

回帰不連続デザイン による実証分析

直井道生

慶應義塾大学経済学部 准教授

1 はじめに

本稿では、統計的因果推論の手法として幅広く用いられるようになった回帰不連続デザイン (Regression Discontinuity Design, 以下 RDD) について、実際の分析への応用を念頭に、その基本的な考え方を解説する。RDDの歴史は古く、Thistlewaite and Campbell (1960) まで遡ることができる。しかしながら、経済学分野で長くこの手法が顧みられることはなく、ようやく1990年代に入って広く実証分析に応用されるようになった¹⁾。その後、実証分析における因果推論の重要性が認識されるにつれ、RDDを利用した分析の事例は爆発的に増加している。実際、代表的な論文検索データベースである Google Scholar で「regression discontinuity design」を検索すると、15,000件を超える検索結果がヒットする (2018年5月)。

以下では、第2節で RDDによる因果推論のイメージを確認し、第3節では実際の分析およびモデルの特定化のチェックの方法について概観する。ここでは、経済学分野における代表的な統計分析用のパッケージである Stata を用いた分析の実施についても補足する。第4節はまとめである。なお、RDDの理論的な詳細や分析上の課題などについては、紙幅の都合で削らなくてはならなかったトピックが多数ある。応用を念頭に置いた RDDの解説論文としては、例えば Imbens and Lemieux (2008)、Lee and Lemieux (2010) などを参照されたい²⁾。

2 RDDによる因果推論

RDDの最も基本的な設定は、ある観察可能な変数 X が既知の閾値を超えたときに、対象となるサンプルに対してなんらかの介入 (トリートメント, D) がなされるような状況で、このことの効果を明らかにしようとするものである。このような状況は、実際の制度においてもしばしばみられる。入学試験では、テストなどの得点 (X) が一定の水準に達していれば、入学が許可 ($D=1$) され、そうでなければ許可されない ($D=0$)。いま、この学校への入学という「介入」が、卒業後の所得水準という「成果」(アウトカム, Y) に与える影響が知りたいとしよう。このとき、介入の因果的な効果はどのように計測できるだろうか。

RDDでは、閾値の前後におけるアウトカムの不連続な変化をみることで、介入の因果的な効果を計測する。さきほどの例では、入学試験の得点が合格点に達したところで介入 (入学許可) の状況が不連続的に変化する。いま、閾値の前後で、介入以外の要因が連続的にしか変化しないのであれば、不連続なアウトカムの変化は介入による因果的な効果を表している可能性が高い。言い換えれば、合格点をわずかに上回った受験者と、わずかに下回った受験者では、この学校への入学を許可されるか否かは大きく異なる一方、将来の所得に影響を与えるそれ以外の要因はほとんど (少なくとも不連続には) 異ならないと考えられる。そのため、合格点の



前後での不連続な所得水準の変化は、入学の効果とみなすことができる。

図1は、この状況を模式的に示したものである。ここで、グラフの横軸は入学試験の得点、縦軸は将来の所得水準を表している。また、2本の曲線 Y_1 および Y_0 は、それぞれ個人が介入を受けた(学校に入学した)場合の所得水準、および介入を受けなかった場合の所得水準を示している(潜在的成果変数)³⁾。いま、任意の入試得点の水準で Y_1 と Y_0 の双方を観察できるのであれば、その差($Y_1 - Y_0$)がこの学校への入学が将来の所得水準に及ぼす因果的な効果であるといえる。しかしながら、個人が実際に経験するのは、介入を受けた状態か受けない状態のいずれか一方であるため、実際に観察可能なのは Y_0 もしくは Y_1 のいずれか一方に限られる。この入学試験の例では、合格点(この例では60点)を下回った受験生については入学しなかった場合の所得水準(Y_0)のみが、合格点を上回った受験生については入学した場合の所得水準(Y_1)のみが観察

