

第8章：回帰不連続デザイン

Jonathan Roth

数理計量経済学 I
ブラウン大学

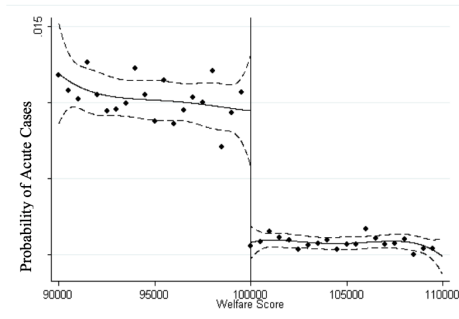
モチベーション

- 因果推論における主要な課題は、処置ステータス以外はあらゆる面で処置群と比較可能な対照群を見つけることです。
- これまで、以下のような場合に因果的効果を推定する方法を見てきました：
 - 処置が文字通りランダムに割り当てられている場合
 - 処置が実質的にランダムに割り当てられている場合（条件付き非交絡性）
 - 選択バイアスが時間を通じて一定である場合（並行トレンド）
 - 実質的にランダムで除外制約を満たす操作変数がある場合
- 本日は、もう一つの戦略である回帰不連続デザイン (RD) を見ていきます。
- RD の背後にある考え方は、スコアがしきい値を上回るか下回るかによって、処置ステータスが決定される場合があるということです。
 - 例えば、学校への入学が、テストの点数があるカットオフ値を超えるかどうかによって依存する場合などです。
- このような場合、RD はスコアがしきい値を「わずかに」上回る人々と、下回る人々の結果を比較します。

例：Hu and Chao (2008)

- Hu and Chao (2008) は以下の問いに興味を持っています：健康保険の欠如は、貧困層が重要な医療を受けることを妨げているか？
- 単に健康保険がある人とない人の医療利用を比較するだけではいけないのはなぜでしょうか？
 - 交絡変数があるからです！ 健康保険を持っている人はより裕福である可能性があり、それが直接健康状態に影響を与えているかもしれません。
- 背景：グルジア共和国（米国のジョージア州ではありません！）は2006年に健康保険プログラムを創設しました。各世帯は80の変数から導出された「貧困スコア」を付与され、スコアが100,000以下の世帯が健康保険を受け取りました。
- 鍵となるアイデア：スコアが100,000をわずかに上回る世帯は、100,000をわずかに下回る世帯と非常に似ているはずですが。
 - 唯一の違いは、100,000以下の世帯が健康保険という「処置」を受けたことです！

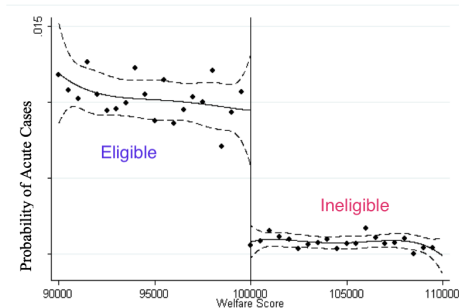
Figure 3: The Effect of MAP on Utilization of Acute Surgeries/In-Patient Services



Note: This figure plots probability of utilization of acute surgeries/inpatient services against welfare scores. Each dot is the average probability within 500 intervals of welfare scores. Solid lines are fitted values from 4th order polynomial regressions on either side of the discontinuity. Dotted lines are 95% confidence intervals.

- このプロットは、スコアの関数としての結果（急激な手術の件数）を示しています。

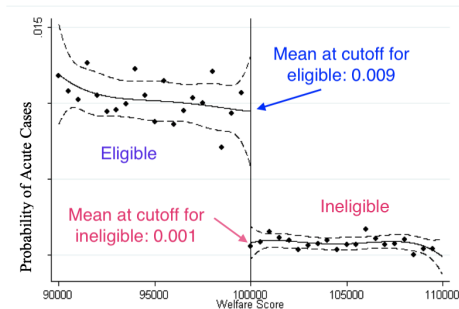
Figure 3: The Effect of MAP on Utilization of Acute Surgeries/In-Patient Services



Note: This figure plots probability of utilization of acute surgeries/inpatient services against welfare scores. Each dot is the average probability within 500 intervals of welfare scores. Solid lines are fitted values from 4th order polynomial regressions on either side of the discontinuity. Dotted lines are 95% confidence intervals.

