

川口康平

Kawaguchi Kohei

澤田真行

Sawada Masayuki

＝著

因果推論の 計量経済学



The Econometrician's Guide to
Causal Inference

はしがき

本書は、「因果推論」のフレームワークでミクロ経済学の実証分析を行うために現在必要とされている知識を、半ば暗黙知となっている実践も含めて、余すことなく、統一的に解説するために書かれた計量経済学の教科書である。主に学部上級から大学院レベルの学生で、ミクロ経済学の実証研究で論文を書きたいと考えている人たちを対象に執筆した。このような教科書を書いたのは、複数の因果推論の教科書や論文を漁らなければ、現在必要とされる知識のすべてをカバーするのが難しいということに、筆者らがそれぞれの勤務校で大学院レベルの講義を準備する際に気づいたからである。

計量経済学の分野には、理論家の手による優れた教科書がすでにたくさん存在する。しかし、それらの教科書では、経済学における伝統的な「構造推定アプローチ」、すなわち「経済理論による均衡条件の構造型の導出 → その誘導型の導出 → 最小2乗推定量 → 操作変数法 → 一般化積率法」へと進む流れで体系付けられており、現在の「潜在結果アプローチ」による因果推論は、その体系の中に埋め込まれる形で解説されるスタイルとなっている。しかしそれは、制度の記述から始まる近年のミクロ経済学の実証研究における因果推論の実践とは必ずしも一致する形とはなっていない。また、計量経済学の理論を学習するうえで欠かせない、推定量の妥当性を保証するための漸近論に紙幅を割かざるをえないという面もある。

加えて、実際に現実の制度とデータを前にして因果推論に基づく実証研究を行う際には、ある程度定型的に用いられる実践の「型」がある。分析方法の仮定の妥当性を検証したり、結果の解釈の妥当性を議論したりするうえで、理論の枠内には収まり難いものの、実証研究を行って論文を学術誌に出版するためには、こうした「型」の習熟が必須である。また、プログラミングを通じて実際に手法を実装する方法を学ぶことも重要である。しかし、理論家の手による教科書は理論を学ぶことを目的に書かれていることが多く、これらの点にはあ

まり言及されていない。

一方、応用家の手による優れた教科書もすでに存在する。それらの教科書では、分析方法の仮定の妥当性を検証したり、結果の解釈の妥当性を議論したりするための実践に関する解説や、プログラミングによる実装の解説が厚く提供されている。しかし、理論的な内容の解説は必ずしも十分とはいえないことが多い。たとえば「差の差法」では、 2×2 の設定のような限定的なシチュエーションにおける直観的な解説にとどまっており、識別や推定量の性質に関する厳密な論証は行われないなどの傾向がある。

このことは、応用家による教科書において、統計的推測における多重検定問題、回帰非連続デザインにおける最適バンド幅の選択、差の差法における一様信頼区間の構成など、実際に論文で頻繁に用いられているにもかかわらず、既存の教科書には解説がなく、元論文に当たるしかないような最新の手法について、厳密な議論が避けられがちになるという問題にもつながっている。また、こちらのタイプの教科書でも、意識的か無意識的にかかわらず、やはり「構造推定アプローチ」の流れを引き継いでいることが多い。

因果推論の活用は経済学以外にも多くの分野で広がっており、各分野で優れた教科書が複数出版されている。しかし、そうした教科書では、経済学で頻繁に用いられる各種手法の解説が手薄になっている。たとえば、「疑似実験」に含まれる回帰非連続デザインや差の差法、特にその派生的な手法が扱われていなかったりする。その代わりに、マッチングなどの実験を近似することを目的とした手法が手厚く議論されていたりする。こうした違いは、因果推論の世界には分野固有の課題に 대응するために生まれた「訛り」があるため生じている。そのため、経済学が直面する課題に向き合うためには、「経済学訛りのある因果推論」の知識が必要になるのである。

計量経済学の講義は、従来は計量経済学の理論家によって教えられることが多かったが、近年では著者の1人である川口のように応用を専門とする研究者によって教えられることも多い。そうすると、因果推論の手法の単純な局面での直観的な側面ばかりが教えられる一方、その理論的な側面や、最新の手法の裏側にある議論などがごっそりと抜け落ちてしまうおそれがある。プログラミングによる実装を教える場合にも、パッケージのどのオプションが理論のどの側面に対応するかは解説されるものの、その厳密な内容の理解が不十分なまま

になってしまうことがある。たとえば、R の `rdrobust` というパッケージにおける“robust”な標準誤差は何を意味するかといった理解があやふやなままとすることすらある。

とはいえ、もう1人の著者である澤田のような理論家が教えればこうした問題が解決されるというわけでもない。今度は、仮定の検証や解釈の妥当性の議論といった、実践的な側面の解説が不十分になってしまう可能性がある。

