

学位論文の要約

高等教育分野における統計分析と意思決定支援
—EBPM と因果推論を契機として

広島大学大学院教育学研究科

教育学習科学専攻 高等教育学分野

D184298 中尾 走

I. 論文題目

高等教育分野における統計分析と意思決定支援：EBPM と因果推論を契機として

II. 論文構成

第Ⅰ部 本研究の課題

序章 背景と目的

第1節 問題の所在

第2節 先行研究の検討

第3節 研究の枠組みと目的

第Ⅱ部 EBPM とは何か：政策評価論から

第2章 政策評価の方法からみた EBPM

第Ⅲ部 高等教育研究における統計分析

第3章 異なる目的の統計分析

第4章 高等教育研究における統計分析の目的の違い：学習成果論を事例に

第Ⅳ部 社会科学における因果推論の課題：高等教育を事例に

第5章 因果効果の推定における外的妥当性

第6章 因果効果の推定における構成概念の統制

第Ⅴ部 まとめ

第7章 統計分析の整理と政策立案への繋がり

第8章 結果の要約と課題

引用参考文献

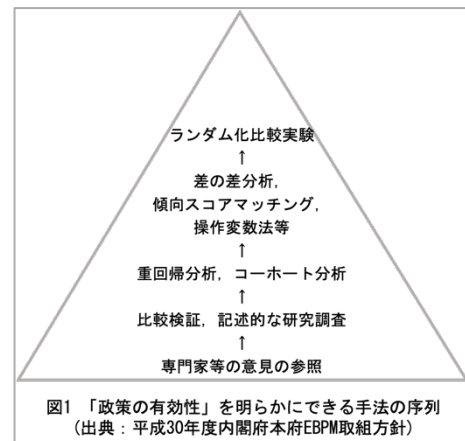
III. 論文要旨

序章 背景と目的

本研究の目的は、EBPM を事例として、高等教育研究の統計分析を政策立案の意思決定支援にどのように応用することが出来るのかを明らかにしていくことである。序章では、先行研究で指摘されてきた高等教育研究の特徴と EBPM に関する先行研究を整理することで、本研究の課題を明確にした。

高等教育研究は、その特徴の一つとして、政策との密接な関わり（金子 1993）や政策的・実践的問題との関連の深さ（中村 2007）が挙げられる。このような特徴は、近年、俄に要請されている根拠に基づく政策立案（Evidence-based policy making, 以下 EBPM）によって、より密接な関わりを生むのではないか。特に、高等教育政策においては、根拠無き政策提言（羽田 2018）と言われており、EBPM が理想とする根拠に基づく政策立案が求められる土壌であり、政策と研究の強い紐帯（小方 2014）を求める動きもある。このように、高等教育政策に対しては、根拠なき政策立案の改善が求められ、高等教育研究に対しては、今以上に政策との関係性を強くしようとの動きがあるが、高等教育政策に対して EBPM の理念は有効だと言えるのか、これが本研究の出発点である。

このような議論は、隣接分野では肯定派と批判派という二つの立場から議論がなされている。しかしながら、注意しなければならないのは、「EBPM をめぐる議論は、それが何なのか明らかでないまま批判することで混乱する」（Cairney 2016）と言われるように、EBPM 批判が議論として噛み合っていないことである（今井 2015）。具体的には、EBPM の批判派は用いられる文脈からエビデンスを捉えているのに対し、肯定派は図 1 に示されるようなエビデンスピラミッドの因果推論の質そのものをエビデンスとして捉えている点である。このような齟齬は、そもそも「EBPM とは何か」という根源的な問いの不在によって生じている。加えて、議論が噛み合っていないために、肯定派の主張するエビデンスを因果推論の質として捉えることについても議論が及んでいないのが現状である。



例えば、林（2019）、Kano & Hayashi（2021）が EBPM のエビデンス概念が狭いことを指摘しているが、この指摘は高等教育研究から EBPM のエビデンス概念を捉えたときにも同様のことが言えるのだろうか。実際に、高等教育研究に掲載されている統計手法のレビューを行った村澤・立石（2018）を見ても、エビデンスピラミッドの中にある手法は、重回帰分析だけである。それでは、高等教育研究は因果推論の質が低い方法を用いてきたのか、それとも、異なる目的で統計分析を用いてきたために比較ができないということなのか、この点を明らかにするためには、高等教育研究でどのような目的で統計分析を用いてきたのかについての整理が必要

となる。

類似の指摘として、行政学・公共政策論の立場から、「どのような統計手法が政策立案のどのような段階で用いることが出来るか」を整理した先行研究がある（山谷 2012, Nutley, Powel & Davies 2013, 小野 2018, 佐藤 2019 など）。けれども、高等教育研究のように複数の専門分野の統計分析が用いられる学際領域では、同じ手法でも異なる目的で用いられており（中尾・樊 2021b, 林 2021a, 2021b）¹、高等教育研究の学際性を考えれば、統計手法との結びつきで整理を行うことは誤用を招きかねない。つまり、より広く分野横断的な統計分析の整理が必要となる。