されることになる。したがって、実際に観察される入学試験の得点(X)と将来の所得水準(Y)との関係は、図の実線($Y = (1 - D)Y_0 + DY_1$)になる。

ここで重要なのは、介入の状況が不連続に変化する閾値においては、 Y_1 と Y_0 の差が、精度よく予測可能であるという点である。RDDは、まさにこのような制度上の特徴を利用して、閾値における介入の因果的な効果(τ)を識別している。

2.1 RDDのメリットとデメリット

図1からもわかるとおり、RDDによって因果効果を識別するためには、 Y_0 および Y_1 が $X=c$ において連続である必要がある(Hahn, Todd and VanderKlaauw, 2001)。Lee (2008)は、対象者が X を完全にはコントロールできない場合、上記の連続性の仮定が満たされ、 $X=c$ の近傍における介入の有無は、局所的な無作為割り当て(local randomization)とみなせることを示した。この結果は、RDDが典型的な無作為割り当て実験

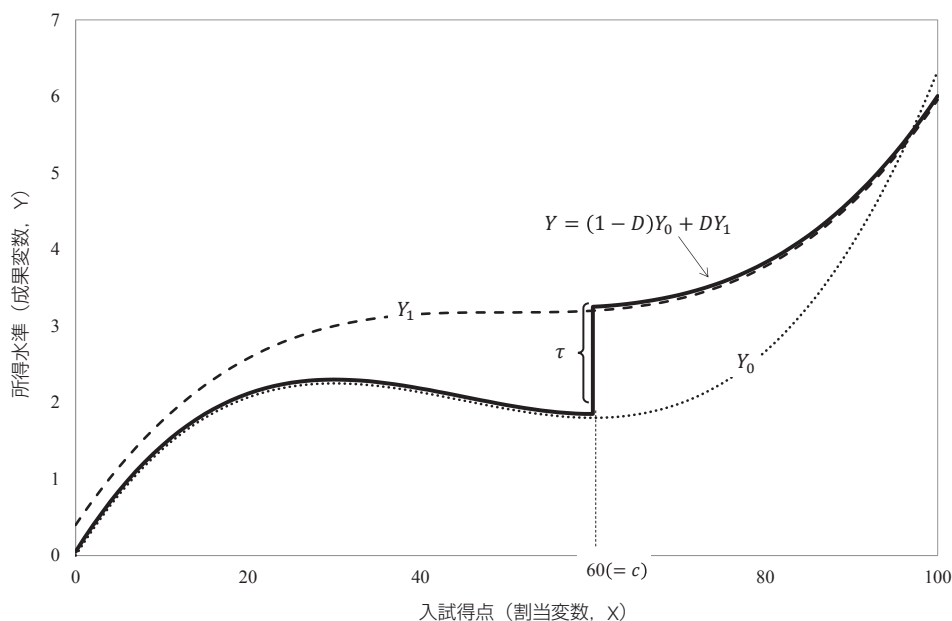


図1 RDDによる因果推論のイメージ

(Randomized Controlled Trial, RCT) と同等の
内的妥当性を持つことを意味しており, RDDの
大きな利点といえる (Lee and Lemieux, 2010)。
また, あとで説明する通り, こうした仮定が満
たされるか否かをチェックするために, 複数の
統計的な手法が提案されている。

一方で, RDDを用いた分析では, $X=c$ の閾値
における因果的な効果のみが明らかになるため,
この結果を異なる制度や母集団に適用するこ
とが難しいという問題も指摘される (外的妥当
性)⁴⁾。具体的に, 図1に示した入学試験の例で
は, 合格点が60点である場合の入学の因果的な
効果は識別可能であるものの, 合格点を大きく
変化させた場合の因果効果 (例えば合格点が80
点であるような場合の Y_1 と Y_0 の差) は, 追加的
な分析上の仮定なしには識別できない。