また、最新の手法について解説しようにも、教科書に記述がないので、学生には元論文を参照してもらわなければならない。さらに、このような講義を準備したことのある研究者の多くは実感されていると思うが、元論文はそれぞれ記法や仮定が異なっており、講義の中の統一的な解説に落とし込むのが意外に難しい。

このような、何をとも「帯に短し褌に長し」といった現状を補完するために書かれたのが本書である。理論家が教えるにしても、応用家が教えるにしても、また、理論を志す学生が読むにしても、応用を志す学生が読むにしても、あるいはビジネスや公共政策の世界のデータサイエンティストなどといった、そのどちらの立場にもない人が読むにしても、既存の因果推論の教科書では扱いの薄かった側面まで分厚く記述されていて読み応えのある内容を目指して、応用家である川口と理論家である澤田が共同で本書を執筆した。

本書が実現しようとしている教育効果を達成するためには、従来、計量経済学の古典的な教科書に因果推論に関する各種教科書をあわせ、元論文の講読や輪講、あるいはデータ分析の実践を通じて、教育を行う必要があった。しかし、その結果として得られる知識は、今や最先端の実証研究の世界ではただのコモディティである。ただのコモディティをもったいぶって学んでも仕方がない。本書は、「そのような既存の知識は暗黙知も含めてドキュメント化してさっさと共有してしまおう。人々の貴重な時間はもっと新しく未解決の問題について考えるために使おう」という構想のもとで書かれている。

そうした構想に従い、本書の内容はウェブ付録を含めて主に4つのパートから成り立っている。まずは、潜在結果モデルや無作為化実験のもとでの因果推論、回帰非連続デザインや差の差法などの疑似実験のもとでの因果推論に関する基礎知識をまとめた「**基礎編**」、それらの手法の実用例について議論する「**応用編**」、多重検定、クラスター頑健標準誤差、回帰非連続デザインと差の差法の最新の手法や理論的な結果についてまとめた「**発展編**」、最後に、それらの手法のRによる実装を

紹介したウェブ付録の「実装編」(本書内の該当箇所に**R**)である。また、数学的な証明に関する「テクニカルノート」(本書内の該当箇所に**T**)もウェブ上で提供している。なお、ウェブ付録は本書のサポート用 GitHub リポジトリ (https://github.com/keisemi/EconometriciansGuide_CausalInference) で公開している。扱う手法は、既存の教科書や公刊論文で導入されており、かつ(未公刊論文を含む)最新の実証研究の論文でいくつかの応用例が認められるものに限った。したがって、最近未公刊論文として公開された手法や、学会やセミナーで発表されたような手法などは含んでいない。こうした手法は、元論文を読んでその意義と限界を自分で理解できる研究者だけが使用すべきだと考えるからである。

本書を読んだ初学者が短い期間で既存の研究者を乗り越え、それを上回る成果を発信し始めることを期待している。

* * *

本書は著者の川口と澤田が構成から内容まで、編集者とともに議論を重ねながら執筆した。その過程で、当時一橋大学の大学院生だった中井絵理奈氏には、文献調査などの面で大いに手助けいただいた。また、東京大学の奥井亮氏、重岡仁氏には、第1稿を批判的な観点から読んでいただき、さまざまな有益なコメントを頂戴した。日本評論社の編集者である尾崎大輔、杉田壮一郎の両氏には、資料の準備、文章の推敲、校正などのさまざまな面でお世話になった。彼・彼女らにここで深く感謝を捧げたい。その結果としてできあがった本書の内容の文責は、すべて川口・澤田の両著者にある。

2024年5月

川口康平・澤田真行

目次

序章	経済学の因果推論アプローチ	1
1	本書のねらいと構成	1
2	「経済学詛りの強い」因果推論	3
3	なぜ潜在結果モデルで一貫させるのか？	7
第 1 部 因果推論の基礎と無作為化実験		
第 1 章	潜在結果モデルと因果関係	13
	イントロダクション	13
1.1	潜在結果	14
1.2	因果関係	14
1.3	顕在結果	16
1.4	因果推論の根本問題	17
1.5	SUTVA	18
1.6	2 項処置モデル	20
1.7	処置割当メカニズム	20
1.8	抽出・割当に起因する不確実性	23
1.9	興味のあるパラメータ	26
1.10	パラメータの識別	27
	1.10.1 識別が成立するケース	27
	1.10.2 識別が成立しないケース	31
	1.10.3 識別の一般的な定義	33
	文献ガイド	33
第 2 章	無作為化実験	35
	イントロダクション	35

