ECON 626: Applied Microeconomics

Lecture 5:

Regression Discontinuity

Professors: Pamela Jakiela and Owen Ozier

Regression discontinuity - basic idea

A **precise** rule based on a **continuous** characteristic determines participation in a program.

Regression discontinuity - basic idea

A **precise** rule based on a **continuous** characteristic determines participation in a program.

When do we see such rules? Five example categories, but surely more:

Regression discontinuity - basic idea

A **precise** rule based on a **continuous** characteristic determines participation in a program.

When do we see such rules? Five example categories, but surely more:

- **Academic test scores:** scholarships or prizes, higher education admission, certificates of merit

Regression discontinuity - basic idea

A **precise** rule based on a **continuous** characteristic determines participation in a program.

When do we see such rules? Five example categories, but surely more:

- **Academic test scores:** scholarships or prizes, higher education admission, certificates of merit
- **Poverty scores:** (proxy-)means-tested anti-poverty programs (generally: any program targeting that features rounding or cutoffs)

Regression discontinuity - basic idea

A **precise** rule based on a **continuous** characteristic determines participation in a program.

When do we see such rules? Five example categories, but surely more:

- **Academic test scores:** scholarships or prizes, higher education admission, certificates of merit
- **Poverty scores:** (proxy-)means-tested anti-poverty programs (generally: any program targeting that features rounding or cutoffs)
- **Land area:** fertilizer program or debt relief initiative for owners of plots below a certain area

Regression discontinuity - basic idea

A **precise** rule based on a **continuous** characteristic determines participation in a program.

When do we see such rules? Five example categories, but surely more:

- **Academic test scores:** scholarships or prizes, higher education admission, certificates of merit
- **Poverty scores:** (proxy-)means-tested anti-poverty programs (generally: any program targeting that features rounding or cutoffs)
- **Land area:** fertilizer program or debt relief initiative for owners of plots below a certain area
- **Date:** age cutoffs for pensions; dates of birth for starting school with different cohorts; date of loan to determine eligibility for debt relief

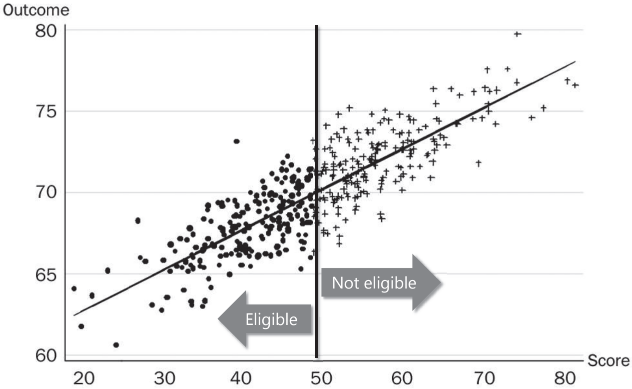
Regression discontinuity - basic idea

A **precise** rule based on a **continuous** characteristic determines participation in a program.

When do we see such rules? Five example categories, but surely more:

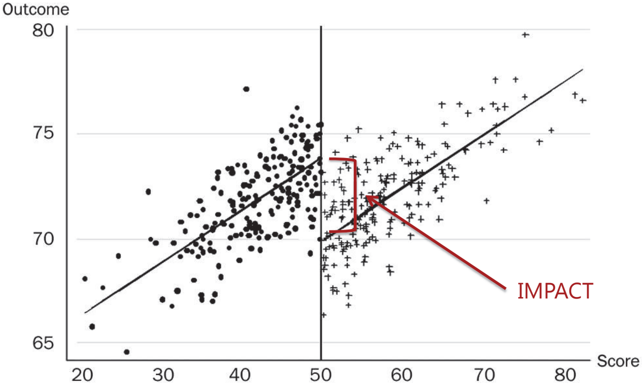
- **Academic test scores:** scholarships or prizes, higher education admission, certificates of merit
- **Poverty scores:** (proxy-)means-tested anti-poverty programs (generally: any program targeting that features rounding or cutoffs)
- **Land area:** fertilizer program or debt relief initiative for owners of plots below a certain area
- **Date:** age cutoffs for pensions; dates of birth for starting school with different cohorts; date of loan to determine eligibility for debt relief
- **Elections:** fraction that voted for a candidate of a particular party

Regression discontinuity - basic idea ("sharp")



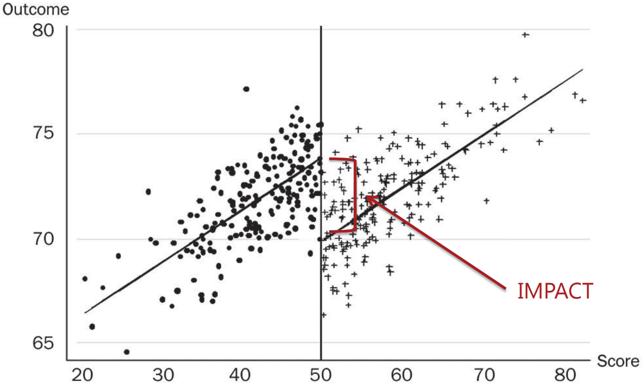
Source: Gertler, P. J.; Martinez, S., Premand, P., Rawlings, L. B. and Christel M. J. Vermeersch, 2010, Impact Evaluation in Practice: Ancillary Material, The World Bank, Washington DC (www.worldbank.org/ieinpractice)

Regression discontinuity - basic idea ("sharp")



Source: Gertler, P. J.; Martinez, S., Premand, P., Rawlings, L. B. and Christel M. J. Vermeersch, 2010, Impact Evaluation in Practice: Ancillary Material, The World Bank, Washington DC (www.worldbank.org/ieinpractice)

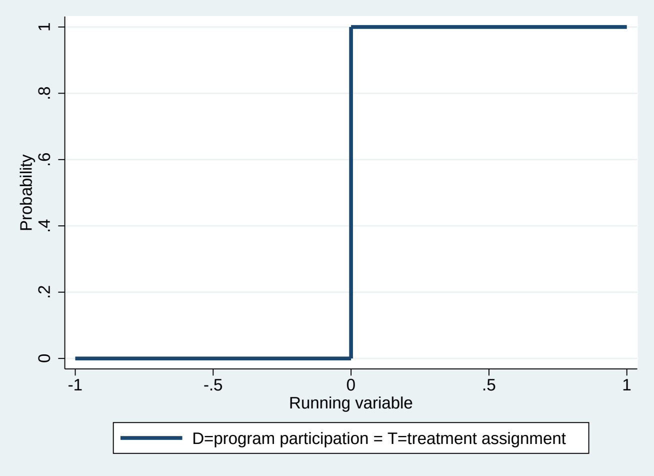
Regression discontinuity - basic idea ("sharp")