すなわち、EBPM が高等教育政策に有効か否かを検討する為には、「EBPM とは何か」という根源的な問いと高等教育研究では統計分析を用いて何をしてきたのかを整理することが必要ではないだろうか。そこで本研究では、以下のような3つの研究課題を設定する。一つ目に、行政学・公共政策論におけるアプローチを踏襲して EBPM とは何かを政策評価の方法から捉え直すこと。二つ目に、分野横断的に統計分析を目的ごとに整理を行い、異なる目的の統計分析が高等教育研究ではどのように用いられてきたかを再考すること。三つ目に、政策評価の方法と目的の異なる統計分析を結びつけ、どのような目的の統計分析を政策評価の方法として用いることが可能かを考察していくことである。これらの研究課題に答えることで、EBPM を事例に、高等教育研究の統計分析を意思決定支援にどのように応用することが出来るのかを明らかにしていく。そのため、次章ではまず、行政学・公共政策論の政策評価の方法から、EBPM とは何かを捉え直していく。

第2章 評価の理念類型からみた EBPM

本章では、政策評価方法の理念的な類型である①政策分析（policy analysis）、②プログラム評価（program evaluation）、③業績測定（performance measurement）の3つの視点²から EBPM を捉える南島（2019）の枠組みを用いて、政策評価の方法における EBPM の位置付けを明らかにした。

まず、政策分析から EBPM を捉えたときに、EBPM は政策分析と同様に合理的な意思決定を支援するものとして期待されている。一方で、政策分析では1ドル当たりの価値（効果）を高める Value for Money を、EBPM では What Works（何が有効か）を重要視しており、合理性の価値基準が異なる点が挙げられる。次に、プログラム評価から EBPM を捉えると、因果効果を明らかにする手法と密接に結びついており、理念上は古くからランダム化比較実験（Randomized Control Test, 以下 RCT）が必要とされていたが、目的と手法の乖離がみられたプログラム評価（総合評価方式）に対して、同様の目的を擁した EBPM は、エビデンスピラミッドという手法

¹ 例えば、小野（2018）や小倉（2020）では回帰分析を因果効果のエビデンスピラミッドの下方に位置付けているが、予測や関連の探索、説明が目的であれば準実験的な方法や実験研究よりも望ましい場合があり得る。

² 日本の制度で言えば、それぞれ事業評価方式、総合評価方式、実績評価方式と親和的である。

も同様に輸入し、手法を適用するという側面がより強く前面に出ている点が挙げられる。業績測定から EBPM を捉えると、政策評価制度では、プログラム評価を志向しながらも業績測定が大きな役割を果たすようになり、EBPM でも同様の失敗を繰り返す可能性が指摘できる。このような整理から EBPM は、良いかどうかは別として3つの評価の理念に基づく類型を組み合わせた概念であるということが言える。一方で、EBPM の固有性という面では、エビデンスピラミッドが大きな存在であり、政策評価制度とはエビデンスの質が異なる（内山 2019）との指摘もあり、初中等教育ではエビデンスの質が低かったことの見直し（寺沢 2015）が指摘されている。しかしながら、高等教育研究は経済学・心理学・社会学などの隣接領域の統計分析を応用してきたため、より多様な目的の統計分析を用いてきた可能性があるのではないかと。そしてその場合、一概にエビデンスの質が低かったと言えないのではないかと。そこで、次章では高等教育研究で用いられることの多かった経済学・心理学・社会学の専門分野を中心に社会科学ではどのような目的で統計分析を用いてきたのかについてレビューを行うこととする。

第3章 異なる目的の統計分析

本章では、社会科学の統計分析を分野横断的に整理するために、目的に焦点を当てて整理を行った。

高等教育研究を含む社会科学の多くの研究では、線形モデルまたはその発展系が用いられており、その基礎となっているのは最小二乗法（Ordinary Least Square, 以下 OLS）による回帰分析（以下、古典的回帰分析）である。一方で、経済学・心理学・社会学では古典的回帰分析を異なる目的で用いている（中尾・樊 2021b）。そして、それらを整理すると社会科学で用いられる古典的回帰分析は、大きく分けて5つの目的に分けられる。具体的には、①記述・測定（得られたデータから情報を得る）、②関連の探索（関連のあるものを探索する）、③説明（なぜその現象が生じたかを説明する）、④予測（得られたデータから予測する）、⑤因果効果の推定（その処置によって結果がどれだけ変化したか明らかにする）である。このように整理すると、高等教育研究は学際領域であるが故に、様々な専門分野を背景とする統計分析が用いられており、混乱を招いてきたと言える。実際に、異なる目的で古典的回帰分析を用いる場合、必要な仮定や係数の解釈の仕方、モデル評価の方法が異なる（林 2021a, 2021b）。また、古典的回帰分析からの発展系である多くの線形モデル³も目的ごとに異なる発展をしているため、目的ごとに理解することで、混乱を整理することが可能となる。それでは、このような多様な目的の統計分析を高等教育研究ではどのように用いてきたのだろうか。この点について、次章で学習成果論を事例にして、詳細に述べていくこととする。