2.2 Sharp RDDとFuzzy RDD

ここまでの説明では, 介入の有無が閾値によっ
て完全に決定される (閾値の前後で介入の確率
が0から1に不連続に変化する) ことを前提に
議論を行ってきた。しかしながら, 実際の制度
においては, 介入の有無は個人の選択に一部が
ゆだねられているケースも多い。実際, 入学試
験の例でも, 合格点に達した受験生のすべてが
この学校に入学するとは限らず, 介入 (学校へ
の入学) を受けるのは合格者の一部に限られる。
通常, 前者は「シャープ (Sharp)」な RDD, 後者
は「ファジー (Fuzzy)」な RDD と呼ばれるが,
本稿では, シャープな RDD に限って議論を行う。

3 RDDによる分析

この節では, RDDによる分析の具体的な方法
を概観する。なお, 以下では, 標準的な用語法
に従い, Y を成果変数, D を介入変数, X を割当変
数と呼ぶことにする⁵⁾。

3.1 図による分析

前節でみた通り, RDDによる分析においては,
割当変数の閾値における成果変数の不連続的
な変化を計測することが目的となる。このた
めの第一ステップとしては, 実際のデータを用
いて図1と対応するような作図を行い, 割当変
数と成果変数の関係性を視覚的に確認するこ
とが重要である。これには, 閾値の左右それ
ぞれで割当変数の水準によってデータを分割し,
分割された各区間における成果変数の平均を
プロットする方法が採られることが多い。図2
は, 選挙における現職効果の検討を行った Lee
(2008) の分析データを用いた作図例である⁶⁾。

図の横軸は割当変数 (前回の選挙における対
立候補との得票率の差), 縦軸は成果変数 (次の
選挙における得票率) であり, 各プロットは区
間ごとに計算された成果変数の平均値になっ
ている⁷⁾。また, 両者の関係性を把握しやすく
するために, しばしばデータにフィットする回
帰曲線も追加される。図2からは, 割当変数の
閾値 ($X=0$) において成果変数が不連続的に変化
しており, 現職効果が存在することが分かる。
このような単純な視覚化によって介入の効果
を確認することができるのも RDD の利点の一
つである。

3.2 回帰分析

図2のようなグラフによる確認を踏まえて, 回
帰モデルによって介入の因果的な効果を計測す
る方法を考える。閾値における成果変数の不
連続な変化を計測するためには, 割当変数の直
接的な影響を適切に取り除く必要がある。割
当変数に関するモデルの特定化の誤りは, 介入
の因果効果の計測にバイアスをもたらす可能性
がある。

いま, 割当変数と成果変数の関係が図3のよ
うに与えられるとしよう。ここで, 図の実線は
両者の真の関係を表し, プロットされた点は観
察されるデータである。このとき, 分析者が割