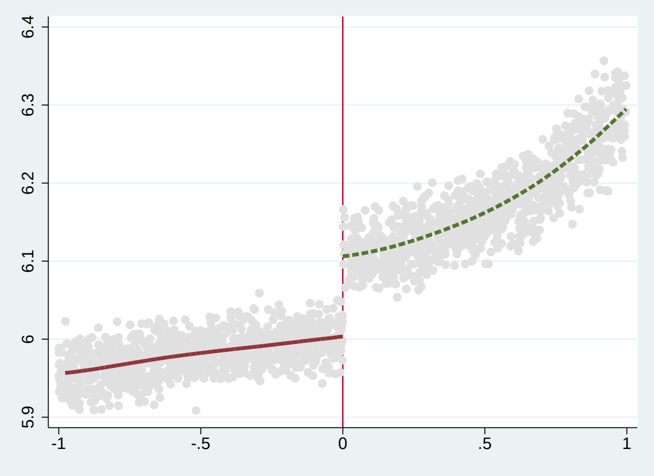
Source: Gertler, P. J.; Martinez, S., Premand, P., Rawlings, L. B. and Christel M. J. Vermeersch, 2010, Impact Evaluation in Practice: Ancillary Material, The World Bank, Washington DC (www.worldbank.org/ieinpractice)

Note: Local Average Treatment Effect

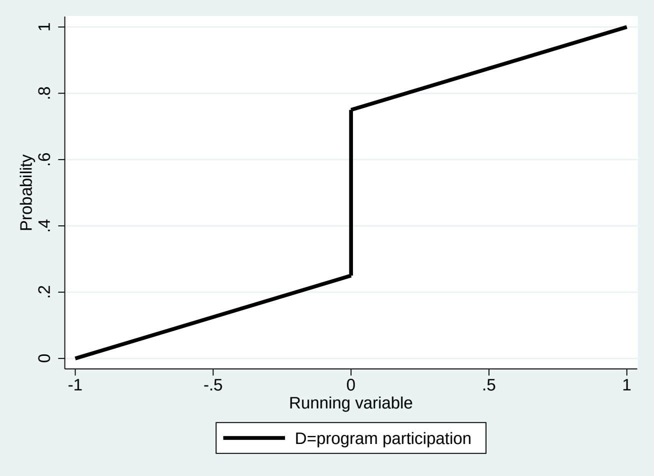
Regression discontinuity - basic idea ("sharp")



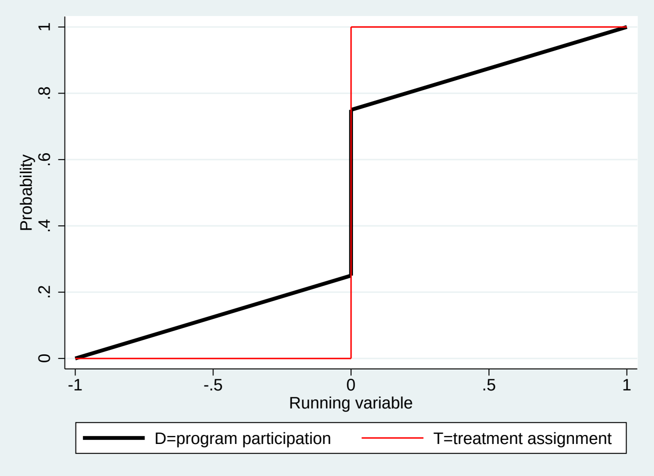
Regression discontinuity - outcome



Regression discontinuity - basic idea ("fuzzy")



Regression discontinuity - basic idea ("fuzzy")



History of the RD design - Cook (2008)

"Several themes stand out in the half century of RDD's history. One is its repeated independent discovery. ...

History of the RD design - Cook (2008)

"Several themes stand out in the half century of RDD's history. One is its repeated independent discovery. ...

- Campbell (1960; psychology / education) first named the design **regression-discontinuity**;

History of the RD design - Cook (2008)

"Several themes stand out in the half century of RDD's history. One is its repeated independent discovery. ...

- Campbell (1960; psychology / education) first named the design **regression-discontinuity**;
- Goldberger (1972; economics) referred to it as **deterministic selection on the covariate**;

History of the RD design - Cook (2008)

"Several themes stand out in the half century of RDD's history. One is its repeated independent discovery. ...

- Campbell (1960; psychology / education) first named the design **regression-discontinuity**;
- Goldberger (1972; economics) referred to it as **deterministic selection on the covariate**;
- Sacks and Spiegelman (1977,78,80; statistics) studiously avoided naming it;

History of the RD design - Cook (2008)

"Several themes stand out in the half century of RDD's history. One is its repeated independent discovery. ...

- Campbell (1960; psychology / education) first named the design **regression-discontinuity**;
- Goldberger (1972; economics) referred to it as **deterministic selection on the covariate**;
- Sacks and Spiegelman (1977,78,80; statistics) studiously avoided naming it;
- Rubin (1977; statistics) first wrote about it as part of a larger discussion of **treatment assignment based on the covariate**;

History of the RD design - Cook (2008)

"Several themes stand out in the half century of RDD's history. One is its repeated independent discovery. ...

- Campbell (1960; psychology / education) first named the design **regression-discontinuity**;
- Goldberger (1972; economics) referred to it as **deterministic selection on the covariate**;
- Sacks and Spiegelman (1977,78,80; statistics) studiously avoided naming it;
- Rubin (1977; statistics) first wrote about it as part of a larger discussion of **treatment assignment based on the covariate**;
- Finkelstein et al (1996; biostatistics) called it the **risk-allocation design**;

History of the RD design - Cook (2008)

"Several themes stand out in the half century of RDD's history. One is its repeated independent discovery. ...

- Campbell (1960; psychology / education) first named the design **regression-discontinuity**;
- Goldberger (1972; economics) referred to it as **deterministic selection on the covariate**;
- Sacks and Spiegelman (1977,78,80; statistics) studiously avoided naming it;
- Rubin (1977; statistics) first wrote about it as part of a larger discussion of **treatment assignment based on the covariate**;
- Finkelstein et al (1996; biostatistics) called it the **risk-allocation design**;
- and Trochim (1980; statistics) finished up calling it the **cutoff-based design**."

History of the RD design - Cook (2008)

"Several themes stand out in the half century of RDD's history. One is its repeated independent discovery. ...

