



RIETI Policy Discussion Paper Series 21-P-002

政策評価と不連続回帰分析(RDD)
- 前提条件、適用の限界と結果解釈の留意点 -

戒能 一成
経済産業研究所



Research Institute of Economy, Trade & Industry, IAA

独立行政法人経済産業研究所

<https://www.rieti.go.jp/jp/>

政策評価と不連続回帰分析(RDD)

- 前提条件、適用の限界と結果解釈の留意点 -

戒能 一成(C)*

- 要 旨 -

近年注目を集めている処置効果手法の1つに不連続回帰分析(RDD)があるが、これは非常に特殊な性質を持った評価手法であり、適用に際しての前提条件や結果解釈における留意点などが他の評価手法とは大きく異なっている。現状でこうした点が一般に理解されているとは言難く、政策評価への応用において問題を抱えた事例が散見される。

本研究においては、RDD に関する主要な先行研究を基礎として、RDD を政策評価の手法として用いる際の前提条件、手法としての限界と結果の解釈における留意点につき説明する。

具体的には、RDD の応用における限界の多くがその横断面分析性・局所平均処置効果評価性に起因していることを示す。また、こうした限界と前提条件に関連して変数の相対指標性への脆弱性、試料の秘匿・欠測への脆弱性、処置の従属性・内生性への脆弱性、「遵守者」の不均一な分布への脆弱性、処置の二次的影響による偏差への脆弱性など結果解釈における留意点について具体例を用いて議論する。更にこれらの問題に対処するための幾つかの方法や新たな方向性について紹介する。

本研究はこうした議論を通じて、RDD の政策評価への応用を支援し、手法としての発展に寄与しようとするものである。

Key-Words: 政策評価、不連続回帰分析(RDD)

JEL Classification: C21, C54

RIETI ポリシー・ディスカッション・ペーパーは、RIETI の研究に関連して作成され、政策をめぐる議論にタイムリーに貢献することを目的としています。論文に述べられている見解は執筆者個人の責任で発表するものであり、所属する組織及び（独）経済産業研究所としての見解を示すものではありません。

** 本研究中の分析・試算結果等は筆者個人の見解を示すものであって、筆者が現在所属する独立行政法人経済産業研究所、UNFCCC CDM Executive Board などの組織の見解を示すものではないことに注意ありたい。

政策評価と不連続回帰分析(RDD)
- 前提条件、適用の限界と結果解釈の留意点 -

- 目 次 -

1. 不連続回帰分析(RDD)の推計原理と前提条件	… 1
1.1 不連続回帰分析(RDD)の推計原理: 横断面分析性と局所平均処置効果評価性	… 1
1.2 不連続回帰分析(RDD)の前提条件	… 2
1.3 不連続回帰分析(RDD)の前提条件と一般の処置効果評価の前提条件の比較	… 3
2. 不連続回帰分析(RDD)の適用と限界	… 4
2.1 不連続回帰分析(RDD)の適用と限界についての先行研究	… 4
2.2 主に説明に用いる事例(第二次大戦中のフランスの分割統治)	… 5
2.3 横断面分析(CS)としての性質から来る限界	… 5
2.4 局所平均処置効果評価(LATE)としての性質から来る限界	… 6
3. 不連続回帰分析(RDD)と結果解釈の留意点	… 8
3.1 変数の相対指標性への脆弱性	… 8
3.2 試料の秘匿・欠測への脆弱性	… 8
3.3 処置の従属性・内生性への脆弱性	…10
3.4 「遵守者」の不均一な分布への脆弱性	…11
3.5 処置の二次的影響による偏差への脆弱性	…12
4. 結果のまとめと新たな方向性	…14
4.1 適用の限界と結果解釈の留意点のまとめ	…14
4.2 包括的な問題解決策としての CIC や Q-DID と新たな方向性	…14
参考文献	…15

2020年8月
戒能 一成(C)

1. 不連続回帰分析(RDD)の推計原理と前提条件

1.1 不連続回帰分析(RDD)の推計原理: 横断面分析(CS)性と局所平均処置効果評価(LATE)性

1.1.1 不連続回帰分析(RDD)の推計原理の概要^{*1}

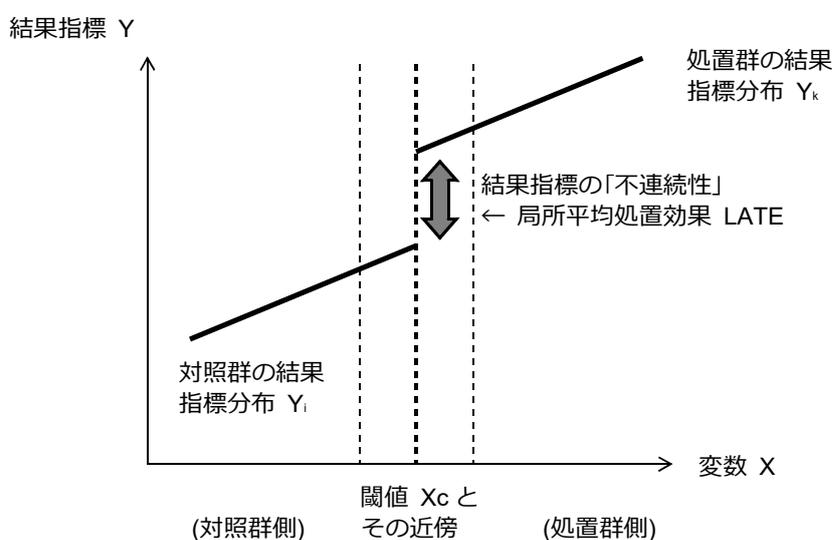
不連続回帰分析(RDD^{*2})とは、処置の選択がある変数 X に従い閾値 X_c の前後で分かれている場合に、 X の処置群側の区間と対照群側の区間での結果指標 Y の「不連続性」を回帰分析などを行って検出することにより、閾値 X_c の近傍での局所平均処置効果(LATE^{*3})を評価する手法である。

処置を実施するか否かが選択され処置群と対象群が分かれる X_c の近傍においては、処置が行われていない状態では処置群側の対象の結果指標 Y_k は対照群側の対象の結果指標 Y_i と同じであったと見なすことができる。この際に X_c の近傍にある処置群・対象群の対については、これらが「あたかもランダムに」処置群と対象群に選択されたものと見なすことができ、対照群の結果指標は処置群の結果指標の「仮想現実」であると見なすことができる。

このため処置が行われて X_c の近傍で両者の結果指標 Y_k と Y_i に差ができ不連続となった場合に、図 1.1 に示すとおりこの結果指標の不連続性を回帰分析などで統計的に検出できれば、 X_c 近傍での LATE を評価できることとなる。

ここで処置群・対象群の選択は X によって必ずしも完全に区分される必要はなく、 X_c を境界として処置を行う確率が変化している場合や、 X_c の前後で処置の選択が不完全である場合^{*4}であっても LATE の評価が可能である。

[図 1.1 RDD による処置効果の推計原理の概要]



1.1.2 不連続回帰分析(RDD)と横断面分析(CS)性・局所平均処置効果評価(LATE)性

RDD が一般の処置効果評価の手法と異なる点として、1)時間方向の情報を用いない横断面

*1 RDD の推計原理や応用事例については、Lee and Lamieux(2010)が先行研究調査による整理を行っている。本研究の 1.についての説明の多くは Lee and Lamieux(2010)を参考としている。