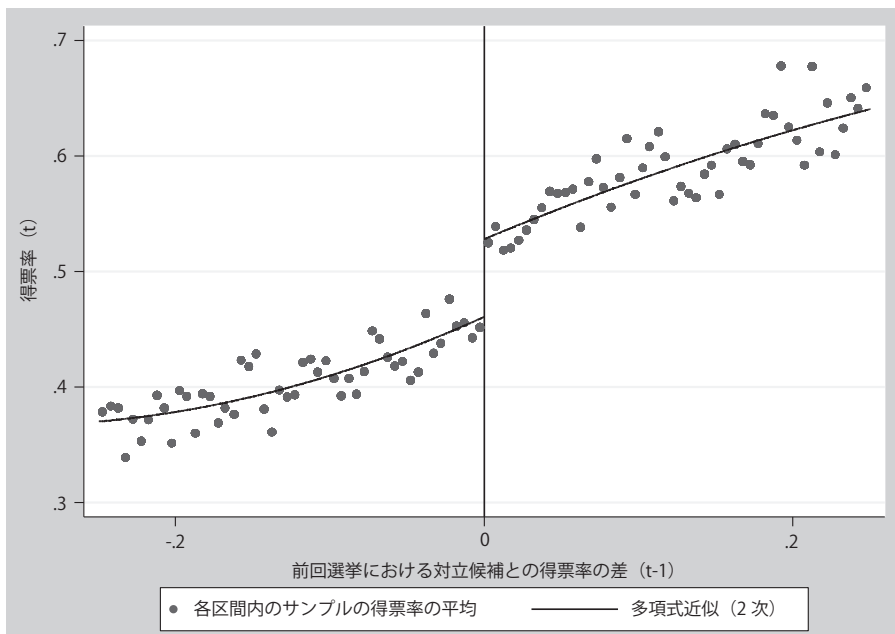


図2 割当変数と成果変数の関係
Lee (2008, Figure 4a) の筆者による再現

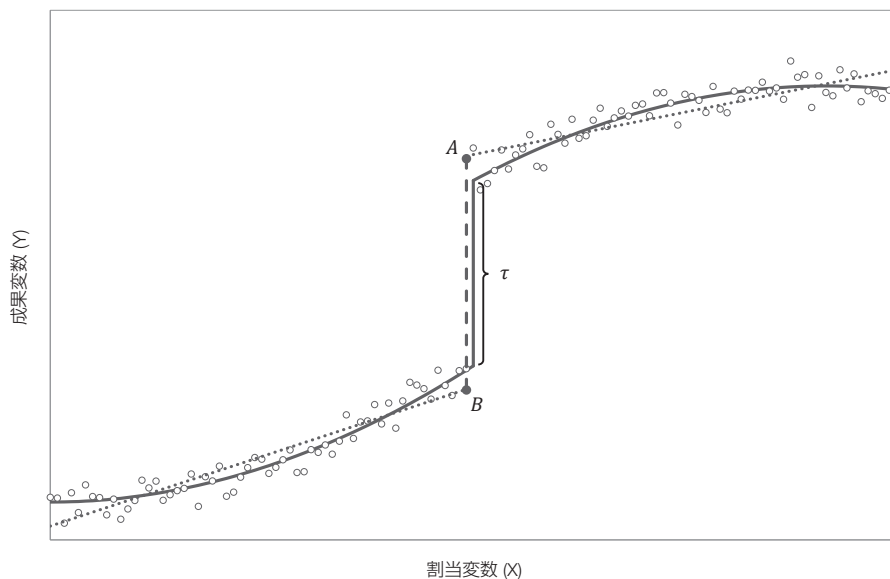


図3 関数形の特定化の誤りとバイアス

当変数と成果変数の間に（誤って）線形関係を仮定して分析すると（図の点線），このモデルに基づく因果効果の推計結果（図のAB）は，真の効果（ τ ）を過大に推計してしまうことになる。

このような問題に対して，初期のRDDによる分析では，割当変数に関してできるだけフレキシブルな関数形を特定化することで対処してきた。典型的には，閾値の左右それぞれで割当変

数の影響を適当な多項式でコントロールしたうえで、閾値を超えたことを示すダミー変数を導入する方法が考えられる。

$$Y_i = \tau \cdot D_i + \sum_{k=1}^{K'} \beta_k^l \cdot (1 - D_i) \cdot (X_i - c)^k + \sum_{j=1}^{K''} \beta_j^r \cdot D_i \cdot (X_i - c)^j + \varepsilon_i \quad (1)$$

ここで、 D_i は $X_i \geq c$ であれば1、そうでなければ0を取るようなダミー変数である。(1)式の右辺第2項および第3項では、閾値の左側および右側における割当変数の影響を、それぞれ K' 、 K'' 次の多項式によってコントロールしている。いま、閾値ではこれらの項はゼロになる($X_i - c = 0$)ため、第1項の係数 τ が閾値における成果変数の不連続な変化をとらえることになる。