- Campbell (1960; psychology / education) first named the design **regression-discontinuity**;
- Goldberger (1972; economics) referred to it as **deterministic selection on the covariate**;
- Sacks and Spiegelman (1977,78,80; statistics) studiously avoided naming it;
- Rubin (1977; statistics) first wrote about it as part of a larger discussion of **treatment assignment based on the covariate**;
- Finkelstein et al (1996; biostatistics) called it the **risk-allocation design**;
- and Trochim (1980; statistics) finished up calling it the **cutoff-based design**."

Boom since 1990s in economics: applications and methodology.

History of the RD design - Cook (2008)

"Several themes stand out in the half century of RDD's history. One is its repeated independent discovery. ...

- Campbell (1960; psychology / education) first named the design **regression-discontinuity**;
- Goldberger (1972; economics) referred to it as **deterministic selection on the covariate**;
- Sacks and Spiegelman (1977,78,80; statistics) studiously avoided naming it;
- Rubin (1977; statistics) first wrote about it as part of a larger discussion of **treatment assignment based on the covariate**;
- Finkelstein et al (1996; biostatistics) called it the **risk-allocation design**;
- and Trochim (1980; statistics) finished up calling it the **cutoff-based design**."

Boom since 1990s in economics: applications and methodology. See **Journal of Econometrics, 2008 Vol.142 (2) - special issue on RD.**

Thistlethwaite and Campbell (1960)

THE JOURNAL OF

EDUCATIONAL PSYCHOLOGY

Volume 51 | December 1960 | Number 6

REGRESSION-DISCONTINUITY ANALYSIS:

AN ALTERNATIVE TO THE EX POST FACTO EXPERIMENT[1]

DONALD L. THISTLETHWAITE AND DONALD T. CAMPBELL

National Merit Scholarship Corporation — *Northwestern University*

Thistlethwaite and Campbell (1960)

Observation: scholarship winners have different attitudes.

Thistlethwaite and Campbell (1960)

Observation: scholarship winners have different attitudes.
Are attitudes changed by the scholarship? (Is it a causal link?)

Thistlethwaite and Campbell (1960)

Observation: scholarship winners have different attitudes.
Are attitudes changed by the scholarship? (Is it a causal link?)

Two groups of near-winners in a national scholarship competition were matched on several background variables in the previous study in order to study the motivational effect of public recognition. The results suggested that such recognition tends to increase the favorableness of attitudes toward intellectualism, the number of students planning to seek the MD or PhD degree, the number planning to become college teachers or scientific researchers, and the number who succeed in obtaining scholarships from other scholarship granting agencies.

Thistlethwaite and Campbell (1960)

Observation: scholarship winners have different attitudes.
Are attitudes changed by the scholarship? (Is it a causal link?)

Outcome: scholarships

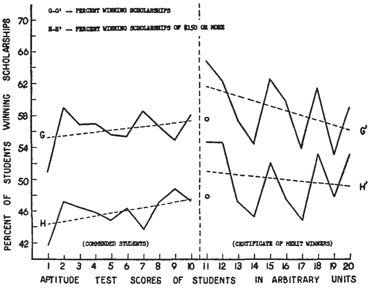


FIG. 2. Regression of success in winning scholarships on exposure determine

Thistlethwaite and Campbell (1960)

Observation: scholarship winners have different attitudes.
Are attitudes changed by the scholarship? (Is it a causal link?)

Outcome: scholarships

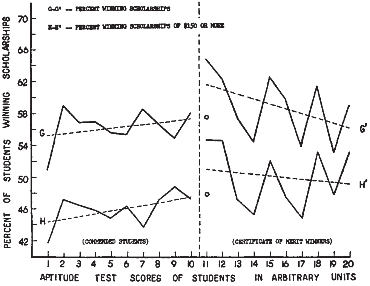


Fig. 2. Regression of success in winning scholarships on exposure determine

Outcome: attitudes

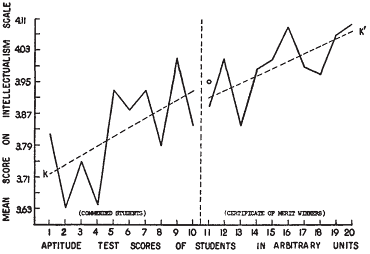


Fig. 4. Regression of attitudes toward intellectualism on exposure determine

Thistlethwaite and Campbell (1960)

Two groups of near-winners in a national scholarship competition were matched on several background variables in the previous study in order to study the motivational effect of public recognition. The results suggested that such recognition tends to increase the favorableness of attitudes toward intellectualism, the number of students planning to seek the MD or PhD degree, the number planning to become college teachers or scientific researchers, and the number who succeed in obtaining scholarships from other scholarship granting agencies. The regression-discontinuity analysis to be presented here confirms the effects upon success in winning scholarships from other donors but negates the inference of effects upon attitudes and is equivocal regarding career plans.

RD, a little more formally

We can (locally) approximate any smooth function:

RD, a little more formally

We can (locally) approximate any smooth function:

$$Y_i = f(x_i) + \rho D_i + \eta_i \tag{1}$$

Substitute:

$$f(x_i) \approx \alpha + \beta_1 x_i + \beta_2 x_i^2 + ... + \beta_p x_i^p \tag{2}$$

RD, a little more formally

We can (locally) approximate any smooth function:

$$Y_i = f(x_i) + \rho D_i + \eta_i \tag{1}$$

Substitute:

$$f(x_i) \approx \alpha + \beta_1 x_i + \beta_2 x_i^2 + ... + \beta_p x_i^p \tag{2}$$

And thus:

$$Y_i = \alpha + \beta_1 x_i + \beta_2 x_i^2 + ... + \beta_p x_i^p + \rho D_i + \eta_i \tag{3}$$

RD, a little more formally

We can (locally) approximate any smooth function:

$$Y_i = f(x_i) + \rho D_i + \eta_i \tag{1}$$

Substitute:

$$f(x_i) \approx \alpha + \beta_1 x_i + \beta_2 x_i^2 + ... + \beta_p x_i^p \tag{2}$$

And thus:

$$Y_i = \alpha + \beta_1 x_i + \beta_2 x_i^2 + ... + \beta_p x_i^p + \rho D_i + \eta_i \tag{3}$$

But because the smooth function may behave differently on either side of the cutoff, we will expand on this. First, transform x_i notationally (and for ease of regression). Let

$$\tilde{x}_i = x_i - x_0 \tag{4}$$

RD, a little more formally

We can (locally) approximate any smooth function:

$$Y_i = f(x_i) + \rho D_i + \eta_i \tag{1}$$

Substitute:

$$f(x_i) \approx \alpha + \beta_1 x_i + \beta_2 x_i^2 + ... + \beta_p x_i^p \tag{2}$$

And thus:

$$Y_i = \alpha + \beta_1 x_i + \beta_2 x_i^2 + ... + \beta_p x_i^p + \rho D_i + \eta_i \tag{3}$$