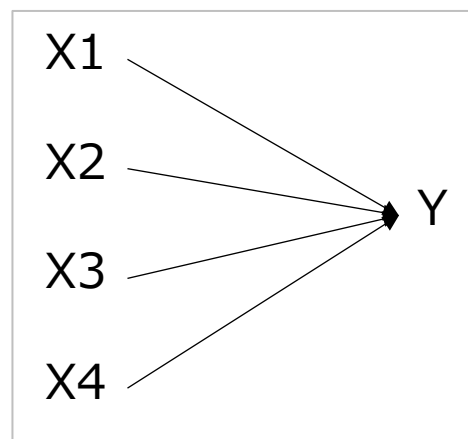
第4章 高等教育研究における統計分析の目的の違い：学習成果論を事例に

本章では、高等教育研究の学習成果論を事例に、これまでどのような目的で統計分析が用い

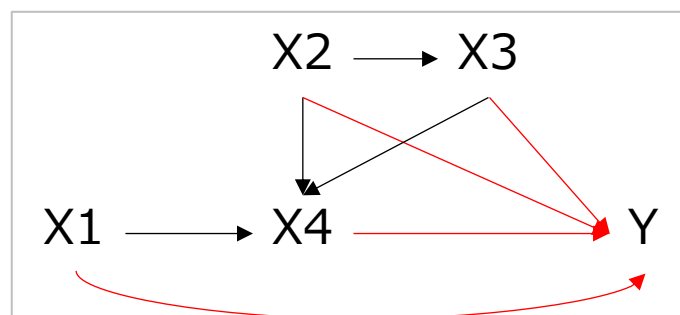
³ 例えば、マルチレベル分析や一般化線形モデル、パネルデータ分析、構造方程式モデリングなどである。

られており、どのような課題があるかを指摘した。高等教育研究で学習成果を論じる際に用いられてきたのは、大規模な学生調査であり、学生の状態や変化を基礎的な集計を「記述」し（藤村 2013）、線形モデルなどの統計分析を用いて「関連の探索」や「説明」が行われてきた（中尾ほか 2021）。例えば、「学習成果が高いのはどのような学生か」といった問題設定は全て「関連の探索」や「説明」という目的で統計分析を用いてきたと言える。

一方で、「学生の学習や大学の教育的介入が、学習成果を高めたのか」という因果効果の推定はほとんど行われていない。もしくは、行われていたが、手法としては「関連の探索」や「説明」の手法が間違いだと認識されないまま用いられてきた（小方 2008, 岡田・鳥居 2011, 三好 2014 など）。例えば、学習成果論で主に分析枠組みとして用いられてきた IEO (Input-Environmental-Output) モデルを用いて分析を行った小方 (2008) では、家庭背景と入学前の特性・学習動機、学部の組織構造、教育プログラム、学生のエンゲージメント（主体性など）4 つに分類される変数を投入し、規定要因分析を行っている。その上で、有意水準が満たされたものを解釈しているが、このような分析を行う場合、それぞれの変数を X1~X4 と置くと、多くの分析者は以下の図のような分析モデルを想定する。



しかしながら、背後に以下の図のような因果モデルが想定される場合、推定されている係数は、図の直接効果（赤線部分）のみとなる。



因果効果の推定として線形モデルを用いるのであれば、計量経済学の古典的回帰分析の仮定

(Classical Assumption, 以下 CA) (浅野・中村 2009, 鹿野 2015) や変数選択をバックドア基準 (林・黒木 2016) に基づいて行い, 得られた係数の評価を不偏性, 一致性, 有効性を評価する必要がある。また, 因果効果の推定が目的であれば大規模な学生調査ではなく, 同質な集団だと見做せるサンプルの方が望ましい場合もある。このように, 高等教育研究では目的を混乱させたまま統計分析を用いてきた。一方で, 因果推論については, 手続き化されているため, 「関連の探索」や「説明」と異なり, 係数の解釈の仕方やそのための仮定, 係数の解釈の仕方, モデル評価の方法が明確である。しかしながら, 因果推論の方法を社会科学に应用する際にも社会科学固有の課題があるのではないか。そこで, 次章以降では, 社会科学の文脈で因果推論を行う際の課題を高等教育のデータを用いた分析事例から明らかにしていく。