- 垂直線（100,000）の左側のグループが政府から健康保険を受け取ります。
- 100,000 をわずかに下回る人々と、わずかに上回る人々は、保険の受給資格以外のすべての要因において非常に似ていることが期待されます。

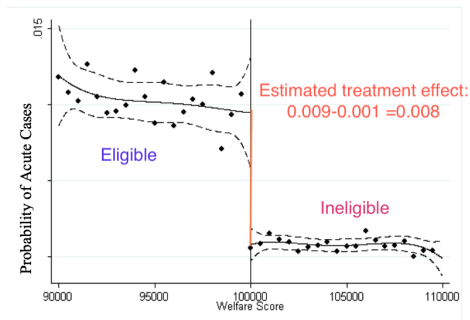
Figure 3: The Effect of MAP on Utilization of Acute Surgeries/In-Patient Services



Note: This figure plots probability of utilization of acute surgeries/inpatient services against welfare scores. Each dot is the average probability within 500 intervals of welfare scores. Solid lines are fitted values from 4th order polynomial regressions on either side of the discontinuity. Dotted lines are 95% confidence intervals.

- 100,000 をわずかに下回る人々と、わずかに上回る人々は、保険の受給資格以外のすべての要因において非常に似ていることが期待されます。
- しかし、しきい値のすぐ左側の人々は、右側の人々よりも明らかに多く手術を受けているようです！

Figure 3: The Effect of MAP on Utilization of Acute Surgeries/In-Patient Services



Note: This figure plots probability of utilization of acute surgeries/inpatient services against welfare scores. Each dot is the average probability within 500 intervals of welfare scores. Solid lines are fitted values from 4th order polynomial regressions on either side of the discontinuity. Dotted lines are 95% confidence intervals.

- しかし、しきい値のすぐ左側の人々は、右側の人々よりも明らかに多く手術を受けているようです！
- もし他のすべての要因がしきい値において連続的であれば、この差が（しきい値付近の人々における）因果的効果となります！

RD の定式化

- D_i を処置のバイナリ指標、 R_i を実行変数 (running variable) とします。
- シャープな (sharp) RD では、ユニット i は $R_i \geq c$ の場合にのみ処置を受けます：

$$D_i = 1[R_i \geq c]$$

- D_i は交絡因子の可能性が高い変数によって決定論的に決まります。
- RD の考え方は、カットオフ c の周囲における CEF の極限を比較することです：

$$\tau_{RD} = \underbrace{\lim_{r \downarrow c} E[Y_i | R_i = r]}_{\text{上からの極限}} - \underbrace{\lim_{r \uparrow c} E[Y_i | R_i = r]}_{\text{下からの極限}}$$

- どのような場合に τ_{RD} が因果的効果に対応するのでしょうか？

極限を考える

- $R_i < c$ の人々は $D_i = 0$ です。したがって：
 $\lim_{r \uparrow c} E[Y_i | R_i = r] = \lim_{r \uparrow c} E[Y_i(0) | R_i = r]$
- $R_i \geq c$ の人々は $D_i = 1$ です。したがって：
 $\lim_{r \downarrow c} E[Y_i | R_i = r] = \lim_{r \downarrow c} E[Y_i(1) | R_i = r]$
- 仮定： $E[Y_i(d) | R_i = r]$ が $d = 0, 1$ について c で連続ならば：
 - $\lim_{r \uparrow c} E[Y_i(0) | R_i = r] = E[Y_i(0) | R_i = c]$
 - $\lim_{r \downarrow c} E[Y_i(1) | R_i = r] = E[Y_i(1) | R_i = c]$
- したがって：

$$\tau_{RD} = \lim_{r \downarrow c} E[Y_i(1) | R_i = r] - \lim_{r \uparrow c} E[Y_i(0) | R_i = r] = E[Y_i(1) - Y_i(0) | R_i = c]$$

これは、カットオフ地点における平均処置効果です！

連続性の評価

- RD における最も重要な仮定は、潜在的結果の期待値 $E[Y_i(d)|R_i = r]$ がカットオフにおいて連続である（ $d = 0, 1$ について）ことです。
- なぜ連続性の仮定が満たされない可能性があるのでしょうか？
- 理由 1：交絡因子がカットオフにおいて不連続に変化する場合。
- 理由 2：人々がカットオフの直上/直下に入るようにスコアを操作 (manipulate) できる場合。