But because the smooth function may behave differently on either side of the cutoff, we will expand on this. First, transform x_i notationally (and for ease of regression). Let

$$\tilde{x}_i = x_i - x_0 \tag{4}$$

Angrist and Pishke, Chapter 6, pp. 251-267

RD, a little more formally

Angrist and Pishke, Chapter 6, pp. 251-267

Then, allowing different trends (and indeed, completely different polynomials) on either side of the cutoff (with and without the program), we can write the conditional expectation functions:

$$E[Y_{0i}] = f_0(x_i) = \alpha + \beta_{01}\tilde{x}_i + \beta_{02}\tilde{x}_i^2 + ... + \beta_{0p}\tilde{x}_i^p \tag{5}$$

$$E[Y_{1i}] = f_1(x_i) = \alpha + \rho + \beta_{11}\tilde{x}_i + \beta_{12}\tilde{x}_i^2 + ... + \beta_{1p}\tilde{x}_i^p \tag{6}$$

RD, a little more formally

Angrist and Pishke, Chapter 6, pp. 251-267

Then, allowing different trends (and indeed, completely different polynomials) on either side of the cutoff (with and without the program), we can write the conditional expectation functions:

$$E[Y_{0i}] = f_0(x_i) = \quad \alpha + \beta_{01}\tilde{x}_i + \beta_{02}\tilde{x}_i^2 + ... + \beta_{0p}\tilde{x}_i^p \tag{5}$$

$$E[Y_{1i}] = f_1(x_i) = \alpha + \rho + \beta_{11}\tilde{x}_i + \beta_{12}\tilde{x}_i^2 + ... + \beta_{1p}\tilde{x}_i^p \tag{6}$$

And because D_i is a deterministic function of x_i (this is important for writing the conditional expectation):

$$E[Y_i|X_i] = E[Y_{0i}] + (E[Y_{1i}] - E[Y_{0i}])D_i \tag{7}$$

RD, a little more formally

Angrist and Pishke, Chapter 6, pp. 251-267

Then, allowing different trends (and indeed, completely different polynomials) on either side of the cutoff (with and without the program), we can write the conditional expectation functions:

$$E[Y_{0i}] = f_0(x_i) = \qquad \alpha + \beta_{01}\tilde{x}_i + \beta_{02}\tilde{x}_i^2 + ... + \beta_{0p}\tilde{x}_i^p \tag{5}$$

$$E[Y_{1i}] = f_1(x_i) = \quad \alpha + \rho + \beta_{11}\tilde{x}_i + \beta_{12}\tilde{x}_i^2 + ... + \beta_{1p}\tilde{x}_i^p \tag{6}$$

And because D_i is a deterministic function of x_i (this is important for writing the conditional expectation):

$$E[Y_i|X_i] = E[Y_{0i}] + (E[Y_{1i}] - E[Y_{0i}])D_i \tag{7}$$

So, substituting in for the regression equation, we can define $\beta_j^* = \beta_{1j} - \beta_{0j}$ for any j, and write:

$$Y_i = \alpha + \beta_{01}\tilde{x}_i + \beta_{02}\tilde{x}_i^2 + ... + \beta_{0p}\tilde{x}_i^p + \tag{8}$$

$$\rho D_i + \beta_1^* D_i\tilde{x}_i + \beta_2^* D_i\tilde{x}_i^2 + ... + \beta_p^*\tilde{x}_i^p + \eta_i \tag{9}$$

RD, a little more formally

Angrist and Pishke, Chapter 6, pp. 251-267

But this can all really be simplified in many practical cases. For **small** values of Δ:

RD, a little more formally

Angrist and Pishke, Chapter 6, pp. 251-267

But this can all really be simplified in many practical cases. For **small** values of Δ:

$$E[Y_i|x_0 - \Delta < x_i < x_0] \approx E[Y_{0i}|x_i = x_0] \tag{10}$$

$$E[Y_i|x_0 \leq x_i < x_0 + \Delta] \approx E[Y_{1i}|x_i = x_0] \tag{11}$$

RD, a little more formally

Angrist and Pishke, Chapter 6, pp. 251-267

But this can all really be simplified in many practical cases. For **small** values of Δ:

$$E[Y_i|x_0 - \Delta < x_i < x_0] \approx E[Y_{0i}|x_i = x_0] \tag{10}$$
$$E[Y_i|x_0 \leq x_i < x_0 + \Delta] \approx E[Y_{1i}|x_i = x_0] \tag{11}$$

and then, in the most extreme case, we can take the limit:

RD, a little more formally

Angrist and Pishke, Chapter 6, pp. 251-267

But this can all really be simplified in many practical cases. For **small** values of Δ:

$$E[Y_i|x_0 - \Delta < x_i < x_0] \approx E[Y_{0i}|x_i = x_0] \quad (10)$$

$$E[Y_i|x_0 \leq x_i < x_0 + \Delta] \approx E[Y_{1i}|x_i = x_0] \quad (11)$$

and then, in the most extreme case, we can take the limit:

$$\lim_{\Delta \to 0} E[Y_i|x_0 \leq x_i < x_0 + \Delta] - E[Y_i|x_0 - \Delta < x_i < x_0] = E[Y_{1i} - Y_{0i}|x_i = x_0] \quad (12)$$

So the difference in means in an extremely (vanishingly!) narrow band on each side of the cutoff might be enough to estimate the effect of the program, ρ.

In practice, usually include linear terms and use a narrow region around the cutoff.

RD, a little more formally

Angrist and Pishke, Chapter 6, pp. 251-267

What if the assignment rule is discontinuous, but does not completely determine treatment status?

$$Prob(D_i = 1|x_i) = \begin{cases} g_1(x_i) & \text{if } x_i \geq x_0 \\ g_0(x_i) & \text{if } x_i < x_0 \end{cases}, \text{where } g_1(x_0) \neq g_0(x_0) \quad (13)$$

RD, a little more formally

Angrist and Pishke, Chapter 6, pp. 251-267

What if the assignment rule is discontinuous, but does not completely determine treatment status?

$$Prob(D_i = 1|x_i) = \begin{cases} g_1(x_i) & \text{if } x_i \geq x_0 \\ g_0(x_i) & \text{if } x_i < x_0 \end{cases}, \text{where } g_1(x_0) \neq g_0(x_0) \quad (13)$$

We need a different notation for being on the left or the right of the cutoff, now that D_i doesn't jump from zero to one. Let $T_i = \mathbb{I}(x_i \geq x_0)$.