2.1	実験データ、観察データ、自然実験、疑似実験	36
2.2	無作為化実験とは？	38
2.3	フィッシャーの p 値	41
2.4	平均処置効果の推定	44
2.4.1	割当に起因する不確実性のみの場合	44
2.4.2	抽出に起因する不確実性もある場合	47
2.5	共変量の「統制」と回帰分析	49
2.5.1	線形回帰モデルと潜在結果モデルの関係	50
2.5.2	線形回帰モデルの最小 2 乗推定量の性質	51
2.5.3	共変量を含む線形回帰モデル	56
2.6	共変量を利用した実験デザイン	59
2.6.1	層化無作為化実験	59
2.6.2	クラスター化無作為化実験	63
	文献ガイド	65
<hr/>		
第 3 章	推測・検定の諸問題 クラスター相関と多重検定問題への対処	67
	イントロダクション	67
3.1	クラスター相関とクラスター頑健分散推定量	68
3.1.1	クラスター頑健分散推定量	68
3.1.2	クラスター頑健分散推定量の少数クラスターに対する補正	70
3.1.3	なぜブートストラップはうまくいくのか？	72
3.1.4	少数クラスターに対するワイルドブートストラップ法	73
3.1.5	母集団のクラスターの数が有限の場合	76
3.1.6	2 段階クラスターブートストラップ	77
3.1.7	2 方向クラスタリング	78
3.2	多重検定問題とその対処	78
3.2.1	多重検定問題とは？	78
3.2.2	FWER を制御する検定手法	80
3.2.3	ボンフェローニ検定の改善手法	82
3.2.4	偽検出率制御による検定手法	87
	文献ガイド	90
<hr/>		
第 4 章	非遵守者	91
	イントロダクション	91
4.1	非遵守者がいる場合の潜在結果モデル	92

4.1.1	処置の受取を考慮した潜在結果	92
4.1.2	平均処置効果	94
4.1.3	処置割当効果	95
4.2	片側非遵守者	97
4.3	両側非遵守者	103
4.4	局所平均処置効果と操作変数法の関連	107
	文献ガイド	110

第 5 章 無作為化実験の実践 113

	イントロダクション	113
5.1	職業訓練は労働者の雇用と収入向上に寄与するか？	114
5.1.1	研究の背景と目的	114
5.1.2	研究デザインとデータ	114
5.1.3	推定方法とその留意点	115
5.1.4	推定結果	118
5.2	与信枠の設定による柔軟な貸付は零細経営者を助けるか？	119
5.2.1	研究の背景と目的	119
5.2.2	研究デザインとデータ	120
5.2.3	推定方法	122
5.2.4	推定結果	123
5.2.5	追加的な分析	124
5.3	実験の事前登録制度と事前査読制度	125
5.3.1	事前登録制度	125
5.3.2	事前査読制度	127
5.4	情報提供によって健康被害の少ない燃料の利用を促進できるか？	127
5.4.1	研究の背景と目的	127
5.4.2	研究デザインとデータ	128
5.4.3	推定方法	130
5.4.4	推定結果	131
5.4.5	事前査読制度の観点からの評価	132
	文献ガイド	134

第 II 部 疑似実験の因果推論

第 6 章	回帰非連続デザインの基礎	141
	イントロダクション	141
6.1	回帰非連続デザインとは？	142
6.1.1	「合格点をとれるかどうかは運次第」を使ったデザイン	142
6.1.2	回帰非連続デザインにおける識別問題と識別が成立する条件	143
6.1.3	条件付き期待値の連続性による識別	146
6.2	無作為化と対応する近傍無作為化	147
6.2.1	近傍無作為化の帰結	148
6.2.2	近傍無作為化における識別	150
6.3	パラメトリック推定の問題点	151
6.4	局所線形推定量による回帰非連続推定	155
6.4.1	局所近似による推定	155
6.4.2	「最適」な推定量の選択とは？	157
6.4.3	カーネル推定量から局所線形推定量へ	159
6.5	回帰非連続推定量のためのパラメータ選択と統計的推測	161
6.5.1	最適な推定のためのパラメータ選択	161
6.5.2	回帰非連続推定量としてのバンド幅選択	163
6.5.3	信頼区間の構成とバイアス除去	167
6.6	回帰非連続デザインの検証プロセス	170
6.6.1	密度検定	171
6.6.2	共変量バランス検定、プラセボ検定	173
	文献ガイド	174
第 7 章	回帰非連続デザインの発展的トピック	175
	イントロダクション	175
7.1	ファジー回帰非連続デザイン	175
7.1.1	ファジー回帰非連続デザインとは？	175
7.1.2	ファジー回帰非連続デザインの検定可能制約	180
7.2	回帰屈折デザイン	181
7.3	離散スコアへの対応	185
7.3.1	関数集合とその関数集合に関して「正直な」信頼区間	186
7.3.2	ヘルダークラス関数に対して「正直な」信頼区間の構成	188
7.3.3	既存の対処とその問題点	189