*2 Regression Discontinuity Design

*3 Local Average Treatment Effect

*4 Fuzzy Regression Discontinuity Design による分析がこれに該当する。本研究では RDD としてどのような手法を用いるかは本質的な問題ではないため、必要がなければ最も単純な Sharp regression Discontinuity Design について説明する。

分析(CS^{*5})の一種であること、2)評価結果は必ず X_c 近傍での LATE となることの 2 点が挙げられる。

(1) RDD の横断面分析(CS)性

RDD では X_c 近傍の処置群の対象の結果指標 Y_k について、仮に処置を受けなかった場合には X_c の反対側の近傍にある対照群の対象の結果指標 Y_l とほぼ同じであったことを前提条件(「 X に対する連続性条件」)とした上で推計を行う。

このため RDD では時間方向の情報を利用する必要がないとされており、処置後の 1 時点での試料を用いた横断面分析(CS)が行われる。

(2) RDD の局所平均処置効果評価(LATE)性

RDD では X_c から離れた場所での対照群が仮に処置を受けた場合の結果や、同様に X_c から離れた場所での処置群が仮に処置を受けなかった場合の結果を知ることができない。

また RDD では処置効果を多数の処置群・対象群の対象が X に従ってどのように分布しているか、その分布が X_s 近傍で不連続となっているか否かを基礎とした分析を行うため、処置群の中の個別対象についての処置効果を知ることができない。

このため RDD による処置効果の評価結果は必ず X_c 近傍での LATE となり、一般の処置効果評価とは異なり平均処置効果(ATE)、処置群平均処置効果(ATET)や処置群処置効果(TET^{*6})を推計することはできない。

1.2 不連続回帰分析(RDD)の前提条件

RDD を用いて偏差のない LATE を推計するための前提条件としては、Hahn, Todd and VanderKlaauw(2001)により以下の 3 つの前提条件の充足が必要とされている。

1.2.1 前提条件 1: 連続性条件

RDD では評価の対象とする処置群・対象群の対象は、処置を受けていない状態では X_c の近傍で X に従って連続的に分布していることが評価の前提である。これを「連続性条件」(前提条件 1)とする。この X に対する連続性条件により、 X_c の近傍では処置が行われていない状態では処置群側の対象の結果指標 Y_k は対照群側の対象の結果指標 Y_l と概ね同じであったと見なすことができる。このため処置群・対象群の対についてはこれらが「あたかもランダムに」選択されたものと見なすことができ、対照群の結果指標は処置群の結果指標の「仮想現実」であったと見なすことができる。

ここで X_c 近傍で処置群・対象群の対象が連続的に分布していることを担保するためには、必然的に X_c 近傍で処置群・対象群の対象の試料が十分に入手可能であることが必要となるが、この要件も連続性条件の一部として扱われる。

1.2.2 前提条件 2: 独立性条件

RDD で LATE を評価するためには、評価の対象とする処置群・対象群の対象が X に応じた処置実施の選択に従って、処置群・対象群に受動的に振分けられていることが必要である。

各対象が内生的選択によって X の値を自分で変化させるなど、処置群となるか対照群となるかを能動的に選択できる場合には、観察される X_c 近傍での不連続な差はこうした内生的選択の影響を受け、本来の処置効果と観察される不連続な差の大きさは一致しなくなる。

また処置群・対象群が何かの因子 Z に応じて決定されており、処置群・対象群への振分けが処置の実施・不実施に関する X の値により決定されていない場合も同様の問題が生じる。

*5 Cross Section (-Analysis)

*6 ATE: Average Treatment Effect, ATET: Average Treatment Effect on Treated, TET: Treatment Effect on Treated

こうした処置の選択に対する内生的選択や他の因子の影響がなく、処置の選択が処置群・対象群の対象の意志や特性などから独立であるとする条件を「独立性条件(前提条件 2)」とする。

1.2.3 前提条件 3: 「遵守者」の存在条件

RDD では処置が行われて処置群・対象群の対象の結果指標 Y_k と Y_i に X_c 近傍で不連続な差ができた場合に、この不連続な差を検出して処置効果と見なしている。しかし、この不連続な差と処置の関係についての因果を直接的に評価分析している訳ではないため、この不連続な差が処置に従って対象の行動が変化する「遵守者」⁷ についての効果であったことを前提条件とする必要がある。

この前提条件が成立せず、この不連続な差が処置に反対して対象の行動が変化する「違背者」についての効果である場合や「遵守者」より「違背者」が圧倒的に多い場合には、推計された不連続な差には処置効果との直接的な関係性がなくなってしまう。

この推計された不連続な差と処置との関係に関する前提条件を「「遵守者」の存在条件(前提条件 3)」とする。

1.3 不連続回帰分析(RDD)の前提条件と一般の処置効果評価の前提条件の比較

1.3.1 一般の処置効果評価の前提条件と RDD の前提条件

一般の処置効果評価では、例えば Rubin 因果モデル(RCM⁸)の 3 条件(CIA, OVLA, SUTVA)に系列相関の不存在性条件(NACA)を加えた 4 つの前提条件の充足を確認する必要がある。

1.2.1 で説明した連続性条件と独立正条件(前提条件 1 と 2)は、一般の処置効果評価における結果指標と処置の選択の独立性条件(CIA⁹)に相当するものと見なすことができる。一般の処置効果評価では、横断面分析(CS)の場合には実験的試料でのランダム化や統計的試料でのマッチングにより CIA を担保するが、RDD ではこれらの 2 つの前提条件が CIA を担保することとなる。

1.1.2 で説明したとおり、RDD は X_c の近傍での LATE を推計するため、評価分析に用いる試料としては X_s 近傍での試料の入手可能性条件と「遵守者」の存在条件(前提条件 1 の一部と前提条件 3)が充足されていればよい。このため、一般の処置効果評価では必ず確認が必要な処置群・対象群の同時存在性(OVLA¹⁰)の問題は発生しない。

また 1.1.2 で説明したとおり、RDD は横断面分析の一種であるため、一般の処置効果評価で問題となる系列相関の不存在性(NACA¹¹)の問題は発生しない。

1.3.2 処置効果の安定性条件(SUTVA)の問題の欠落

他方で Hahn, Todd and VanderKlaauw(2001)による 1.2.1 から 1.2.3 の 3 つの前提条件では、処置効果の安定性条件(SUTVA¹²)に関連する問題については何も触れられていない。

このため RDD による評価結果については、SUTVA に起因した偏差が含まれている可能性が指摘できる¹³。

*7 「遵守者」「違背者」などの概念は、RDD と同じ LATE を推計する操作変数(IV: Instrumental Variable)に関する Imbens and Angrist(2004)他の議論における概念を類推適用したものである。Imbens and Angrist(2004)他では遵守者は"Complier"、違背者は"Defier"とされている。