第5章 因果効果の推定における外的妥当性

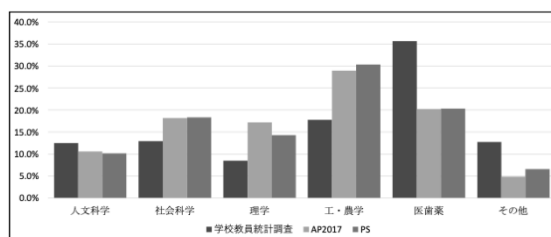
本章では, 社会科学の因果推論の課題としてよく指摘される外的妥当性問題を取り上げ, 具体的な分析事例から先行研究の外的妥当性問題の指摘が間違いである点を指摘し, 実験データ, 観察データを問わず, 因果推論の外的妥当性問題は効果修飾 (effect modification) の問題であること, その分析事例を示した。

そもそも, なぜ根拠に基づく医療 (Evidence-based medicine, 以下 EBM) では内的妥当性がエビデンスの質として捉えられてきたのか。それは, EBM は生態学的斉一性が前提であるという林 (2019a) の指摘が一つの答えであろう。斉一性とは, 対象が均質であるという原理である。一部のサンプルの知見が外的妥当性を満たすかどうかは, 対象が均質であれば満たされて当然であるが, 星野 (2009) は, 社会科学では斉一性のような仮定を置くことは出来ないとしている。この指摘を踏まえると, 政策立案という文脈では RCT によるエビデンスが, そのまま質の高いエビデンスとなり得るとは限らないことが分かる。また, 実験研究に対する外的妥当性問題の指摘は多くある (中澤 2016, 大久保 2019) が, 観察研究でもあまり関係はない。何故なら, 因果推論における外的妥当性

は, 効果修飾によって決まるからである

(Hernan & VanderWeele 2011)。効果修飾とは, 共変量がとる値ごとに因果効果の大きさが異なること (黒木 2017) であり

4, もし仮に効果が全員同じであれば, 外的妥当性は自然と満たされる。また観察研究において外的妥当性が満たされないのは, 交絡調整した後のデータからも導かれるため, ランダムサンプリングによってデータを取得できたとしても, 交絡調整段階で, 異なる集団となっていることもある (図 2 参照)。



4 効果修飾と交互作用の違いについては VanderWeele (2009) 参照のこと。

それではこのような効果修飾は高等教育研究ではどのような事例があるか。一つが専攻分野である。実際に、内定獲得確率に対する大学の選抜度の効果は専攻分野で大きく異なることが実証された（図3参照）。そのため、この事例であれば、未観測の効果修飾がなければ、母集団分布と手元のデータの専攻分野の分布が同じであれば外的妥当性が満たされる可能性がある。本章では、社会科学における因果推論でよく指摘される外的妥当性の問題を取り上げ、その事例分析を行ってきたが、そのほかに大きな課題はないだろうか。社会科学固有の問題として、構成概念の統制の問題を取り上げ、次章ではその因果推論の問題について説明する。

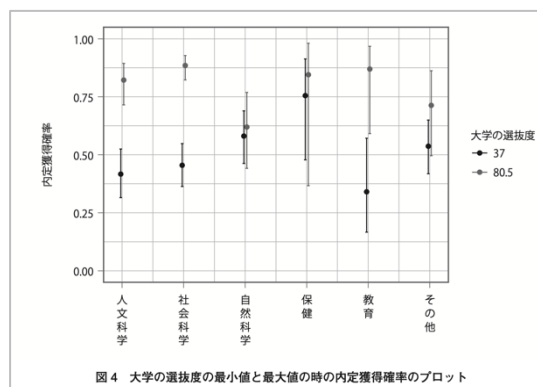


図4 大学の選抜度の最小値と最大値の時の内定獲得確率のプロット

第6章 因果効果の推定における構成概念の統制

本章では、社会科学における因果推論の固有の問題として、構成概念の統制という問題を取り上げ、構成概念と代理指標間に生じる測定誤差によって、バイアスの挙動がどのようなものかを明らかにした上で、構成概念の統制を代理指標で行った場合の事例分析を行なった。