文献ガイド	190
第 8 章 回帰非連続デザインの実践	191
イントロダクション	191
8.1 滞在の合法化は移民の犯罪率を下げるか?	192
8.1.1 研究の背景と目的	192
8.1.2 研究デザインとデータ	192
8.1.3 研究デザインの補足情報	194
8.1.4 推定方法とその留意点	195
8.1.5 推定結果	196
8.1.6 頑健性チェックと妥当性チェック	199
8.1.7 補足的な議論	202
8.2 雇用における割当の効果とその代償は?	204
8.2.1 研究の背景と目的	204
8.2.2 研究デザインとデータ	205
8.2.3 推定方法とその留意点	207
8.2.4 推定結果	207
8.3 義務教育期間を延長すると年取は増えるか?	210
8.3.1 研究の背景と目的	210
8.3.2 研究デザインとデータ	211
8.3.3 推定方法とその留意点	212
8.3.4 推定結果	213
文献ガイド	214
第 9 章 差の差法の基礎	217
イントロダクション	217
9.1 疑似実験の評価としての差の差法	218
9.2 差の差法の基本的発想と識別	220
9.3 差の差法の回帰推定量	224
9.4 共変量と条件付き平行トレンドの仮定	225
9.4.1 条件付き独立の仮定における推定問題	229
9.4.2 1 期間のデータにおける逆確率重み付け推定量	229
9.4.3 1 期間のデータにおける 2 重頑健推定量	231
9.4.4 差の差法における逆確率重み付け推定量	233
9.5 処置群の数が限られる場合における手法：合成コントロール法	235

文献ガイド	241
第 10 章 差の差法とその周辺の発展的トピック	243
イントロダクション	243
10.1 より複雑な処置タイミングを伴う差の差法	244
10.1.1 2方向固定効果回帰における処置群平均処置効果	244
10.1.2 2方向固定効果回帰の問題点	246
10.2 異なる時点で処置割当が生じる場合の分解・集計手法	251
10.3 差の差法の検証とその他の留意事項	254
10.3.1 事前トレンドの検定	254
10.3.2 一様検定の必要性	255
10.3.3 時間に応じて変化する共変量の扱い	256
10.4 複数の処置群に対する合成コントロール法	257
10.4.1 処置群が複数の場合の合成コントロール法	257
10.4.2 合成差の差法	259
10.5 ファジー差の差法	262
文献ガイド	266
第 11 章 差の差法の実践	267
イントロダクション	267
11.1 病院合併による雇用集中度上昇は賃金成長率を低下させるか？	268
11.1.1 研究の背景と目的	268
11.1.2 研究デザインとデータ	269
11.1.3 推定方法とその留意点	271
11.1.4 推定結果	273
11.1.5 妥当性チェック	274
11.1.6 頑健性チェック	275
11.2 住宅シェアリングは不動産投資を促進するか？	276
11.2.1 研究の背景と目的	276
11.2.2 研究デザインとデータ	277
11.2.3 推定方法	278
11.2.4 推定結果	278
11.2.5 頑健性チェック	279
11.3 マルクスの地位はロシア革命あつてのものか？	280
11.3.1 研究の背景と目的	280

11.3.2	研究デザインとデータ	281
11.3.3	推定結果	283
11.3.4	頑健性チェック	285
11.3.5	妥当性チェック	286
	文献ガイド	286
	参考文献	289
	索引	300

序

章

経済学の因果推論アプローチ

1 本書のねらいと構成

本書は、主に学部上級生から大学院生の読者を想定した、計量経済学における因果推論を学ぶための教科書である。一貫して潜在結果モデル (potential outcome model) に基づいて因果推論の議論を進めること、および直観的な議論と数学的に厳密な議論の両方を提示することの2点に留意して執筆した。紹介する手法は、「無作為化実験」「回帰非連続デザイン」「差の差法」など、経済学の実証研究でよく用いられるものに限定している。ただし、その基本的な理論にとどまらず、最新の实証研究でよく用いられる発展的な方法や、その実践についても詳しく解説している。また、「どのような仮定のもとで、因果関係についてどのような結論が得られるか」という理論を解説するだけでなく、それらの理論を使って分析を進める際に直面するであろう「その仮定の確からしさをどのように判断すればよいのか」「因果推論の結果をどう解釈すればよいのか」といった実践的な課題についても、最新の实証研究を紹介しつつ議論している。

本書は、2つの部・11の章で構成されている。第1～5章からなる第I部では、潜在結果モデルを導入したうえで、理想的な無作為化実験 (randomized experiment) のもとでの因果推論について議論する。第1章では、本書の基礎となる潜在結果モデルに基づく因果推論のフレームワークを導入する。続く2～5章では、主に無作為化実験を扱う。第2章では、さまざまな処置割当メカニズム、特に無作為化実験と呼ばれる処置割当メカニズムのもとでの平均処置効

果の推定について解説する。第3章では、そこまでの章で言及される多重検定問題やクラスター頑健標準誤差など、統計的推測に関するいくつかの論点を詳しく紹介する。第4章では、無作為化実験を行うだけでは推定の目的が達成できない場合、すなわち**非遵守者** (noncomplier) と呼ばれる、分析者の想定通りに処置に従わない者がデータに含まれる場合に、分析者が関心の対象とする処置割当効果や局所平均処置効果などのパラメータ (推定対象) を適切に推定するための仮定について議論する。第5章では、無作為化実験を用いた実証研究に基づいて、非遵守者の存在もふまえた仮定の妥当性の判断や、結論の解釈の仕方について論じる。