*8 Rubin Causality Model

*9 Conditional Independence Assumption

*10 Overlap Assumption

*11 No Auto-Correlation Assumption

*12 Stable Unit Treatment Value Assumption

*13 RDD における SUTVA に起因した偏差については、3.5 で議論する。

2. 不連続回帰分析(RDD)の適用と限界

2.1 不連続回帰分析(RDD)の適用と限界に関する先行研究

Hahn, Todd and VanderKlaauw(2001)は、処置効果が対象を問わず一定の場合や対象毎に異なる場合、処置効果の影響が完全である場合(Sharp RDD)や内生的選択などにより不完全である場合(Fuzzy RDD)に分類した上で処置効果の推計に必要な前提条件を議論している。当該議論の結果から、RDDにより閾値近傍での処置効果が正しく推計できるために必要な前提条件として、連続性、独立性及び「遵守者」の存在性の3つの前提条件を整理している。特にImbens and Angrist(2004)による操作変数のLATEに関する議論からの類推に基づき独立性と「遵守者」の存在が必要であることを詳しく論証している。他方でRDDによる処置効果の推計においては操作変数による推計と異なり、観察指標とその誤差や変数の間での独立性は問題とならないことを示している。

Brundell and Costa Dias(2009)は、処置効果評価などミクロ経済分野での計量分析に用いられる代表的手法を6分類し、各手法について前提条件・適用性・必要試料の3つの要件に着目した包括的評価を試みている。その中でRDDについては、閾値近傍での処置群・対照群の存在性、閾値近傍での処置・対照群の連続性、閾値近傍での処置効果の均一性の前提条件の充足が必要と説明している。RDDでは閾値近傍でのLATEの推計が可能であるが、閾値近傍での推計しかできない点や処置・対照群の内生的選択への脆弱性、閾値近傍の試料を多数用いる必要性などを指摘している。

Lee and Lemieux(2010)は、RDDの基本的な考え方と推計に必要な前提条件について整理するとともに、近年の主要な応用事例を分野・手法別に整理し先行研究調査を行っている。RDDは閾値近傍での処置・対照群の分布が「ランダム化に近い状況」にあることを利用して閾値直前・直後の結果指標Yの差を処置効果として推計する手法であり、DIDなど他の統計的(非実験的)な処置効果手法での前提条件などと比較して処置・対象群間でのCIAやOVLAの前提条件が必要ない反面で、連続分布性(連続性・関数形正確性)、独立性(非内生性)などの前提条件が充足される必要があるとしている。更に先行研究を基礎としてRDDの種類、マッチングや操作変数など他の推計手法との比較、推計結果の説明・表示手法と手順、分析事例と類型化・分類について概観し説明している。またRDDの今後の課題として1)処置に対する遵守者・不遵守捨・違背者別の挙動差と調査誤差の問題、2)処置・対照群の整序化・試料操作への脆弱性と内生的選択による異質な挙動の識別の問題の2つを指摘している。

Sekhon and Titiunik(2012)は、計量政治学の分野において自然現象・社会現象の不連続な変化が処置の有無に差異を生じる機会を利用した「自然実験」による評価手法について、RDDなどの手法を用いた「自然実験」に関する先行研究を4例引用した上でその問題点について指摘し改善策を提示している。このうちRDDに関連するものとしては、米国下院議員選挙で両政党の得票率が50%となる近傍でのRDDと現職有利度の研究(Lee(2008)・Butler(2009))、RDDで試料とする「接戦選挙区」での前回敗者側の「復讐」効果や敗北が予想される現職の「自主引退」効果(Caughey 他(2011))などによる内生的選択の可能性など評価分析に問題を生じる事例について議論している。

Athey and Imbens(2017)は、処置効果評価を中心とする最近の応用計量経済学分野の動向について概観している。このうちRDDについての応用事例について説明した上で、局所線形解析の必要性、閾値近傍での帯域選択の問題、LATEに起因した結果解釈の一般化困難性など方法論としてなお問題があり更なる研究が必要である点を指摘している。

他に一般的な解説としてCameron and Trivedi(2005)、Angrist and Pischke(2008)、Imbens and Wooldridge(2009)がRDDを取上げて説明している。

2.2 主に説明に用いる事例(第二次大戦中のフランスの分割統治)

2.2.1 Ferwarda and Miller(2014)による第二次大戦中のフランス分割統治の事例

本研究では、RDD の適用と限界などを議論するための具体的な題材として、Ferwarda and Miller(2014)による RDD を応用した分析の事例を取上げる。

Ferwarda and Miller(2014)は、アフガニスタンやイラクなど政治的分権が行われておらず戦時統制下にある地域への応用を念頭に、テロなどの抵抗運動と外部統治体制の関係について第二次世界大戦中のフランスを題材とした分析を行っている。具体的には、テロなどが多い地域と外部統治体制の強度の間には内生的関係があり直接的な識別ができない問題に対処するため、1940 ~ 42 年の間にフランスがナチス直轄地域とビシー政権統治地域に分割されたことを自然実験とした分析を行っている。分割は概ねリヨンからボルドーに至る線で行われ、パリやストラスブールを含む北側がナチス直轄地域となっていた。

2.2.2 Ferwarda and Miller(2014)による RDD を応用した分析

Ferwarda and Miller(2014)は、両統治地域の分割線で分断された郡であって国境に接しておらず両側に十分な数の市町村が含まれる 4 郡でのレジスタンスによる破壊活動・怠業発生件数の統計を用い、分割線を閾値とした RDD により比較して政治的分権の有無による抵抗運動と外部統治体制の関係を分析している。

最初にこれら 4 郡での市町村毎に鉄道駅からの距離・郵便局の有無・農地面積比率・人口密度などを変数とした比較を行い、分割線から両側 5km 以内の市町村については分割がランダムに行われていたと見なせることを確認している。その上で、破壊活動・怠業発生件数を RDD により比較し破壊活動は 43:15、怠業は 83:22 といずれもナチス直轄地域の市町村の方が有意に多くビシー政権地域、特に与党であった右翼政権支持層の多い市町村で有意に少なかったことを示している。

更に当該結果の頑健性を確認するため、両地域において統計的に見てランダムと見なせる分割線を設けて市町村を分割した偽薬試験や、分割線近傍での市町村のマッチングを用いた比較による検証・確認を実施している。

2.3 横断面分析(CS)としての性質から来る限界

2.3.1 RDD の横断面分析(CS)性と因果の問題

1.1.2 で説明したとおり、RDD は時間方向の情報を用いない横断面分析の一種であると位置づけられる。横断面分析であることは、試料が処置後の 1 時点でのみ入手できかつ 1.2 での 3 つの前提条件が充足されれば評価分析が可能である、という RDD の大きな利点となっている。

しかしその反面、横断面分析の一種であることは観察される効果と処置との間の因果性について直接的な立証が困難であるという、RDD を用いた評価分析の限界の原因となっている。

政策評価の手法として見た場合、この横断面分析性に起因した問題は致命的である。この問題は政策措置が「政策措置をしてもしなくても結果指標が増加(・減少)していた対象」に対して行われていた可能性が排除できないことを意味するからである。