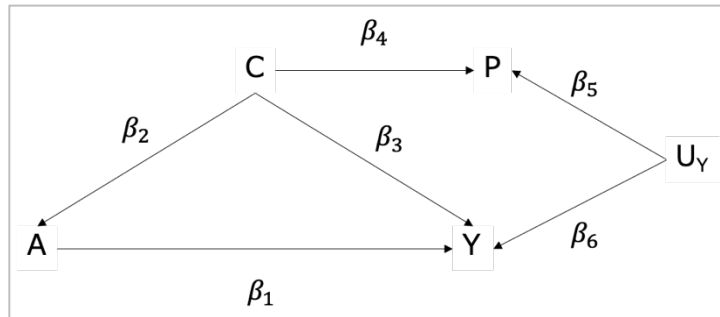
高等教育研究でよく用いられる構成概念として、能力や社会経済的地位 (Social Economic Status)、コンピテンシーなどが挙げられる。このような構成概念はあくまでも構成概念であるため、完璧に測定することは困難であり、統計分析上は代理指標によって測定され、分析に用いられる。そのため、必ず測定誤差が生じると言っても過言ではない。それでは、このような測定誤差が生じる場合、因果推論で得られた結果のバイアスはどのような挙動をするのであろうか。先行研究は、主に3つの視点から測定誤差と因果推論の問題について言及されている。一つ目が処置変数の測定誤差によって生じるバイアスであり、減衰バイアス (attenuation bias) と呼ばれ、内生性問題として多くの蓄積がある (浅野・中村 2009, Jerrim and Micklewright 2014, Engzell and Jonsson 2015 など)。二つ目が、従属変数の測定誤差が挙げられる (VanderWeele and Hernan 2012)。三つ目は、交絡変数の測定誤差である (Frost 1979, Ogburn and VanderWeele 2012, Miao et al. 2017)。これらの先行研究に共通しているのは、以下のように偶然誤差として測定誤差を捉えている点である。

$$x = x^* + e$$

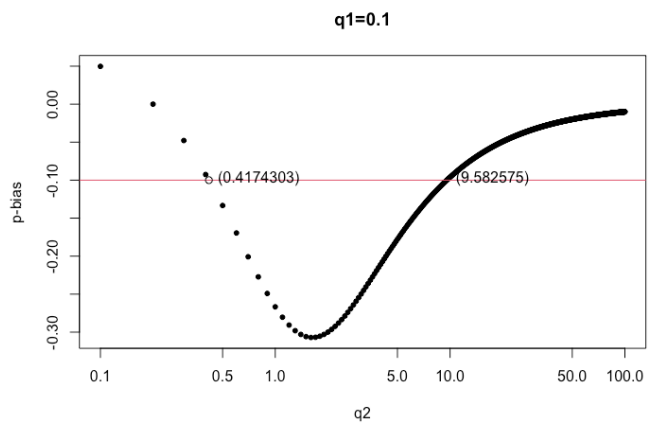
一方で、下記式のように偶然誤差と系統誤差の両方が生じる場合、推定量は異なる挙動となることが予想される。

$$x = x^* + u + e$$

特に、交絡変数の測定誤差が系統誤差と偶然誤差によって生じる場合、以下の因果モデル（図4参照）となり、代理指標であるPを統制した場合に新たに合流点バイアスが生じるため、代理指標を統制するか否かについての判断が難しい。



そのため、数式展開とシミュレーションによってバイアスの挙動を明らかにしたところ、測定誤差が小さいとしても必ずバイアスが小さくなるわけではないことを明らかにした（図5参照）。



このような分析事例を見ても、因果推論の手法を高等教育という対象にそのまま適用可能なわけではない。ここまで、因果推論においては社会科学固有の課題が、そして関連の探索や説明

は、そもそも統計学的方法が確立していないという課題を指摘した。それでは、EBPMで推奨されるように、意思決定に統計分析をどのように応用していくことは不可能なのか。次章では、統計分析の目的の整理と政策評価の方法を結びつけて整理することで、この点について考察を行なっていく。

第7章 統計分析の整理と政策立案への繋がり

本章では、EBPMを事例に社会科学の異なる目的の統計分析が意思決定支援としてどのように応用できるのかの考察を行った。

ここまでの整理より、EBPMで求められるエビデンスは因果推論だけであり、あくまでも社会科学で用いられる統計分析の一つの目的でしかない点が明らかになった。例えば、エビデンスピラミッドで最も質の高いRCTは、因果推論の目的であれば、質の高さは保証できるが、関連の探索や説明など、その他の目的としては相応しい手法ではない。加えて、第3章の整理に基づけば、EBPMは3つの評価の理念に基づく類型を組み合わせた概念であり、その他の目的の統計分析も用いることが可能であろう。このような政策評価方法の理念的な類型に基づいて、統計分析の目的がどの政策評価の方法として用いることが可能かを整理すると表1の通りとなる。

表 1 評価の理念類型と統計分析の目的ごとの整理

	政策分析	プログラム評価	業績測定
評価の方法	事前	事後	定点観測
評価の観点	効率性	有効性	改善傾向
日本の制度	事業評価方式	総合評価方式	実績評価方式
目的	合理的判断への貢献	事後検証	点検
統計分析との接続	関連の探索 説明 予測	因果推論	記述・測定