第6～11章からなる第II部では、理想的な無作為化実験の想定からは逸脱しているものの、制度の知識などに基づいて実験が持つ一部の望ましい性質が成り立っていることを正当化できそうな状況、すなわち**疑似実験** (quasi experiment) のもとでの因果推論について議論する。第6～8章では、**回帰非連続デザイン** (regression discontinuity design: RDD) について議論する。まず、第6章において回帰非連続デザインの基本的な理論を導入する。そのうえで、第7章ではファジー回帰非連続デザイン、スコア操作の可能性への対応、離散スコアへの対応などといった発展的な内容を解説する。それらをふまえて、第8章では回帰非連続デザインの実証研究を紹介しつつ、実際に分析を進めるうえで直面しうる課題と、それらへの最新の対処方法、仮定の妥当性や議論の頑健性のチェックなどについて議論する。

第9～11章では、近年も新しい論点や手法が登場し続けている**差の差法** (difference in differences: DID) について議論する。第9章では、差の差法の基本的な理論を導入する。第10章では、異なる時点で処置の割当が生じる場合を中心に、差の差法、およびそれに関連する手法に関する近年の理論的発展を解説する。第11章では、差の差法を使った最新の実証研究を紹介しつつ、実際の分析の進め方を解説する。

QR



本書のサポート用 GitHub リポジトリ (https://github.com/keisemi/EconometriciansGuide_CausalInference) では、本書で紹介した手法を実装するための R のサンプルコード や、数学的な証明に関する「テクニカルノート」などを別途提供している。本書の記述のどの部分がサンプルコードやテクニカルノートのどの部分に対応しているかは、本書の中でその都度、R コードの

対応箇所には **R**、テクニカルノートの対応箇所には **T** を添えて記載している。

2 「経済学訛りの強い」因果推論

近年、因果推論に関するさまざまな教科書が出版されている。経済学では Angrist and Pischke (2009) や Cunningham (2021) が有名だが、本書ではこれらの本よりも最新の手法を多く盛り込み、数学的な導出もより丁寧かつ厳密に行っている。その他にも計量経済学の教科書で因果推論を扱ったものはいろいろあるものの、標準的な教科書はどれも後述する**構造推定** (structural estimation) アプローチの伝統の中に潜在結果モデルに基づく因果推論を埋め込むような書き方になっている。しかし筆者らは、本書のように潜在結果モデルだけに依拠して議論を進めたほうが初学者にとっては理解しやすいのではないかと考えている。その理由については、本章の第3節で学説史上および教育上の意義もふまえて詳しく議論する。

経済学以外の分野の研究者が書いた教科書としては、たとえば社会学者による Morgan and Winship (2014) や、疫学・生物統計学者による Hernán and Robins (2023) が挙げられる。これらの教科書では、経済学ではあまり用いられない考え方や手法や応用例を学ぶことができる。特に、Pearl (2001) によって導入された「有向非巡回グラフ (directed acyclic graph: DAG)」を用いた因果効果の表現、有向非巡回グラフと潜在結果モデルの関係や、それらに基づいた因果推論の考え方が詳しく解説されている。統計学者による Rosenbaum (2017) や、計量経済学者と統計学者の共著である Imbens and Rubin (2015) も、潜在結果アプローチに基づく因果推論を幅広く厳密に学ぶことのできる教科書である。しかし、これらの教科書は、それぞれの分野の関心に基づいて書かれているため、経済学でよく用いられる手法を十分に学ぶことができない。回帰非連続デザインや差の差法といった経済学で頻繁に用いられる手法に関する議論が特に弱い。この傾向は、計量経済学者であるインベンス (Guido W. Imbens) が執筆した Imbens and Rubin (2015) についても当てはまる。

それに対して、本書は経済学が直面する問題に対して、経済学者が納得する

形で答えるための、極めて「経済学訛りの強い」因果推論を学ぶための本である。

異なる学問分野間で因果関係の捉え方や手法の選択に「訛り」が出る根本的な理由は、対象としている問題がそれぞれで異なるからである。経済学はおおむね、「ある制度やルールの中で、個人が将来や他人の行動を予測しつつ、自分にとって望ましい結果を得るべく意思決定している状況で、特定の仕方では制度やルールや個人に介入すると何が起こるかを予測する」という課題に答えることを目的としている。これらの問題に対する理論の検証や介入の効果の測定には、他の分野の検証や測定とは異なる要因の考慮が必要となる。たとえば、個人が将来や他人の行動を予測するし、その個人にとって望ましい結果を得ようとする、その結果として処置が選択されたり、処置に従わなかったりする、といった分野固有の事情を織り込む必要がある。