2.3.2 処置前から不連続な差が存在した可能性

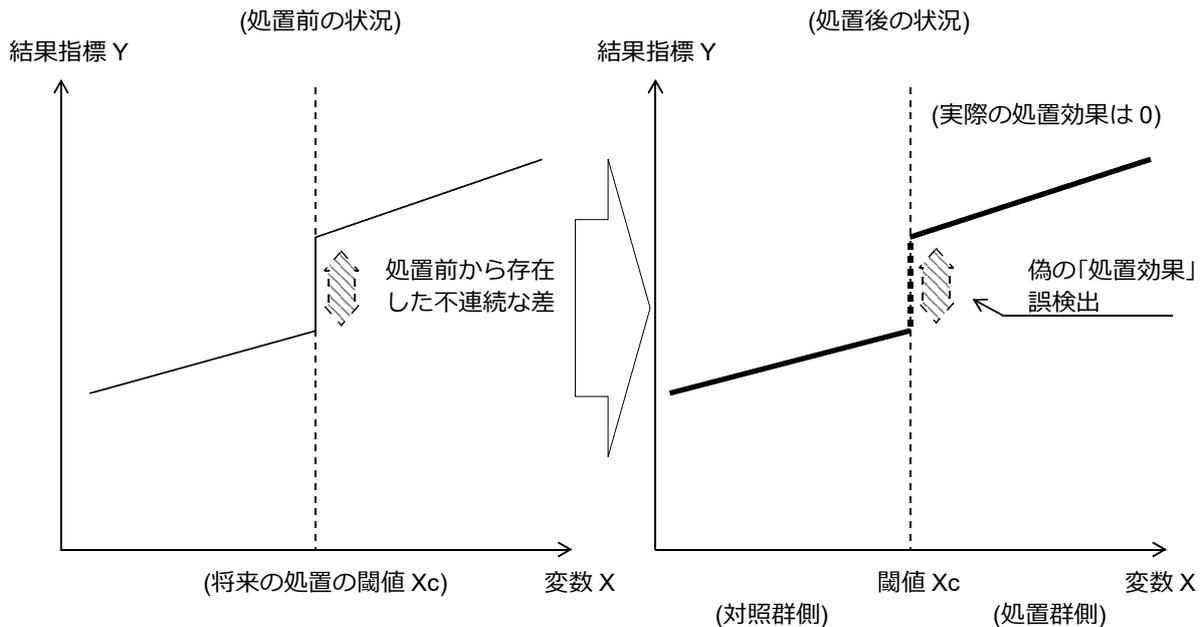
Ferwarda and Miller(2014)の事例では、第二次世界大戦中のナチス直轄地域とビシー政権地域への分割を処置と見なし、分割線近傍での破壊活動などの発生件数の比較を行っている。

しかし当該比較の限界として、分割前からナチス直轄地域とビシー政権地域で、犯罪や怠業の発生程度など「潜在的な破壊活動」の水準にそもそも差があり、分割後にこの分割前から存在した差が「破壊活動」の不連続な差として観察された可能性がある。

この問題は図 2.1 に示すとおり、処置前の時点での情報を利用せず、処置後の時点での横断面分析(CS)を行うという RDD の本質に起因したものである。

Ferwarda and Miller(2014)では、4 郡の市町村の経済的・社会的特性の比較から分割がランダムに行われたと見なせることの確認や、偽薬試験やマッチングによる推計を用いた検証・確認を行っている。しかし分割前の時点での状況についての試料を何も用いていないため、犯罪や怠業の発生程度など「潜在的な破壊活動」の水準にそもそも差があったとする説に対して反証ができない。

[図 2.1 処置前から不連続な差が存在し処置効果が誤検出となる場合]



2.3.3 考えられる対処

処置後の時点での横断面分析(CS)においては、処置前の時点での情報を利用しないため、「そもそも不連続な差が存在した可能性」を否定することは困難であると考えられる。

従ってこの事例では分割が行われた 1940 年以前での犯罪や怠業の発生程度など、これら 4 郡の市町村に関する「潜在的な破壊活動」の水準を示す試料を用い、横断面前後差分分析(DID^{*14})などを行って傍証とすることが考えられる。

2.4 局所平均処置効果評価(LATE)としての性質から来る限界

2.4.1 RDD の局所平均処置効果評価(LATE)性と敷衍・外挿の問題

1.1.2 で説明したとおり RDD は X_c 近傍での LATE を推計する手法であり、 X_c から離れた場所での処置効果については何の情報も得られない。

従って仮にある試料について RDD により明確な処置効果が観察されたとしても、同じ試料で X_c から離れた場所にある試料や、他の試料について敷衍・外挿して何かの情報が見られる訳ではない。

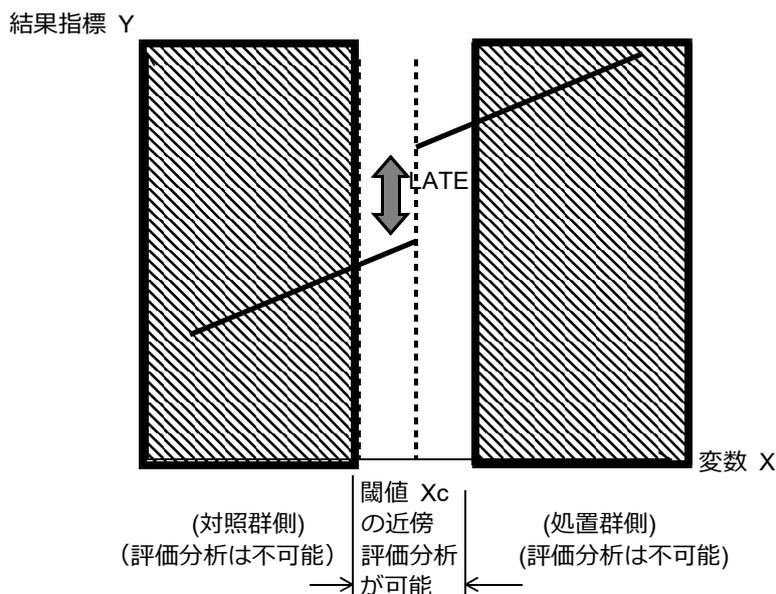
つまり図 2.2 に示すとおり、RDD による評価分析は X_c 近傍という非常に狭い範囲に限っては「処置効果が存在した」ことを論証することはできるが、 X_c から離れた場所については何

*14 Difference-In-Difference

の情報も得られないため評価分析が不可能なのである。

政策評価への応用という視点から見た場合、この RDD の LATE しか推計できないという性質は致命的ではないにせよ、手法としての適用が非常に限定されたものとなる。一般に政策評価においては政策の対象のうち限られた一部の対象で効果が見られるという「限定的な対象での効果の存在証明」が求められることは稀であり、通常は処置を行った対象全般について「平均的に見た効果の存在証明」が求められるためである。

[図 2.2 RDD の局所平均処置効果評価(LATE)性と評価分析ができる範囲]



2.4.2 分割線から離れたパリでは全く状況が違った可能性

Ferwarda and Miller(2014)の事例では、分割線近くの4郡の市町村で線の両側で差があったことを論証しているが、(現実がどうであったかは別として)分割線から遠く離れたパリやマルセイユでは全く状況が違っていた可能性が考えられる。

従って RDD による推計の性質上から言えば、この結果からは第二次世界大戦当時のフランスでの他の地域についてすら何も言えない訳であり、この結果を現在のアフガニスタンやイラクに対して直接敷衍するというのはそもそも不可能な試みである。勿論、何らかの政治的含意や示唆を得ることは不可能ではないが、そうした含意や示唆は処置効果の評価分析結果とは切分けて考えることが必要である。