RD, a little more formally

Angrist and Pishke, Chapter 6, pp. 251-267

What if the assignment rule is discontinuous, but does not completely determine treatment status?

$$Prob(D_i = 1|x_i) = \begin{cases} g_1(x_i) & \text{if } x_i \geq x_0 \\ g_0(x_i) & \text{if } x_i < x_0 \end{cases}, \text{where } g_1(x_0) \neq g_0(x_0) \quad (13)$$

We need a different notation for being on the left or the right of the cutoff, now that D_i doesn't jump from zero to one. Let $T_i = \mathbb{I}(x_i \geq x_0)$. Now, following the equations in the text, we arrive at two (piecewise) polynomial approximations:

$$Y_i = \mu + \kappa_1 x_i + \kappa_2 x_i^2 + ... + \kappa_p x_i^p + \pi\rho T_i + \zeta_{2i} \quad (14)$$

RD, a little more formally

Angrist and Pishke, Chapter 6, pp. 251-267

What if the assignment rule is discontinuous, but does not completely determine treatment status?

$$Prob(D_i = 1|x_i) = \begin{cases} g_1(x_i) & \text{if } x_i \geq x_0 \\ g_0(x_i) & \text{if } x_i < x_0 \end{cases}, \text{where } g_1(x_0) \neq g_0(x_0) \quad (13)$$

We need a different notation for being on the left or the right of the cutoff, now that D_i doesn't jump from zero to one. Let $T_i = \mathbb{I}(x_i \geq x_0)$. Now, following the equations in the text, we arrive at two (piecewise) polynomial approximations:

$$Y_i = \mu + \kappa_1 x_i + \kappa_2 x_i^2 + ... + \kappa_p x_i^p + \pi\rho T_i + \zeta_{2i} \quad (14)$$

$$D_i = \gamma_0 + \gamma_1 x_i + \gamma_2 x_i^2 + ... + \gamma_p x_i^p + \pi T_i + \zeta_{1i} \quad (15)$$

RD, a little more formally

Angrist and Pishke, Chapter 6, pp. 251-267

What if the assignment rule is discontinuous, but does not completely determine treatment status?

$$Prob(D_i = 1|x_i) = \begin{cases} g_1(x_i) & \text{if } x_i \geq x_0 \\ g_0(x_i) & \text{if } x_i < x_0 \end{cases}, \text{where } g_1(x_0) \neq g_0(x_0) \quad (13)$$

We need a different notation for being on the left or the right of the cutoff, now that D_i doesn't jump from zero to one. Let $T_i = \mathbb{I}(x_i \geq x_0)$. Now, following the equations in the text, we arrive at two (piecewise) polynomial approximations:

$$Y_i = \mu + \kappa_1 x_i + \kappa_2 x_i^2 + ... + \kappa_p x_i^p + \pi \rho T_i + \zeta_{2i} \quad (14)$$

$$D_i = \gamma_0 + \gamma_1 x_i + \gamma_2 x_i^2 + ... + \gamma_p x_i^p + \pi T_i + \zeta_{1i} \quad (15)$$

So to estimate ρ, we use instrumental variables, and in essence divide the coefficient estimate on T_i in the "first stage" regression (variations on Equation 15) by the coefficient estimage on T_i in the "reduced form" regression (variations on Equation 14).

RD, a little more formally

Angrist and Pishke, Chapter 6, pp. 251-267

What if the assignment rule is discontinuous, but does not completely determine treatment status?

$$Prob(D_i = 1|x_i) = \begin{cases} g_1(x_i) & \text{if } x_i \geq x_0 \\ g_0(x_i) & \text{if } x_i < x_0 \end{cases}, \text{where } g_1(x_0) \neq g_0(x_0) \quad (13)$$

We need a different notation for being on the left or the right of the cutoff, now that D_i doesn't jump from zero to one. Let $T_i = \mathbb{I}(x_i \geq x_0)$. Now, following the equations in the text, we arrive at two (piecewise) polynomial approximations:

$$Y_i = \mu + \kappa_1 x_i + \kappa_2 x_i^2 + ... + \kappa_p x_i^p + \pi\rho T_i + \zeta_{2i} \quad (14)$$

$$D_i = \gamma_0 + \gamma_1 x_i + \gamma_2 x_i^2 + ... + \gamma_p x_i^p + \pi T_i + \zeta_{1i} \quad (15)$$

So to estimate ρ, we use instrumental variables, and in essence divide the coefficient estimate on T_i in the "first stage" regression (variations on Equation 15) by the coefficient estimage on T_i in the "reduced form" regression (variations on Equation 14). Again, as in IV: **Exclusion restriction, standard errors**

RD, a little more formally

Angrist and Pishke, Chapter 6, pp. 251-267

What if the assignment rule is discontinuous, but does not completely determine treatment status?

$$Prob(D_i = 1|x_i) = \begin{cases} g_1(x_i) & \text{if } x_i \geq x_0 \\ g_0(x_i) & \text{if } x_i < x_0 \end{cases}, \text{where } g_1(x_0) \neq g_0(x_0) \quad (13)$$

We need a different notation for being on the left or the right of the cutoff, now that D_i doesn't jump from zero to one. Let $T_i = \mathbb{I}(x_i \geq x_0)$. Now, following the equations in the text, we arrive at two (piecewise) polynomial approximations:

$$Y_i = \mu + \kappa_1 x_i + \kappa_2 x_i^2 + ... + \kappa_p x_i^p + \pi\rho T_i + \zeta_{2i} \quad (14)$$

$$D_i = \gamma_0 + \gamma_1 x_i + \gamma_2 x_i^2 + ... + \gamma_p x_i^p + \pi T_i + \zeta_{1i} \quad (15)$$

So to estimate ρ, we use instrumental variables, and in essence divide the coefficient estimate on T_i in the "first stage" regression (variations on Equation 15) by the coefficient estimage on T_i in the "reduced form" regression (variations on Equation 14). Again, as in IV: **Exclusion restriction, standard errors**

Practical considerations

Five basic issues are highlighted by Guido Imbens and Thomas Lemieux in their paper, *Regression discontinuity designs: A guide to practice*:

- Specification tests: density, covariates, other jumps
- **Density: analogy to attrition. This is conceptually important.**

Practical considerations

Five basic issues are highlighted by Guido Imbens and Thomas Lemieux in their paper, *Regression discontinuity designs: A guide to practice*:

- Specification tests: density, covariates, other jumps
- **Density: analogy to attrition. This is conceptually important.**
- Visualization

Practical considerations

Five basic issues are highlighted by Guido Imbens and Thomas Lemieux in their paper, *Regression discontinuity designs: A guide to practice*:

- Specification tests: density, covariates, other jumps
- **Density: analogy to attrition. This is conceptually important.**
- Visualization
- Specification: polynomial order (linear in many cases), "kernel"