経済学では、フリッシュ (Ragnar A. K. Frisch) とともに 1969 年に最初のノーベル経済学賞を受賞した計量経済学者のティンバーゲン (Jan Tinbergen) 以来、このような問題に答えるための適切な方法は何かということが模索されてきた (Tinbergen, 1930)。そのために伝統的に採用されてきたのは、1989 年のノーベル経済学賞受賞者であるハーヴェルモ (Trygve M. Haavelmo) が提唱した「構造推定」という方法である (Haavelmo, 1943)。これは「まずその個人の意思決定や相互作用を経済理論で定式化し、その定式化の中のパラメータをデータから推定し、特定の処置の効果を定式化の一部を変化させる反実仮想 (counterfactual) として求める」という方法である。

このアプローチでは、意思決定や相互作用の複雑さゆえに、複数の経済理論の定式化やパラメータが観測と整合的になってしまい、それらを互いに区別できない観測同値性 (observational equivalence) という問題が生じる。そのため、1930~40 年代から、どのような制約や特殊な観測があれば経済理論の定式化が一意に定まるかという問題、すなわち識別 (identification) の問題が盛んに議論されてきた。このように構造推定アプローチでは、経済理論に基づいて設定された関心の対象の識別の成立が自明ではない状況を出発点として、そこにどのような制約を置けば識別が達成されるかが議論の焦点となってきた。

しかし、過去 30 年の間に、経済学者の因果推論に関する考え方は大幅に変化した。その 1 つのきっかけは、LaLonde (1986) だといわれている。この論文では、不利な立場にある労働者に対して提供された職業訓練がその後の所得

に与える影響をフィールド実験で推定した場合と、それまでに知られていた計量経済学の非実験的手法を用いた推定結果の間に大きな乖離があることを示した。それ以来、経済学の対象領域において因果推論を行う場合は、実験あるいは実験に近い性質を制度などに関する客観的な情報によって正当化できる「自然実験」や「疑似実験」に着目する必要があるのではないかと考える経済学者が増えていった。この変化を**信頼性革命** (credibility revolution) という (Angrist and Pischke, 2010)。

本書で最新の実証研究でよく用いられる発展的な内容について詳しく解説することにしたのは、まさにこうした状況に対応するためである。経済学の論文で求められる議論の仕方は日進月歩で変化している。論文が書かれるたびに、この分野固有の新しい問題が認知され、その問題に対処するための方法が提案され、研究者の間で周知され、受け入れられ、広まり、一部は定着し、一部は廃れていくからである。詳しくは後の章で議論するが、たとえば、回帰非連続デザインを局所線形モデルで推定するとき、推定の対象となる範囲を定めるためのバンド幅はどのように選べばよいだろうか？ また、処置群と統制群を分けるスコア変数が離散的なときはどのように対処すべきだろうか？ 差の差法で因果効果を推定しようとするとき、処置が個体間で一斉に行われるのではなくタイミングがずれている場合にはどのような定式化を用いるべきだろうか？ また、なぜそのような方法をとるべきかを説得的に説明できるだろうか？

学術研究は研究者のコミュニティで進められるものである。論文を書いて認められるためには、コミュニティの他のメンバーと対話して説得しなければならないが、その際に最新の実践に通じていなければ他のメンバーを説得するのは難しい。その際に、コミュニティの最新の実践を「単なるお作法」として軽んじるべきではない。あるお作法の背後には、必ずそのお作法が生まれるに至った問題がある。そのお作法は、その問題に対するその時点での最良の対応であるとみなされているからこそ、コミュニティで採用されているのである。そのため、お作法に従うにせよ退けるにせよ、自分がそれをなぜ、どのように扱うかを明確に説明できなければ、説得的な議論はできない。

本書で、無作為化実験、回帰非連続デザイン、差の差法のそれぞれにおいて、実証研究の詳細な解説を通じて、仮定の妥当性についてどう判断すればよいのか、結果をどう解釈すればよいのかについて議論することにしたのも、同様の

問題意識に基づいている。

理論は、仮定が満たされる限りは正しい結論を与えてくれる。問題は、多くの場合その仮定が満たされるか否かが定かではないことである。そのため、実証研究では仮定から結論を導く論理的な「推論」とは別に、仮定や解釈の妥当性に関する「判断」が必要になってくる。たとえば、非遵守者が存在する無作為化実験で局所平均処置効果を識別するためには、処置が除外制約を満たしている必要がある。回帰非連続デザインでは、処置が非連続に変化する点で潜在結果の回帰関数が連続的である必要がある。分析を行ううえで、これらの理論が満たしているべき仮定が確からしいか否かをどのように判断すればよいだろうか？

出版された論文を読むと、実験や制度のデザインに依拠してこれらの仮定が確からしいことを主張したり、補足的な分析を行うことでその確からしさの傍証を行ったりすることに多大な紙幅が割かれていることに気付く。こうした議論の仕方もまた、単なる「お作法」だと捉えることもできる。しかし、推論の仮定が常に成り立つことはありえない以上、多くのケースで仮定の妥当性に関する判断が必要となるし、そのための基準も必要である。そうした判断基準を育むことこそが、コミュニティの重要な役割である。本書ではこうした点についても、実践的に学べるようになっていく。