この問題も X_c 近傍での LATE を推計するという、RDD の本質に起因した問題である。

2.4.3 考えられる対処

X_c 近傍での LATE を推計している限り、 X_c から離れた場所での処置効果については何の情報も得られない訳であり、「離れた場所では全く状況が違った可能性」を否定することは困難であると考えられる。

従ってこの事例では実際に Ferwarda and Miller(2014)が行ったとおり、フランス全土に無作為に仮想的な分割線を轆いた偽薬試験を行う、分割線から離れた他の市町村について社会的・経済的特性を示す試料を使ってマッチングを行った上で横断面前後差分析を行うなど、RDD とは別の手法を用いた検証・確認を行うことが考えられる。

3. 不連続回帰分析(RDD)と結果解釈の留意点

3.1 変数の相対指標性への脆弱性

3.1.1 変数 X や結果指標 Y の相対指標性の問題

1.2.1 で説明したとおり、RDD により処置効果を評価するためには連続性条件(前提条件 1)の充足が必要である。

しかし、処置群・対象群の対象が X に従って連続的に分布していることという条件が充足されるためには、 X_c 近傍で試料が十分に入手可能であることの他に、指標 X が絶対的指標^{*15}であることが必要である。同様に結果指標 Y についても絶対的指標であることが必要である。

X や Y が評価者や記入者の主観に左右されたり、評価・調査時点での周囲の状況などの要因が影響する相対的指標である場合には、仮に処置群・対象群が X に対して連続的に分布していたとしても、RDD により推計された Y の不連続な差である LATE には相対性に起因した偏差が入込んでいる可能性が考えられる。

一般の処置効果評価では、評価に際して処置群・対照群の境界を定義する X や X_c を用いないため、このような X の相対指標性に起因した問題は発生しない。また一般の処置効果評価では結果指標 Y の局所的な差を処置効果と見なすことはなく、評価対象全体での平均的な差を処置効果と見なすため Y の相対指標性に起因した問題も起きにくいと考えられる。

3.1.2 「破壊活動」の定義や数え方がナチス直轄地域とビシー政権地域で違った可能性

Ferwarda and Miller(2014)の事例では、分割線のナチス直轄地域側とビシー政権地域側で結果指標 Y である「破壊活動」の定義や数え方が違った場合、分割線近傍での比較を行った際の精度は大きく損なわれてしまい、比較すること自体が無意味なものとなってしまふ。

例えばナチスによる介入を嫌いビシー政権地域では破壊活動を少なめになるよう解釈して報告していた場合や、ナチス直轄地域では出先の部署が「取締りの実績作り」のためその逆を行っていた場合には、破壊活動の統計値に不連続な差があることは外部統治体制という処置の効果であるとは言えなくなってしまう。

こうした事例以外でも、例えばコンテストの順位やアンケートの回答など、試料が作成される際に審査員や記入者の主観や他の参加者との関係などの要因が入込む余地がある相対的指標を X や Y として用いることは、RDD の推計原理から見て避けるべきと考えられる。

3.1.3 考えられる対処

相対指標性の問題への対策としては、1.2.1 で説明した連続性条件の充足を確認する際に、X や Y が絶対的指標かこれに近いものであることを併せて確認すべきである。例えば、Ferwarda and Miller(2014)の事例では破壊活動といった広い定義で Y を比較するのではなく、殺人など定義が明瞭で主観が入る余地がない Y を用いて比較すべきである。

また上記のとおりコンテストの順位やアンケートの回答など絶対的指標とは見なせないものを RDD での X や Y として用いることは、同様の理由から避けるべきである。

3.2 試料の秘匿・欠測への脆弱性

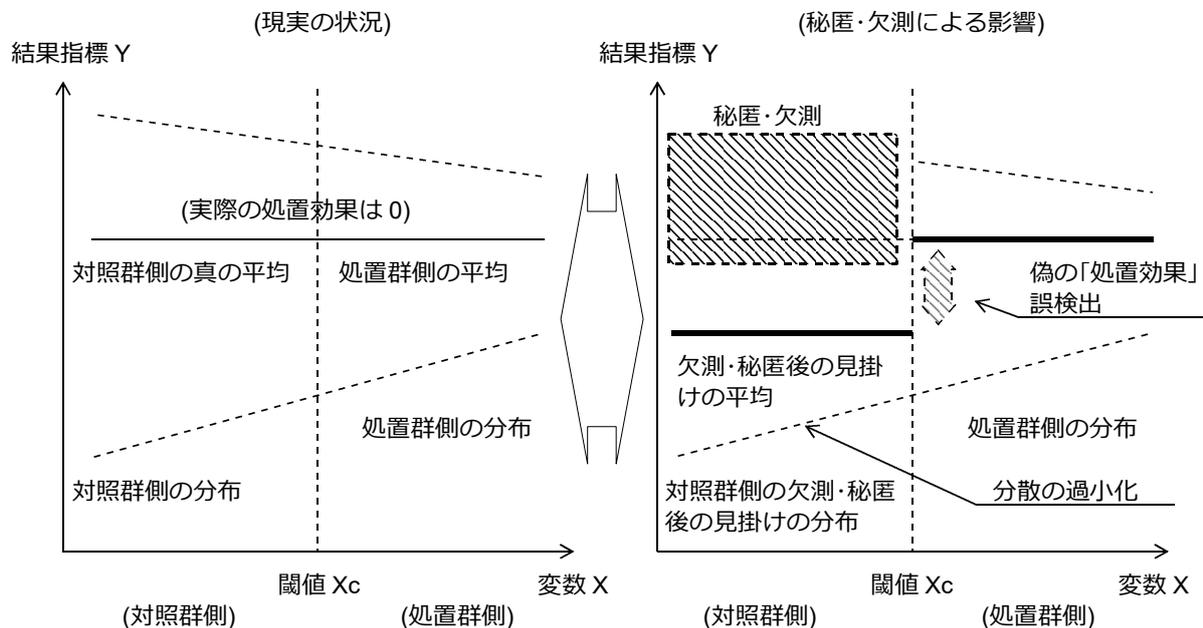
3.2.1 統計的試料における秘匿・欠測の問題

1.2.1 で説明したとおり、RDD により処置効果を評価するためには連続性条件(前提条件 1)の充足が必要であり、その一環として X_c 近傍で試料が十分に入手可能であることについて

*15 例えば駅伝であれば、区間タイムが絶対的指標であり、区間順位が相対的指標である。

説明した。しかし、 X_c の近傍かどうかにかかわらず RDD では試料が特定の領域で秘匿・欠測していないかどうかを確認することが必要である。

[図 3.1 統計的試料の秘匿・欠測による誤検出の可能性]



(1) 意図せざる秘匿・欠測の場合

RDD では処置群・対象群の対象の結果指標が X に対して連続的に分布していることを前提として不連続な差を検出しているため、特定の領域での秘匿・欠測により処置群側・対照群側での分布が歪み平均値が変わってしまう場合には、図 3.1 に示すとおり処置効果が誤検出される可能性がある。また分布の歪みに伴い分散も影響を受けるので、統計検定に用いる分散が過小となり誤検出が起きやすくなることも懸念される。