例：交絡因子

- 政策はしばしば州境において不連続に変化します。
- 例：Holmes (1998) は、労働組合を弱める「労働権法 (right-to-work laws)」がビジネスにどのように影響するかに興味を持っています。
- 彼は RD を用いて、労働権法がある州とない州の境界の両側で、製造業の雇用密度を比較しました。

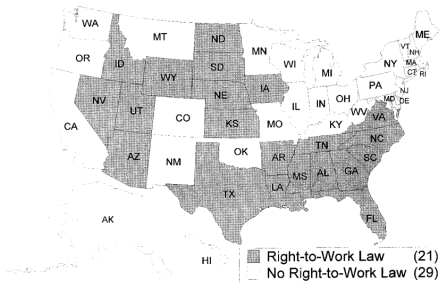


FIG. 1.—Geography of right-to-work laws

680

JOURNAL OF POLITICAL ECONOMY

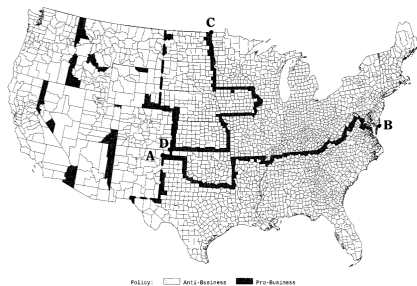


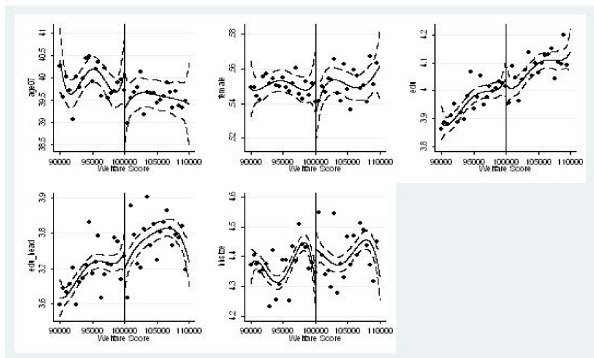
FIG. 3.—Counties within 25 miles of the policy change border

TABLE 1
 MANUFACTURING EMPLOYMENT SHARES AND GROWTH RATES: CROSS-COUNTY
 AVERAGES BY DISTANCE FROM BORDER AND SIDE OF BORDER

MILES FROM BORDER	COAL REGION INCLUDED		COAL REGION EXCLUDED	
	Share of 1992 Total (1)	Growth Rate, 1947-92 (2)	Share of 1992 Total (3)	Growth Rate, 1947-92 (4)
A. Antibusiness Side of Border				
75-100	25.9	67.5	25.0	68.2
50-75	23.1	62.7	25.0	80.9
25-50	23.2	82.0	24.7	88.8
0-25	21.0	62.4	22.1	77.2
B. Probusiness Side of Border				
0-25	28.6	100.7	27.9	104.2
25-50	26.7	89.1	25.5	88.3
50-75	26.7	92.9	24.5	90.1
75-100	25.4	91.8	23.1	93.5

- 境界線の労働権法（RTW）がある側で、製造業が多いように見えます。
- しかし、州境で変化するのは労働権法だけでしょうか？！労働権法が、製造業に影響を与える他の政策と相関している可能性はないでしょうか？

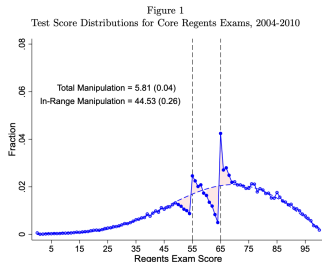
- このような懸念に部分的に対処するため、観察可能な特徴がカットオフにおいて不連続に変化していないことを示すのが一般的です。
- 下の図は、Hou and Chao の論文において、年齢、性別、教育、世帯人数に不連続性が見られないことを示しています。



- しかし、いつものように、他の未観察の交絡因子がカットオフにおいて変化している可能性は依然として残ります。

例：操作

- ニューヨーク市では、高校卒業資格を得るために Regents 試験で少なくとも 55 点を取る必要があります。
- これを利用した卒業資格の効果の研究は魅力的ですが... 実際には、しきい値のすぐ上に非常に多くの生徒が集中しています。



- 考えられる理由：教師が、生徒をしきい値に乗せるために「おまけ」をしている可能性があります。
- しきい値をギリギリ超えた生徒と、そうでない生徒が似ているとは言えません。