例えば、研究生産性を事例に、意思決定支援に統計分析を応用することを考えると以下のよう整理できる。

課題発見から意思決定の前段階の分析としては、予測が最も親和性が高い。なぜなら、論文数の予測値は、今後どのような取り組みをすれば良いかを考える機会の提供として課題の共有を行うのに適しているからである。一方で、予測として統計分析を用いる場合は、解釈性が低いため何に組みれば良いかは不明である。加えて、多くの予測モデルは予測誤差を最小化することが最も良いモデルとされているが、予測精度が高いモデル＝望ましい意思決定を導けるモデルではない点には注意が必要となる。

次に、意思決定の前段階の分析として親和性が高いのは、関連の探索（研究生産性はどのような人が高いのか）や説明（研究生産性の違いはどのように説明できるのか）である。予測は、解釈性が低く何に組みればよいか不明である点が欠点であったが、社会科学分野で関連の探索や説明を行う場合、解釈できるような分析枠組みの中で統計分析を用いるため（Miller 2019）、何に組みれば良いかを考える場面で、意思決定を支援することに応用することが向いている。一方で、相関関係と因果関係は異なるため、何に組みれば良いかを考える場合に参考程度にはなるが、関連のある項目に取り組んだとしても、改善するとは限らず、関連の探索や説明として統計分析を用いる場合、確立した統計学の方法があるわけではない点には注意が必要である。

最後に、意思決定後の事後的な検証の分析としては、因果効果の推定（取り組んだ結果、その効果はどうだったのか）が親和的である。このような整理を行うと、EBPM で推奨されている因果推論は、何か取り組んだ結果の事後的な検証として用いる場面に限られる。

第 8 章 結果の要約と課題

本章では、EBPM を事例として統計分析の結果を政策立案の意思決定支援にどのように応用することが出来るのかという本研究の議論を総括し、本研究で得られた知見の整理と、今後の課題について述べた。

まず、「EBPM とは何か」を政策評価の方法から捉え直した時に、政策分析、プログラム評価、業績測定を組み合わせた概念であり、固有性として因果推論の質をエビデンスの質と捉えている点にある。そのため、EBPM で推奨されている因果推論は、プログラム評価の手法の一つとして捉えれば、機能する場合もある。学習成果論の文脈では「学生の学習や大学の教育的介入が、学習成果を高めたのか」という問いに対して適切に因果推論が行われていたわけではないため、再分析によって結果が異なることはあり得るだろう。一方で、因果推論が適切に用いられてこなかったことで、これまでの高等教育研究はエビデンスの質が低かったかという点、そうではなく、異なる目的（関連の探索や説明）で統計分析を用いてきたとも考えられる。しかしながら、関連の探索や説明という目的で統計分析を用いる場合、統計学的に確立した方法があるわけではない点には注意が必要となる。逆に言えば、因果推論は統計学的に確立した方法があるため、応用しやすいが、社会科学固有の課題に対して豊富な研究蓄積があるわけではないと言える。

以上が得られた主な知見であるが、本研究を踏まえて、以下の4点が今後の課題として考えられる。一つ目が、高等教育政策や高等教育研究の過去の事例を基に適切に因果推論を応用し、事後検証を行うこと。二つ目が、関連の探索、説明という目的で統計分析を用いる際の必要な仮定や係数の解釈の仕方、モデル評価の方法を確立すること。三つ目が、因果推論を社会科学分野へ応用する際に生じるバイアスや仮定違反することによる推定量の性質を明らかにすること。四つ目が様々な目的の統計分析だけでなく、質的研究までを含めて意思決定支援へどのように応用できるかを明らかにすること、である。これらの課題については今後、改めて取り組んでいくこととしたい。