統計的試料で年収や健康状態を聞いた場合など X の回答に対し何らかの利害関係が存在する場合や、 X の分布の背後に政治信条、経済格差や組織活動など直接観察できない異質性が存在する場合は少なくない。こうした場合に、変数 X と結果指標 Y の特定の組合せに該当する対象が回答を返さないなどの理由から、秘匿・欠測の影響を受けて試料が本来の分布とは異なる分布を示す場合がある。

一般の処置効果評価では、処置群・対照群の同時存在性(OVLA)が前提条件として置かれているため、試料の秘匿・欠測は処置群・対照群の両方の試料が同時に減少するだけであり、処置効果の評価結果には大きな影響を及ぼさない。しかし RDD では処置群・対照群の分布が X_c を境に分かれているため、特定の領域での秘匿・欠測に対し評価結果が脆弱となってしまう。

(2) 意図的な秘匿・欠測などの場合

本来は絶対にあってはならない事であるが、Lee and Lamieux(2010)は分析者による恣意的な試料の整序化や取捨選択などの不正な試料操作に対しても RDD が脆弱であることを述べている^{*16}。本研究では細かく立ち入らないが、意図的な整序化や取捨選択などの不正な試料操作の影響についても概略上記の(1)の場合と同じと考えられる。

*16 通常の経済学系の論文では、自らの立場に不利な結果や内容を記載するなどという事は滅多に行われず、博士課程では論文の指導教官からこうした内容に触れないよう注意される。ところが Lee and Lamieux(2010)では RDD の問題点を不正行為の可能性を含めて包み隠さず整理して議論している。筆者はこうした両先生の科学的姿勢に大いに賛同し敬意を表するものである。

政策評価の手法として見た場合、この試料の秘匿・欠測への脆弱性の問題もまた致命的である。この問題は分析者が公的な統計的試料を用いた分析を行っていても、意図する・せざるに関わらず評価結果に誤検出の可能性が残ることを意味するからである。

3.2.2 「破壊活動」が両地域で組織的に秘匿されていた可能性

Ferwarda and Miller(2014)の事例では、ナチスによる政治的介入を嫌い、ビシー政権地域では怠業の大部分を「大目に見て」秘匿してしまい数えていなかった可能性が考えられる。逆にナチス直轄地域で大規模な破壊活動が続いた場合、出先の部署が監督責任を問われるためこれを秘匿していた可能性もある。

他の事例では、倒産防止の支援制度についての評価するため、これを受給した処置群と受給しなかった対照群の企業の売上をアンケート調査などで聞く場合を考える。対照群のうちそもそも業績優秀で受給資格のない企業は回答をせず欠測となり、将来受給の可能性のある業績不振の企業だけが熱心に回答を返してくる可能性がある。こうした問題を考慮せず欠測がある試料をそのまま RDD に掛けた場合には、図 3.1 のように対照群側の一部が欠測して分布が歪み、平均・分散が影響を受けて処置効果を誤検出する可能性がある。

3.2.3 考えられる対処

RDD を用いた評価分析において、試料の秘匿・欠測が評価結果に影響を与えていないことを担保するためには、特定の X と Y の組合せについての秘匿・欠測がなかった状態での試料の分布についての情報を使った評価が必要である。例えば処置前と比較して処置後の試料の分布が歪んでいないかどうかを確認するなどの方法が考えられる。

見方を変えれば、3.1 で述べたコンテストの順位やアンケートの回答などを用いる場合は言うに及ばず、公的統計調査であっても回収率が低く秘匿・欠測が多いことが知られているようなものから得られた試料を RDD の X や Y として用いることは避けるべきである。

また分析者による不正な試料操作についても、こうした対処が適切に採られているか否かを確認することによって、ある程度の判別ができると思われる。

3.3 処置の従属性・内生性への脆弱性

3.3.1 内生的選択による偏差の問題

1.2.1 で説明したとおり、RDD により処置効果を評価するためには、独立性条件(前提条件 2)の充足が必要である。

しかし現実の処置効果評価では、評価の対象とする処置群・対象群の対象が X に応じて受動的に振分けられているとは限らない。各対象が内生的選択によって処置群となるか対照群となるかを能動的に選択できる可能性がある場合には、この条件の充足について何らかの方法による検証・確認が必要である。

RDD の手法においても、Xc 近傍の限られた範囲で内生的選択により処置群・対照群の結果指標が変化する場合についてはこれを想定した手法が開発されている。しかし、こうした内生的選択が広範囲に及ぶ場合については対処が不可能であるため、この独立性条件が前提条件として置かれている訳である。

一般の処置効果評価では、処置の選択の独立性条件(CIA)を充足する手法としてランダム化・マッチングなどの処理により、処置群・対照群が能動的に処置の有無を別々に選択した可能性がなくなるよう、処置前の時点で結果指標や説明変数などに殆ど差がない処置群・対照群の対象を抽出して比較する。処置前に結果指標や説明変数に殆ど差がなければ、各対象には内生的選択を行い処置群・対照群に分かれる動機がないはずだからである。

しかし RDD では処置群・対照群が Xc を境に分かれており、時間方向での情報を用いない

ため、このような「処置前に結果指標や説明変数に殆ど差がない」処置群・対照群による比較は不可能であり、内生的選択に対して評価結果が脆弱となってしまふ。

政策評価の手法として見た場合、この処置の従属性・内生性への脆弱性の問題もまた致命的である。この問題は RDD により検出された処置効果が政策措置の結果ではない可能性を示唆するものであり、政策評価を行った意味自体が揺らいでしまうためである。

3.3.2 ナチスやレジスタンスが分割線や活動領域を内生的に決定していた可能性

Ferwarda and Miller(2014)の事例では、内生的選択については2つの可能性が考えられる。

(1) 変数 X , 閾値 X_c の内生的選択

ナチス直轄地域とビシー政権地域で、犯罪や怠業の発生程度など「潜在的な破壊活動」の水準に分割前から差があり、ナチス政権側がこうした水準に応じて内生的選択により分割線を引いていた可能性が考えられる。この場合には、2.1 の場合と同様に分割線の両側で破壊活動の水準に不連続な差があるのは当然であり、評価分析の結果は外部統治体制による処置効果とは言えない。

(2) 結果指標 Y の内生的選択

破壊活動を行う側のレジスタンスが、分割後には相対的に軍事的・政治的に重要な拠点の多いナチス直轄地域に対し内生的選択により活動を集中させていたことが考えられる。この場合には、分割線の両側で破壊活動の水準に不連続な差があるのは別の理由によるものであり、外部統治体制による処置効果とは言えない。

3.3.3 考えられる対処

1)の分割線が内生的選択により決定されていた可能性については、Ferwarda and Miller(2014)の分析手順にあるとおり、分割線がランダムと見なすことができ内生的選択の可能性がない地域・領域(この場合4郡が該当)を探した上で評価分析を行うべきである。またより直接的には、2.3 で述べたとおり分割前の試料を用いた横断面前後差分析などを傍証として利用すべきである。

他方で2)のレジスタンスが活動を行う地域が内生的選択により決定されていた可能性については、一連のFerwarda and Miller(2014)の分析では対処されていない様子である。従ってこうした内生的選択による評価結果への影響があった可能性については、結果を解釈する段階において考慮しなければならないと考えられる。