しかしながら、(1)式による τ の推定は、閾値から遠く離れた観測値による影響を受けてしまうため、近年では局所線形回帰 (local linear regression) と呼ばれる代替的な手法が用いられることが多い (Hahn, Todd and van der Klaauw, 2001)⁸⁾。最も基本的なモデルでは、割当変数の値が閾値の左右それぞれ h に入るサンプルのみを用いて、以下の(2)および(3)式を最小化するような推計を行う⁹⁾。

$$\min_{\alpha_l, \beta_l} \sum_{l: c-h < X_i < c} (Y_i - \alpha_l - \beta_l(X_i - c))^2 \quad (2)$$

$$\min_{\alpha_r, \beta_r} \sum_{r: c \leq X_i < c+h} (Y_i - \alpha_r - \beta_r(X_i - c))^2 \quad (3)$$

ここで、(2)式は閾値の左側、(3)式は右側のデータを用いた局所線形回帰モデルとなっている。これは、割当変数の値が閾値に近い($c-h < X_i < c+h$)サンプルのみを用いて、閾値の左右それぞれで通常の線形回帰モデルを推計していることに他ならない。

いま、閾値に十分近いサンプルのみを用いるのであれば、線形回帰は(潜在的には非線形な関係となりうる)真の関係のよい近似になる¹⁰⁾。(2)および(3)式から、介入の因果的効果は、 $\hat{\tau} = \hat{\alpha}_r - \hat{\alpha}_l$ として求められる¹¹⁾。

局所線形回帰モデルの推計に当たっては、分析に用いるデータの範囲を示す h (バンド幅と呼ばれる)をどのように選ぶかが問題となる。一般的なノンパラメトリック回帰でよく知られるように、 h の選択にはトレードオフが存在する。いま、 h を小さく取るほど、(2)および(3)式の線形の特化が、真の関係のよい近似になるため、因果効果の推計におけるバイアスを小さく抑えることができる。一方で、 h を小さく取るほどサンプルサイズは減少するため、因果効果の推定値の分散は増加する。

実際の分析においては、理論的な最適バンド幅に基づく決定 (plug-in method) や実際のデータを用いた交差検証 (cross-validation) に基づく決定などがしばしば用いられる。前者は、全ての変数の分布に依存する最適なバンド幅関数を導出したうえで、データから必要な要素を推計するという方法である。後者は、平均二乗誤差を最小化するような交差検証 (cross-validation) による方法である。

3.3 モデルの特化性のチェック

RDDによって意味のある介入の因果効果を測定するためには、いくつかの前提条件が満たされている必要がある。例えば、割当変数が操作可能である場合、閾値の前後で異質なサンプルを比較している可能性が高く、RDDによる分析は誤った結果を導く可能性がある。ここでは、RDDにおけるモデル特化性のチェックについて、代表的なものを取り上げる。

(1) 共変量や疑似成果変数を用いたチェック

前述の通り、RDDが前提とする理想的な状況下では、閾値の近傍において介入が無作為に割り当てられているとみなせる。この場合、介入を受けたグループとそうでないグループは、よく似た特性を持つことが期待される。そのため、多くの実証分析では、観察可能な共変量を利用し、閾値の前後での平均値の比較や、3.1および3.2節で議論した方法をこれらの共変量に対し



て適用することで、局所的な無作為割り当ての仮定をチェックしている。これと関連して、成果変数に影響を与えうる（割当変数以外の）共変量を追加的な説明変数として加え、因果効果の推計結果に違いが生じるかを確認するというものもある。理想的な状況では、他の共変量の追加は結果に大きな影響を与えないことが予想される。

また、介入によって直接の影響は受けないと考えられる疑似成果変数 (pseudo outcome) を用いたチェックもしばしば行われる。具体的には、介入に先立って決定している変数などが疑似成果変数として用いられる。閾値においてこれらの変数が不連続に変化する場合、割当変数の操作に伴う自己選択の問題や、興味のある介入以外に何らかの構造的な変化が生じている可能性などが疑われる。

