., 2014, "The role of prediction modeling in propensity score estimation: A evaluation of logistic regression, bCART, and the Covariate-Balancing Propensity Score", *American Journal of Epidemiology*, **180**(6),
- 康永秀生・笹渕裕介・道端伸明・山名隼人, 2018, 『できる！傾向スコア分析』金原出版株式会社.
- 吉田文, 2013, 「グローバリゼーションと大学」広田照幸ほか編『グローバリゼーション, 社会変動と大学』岩波書店, pp15-42.