IV. 主要参考文献

- 浅野哲・中村二郎, 2009, 『計量経済学 第2版』有斐閣.
- Cairney, P., 2016, *The Politics of Evidence-Based Policy Making*. Palgrave Pivot.
- Elwert, Felix and Christopher Winship, 2014, “Endogenous Selection Bias: The Problem of Conditioning on a Collider Variable,” *Annual Review of Sociology*, 40(1): 31–53.
- 樊怡舟・中尾走・西谷元・村澤昌崇, 2021, 「交互最小二乗法を用いた大量欠損の成績表データからの因子抽出」『行動計量学』48 (2) : 69-77.
- Frost, P. A., 1979, “Proxy Variables and Specification Bias,” *The Review of Economics and Statistics*, 61(2): 323-325.
- 藤村正司, 2013, 「大規模学生調査から学習成果と学習時間の構造を掴む」『大学論集』44 : 1-17.
- 羽田貴史, 2018, 「混乱にみち、根拠なき最近の高等教育政策」『教育学術新聞』.
- 羽田貴史, 2019, 「高等教育研究の制度化と教育社会学」『教育社会学研究』104 : 7-28.
- 林岳彦, 2019, 「EBPM, “E” から見るか? “PM” から見るか?」『研究者/研究所として “EBPM” とどう関わるとよいのか? 国立環境研究所研究集会発表資料』.
- 林岳彦, 2020, 「統計的因果推論, 構造から見るか? 差分から見るか?」
- 林岳彦, 2021a, 「コメンテーターからの意見」広島大学高等教育研究開発センター2021年度第5回公開研究会 発表資料.
- 林岳彦, 2021b, 「学術的に理解するってどういうこと?」先端課題研究会 発表資料.
- 林岳彦・黒木学, 2016, 「相関と因果と丸と矢印のはなし: はじめてのバックドア基準」岩波データサイエンス刊行委員会『岩波データサイエンス3』岩波書店.
- Hernan, M. A. and Vanderweele, T. J., 2011, Compound Treatments and Transportability of Causal Inference *Epidemiology*, 22(3): 368-377.
- 井伊雅子・五十嵐中・中村良太 (2019) 『新医療経済学』日本評論社.
- 今井康雄, 2015, 「教育にとってエビデンスとは何か」『教育学研究』82 (2) : 188-201.
- Jerrim, John and Micklewright, John, 2014, “Socio-economic Gradients in Children’s Cognitive Skills- Are Cross-Country Comparisons Robust to Who Reports Family Background?,” *European Sociological Review*, 30(6): 766-781.
- 石橋章市朗・佐野亘・土山希美枝・南島和久『公共政策学』ミネルヴァ書房.
- 金子元久, 1993, 「高等教育制度・政策の研究」『大学論集』22 : 187-208.
- Kano Hiroyuki and Hayashi Takehiko, 2021, A framework for implementing evidence in policymaking- Perspectives and phases of evidence evaluation in the science-policy interaction, *Environmental Science and Policy*, 116: 86-95.
- 狩野裕, 2002, 「構造方程式モデリング, 因果推論, そして非正規性」甘利俊一・狩野裕・佐藤俊哉・松山裕・竹内啓・石黒真木夫『多変量解析の展開: 隠れた構造と因果を推理する』

- 岩波書店：65-129.
- 小林雅之・山田礼子，2016，『大学の IR』慶應義塾大学出版会.
- 小林信一，2019，「高等教育政策の研究」『教育社会学研究』104：57-80.
- Looker, E. D., 1989, “Accuracy of proxy reports of parental status characteristics,” *Sociology of Education*, 62(4): 257-276.
- 正木朋也・津谷喜一郎，2010，「エビデンスに基づく医療（EBM）の展開から学ぶもの」『日本評価研究』10（1）：3-16.
- 松宮慎治・中尾走・樊怡舟・宮田弘一・村澤昌崇，2022，「大学の経営・特性からみた教職課程の設置行動」『大学論集』54：(in press) .
- Miao Wang, Geng Zhi, and Tchetgen, E. T., 2018, “Identifying Causal Effects With Proxy Variables of an Unmeasured Confounder,” *Biometrika*, 105(4): 987-993.
- Miller Tim, 2019, Explanation in artificial intelligence- Insights from the social sciences,” *Artificial Intelligence*, 267: 1-38.
- 南島和久，2020，『政策評価の行政学』晃洋書房.
- 内閣府，2017，『第1回 EBPM 推進委員会議事要旨』.
- 中泉拓也，2019，「英国の EBPM（Evidence Based Policy Making）の動向と我が国への EBPM 導入の課題」『経済経営研究所年報』41：3-9.
- 中尾走，2019，「外部資金獲得の効果は？」『広島大学大学院教育学研究科紀要第三部（教育人間科学関連領域）』68：151-159.
- 中尾走，2021a，「EBPM の動向と課題—今後の高等教育研究の蓄積のために」『大学論集』，53，pp.37-52.
- 中尾走，2021b，「データ分析の結果はどのように使えるか」2021 年度河合塾グループ I R セミナー（オンライン）#2 大学教職員のための初歩からの I R 講座 発表資料.
- 中尾走・樊怡舟，2021a，「留学効果の検証」広島大学高等教育研究開発センター2020 年度第 20 回公開研究会 発表資料.
- 中尾走・樊怡舟，2021b，「高等教育研究のための計量手法の整理」広島大学高等教育研究開発センター2021 年度第 5 回公開研究会 発表資料.
- 中尾走・樊怡舟・村澤昌崇，2022，「交絡変数の測定誤差が合流点バイアスをもたらす影響について：合流点バイアスと交絡統制のジレンマに対する考察」『理論と方法』(in press) .
- 中尾走・樊怡舟・西谷元・村澤昌崇，2021，「自然実験環境を用いた留学効果の推定」広島大学高等教育研究開発センター編『Advancement of Higher Education Research: RIHE Monograph Series』3：1-10.
- 中尾走・樊怡舟・村澤昌崇・松宮慎治・西谷元，2021，「「学習成果」論再考」『大学評価研究』20：53-62.
- 中尾走・平尾智隆・梅崎修，2021，「大学での学習内容は新規学卒労働市場で評価されているのか？」『キャリアデザイン研究』17：47-59.