Practical considerations

Five basic issues are highlighted by Guido Imbens and Thomas Lemieux in their paper, *Regression discontinuity designs: A guide to practice*:

- Specification tests: density, covariates, other jumps
- **Density: analogy to attrition. This is conceptually important.**
- Visualization
- Specification: polynomial order (linear in many cases), "kernel"
- Bandwidth

Practical considerations

Five basic issues are highlighted by Guido Imbens and Thomas Lemieux in their paper, *Regression discontinuity designs: A guide to practice*:

- Specification tests: density, covariates, other jumps
- **Density: analogy to attrition. This is conceptually important.**
- Visualization
- Specification: polynomial order (linear in many cases), "kernel"
- Bandwidth
- Standard errors (confidence interval)

Practical considerations

Five basic issues are highlighted by Guido Imbens and Thomas Lemieux in their paper, *Regression discontinuity designs: A guide to practice*:

- Specification tests: density, covariates, other jumps
- **Density: analogy to attrition. This is conceptually important.**
- Visualization
- Specification: polynomial order (linear in many cases), “kernel”
- Bandwidth
- Standard errors (confidence interval)

Practical considerations

Five basic issues are highlighted by Guido Imbens and Thomas Lemieux in their paper, *Regression discontinuity designs: A guide to practice*:

- Specification tests: density, covariates, other jumps
- **Density: analogy to attrition. This is conceptually important.**
- Visualization
- Specification: polynomial order (linear in many cases), “kernel”
- Bandwidth
- Standard errors (confidence interval)

Methodological updates and extensions:

Practical considerations

Five basic issues are highlighted by Guido Imbens and Thomas Lemieux in their paper, *Regression discontinuity designs: A guide to practice*:

- Specification tests: density, covariates, other jumps
- **Density: analogy to attrition. This is conceptually important.**
- Visualization
- Specification: polynomial order (linear in many cases), "kernel"
- Bandwidth
- Standard errors (confidence interval)

Methodological updates and extensions:

- Cattaneo, Calonico, and Titiunik series (SE's, visualization)

Practical considerations

Five basic issues are highlighted by Guido Imbens and Thomas Lemieux in their paper, *Regression discontinuity designs: A guide to practice*:

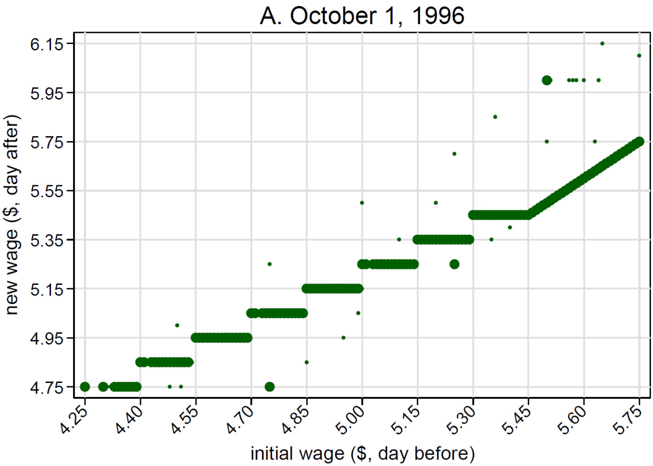
- Specification tests: density, covariates, other jumps
- **Density: analogy to attrition. This is conceptually important.**
- Visualization
- Specification: polynomial order (linear in many cases), "kernel"
- Bandwidth
- Standard errors (confidence interval)

Methodological updates and extensions:

- Cattaneo, Calonico, and Titiunik series (SE's, visualization)
- Card, Lee, Pei, and Weber (Kink design)

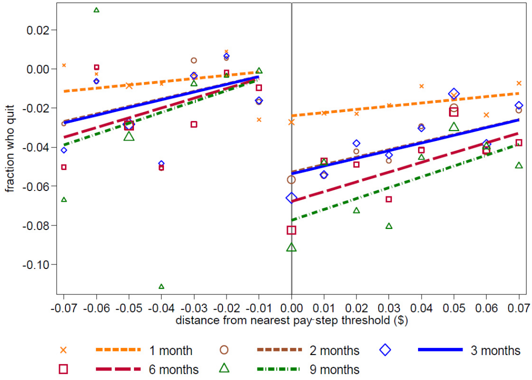
Visualization: Dube, Giuliano, Leonard example

Figure 1. Wages on days before and after each minimum wage increase



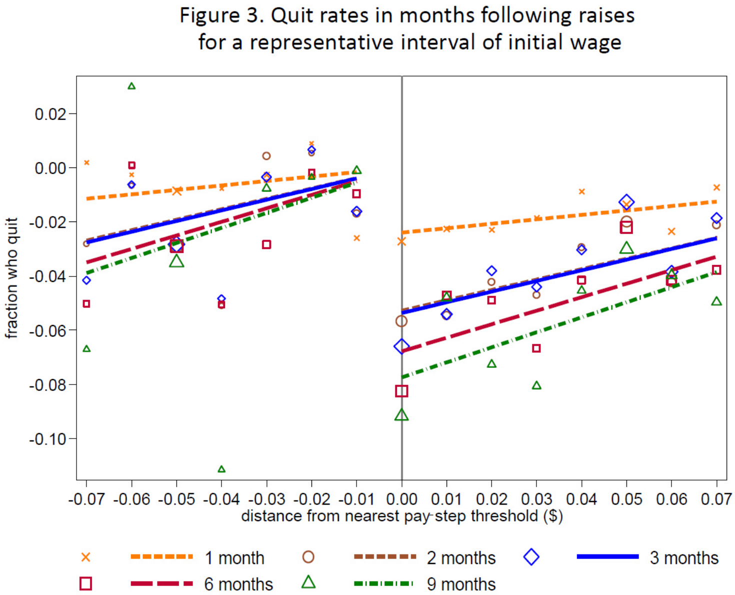
Visualization: Dube, Giuliano, Leonard example

Figure 3. Quit rates in months following raises
for a representative interval of initial wage



Note: The figure shows residuals from the RD model of quits with baseline controls (as in Table 2, row 3) for a representative interval of the running variable (initial own wage). For visual simplification, the running variable is normalized as the distance to the nearest pay-step threshold (see text for details). The lines show the fitted relationship between residualized quit rates and the normalized running variable. For each value of the normalized running variable, the data points are constructed by adding back to fitted values the mean of the residuals taken across all 12 intervals. Marker size is scaled by the number of observations at each value. For all series, the intercepts are normalized to be zero at the left limit of the threshold, so the value at the right limit is the estimated effect of the $.10 discontinuity in the wage. Estimation samples are as in Table 2.