操作のテスト

- このような RD の操作をテストするために、カットオフの両側でユニットの数が同様であるかを確認するのが一般的です。これはしばしば「マクレアリー・テスト (McCrary test)」と呼ばれます。
- カットオフの片側に「集積 (bunching)」が見られる場合、通常は操作の証拠として解釈されます。
- 集積がある場合、連続性の仮定は非常に疑わしいものになります。

RD の推定

- 連続性の下で、識別式は以下の通りです：

$$\lim_{r \downarrow c} E[Y_i | R_i = r] - \lim_{r \uparrow c} E[Y_i | R_i = r] = E[Y_i(1) - Y_i(0) | R_i = c]$$

- この CATE を推定するには、実行変数 R_i が与えられたときのアウトカム Y_i の CEF を推定し、極限をとるだけです。
- CEF を推定するには？ OLS です！

線形回帰を用いた RD

- CEF が区分的に線形であると仮定しましょう：

$$\begin{aligned} E[Y_i | R_i = r] &\approx \alpha_0 + \alpha_1 r && \text{if } r < c \\ E[Y_i | R_i = r] &\approx \beta_0 + \beta_1 r && \text{if } r \geq c \end{aligned}$$

- すると RD の仮定の下で：

$$CATE = \lim_{r \downarrow c} E[Y_i | R_i = r] - \lim_{r \uparrow c} E[Y_i | R_i = r] \approx (\beta_0 + \beta_1 c) - (\alpha_0 + \alpha_1 c)$$

- これらの回帰係数を OLS で推定することで、処置の効果を推定できます。

1つの回帰で実行する

- 一般性を失うことなく $c = 0$ と正規化すれば、実は1つの回帰でこれを行うことができます。
- 以下の回帰を考えます：

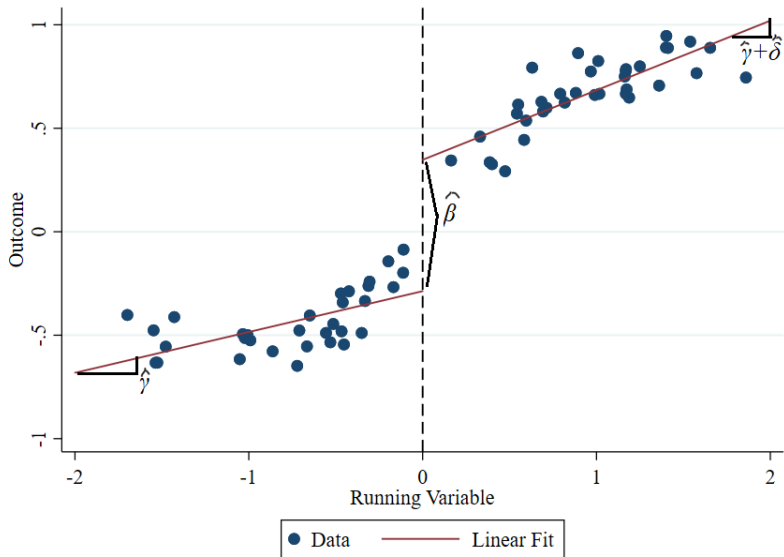
$$Y_i = \alpha + \beta 1[R_i \geq 0] + \gamma R_i + \delta R_i 1[R_i \geq 0] + U_i$$

- すると：

$$\begin{aligned} \text{if } r < 0, & \quad E[Y_i | R_i = r] \approx \alpha + \gamma \\ \text{if } r \geq 0, & \quad E[Y_i | R_i = r] \approx \alpha + \beta + \gamma r + \delta r \end{aligned}$$

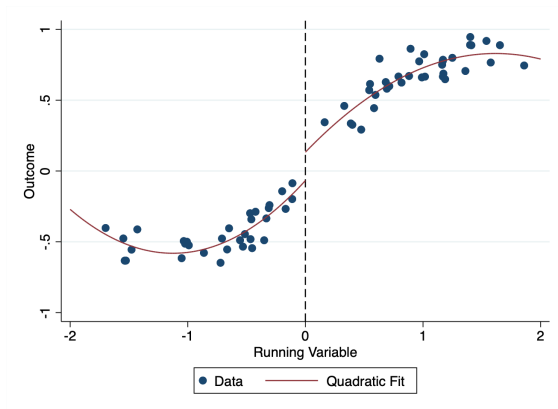
- したがって：

$$\begin{aligned} ATE &= \lim_{r \downarrow 0} E[Y_i | R_i = r] - \lim_{r \uparrow 0} E[Y_i | R_i = r] \\ &\approx (\alpha + \beta + (\gamma + \delta) \times 0) - (\alpha + \gamma \times 0) = \beta \end{aligned}$$