3.4 「遵守者」の不均一な分布への脆弱性

3.4.1 「遵守者」「違背者」による処置との因果の問題

1.2.1 で説明したとおり、RDD により処置効果を評価するためには、「遵守者」の存在条件(前提条件3)の充足が必要である。

RDD により結果指標 Y の不連続な差が X_c 付近で観察された場合に、これが処置に対して正常な反応を示す「遵守者」の行動によるものなのか、処置に対して通常とは反対の反応を示す「違背者」の行動によるものなのかは、外形的には区別することができない。

一般的な処置効果評価では、処置の選択の独立性条件(CIA)を充足する手法としてランダム化・マッチングなどの処理により、処置群の対象が処置に対して反対の行動をした可能性を含めて評価を行う。処置前の時点で結果指標や説明変数などに殆ど差がない処置群・対照群の対象を抽出して比較すれば、観察される処置効果は処置に対して正常に反応した対象と反対に反応した対象の結果指標の差の加重平均となっているはずである。

ところが RDD では処置群・対照群のうち X_c 近傍の試料を用い、時間方向での情報を用いないため、試料全体ではなく X_c 近傍での遵守者と違背者の分布が結果を左右してしまう可

能性がある。このため、遵守者の不均一な分布に対して評価結果が脆弱となってしまう。

3.4.2 破壊活動の大部分がナチス側による「自作自演」であった可能性

Ferwarda and Miller(2014)の事例では、ナチス直轄地域側での破壊活動の大部分が、レジスタンスへの協力者を減らし取締強化の口実を作るためのナチスによる「自作自演」であった場合が該当する。ここでナチスによる自作自演の破壊活動が違背者の行動に相当し、レジスタンスによる破壊活動が遵守者の行動に相当する。

この場合には、破壊活動という結果指標の不連続な差は Ferwarda and Miller(2014)の議論とは別の理由により生じたものであり、外部統治体制による処置効果ではあるものの Ferwarda and Miller(2014)の結論は正当化できないものとなってしまう。特に当時のレジスタンスに近い立場のアフガニスタンやイラクの旧体制派による破壊活動が大部分を占める地域について、この分析の結果から意味のある知見・示唆は何も得られないこととなる。

別の事例では、新聞社の世論調査や無記名のアンケート調査などで回答者が面白半分の本意と異なる回答を行う場合が散見されるが、これらも違背者の行動の一種である。こうした情報源からの試料を RDD で評価分析する場合も、この問題は避けられないものとなる。

3.4.3 考えられる対処

RDD による評価結果が遵守者のものであるか否かが疑われる場合には、Ferwarda and Miller(2014)の事例にあるとおり、鉄道や橋梁の破壊などの派手な活動だけでなく、違背者の行動としては極めて不自然な怠業などの地味な活動を分析対象に混ぜておくなど、複数の異質な結果指標を分析対象として設定しておくことが考えられる。

新聞社の世論調査や無記名のアンケート調査など一定の比率で違背者が混在することが避けられない場合、違背者を識別・排除できるよう設問を工夫するなど入念な措置を採った上で RDD を適用することが考えられる。あるいはより本質的には、こうした情報源からの試料は RDD 以外の別の手法を用いて評価分析を行うべきである。

3.5 処置の二次的影響による偏差の可能性

3.5.1 処置の二次的影響による偏差の問題

2.1 では RDD により処置効果を評価するために、連続性など 3 つの前提条件の充足が必要であることを Hahn, Todd and VanderKlaauw(2001)の議論に基づき説明した。

しかし Hahn, Todd and VanderKlaauw(2001)では正しく認識されていないが、1.3.2 で指摘したとおり、RDD においても SUTVA の問題に起因した偏差が生じる可能性が考えられる。

SUTVA の問題とは、処置群の対象に対する処置の効果が対照群の対象に対して二次的影響を及ぼし、結果として評価される処置効果が過大評価や過小評価となる問題である。

RDD の場合であれば、図 3.2 に示すとおり処置群の対象への処置の結果、対照群の対象にこれらの対象の間での代替・競合関係や人的交流・血縁関係などにより二次的影響が波及してしまい、評価分析した処置効果が偏差を含んだものとなってしまう場合が該当する。

SUTVA の問題は一般の処置効果評価でも起き得る厄介な問題であり RDD に限った問題ではないが、RDD では起きないものと誤解されている可能性があるため、敢えてここで注意喚起しておく。

3.5.2 占領政策によりレジスタンスの担い手の数に差ができた二次的影響の可能性

Ferwarda and Miller(2014)の事例では、SUTVA の問題については幾つかの可能性が考えられるが、レジスタンスの多くが占領前のフランス軍人・警察官であった点を例として考える。

ナチスによる占領開始に伴い、ナチス直轄地域では多くのフランス軍人・警察官が失職したが、彼らの一部は潜伏してレジスタンスとなり破壊活動を担っていたと考えられる。こうし

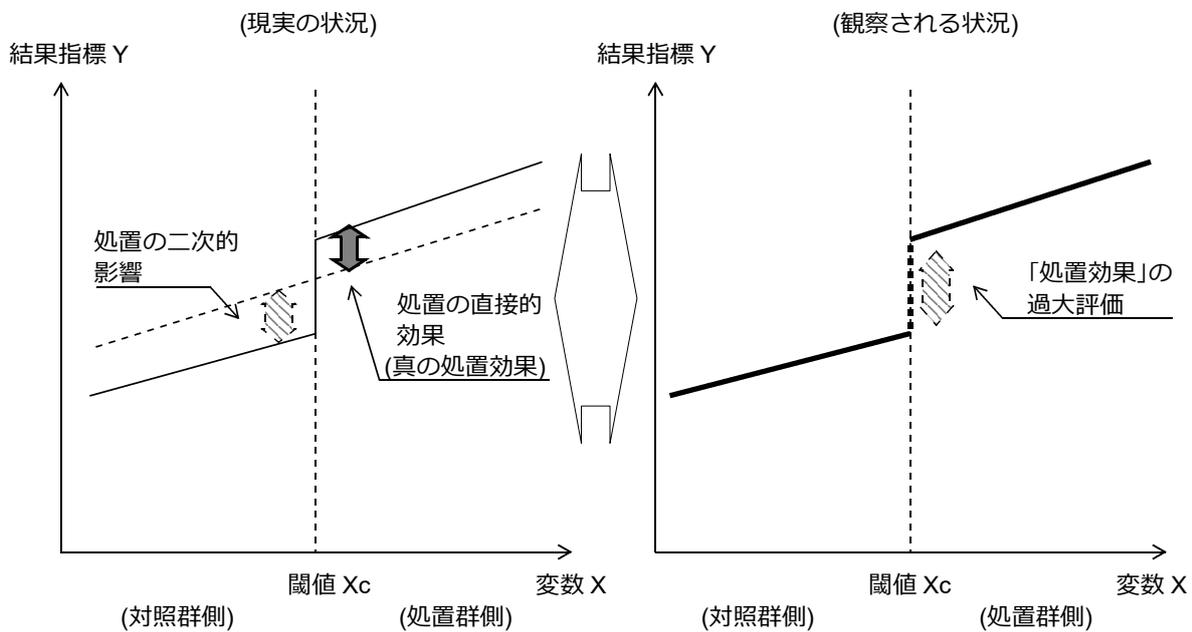
た失職はビシー政権地域では起きなかったため、そもそもレジスタンスの担い手の数が両地域で全く違ったと考えられる。