このように、本書は意図的に「経済学訛りの強い」因果推論の教科書として書かれてはいるが、筆者たちは経済学に直接関心のない人にとっても有益な知見を提供できると考えている。経済学が直面するような問題には、経済学を学ぶつもりのない人でも直面することがあり、その際には本書で議論するようなアプローチが必要とされるためだ。仮定の妥当性を判断したり結果の解釈を行ったりする際の議論の進め方は、データ分析の結果をマネージャーや顧客に的確に伝えることが求められるビジネスの現場でも重要である。説明責任が強く求められる政策の現場であれば一層重要となるだろう。

計量経済学は、人間社会の相互作用に関する実証的な分析を行うための手法を開発する学問である。経済学自体に興味があろうとなかろうと、同じような問題について考えている人は、一度はこの極めて「経済学訛りの強い」因果推論の本を学んでみるべきである。経済学者である筆者たちが「疫学訛りの強い」書籍である Hernán and Robins (2023) を読んで楽しめるように、同じ概念に

ついて少し違った視点から学んでみるのもなかなか楽しいのではないだろうか。

3 なぜ潜在結果モデルで一貫させるのか？

最後に、本書で「一貫して潜在結果モデルの枠組みで議論を進める」ということの経済学の学説史上、および教育上の意義を述べておきたい。

古典的な計量経済学の教科書(たとえば Wooldridge, 2010)では、線形モデルの最小2乗推定量の議論から始まって、同時方程式の操作変数法(instrumental variable method)、一般化積率法(generalized method of moments: GMM)による同時方程式の推定の議論などに進んでいくことが多い。今となつては明示的に意識されることは少ないが、実はこの背後には先に紹介した Haavelmo (1943) 以来の構造推定アプローチが想定されている。

構造推定アプローチでは、まず経済理論を想定し、その均衡条件式を書き下す。これを**構造型**(structural form)の式と呼ぶ。次に、その式を理論の内です定まる「内生変数(endogenous variable)」について解き、理論の外です定まる「外生変数(exogenous variable)」だけが右辺に含まれた式を導出する。これを**誘導型**(reduced form)の式と呼ぶ(Hsiao, 1983; Matzkin, 2007)。誘導型の式のパラメータは、最小2乗法などで推定できる。このパラメータは、背後の構造型の式に含まれる経済理論のパラメータが組み合わさってできたものである。そこで推定された誘導型の式のパラメータから、それと整合的な経済理論のパラメータを一意に定められるかどうかを議論する。

一方、構造型の式をショック(需要ショック、生産性ショックなどの計量経済学者に観測されない外生変数)について解いて、その積率条件(moment condition)を利用して経済理論のパラメータを推定しようとする、一般化積率法を用いることになる。その際、構造型の式のうち興味のある式(たとえば需要関数)のみを定式化して、それ以外の式は定式化せずに経済理論のパラメータを推定しようとする、操作変数法を用いることになる。

標準的な計量経済学の教科書を用いた標準的な計量経済学の講義で学ぶ手法は、教わる学生にも教える教員にも必ずしも明確に意識されているわけではないが、こうした経済学特有の実証分析の歴史を前提とした、因果推論とはまた

別の「経済学訛りの強い」手法なのである。

構造推定自体は、今でも産業組織論を筆頭に、国際経済学、マクロ経済学などの経済学のさまざまな分野で用いられている。しかし、経済学における因果推論の議論の多くは、潜在結果アプローチという構造推定アプローチとはまったく異なる方法によって展開されることが多くなっている。こうしたアプローチの変化に伴って、経済学の実証研究における経済理論の役割は大きく変化した。問題は、それにもかかわらず両者を明示的に区別せずに因果推論が教えられている点にある。

構造推定アプローチの目的は、経済理論から書き下された経済モデルである構造型の式からモデルのパラメータを推定したり、構造型の式から導出された誘導型の式で経済理論の含意を検証したりすることである。まず経済理論があって、そこから推定されるべき統計モデルが導出されて、その統計モデルがデータに当てはまるか否かを検証する。そのようにしてデータに当てはまることが検証された経済理論が得られれば、次からはその経済理論から導出される統計モデルへの制約を前提としてデータの分析を行うというのが構造推定アプローチで想定されている段取りである。

しかし、潜在結果アプローチに基づいて統計モデルを推定することによって政策の効果を分析する際には、経済理論から導出された制約を利用することは恣意的で直接検証不可能な「関数形への制約」として退けられる。そこでは、できる限り実験や制度のデザインなどの経済理論の外側にある知識や情報から「導出」された直接検証可能な制約のみを利用することが求められる。まず実験や制度のデザインがあり、そこから統計モデルが導出されて、データに当てはめられる。