- 中尾走・樊怡舟・宮田弘一・村澤昌崇・松宮慎治, 2022, 「大学教員の意識に関する Age-Period-Cohort 分析」『大学論集』 54 (in press) .
- 中村高康, 2007, 「高等教育研究と社会学的想像力」『高等教育研究』 10 : 97-109.
- Nutley, S., Powel, A and Davies, H., 2013, *What Counts as Good Evidence? Alliance for Useful Evidence*.
- Ogburn, E. L. and VanderWeele, T. J., 2012, “On the Nondifferential Misclassification of a Binary Confounder,” *Epidemiology*, 23(3): 433-439.
- 小方直幸, 2011, 「学生のエンゲージメントと大学教育のアウトカム」『高等教育研究』 11 : 45-64.
- 小方直幸, 2014, 「専門知と政策形成」『大学論集』 47 : 73-88.
- 小倉將信, 2020, 『EBPM とは何か』 中央公論事業出版.
- 岡田有司・鳥居朋子, 2011, 「私立大学における大学生の学習成果の規定要因」『京都大学高等教育研究』 17 : 15-26.
- 大久保将貴, 2019, 「因果推論の工具箱」『理論と方法』 65 : 20-34.
- 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) : 55-66.
- 小野達也, 2018, 「エビデンスに基づく政策評価」『統計』 69 (7) : 8-13.
- 大塚淳, 2020, 『統計学を哲学する』 名古屋大学出版会.
- 大槻大輔, 2018, 「英国、米国における EBPM の動向」『統計』 69 (7) : 28-39.
- 佐藤仁, 2019, 「教育政策においてエビデンスを「つかう」とはどういうことか」 杉田浩崇・熊井翔太編 『「エビデンスに基づく教育」の関を探る』 42-68.
- 鹿野繁樹, 2015, 『新しい計量経済学』 日本評論社.
- 清水裕士, 2021a, 『心理学統計法』 放送大学教材.
- 清水裕士, 2021b, 「心理学における重回帰分析の使い所を考える」 (<https://osf.io/bpws2/> : 2021年12月20日, 最終確認) .
- 田辺智子, 2006, 「エビデンスに基づく教育」『日本評価研究』 6 (1) : 31-41.
- 田中隆一, 2020, 「教育政策における EBPM」 大橋弘編 『EBPM の経済学』 61-87.
- 立石慎治・丸山和昭・速水幹也・松宮慎治・中尾走・村澤昌崇, 2019, 「高等教育研究と統計分析」『教育社会学研究』 104 : 29-55.
- 寺沢拓敬, 2015, 「英語教育学における科学的エビデンスとは？—小学校英語教育政策を事例に」『外国語教育メディア学会 (LET) 中部支部外国語教育基礎研究部会 2014 年度報告論集』 15-30.
- 津田敏秀, 2011, 『医学と仮説』 岩波書店.
- 津田敏秀, 2013, 『医学的根拠とは何か』 岩波書店.
- 津谷喜一郎, 2011, 「日本の EBM の動きからのレッスン」『国立教育政策研究所紀要』 140 : 45-54.

- 筒井淳也, 2021, 『社会学』岩波書店.
- 内山融, 2019, 「EBPM の行政への実装に向けて」『RIETI EBPM シンポジウム エビデンスに基づく政策立案を進展させるために 発表資料』 (<https://bit.ly/3zs2c9F>, 2021年12月20日最終アクセス) .
- VanderWeel, T. J., 2009, On the distinction between interaction and effect modification *Epidemiology*, 20(6): 863-871.
- Vanderweele, T. J. and Hernan, M. A., 2012, “Results on Differential and Dependent Measurement Error of the Exposure and the Outcome Using Signed Directed Acyclic Graphs,” *American Journal of Epidemiology*, 175(12): 1303-1310.
- 山口富子・福島真人, 2019, 『予測がつくる社会』東京大学出版会.
- 山本清, 2016, 「証拠に基づく大学政策の推進の実態と課題」『大学経営政策研究』6 : 1-16.
- 山本清, 2018, 「『証拠に基づく政策立案』の課題と展望」『大学経営政策研究』8 : 217-230.
- 山谷清志, 2012, 『政策評価』ミネルヴァ書房.
- 吉田寿夫・村井潤一郎, 2021, 「心理学的研究における重回帰分析の適用に関わる諸問題」『心理学研究』(in press) .