A perfect discontinuity with nothing behind it

Every child at 124 mm was referred and 2% of children at 125 mm were. The assignment is as sharp as a rule can be, the density shows no sign of anyone being nudged across, and the design still cannot run — because no outcome was ever measured on the children who were not referred.

Pierre Robentz Cassion

Lesson 5 of 8

Impact Evaluation Methods — Rules, and limits

What this lesson covers

- A rule is a design
- What a regression discontinuity needs
- The condition this register fails
- What it would have taken
- Where thresholds hide in programme data
- Report it whole
- What comes next

A rule is a design — In Python

```python
import pandas as pd
import numpy as np

screening = pd.read_csv("muac-screening-artibonite-2024.v1.csv")
screening["referred"] = screening["outcome"] != "no-action"

window = screening[screening["muac_mm"].between(118, 132)]
print(window.groupby("muac_mm")["referred"].agg(["mean", "size"]).round(3))
```

A rule is a design — In R

```
library(dplyr)

screening |>
  filter(between(muac_mm, 118, 132)) |>
  summarise(referred = mean(outcome != "no-action"), n = n(), .by = muac_mm)
```

A rule is a design

MUAC (mm)	Referred	n
122	100.0%	44
123	100.0%	47
124	**100.0%**	**45**
125	**2.0%**	**51**
126	0.0%	78
127	0.0%	70

A rule is a design

- One millimetre moves the referral probability from 1.00 to 0.02 — Children at 124 and 125 mm are, in every way that...
- That is the closest thing to a randomised experiment that routine data produces — and it arrives free, because the...

What a regression discontinuity needs

- One: the assignment must actually be sharp at the cut-off — Checked above
- Two: nobody may manipulate the running variable — If a screener who wants a child treated records 124 instead of 126,...

What a regression discontinuity needs — In Python

```
counts = screening["muac_mm"].value_counts().sort_index()
print(counts.loc[120:130])

below = screening["muac_mm"].between(120, 124).sum()
above = screening["muac_mm"].between(125, 129).sum()
print(f"120-124: {below}   125-129: {above}")

screening["last_digit"] = screening["muac_mm"] % 10
print((screening["last_digit"].value_counts(normalize=True).sort_index() * 100).round(1))
```

What a regression discontinuity needs — In R

```
table(screening$muac_mm)[as.character(118:132)]
table(screening$muac_mm %% 10)
```

What a regression discontinuity needs

- The counts rise smoothly through the cut-off: 45 at 124, 51 at 125, 78 at 126 — There is no pile-up just below 125, and...
- Terminal digits are flat, 9.6% to 10.8% across all ten — No rounding to fives, no avoidance of the cut-off
- Three: an outcome must be measured on both sides — It is not, and that ends the analysis

The condition this register fails — In Python

```
print(screening.columns.tolist())
print(f"children screened once: {(screening['child_id'].value_counts() == 1).sum()}")
print(f"children screened twice: {(screening['child_id'].value_counts() == 2).sum()}")
```

The condition this register fails — In R

```
names(screening)
table(table(screening$child_id))
```

The condition this register fails

- The file records a measurement, a referral, and nothing else — 4,194 of 4,206 children appear exactly once
- The treatment register does not fill the gap — `cmam-admissions-2024` holds outcomes for treated children only, and its...
- An impact estimate needs the untreated side — Without it, the sharpest assignment rule in the sector produces nothing

What it would have taken

- A follow-up measurement on children just above the cut-off — Not all of them — children between 125 and 130 mm,...
- The same outcome on both sides — Recovery is defined differently for treated and untreated children, and a comparison...
- Enough children near the cut-off — The estimate uses a window, and everything outside it is discarded

What it would have taken — In Python

```
for bandwidth in (2, 3, 5, 10):
    band = screening[
        screening["muac_mm"].between(125 - bandwidth, 124 + bandwidth)]
    print(f"±{bandwidth} mm: {len(band)} children"
          f" ({(band['muac_mm'] < 125).sum()} below, "
          f"{(band['muac_mm'] >= 125).sum()} above)")
```

What it would have taken — In R

```
# The bandwidth trade-off, in four lines.
```

What it would have taken

Bandwidth	Children	Below cut-off	Above
$\pm$2 mm	221	92	129
$\pm$3 mm	335	136	199
$\pm$5 mm	556	187	369
$\pm$10 mm	1,210	265	945

What it would have taken

- A narrow bandwidth is more credible and less precise — Children within 2 mm of the cut-off are almost identical to each...
- Choose the bandwidth before seeing the outcome, and report the estimate at two or three others — A result that appears...

Where thresholds hide in programme data

Threshold	Programme	Running variable
MUAC $<$ 125 mm	Supplementary feeding	Arm circumference
MUAC $<$ 115 mm	Therapeutic feeding	Arm circumference
Poverty score below a cut-off	Cash transfer	Proxy means test
IPC Phase 3	Emergency response	Area classification
Below a coverage target	Facility support	Reported coverage

Where thresholds hide in programme data

- Each is a natural experiment nobody designed, and each needs the same three checks — The third — an outcome on both. . .
- That is a data-collection decision, not an analysis one — Deciding to measure a sample of ineligible units costs money. . .

Report it whole — Example (cont.)

```
Supplementary feeding eligibility: assessment of a regression discontinuity

  Assignment: referral to TSFP below MUAC 125 mm.

  Sharpness     100% referred at 124 mm (n=45), 2.0% at 125 mm (n=51).
                The rule is applied without exception.

  Manipulation No pile-up below the cut-off: counts are 45, 51, 78 at 124,
                125 and 126 mm. Terminal digits are uniform (9.5-10.8%).
                No evidence the running variable was adjusted.

  Outcome       Not available. The register records screening and referral
                only; 4,194 of 4,206 children appear once, and no follow-up
                measurement exists on either side of the cut-off.

  Conclusion    The design is available and cannot be estimated. To make it
```

Report it whole — Example (cont.)

```
estimable, re-measure MUAC at 8 weeks on a sample of children
screened between 125 and 130 mm, using the same definition
applied to referred children. 556 children sit within
+/-5 mm of the cut-off, of whom 369 were not referred.
```

Report it whole

- A design assessment that concludes "not estimable" is a deliverable — and this one is more useful than a forced...

What comes next

- The threshold design failed on data that does not exist.

Read the full lesson, with runnable code

https://data-analysis.cassion.dev/courses/impact-evaluation-methods/ie-05-the-threshold-as-a-design