この場合には、同じナチス占領政策の影響ではあるものの、両地域での破壊活動の差は外部統治体制のみによる効果と解釈することはできず、レジスタンスの担い手の数の差というナチス占領政策の二次的影響が混在して過大評価された効果と考えなければならなくなる。

レジスタンス側が破壊活動の集中的・連続的实施を狙うべくビシー政権地域側からナチス直轄地域側に人員や武器を移送していた場合、その逆にナチス側が破壊活動を避けるべく重要な設備や組織・要人をナチス直轄地域側に疎開させていた場合などでも、同じような二次的影響の混在が起きて処置効果が過大評価されることとなる。

他方でナチス直轄地域側ではレジスタンスと疑われる者が無差別に逮捕・投獄され破壊活動が困難になっていた場合などでは、外部統治体制の差による処置効果は過小評価されることとなる。

[図 3.2 処置の二次的影響による偏差の問題 (過大評価の例)]



別の事例では、企業への研究開発助成を官公庁が審査により行う場合を考える。この場合に助成による効果には助成を受給した金銭的效果に加えて、官公庁による審査に合格した「名誉効果」など非金銭的效果が含まれているはずである。こうした非金銭的效果を考慮せずに効果の全部を金銭的效果と考えた場合には、政策効果を過大評価してしまうこととなる。また苦労して回帰分析などを行っても、助成の支給額と結果指標の増加分の間に有意な係数が見られない、などといった「怪現象」も起こり得ることとなる。

3.5.3 考えられる対処

処置の二次的影響など SUTVA の問題については、現状で横断面分析の場合に適用可能な手法は実験的試料を用いた Hudgens and Halloran(2006・2008)の「3群法」のみである。しかし第二次大戦中のフランス分割統治のような事例を実験で再現することは不可能であるため、この手法が適用できる場合は社会的・経済的に実験が可能な場合に限定される。

従ってこうした SUTVA の問題による処置効果への偏差があった可能性については、結果を解釈する段階において考慮しなければならないと考えられる。

4. 結果のまとめと今後の方向性

4.1 適用の限界と結果解釈の留意点のまとめ

2.と3.で説明したとおり、RDD においてはその横断面分析性と局所平均処置効果評価性という2つの性質が、この手法の利点であると同時にその適用の限界の原因となっている。

この2つの性質と3つの前提条件に起因・関連して、RDD では変数の相対指標性への脆弱性、試料の秘匿・欠測への脆弱性、処置の従属性・内生性への脆弱性、「遵守者」の不均一な分布への脆弱性、処置の二次的影響による偏差への脆弱性などの結果解釈における留意点が幾つ也存在していることが理解される。

政策評価という視点から見た場合には、これらの問題のうち特に1)横断面分析性と政策との因果の希薄性、2)試料の秘匿・欠測への脆弱性と誤検出の可能性、3)処置の従属性・内生性への脆弱性と政策と別の要因による効果の可能性の3つが致命的な問題として指摘できる。またこれらの問題に加えて、局所平均処置効果評価性と部分的な効果の存在証明性がRDDの手法としての適用範囲を著しく制約しているものと考えられる。

こうした問題を改善していくためには、本質的に時間方向の情報をを用い、試料全域での比較を行うことなどの対策が必要であり、現状のRDDの手法のままでは対処が困難であると考えられる。

4.2 包括的な問題解決策としてのCICやQ-DIDと新たな方向性

最後にRDDが持つ問題点への包括的な解決策として、時間方向の情報をを用いて試料全域での比較によりATETが推計できるようにした、Athey and Imbens(2006)によるCICやQ-DIDなどの新たな取組みについて紹介しておく。

Athey and Imbens(2006)は、処置群・対照群の挙動が部分毎に異なる場合や処置効果が対象毎に均一でない場合の問題に対応するための新たな手法について提唱している。新たな手法とは1)DIDの特殊な場合として処置群の処置実施前後での試料の確率密度分布の変化と同期間の対照群の試料の確率密度分布の変化を比較し処置効果を推計する「確率密度変化前後差分分析(CIC: "Change-In-Change")」、2)処置群・対照群を処置率の四分値区分毎で横断面前後差分分析する「四分値横断面前後差分分析(Q-DID: "Quantile Difference-In-Difference")」の2つである。

これらの手法による推計に必要な前提条件は通常のDIDと概ね同じCIA、OVLA^{*17}と結果指標の単調性条件("Monotonicity Assumption")である。これらの手法は処置が連続値か離散値かを問わず適用できる。処置群・対照群の処置率に応じた対象の部分毎に処置効果の符号や大きさが異なっている場合や、処置効果が対象毎に均一でない場合の問題の対策として有効であることを説明している。

Athey and Imbens(2006)によるCICやQ-DIDはDIDの特殊な場合とされているが、見方を変えればRDDを処置前・処置後の2時点で行い、Xに沿った確率密度分布の変化を処置効果としてXの全域で推計するものである。従ってこれらの手法は、RDDの新たな可能性と今後の方向性を示したものと見ることができる。

*17 原典では触れられていないが、前提条件としてSUTVAも必要であると考えられる。

参考文献

- Angrist J. and Pischke J. 2008. *"Mostly Harmless Econometrics: An Empiricist's Companion"* Princeton Univ. Press
- Athey S. and Imbens G. 2006. 'Identification and Inference in Nonlinear Difference-In-Difference Models' *Econometrica*, Vol.74 No.2 431-497
- 2017. "The State of Applied Econometrics: Causality and Policy Evaluation" *The Journal of Economic Perspectives*, Vol.31 No.2 3-32
- Brundell R. and Costa Dias M. 2009. "Alternative Approaches to Evaluation in Empirical Microeconomics" *The Journal of Human Resources*, Vol.44 No.3 565-640
- Cameron A. and Trivedi P. 2005. *"Microeconometrics"* Cambridge University Press
- Ferwarda J. and Miller N. 2014. "Potential Devolution and Resistance to Foreign Rule: A Natural Experiment" *The American Political Science Review* Vol.108 No.3 642-660
- Hahn J., Todd P. and VanderKlaauw W. 2001. "Identification and Estimation of Treatment Effects with a Regression Discontinuity Design" *Econometrica* Vol.69 No.1, 201-2309
- Hudgens M. and Halloran M. 2006. "Causal Vaccine Effects on Binary Postinfection Outcomes" *Journal of the American Statistical Association*, Vol.101, No.473, 51-64
- 2008. "Towards Causal Inference with Interference" *Journal of the American Statistical Association*, Vol.103, No.478, 832-842
- Imbens G. and Wooldridge J. 2009. "Recent Development in the Econometrics of Program Evaluation" *Journal of Economic Literature*, Vol. 47 No. 1 5-86
- Lee D. and Lemieux T. 2010. "Regression Discontinuity Design in Economics" *Journal of Economic Literature* Vol.48 No.2 281-355
- Sekhon J. and Titiunik R. 2012. "When Natural Experiments Are Neither Natural nor Experiments" *The American Political Science Review*, Vol.106 No.1 35-57