Visualization: Dube, Giuliano, Leonard example



Manipulation of the running variable

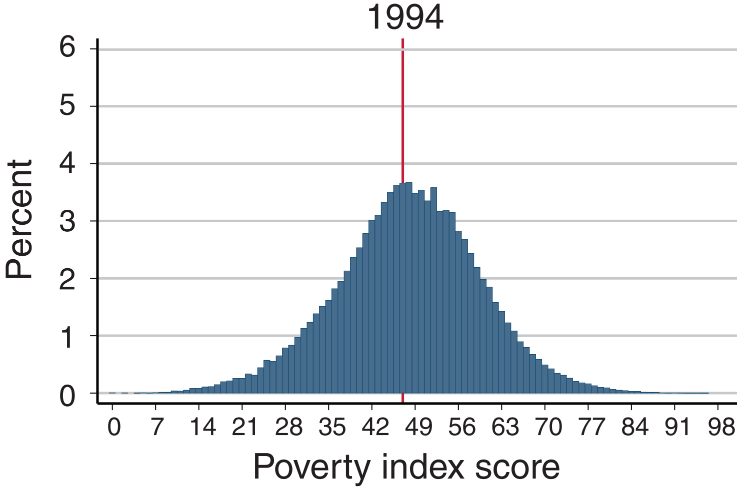
What if the population of potential program participants
is able to precisely influence the running variable,
and knows the program assignment rule?

Manipulation of the running variable

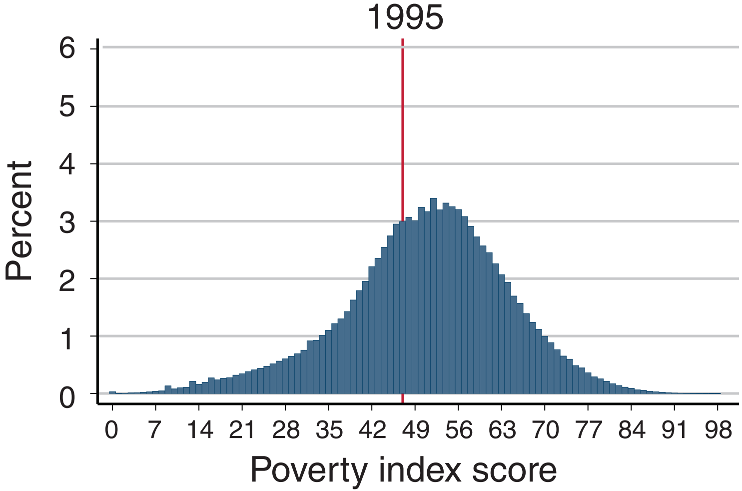
What if the population of potential program participants
is able to precisely influence the running variable,
and knows the program assignment rule?

Example from Camacho and Conover (2011) in Colombia:
program rule became known in 1997;
watch what happens.

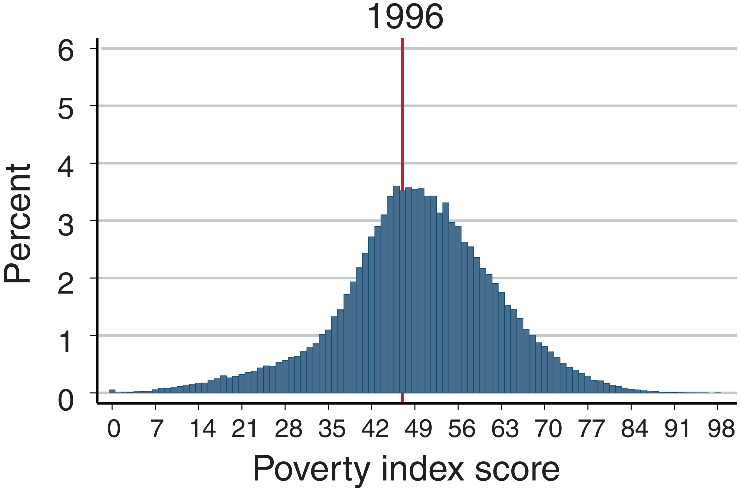
Poverty score distribution - Camacho and Conover (2011) in Colombia



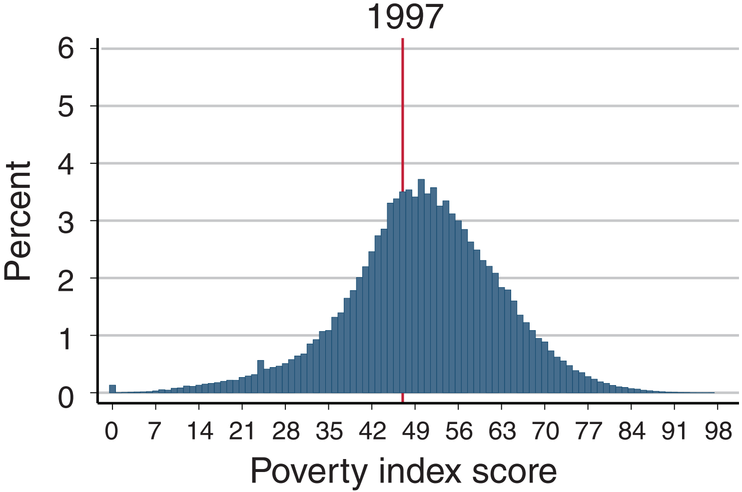
Poverty score distribution - Camacho and Conover (2011) in Colombia



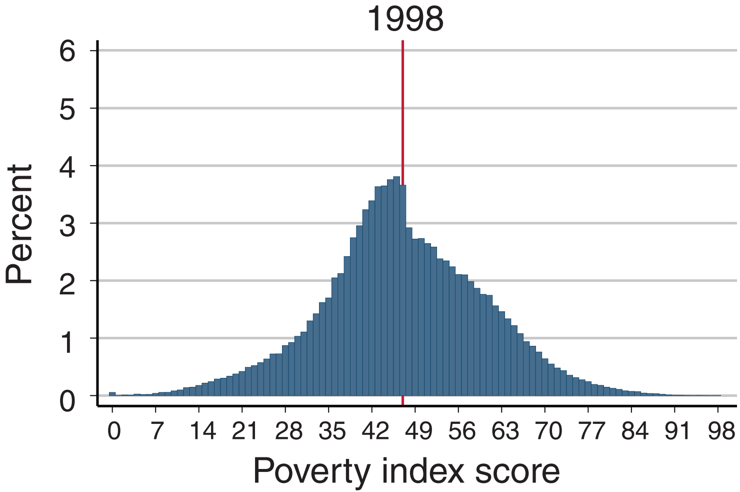
Poverty score distribution - Camacho and Conover (2011) in Colombia