ここでは経済理論は、「このような結果が得られればこういうことを意味し、そうでなければ別のこういうことを意味する」という枠組みを、分析に先立って用意するために用いられる。このときに重要なのは、その知見は統計モデルや推定方法を導出するためには用いないという点である。そして、経済理論抜きにデータから明らかにできることを確定したうえで、結果の背後にあるメカニズムを「解釈」するために用意しておいた経済理論を参照する。これが、経済学における潜在結果アプローチの想定している段取りである。

なお、解釈の枠組みを提示するという役割がある以上は経済理論が前提となっ

ているのだと考えれば、これもあくまで構造推定アプローチの一種だといえなくもない。実際、そのように考えることでこの2つの伝統の交わりを理解している経済学者は少なくない。

加えて、潜在結果アプローチと構造推定アプローチの混同は、単純に術語の混乱も呼び起こしている。たとえば、潜在結果アプローチに基づく因果推論のことを「誘導型の推定」と呼ぶことがあるが、これは必ずしも正しくない。上記の通り、誘導型は特定のクラスの経済理論から導出されるものである一方、潜在結果モデルは必ずしも経済理論から導出されるわけではないからだ。

また、推定する式の中の説明変数と誤差項の間に相関があること全般を「内生性」と呼ぶことが慣習化しているが、そこではしばしば「その変数が経済モデルの均衡条件を解いて得られる内生変数であることの含意を指す概念であって、そうした経済モデル抜きの純粋な確率的な相関を指すものではない」ということも忘れ去られている。

本書で一貫して潜在結果モデルに基づいて因果推論の議論を進めるのは、この種の混乱を避けるためでもある。

* * *

以上、やや経済学内部の込み入った話に立ち入ってしまったが、本書は、人間の意思決定や相互作用を対象とした因果推論を成り立たせるために経済学が実践してきた試みを、できる限り網羅的かつ統一的に、最新の実践もふまえて解説すべく工夫した計量経済学の教科書となっている。本書をきっかけに、経済学的な実証研究やそのビジネス、政策などへの応用を志す方が1人でも増えてくれれば本望である。

■ 著者紹介

川口 康平 (かわぐち・こうへい)

香港科技大学商学院经济学院助理教授

2015年、ロンドン・スクール・オブ・エコノミクス経済学部にてPh.D.を取得。一橋大学大学院経済学研究科専任講師等を経て、2017年より現職。東京大学エコノミックコンサルティング株式会社アドバイザーも務める。専門は、実証産業組織論、計量マーケティング、マイクロ計量経済学で、ビジネスや政策の意思決定に関する問題について、フィールドでの無作為化実験、実務データを用いた構造推定から、オンラインサーベイ実験まで、さまざまな手法を用いた実証分析を幅広く手掛ける。

主 著：“When Will Workers Follow an Algorithm?: A Field Experiment with a Retail Business,” *Management Science*, 67(3), 2021, pp.1670–1695, “Designing Context-Based Marketing: Product Recommendations Under Time Pressure,” (共著) *Management Science*, 67(9), 2021, pp.5642–5659.

澤田 真行 (さわだ・まさゆき)

一橋大学経済研究所講師

2019年、イェール大学経済学部にてPh.D.を取得。2019年より現職。

専門は、計量経済学。なかでも、因果推論と、産業組織論や労働経済学などの応用マイクロ経済学において、手法開発と実証分析を行っている。学生時代に識別の概念に触れてからというもの、識別の面白さと奥深さに魅了されて今に至る。

主著：“Privatization and Productivity in China,” (共著) *RAND Journal of Economics*, 52 (4), 2021, pp.884–916, “Win/Loss Data and Consumer Switching Costs: Measuring Diversion Ratios and the Impact of Mergers,” (共著) *Journal of Industrial Economics*, 72 (1), 2024, pp.327–355.

いん が ずいろん けいりようけいざい がく 因果推論の計量経済学

2024年9月20日 第1版第1刷発行

著 者 ————— 川口康平・澤田真行

発行所 ————— 株式会社日本評論社

〒170-8474 東京都豊島区南大塚3-12-4

電 話 ————— (03)3987-8621[販売]

(03)3987-8595[編集]

印 刷 ————— 藤原印刷

製 本 ————— 難波製本

装 幀 ————— 図工ファイブ

JCOPY (社)出版者著作権管理機構委託出版物)

本書の無断複写は著作権法上での例外を除き禁じられています。

複写される場合は、そのつと事前に、(社)出版者著作権管理機構(電話03-5244-5088, FAX 03-5244-5089, e-mail:info@jcopy.or.jp)の許諾を得てください。また、本書を代行業者等の第三者に依頼してスキヤニング等の行為によりデジタル化することは、個人の家内での利用であっても、一切認められておりません。

©Kohei Kawaguchi, Masayuki Sawada 2024 Printed in Japan

ISBN978-4-535-54043-9