(2) 割当変数の分布によるチェック

前述の方法は、疑似実験的な枠組みに基づく他の手法（差分の差分法やマッチング法など）でも広く用いられるものであるが、よりRDDに特有の手法として、割当変数の分布に着目したものがあつた。これは、割当変数の分布が閾値の前後で不連続に変化していないかを確認するものである。いま、割当変数の分布が閾値の前後で不連続である場合、割当変数が意図的に操作されている可能性が強く疑われる。分析に当たっては、単純なヒストグラムによる視覚的な確認のほか、McCrary (2008) の密度検定などが用いられる¹²⁾。

(3) 回帰モデルの特定化に関するチェック

上記のRDDによる分析の前提となる仮定についてのチェックに加え、実際の分析に当たっては、一般的な回帰モデルの特定化に関するチェックも必要となる。例えば、局所線形回帰モデルの推計に当たっては、前節で紹介した手法のいずれかを用いてバンド幅を決定することとなるが、実際の分析では、複数の異なるバンド幅に基づく結果を報告することで、バンド幅の

選択に関する感度分析が実施されることが多い。また、図3で示した通り、割当変数と成果変数の間に非線形な関係がある場合には、特定化の誤りに伴うバイアスが懸念されるため、バンド幅の選択と合わせて割当変数の特定化のチェック（例えば、2次、3次の項を考慮するなど）も併せて行うことが望ましい。

4 まとめ

本稿では、近年応用が進んでいる回帰不連続デザインによる実証分析の基本的な考え方とその応用法について概観した。本稿で扱うことのできなかった代表的な論点としては、例えば次のようなものがある¹³⁾。

第一に、多くの実際の制度では、介入の割り当ては必ずしも割当変数のみによっては決まらず、ファジーなRDDの枠組みが必要となる場合がある。ファジーなRDDに基づく分析も、基本的な考え方はシャープなRDDと同様であるが、推計される因果効果は閾値を超えることで介入の有無を変化させるようなサンプルに対してのもの (intention to treat) である点に注意が必要となる。

第二に、介入の有無を決定する割当変数が複数存在するケースが考えられる。具体例としては、入学試験の合否が複数の科目の試験の結果によって決まるケースがそれにあたる。いま、単純に合計点によって合否が決まるのであれば、これを割当変数として考えればよいが、科目ごとに合格最低点が設定されているようなケースでは、これまでの議論を単純に適用することはできない。

第三に、RDDが対象とするような閾値によって介入が決定される状況で、いくつかの関連する手法が提案されている。回帰ねじれデザイン (RKD: Regression Kink Design) と呼ばれる手法は、閾値の前後で（連続変数としての）介入

の程度が不連続に変化する状況で用いられる。例えば、所得税などの税制では、ある（所得水準の）閾値を境に、適用される税率が不連続に変化する。RKDでは、こうした制度上の特徴を利用して、介入の因果効果の識別を行う。また、集積分析 (bunching analysis) では、割当変数が操作可能であるような状況下で、その分布が閾値の前後で不連続に変化することを利用して、介入の因果効果を識別することが検討され

ている。

本稿で見てきた通り、RDDは割当変数の閾値によって介入が決定されるような特定の設定の下で、因果効果を計測するための有力なツールとなっている。上記でみたような拡張まで含めると、RDDとそれに関連する手法がカバーする範囲はかなり広く、今後より一層の応用研究が進むことが期待される。