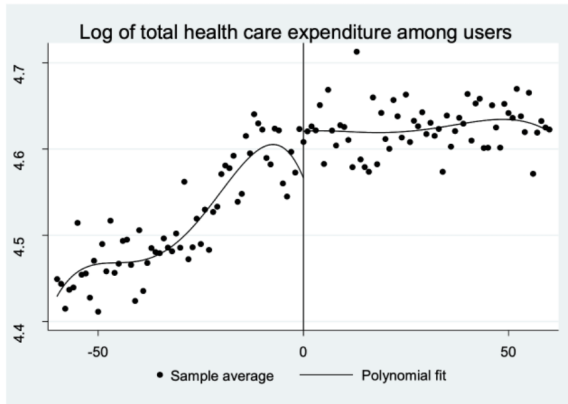
高次の項

- これは、CEF への線形近似が適切である場合にのみうまく機能します。
- これまでと同様に、回帰に高次の項を加えることもできます。
- 例えば、ここでは両側で2次式をフィットさせています：



高次多項式の問題点

- 高次多項式の問題は、データに過剰適合（過学習）してしまう可能性があることです。
 - 特に、データの「境界」に向かって外挿する際に、この問題は深刻になります。
- この図の中に、本当に不連続性は存在しますか？



線形に固執すべきか？

過学習の問題により、あるノーベル賞受賞者（かつブラウン大卒業生）は強い立場をとっています！

Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs

Andrew GELMAN

Department of Statistics and Department of Political Science, Columbia University, New York, NY, 10027
(gelman@stat.columbia.edu)

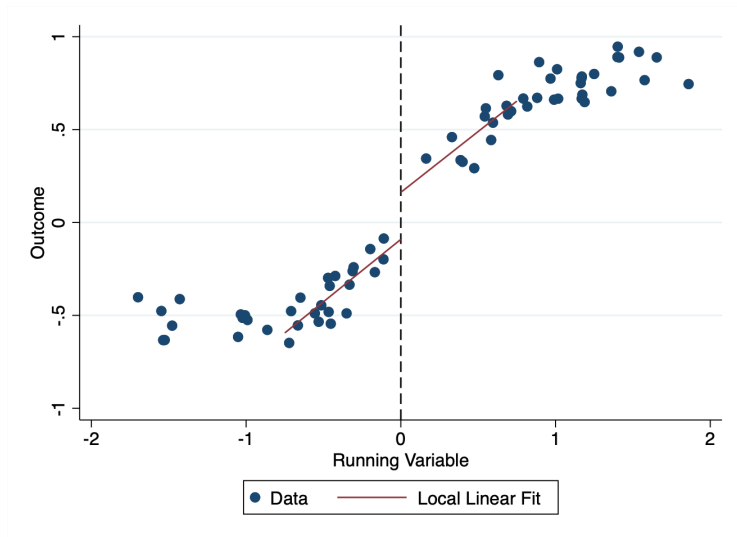
Guido IMBENS

Graduate School of Business, Stanford University, Stanford, CA 94305, and NBER, Stanford University, Stanford, CA 94305 (imbens@stanford.edu)

実務ではどうすればよいか？

- 明らかに、CEF を十分に捉える豊かな近似と、過学習との間にはトレードオフがあります。
- 実務における標準的なアプローチは、局所線形回帰 (local linear regression) を用いることです。
- 基本的な考え方は、線形回帰をフィットさせますが、境界に「近い」点のみを使用することです。
 - 最も単純なケースでは、カットオフの「バンド幅 (bandwidth)」内の点のみを使用します。
 - より複雑なバージョンでは、カットオフに近い観測値ほど高い重みを付けます。
- バンド幅を選ぶためのデータ駆動型の方法もいくつかあります (例：観測値が多い場合はより狭いバンド幅を使う)。
- これは Stata の rd パッケージなどを用いることで、簡単に実装できます。

局所線形回帰



RD の推定は難しい！

- 局所線形回帰は、モデルの豊かさと過学習の間の比較的良い妥協点です。
- しかし、完璧ではありません！ 一般に、非常に非線形な CEF と不連続性を区別するのは困難です。
- 自分の目を信じることも大切です。プロットを見て、不連続性があるように見えますか？
- 最も説得力のある RD は、プロットから一目瞭然であり、高度な計量経済学的手法を必要としません。