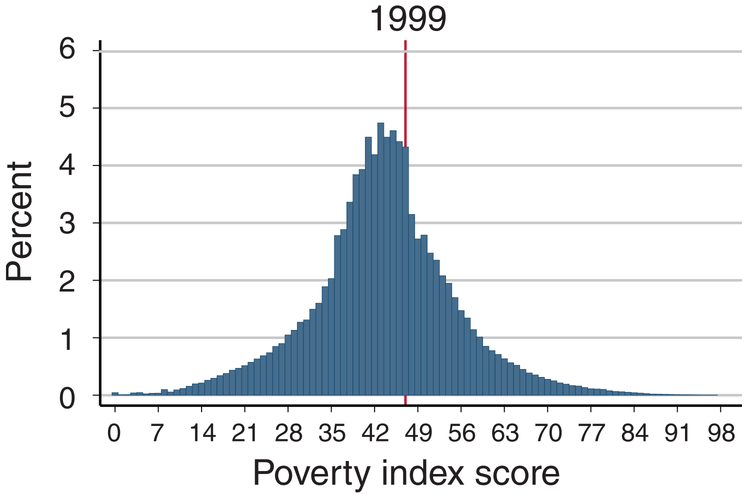
Poverty score distribution - Camacho and Conover (2011) in Colombia



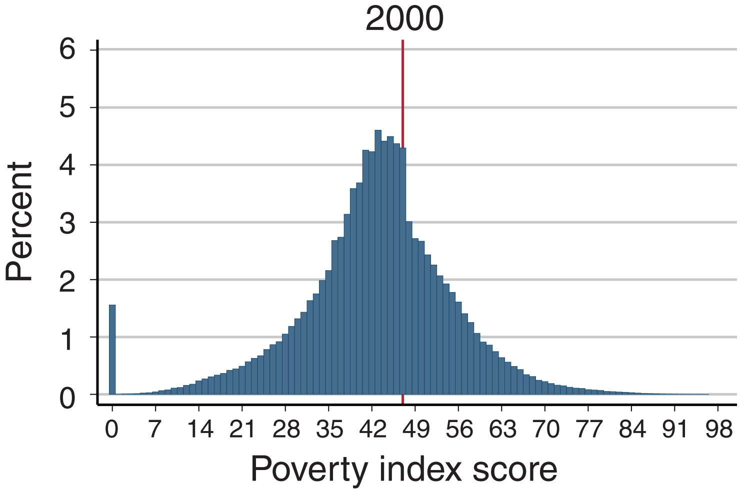
Poverty score distribution - Camacho and Conover (2011) in Colombia



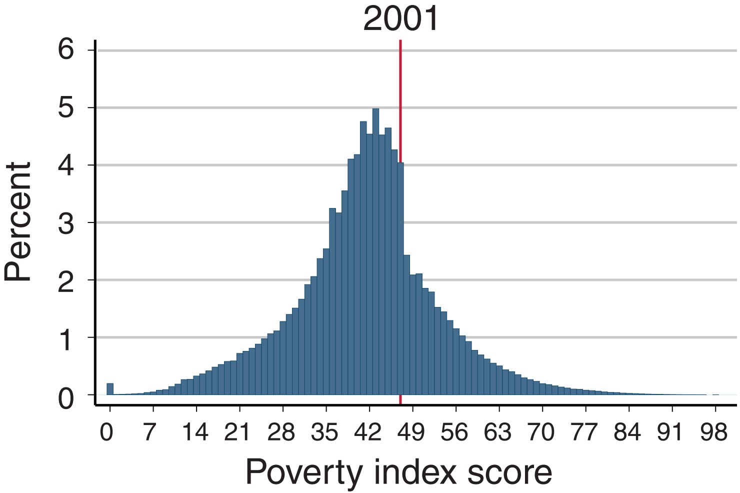
Poverty score distribution - Camacho and Conover (2011) in Colombia



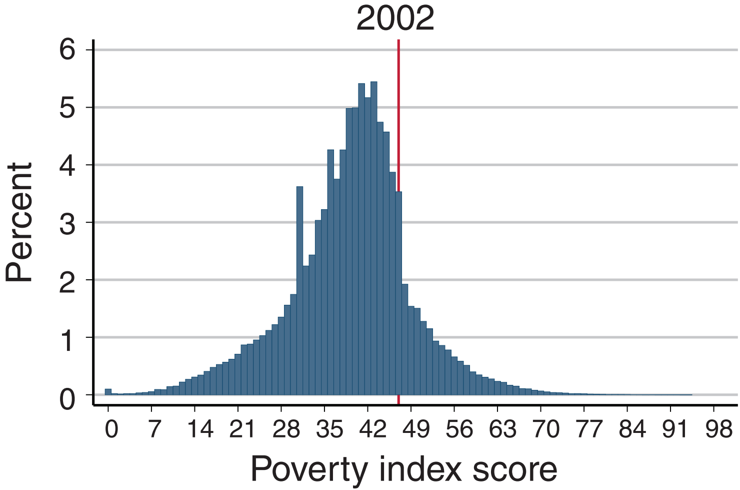
Poverty score distribution - Camacho and Conover (2011) in Colombia



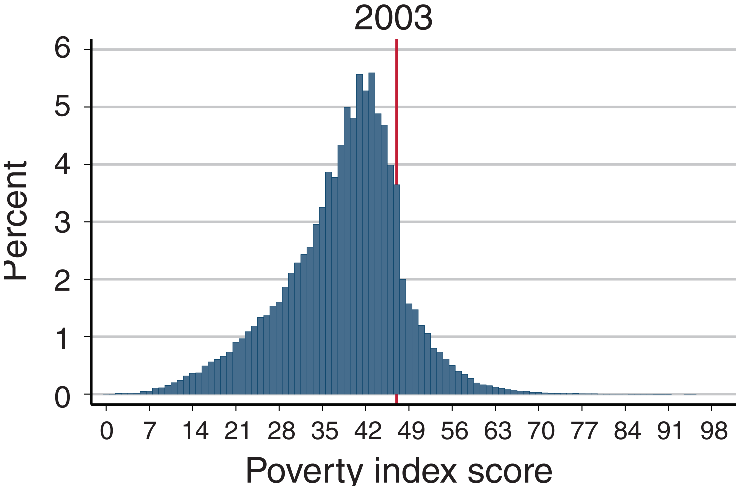
Poverty score distribution - Camacho and Conover (2011) in Colombia



Poverty score distribution - Camacho and Conover (2011) in Colombia



Poverty score distribution - Camacho and Conover (2011) in Colombia



An example.