注

- 1) 統計学や心理学などの分野におけるRDDの利用とその歴史的な経緯については、Cook (2008) による解説を参照されたい。
- 2) 日本語によるRDDの入門的な解説としては、安藤 (2015) なども参考になる。
- 3) ここでの説明は、経済学分野における標準的な統計的因果推論の枠組みである、Rubin Causal Model (RCM) の考え方に基づいている。RCMに関するより一般的な解説については、例えば DiNardo and Lee (2011) などを参照されたい。
- 4) より正確には、RDDによって識別される因果効果は、閾値によって介入の有無が左右される可能性がある個人の因果効果の重み付き平均として解釈できる (Lee and Lemieux, 2010)。
- 5) 割当変数は assignment variable の日本語訳であるが、その他にも文献によって index, score, forcing variable, running variable などと呼ばれることがある。
- 6) データセットは、<http://economics.mit.edu/faculty/angrist/data1/mhe> から入手可能である。なお、作図に当たっては、Stataのrdplotコマンドを用いた。以降で用いるStataのプログラム(adoファイル)については、<https://sites.google.com/site/rdpackages/> から導入が可能である。
- 7) こうした作図を行う場合には、区間の幅 (bin) を事前に設定する必要があるが、そのための統計的な方法および実際上の選択方法については Lee and Lemieux (2010) などを参考のこと。
- 8) Gelman and Imbens (2018) は、特に高次の多項

式を用いた (1) 式に基づく推計のパフォーマンスが悪いことを指摘している。

9) (2) および (3) 式では、単純化のために閾値の左右 h に入るサンプルを全て同じ重み付けで推計に用いるケースを想定している (rectangular kernel)。より一般には、閾値からの距離に応じて重みづけを変えることも可能であるが、通常こうした選択は結果に大きな影響を与えないことが期待される (Imbens and Lemieux, 2008)。

10) ただし、後述する通り、実際の分析においては $(X_i - c)$ に関するより高次の項を説明変数に追加した分析を行い、結果の頑健性をチェックすることが望ましい。また、(2) および (3) 式の推計に当たっては、割当変数以外の成果変数に影響を与える共変量を説明変数に追加することも可能である。こうした追加的な説明変数は、推計精度の向上に役立つ可能性がある一方、モデルの特定化が正しいのであれば、推計される因果効果には大きな影響を与えないことが期待される。

11) Stataを用いた局所線形回帰モデルによる推計は、rdrobustコマンドによって実行可能である。これは、(2) および (3) 式で示した基本的なモデルに加え、異なるカーネル関数を用いた推計や、バンド幅の決定にかかわる複数の手法の適用などにも対応している。

12) 後者については、rddensityコマンドによってStataでも実行することができる。

13) 個別の論点に関する詳細については、例えば Choi and Lee (2017) などを参考にされたい。



文献

- Choi, J. Y. and M. J. Lee, 2017, "Regression Discontinuity: Review with Extensions," *Statistical Papers*, 58(4):1217-1246.
- Cook, T. D., 2008, "“Waiting for Life to Arrive”: A History of the Regression-Discontinuity Design in Psychology, Statistics and Economics," *Journal of Econometrics*, 142(2):636-654.
- DiNardo, J. and D. S. Lee, 2011 "Program Evaluation and Research Designs," *Handbook of Labor Economics*, 4:463-536.
- Gelman, A. and G. Imbens, 2018, "Why Higher-Order Polynomials Should Not Be Used in Regression Discontinuity Designs," *Journal of Business and Economic Statistics*, available online. Doi: 10.1080/07350015.2017.1366909.
- Hahn, J., P. Todd and W. van der Klaauw, 2001, "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica*, 69(1):201-209.
- Imbens, G. W. and T. Lemieux, 2008, "Regression Discontinuity Designs: A Guide to Practice," *Journal of Econometrics*, 142(2):615-635.
- Lee, D. S., 2008, "Randomized Experiments from Non-Random Selection in U.S. House Elections," *Journal of Econometrics*, 142(2):675-697.
- Lee, D. S. and T. Lemieux, 2010 "Regression Discontinuity Designs in Economics" *Journal of Economic Literature*, 48(2):281-355.
- McCrary, J., 2008, "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test" *Journal of Econometrics*, 142(2):698-714.
- Thistlethwaite, D. L. and D. T. Campbell, 1960, "Regression-Discontinuity Analysis: An Alternative to the Ex Post Facto Experiment" *Journal of Educational Psychology*, 51(6):309-317.
- 安藤道人, 2015, 「多重回帰分析と回帰不連続デザイン」, 『日本労働研究雑誌』657:12-13.