ファジーな RD

- しきい値を超えることが、処置ステータスを完全には決定しないこともあります。しかし、処置を受ける確率が不連続に上昇する場合があります。
- これはファジーな (fuzzy) RD と呼ばれます。
- 基本的な考え方は IV と似ています。しきい値を超えることがアウトカムに与える効果を推定し、それを処置ステータスへの効果（すなわち処置確率の変化）で割ります。
- 特定の条件下で、これはカットオフ地点における LATE を識別します。

例：Bleemer and Mehta (2022)

- Bleemer and Mehta は、(皆さんにとって) 非常に重要な問いを投げかけています：経済学を専攻するとお金持ちになれるのか？
- 彼らはこの問題をカリフォルニア大学サンタクルーズ校 (UCSC) の文脈で研究しました。そこでは経済学部が、入門科目の GPA が 2.8 未満の学生に対し、経済学専攻を「学部の裁量」として制限していました。

第一段階

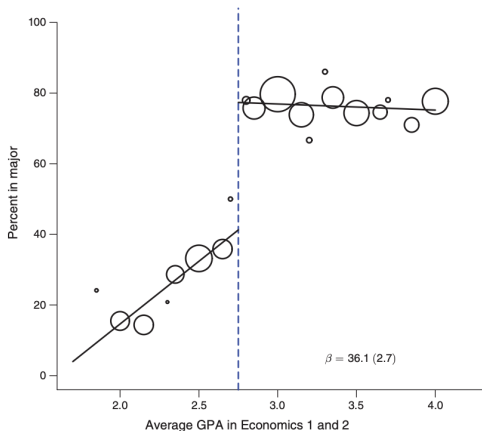


FIGURE 1. THE EFFECT OF THE UCSC ECONOMICS GPA THRESHOLD ON MAJORING IN ECONOMICS

Notes: Each circle represents the percent of economics majors (y-axis) among 2008–2012 UCSC students who earned a given *EGPA* in Economics 1 and 2 (x-axis). The size of each circle corresponds to the proportion of students who earned that *EGPA*. *EGPAs* below 1.8 are omitted, leaving 2,839 students in the sample. Fit lines and beta estimate (at the 2.8 GPA threshold) from linear RD specification; standard error (clustered by *EGPA*) in parentheses.

しきい値を超えた学生は、経済学を専攻する確率が約 36 ポイント高くなっています。

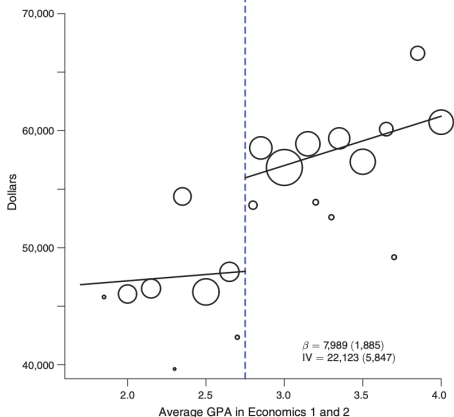


FIGURE 2. THE EFFECT OF THE UCSC ECONOMICS GPA THRESHOLD ON ANNUAL WAGES

- しきい値を超えた学生は、所得が約 8,000 ドル多くなっています。
- この講義を履修してよかったと思いませんか？！
- 経済学専攻の効果のファジー RD 推定値は、 $\$8,000 / 0.36 \approx \$22,000$ となり、これは平均所得の 40

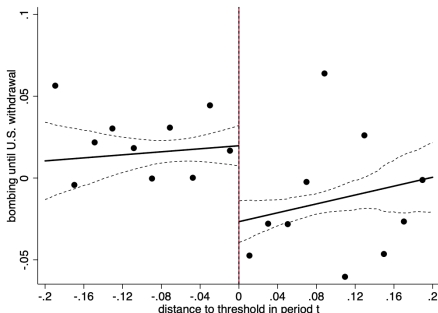
どのような場合にファジー RD は LATE を与えるか？

- IV と同様の仮定の下で、ファジー RD はカットオフ地点におけるコンプライヤーの LATE を与えます。
- 連続性。 $d = 0, 1$ について $E[Y_i(d)|R_i]$ がカットオフで連続であること。また、コンプライヤー/オールウェイズ・テイカー/ネバー・テイカーの割合も連続である必要があります。
- 除外制約。カットオフの直上/直下であることは、処置ステータスへの影響を通じてのみアウトカムに影響を与えること。
- 関連性。カットオフにおいて処置の受け入れに不連続性が存在すること。
- 「局所的」単調性。カットオフを下回る場合にのみ処置を受ける、というディファイアーが存在しないこと。

Dell and Querubin (2017)

- Dell and Querubin は、ベトナム戦争中に米国空軍が、村の 169 の治安・政治・経済的特徴に基づいた「リスクスコア」を用いて爆撃目標を選定していた事実を利用しました。
- アルゴリズムは連続的なスコアを算出しましたが、将軍たちに渡される前に最も近い整数に四捨五入されていました。→ Dell and Querubin は、この四捨五入による不連続性を利用しました。

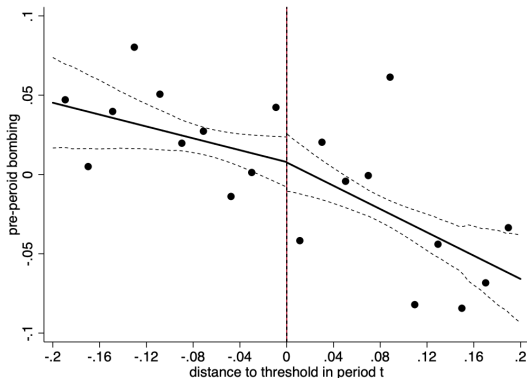
(b) Cumulative First Stage



バランスの確認

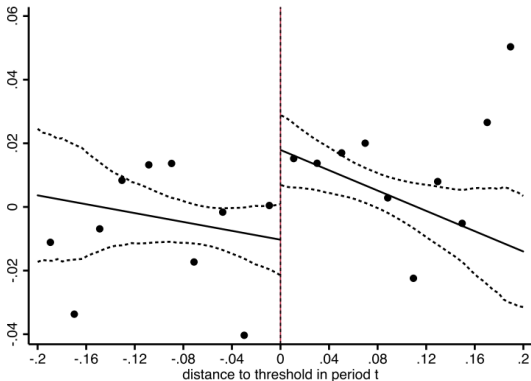
- しきい値の直上と直下の地域は、爆撃の決定が行われる前は同様の特徴を持っていました。
- 例えば、そのスコアが使われるようになる前の爆撃回数は同様でした。

(c) All Prior Quarters Bombing



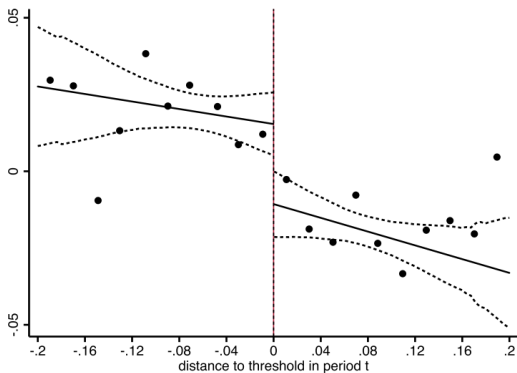
- 爆撃は村々にとって悪影響だったようです → 学校の数が増えませんでした。

(E) *Access to a Primary School (Cumulative)*



- 爆撃は軍事目標にとっても悪影響だったようです → 長期的なベトコン (VC) の活動が活発化しました。

(A) *VC Presence (Cumulative)*

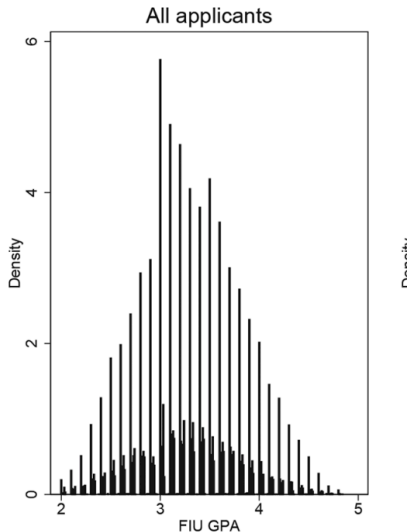


結論として...

- 爆撃は良くない ✓
- 計量経済学は良い ✓

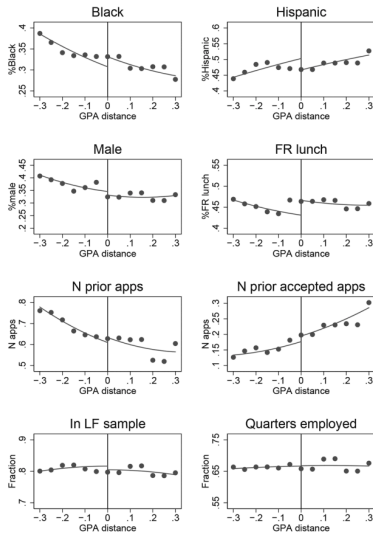
Zimmerman (2009)

- Zimmerman (2009) は、4年制の州立大学に通うことの所得への影響を研究しました。
- フロリダ州立大学の中で最も選択度の低いフロリダ国際大学 (FIU) に焦点を当てました。
- 州法は州立大学の学生に少なくとも 3.0 の GPA を求めていますでしたが、FIU は計算方法が最も寛大でした。
- したがって、FIU のしきい値付近にいる学生は、通常、フロリダの他のどの4年制公立大学の資格も持っていませんでした。
- Zimmerman は、3.0 の GPA カットオフを用いたファジー RD デザインを行いました。



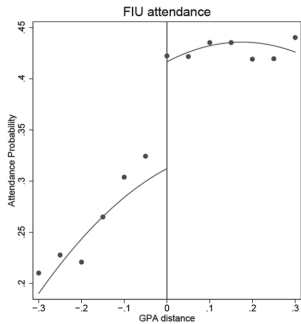
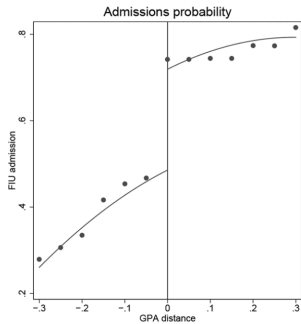
一つの懸念：カットオフの直上に集積が見られます。意欲の高い学生が、必要な GPA をぴったり取るように調整したのでしょうか？

あるいは、2.99 よりも 3.0 になる成績の組み合わせの方が多いのかもしれない。



しかし、カットオフ周辺で観察可能な特徴に不連続性は見られません。

第一段階



誘導型

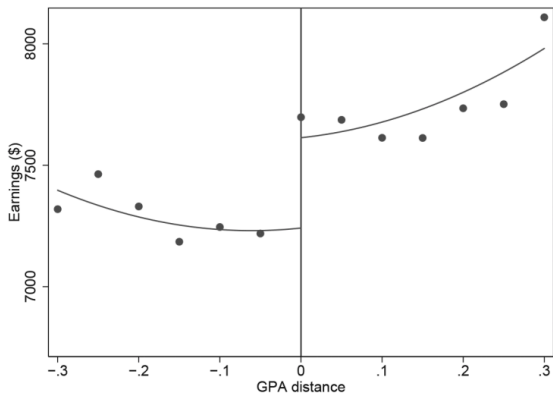


FIG. 8.—Quarterly earnings by distance from GPA cutoff. Lines are fitted val-

Table 5
Earnings Effects 8–14 Years after High School Completion

	Main	Controls	BW=.5	BW=.15	Local Linear
Reduced-form estimates:					
Above cutoff	372* (141)	366** (130)	409** (154)	479** (198)	410** (147)
Instrumental variables estimates:					
FIU admission	1,593* (604)	1,575** (584)	1,665** (645)	1,700** (621)	2,001* (696)
Years of SUS attendance	815** (276)	792** (262)	833** (271)	966*** (305)	977** (306)
BA degree	6,547* (2,496)	6,442* (2,411)	7,366* (2,998)	10,769 (5,726)	5,958** (2,024)
N	6,542	6,542	9,659	3,294	6,542

- FIU への「入学 (admission)」に対する第一段階と比較して、FIU への「進学 (attendance)」に対する第一段階は約半分であることに注意してください。
- これは、進学の LATE が 3,000-4,000 ドル (カットオフ以下の平均所得の約半分) であることを示唆しています。

- 結果は、FIU に進学することの大きな収益を示唆しています。
- 解釈における重要な問い：代替案は何でしょうか？
- 論文は、コミュニティカレッジへの進学が減少していることを示しており、私立大学からの転換もあるかもしれません。
- 底辺校の4年制公立大学におけるこの大きな収益は、「極めて選択度の高い大学」対「依然として選択度の高い大学」の差が小さいとした Dale & Krueger の結果とは対照的です。