Empirical Corporate Finance

Spring 2024

A Quick Guide for the Perplexed PhD Student

Table of contents

- **Part 1:** How to write a (good) paper
  - Key: Contribution, data, methodology, writing
  - Other important parameters
- **Part 2:** Case study: "Incentivizing Financial Regulators"
  - Contribution
  - Identification strategies
  - Alternative strategies
- This is an extremely abbreviated guide; use it as a roadmap to dig deeper into the specific topics

Part 1:
How to write a (good) paper

The key ingredients

- Almost all empirical CF papers have some causal story: x causes y

- The four key ingredients for a good paper are:
  1. Does it matter whether $x \rightarrow y$ or not? <= **contribution**
  2. Do I have the right data for $x \rightarrow y$? <= **data**
  3. Do I have the right strategy to identify $x \rightarrow y$? <= **methodology**
  4. Can I communicate 1-3 effectively? <= **writing**

- A benevolent view of "good"
  - Potentially different from "publishable", "cited", "invited to conferences"; but hopefully correlated
  - We can talk about some of the frictions later

(1/4) Contribution

Contribution

- Probably the main top reason for rejection in top journals

- Common types of **contribution**: the paper addresses –
  1. A new question (emerging literature)
  2. An old question without an answer (anomaly / puzzle)
  3. An old question with wrong answers (mistake in the literature)
  4. An old question with imprecise answers ("identification paper")

- To assess the contribution, you need to know the literature
  – Do not jump into unknown territory; easy way to antagonize
  – Beware of downplaying / mischaracterizing existing literature

Significance and relevance

- The contribution needs to be <u>significant:</u>
  - There is almost always “a” contribution
  - The significance bar depends on the journal and is ill-defined
  - What differentiates “significant” from “insignificant”?

- The contribution needs to be <u>relevant:</u>
  - It should be a “finance” (“accounting”) question
  - However, tastes vary in the cross-section and over time
  - Expanding the boundaries: high risk & reward
  - Specialized journals exist, e.g., JFI and JLE

Related concepts

- You will sometimes hear synonyms for "what is the contribution?"

- **What is the null hypothesis?**
  - Is it reasonable to think that x does not cause y?
  - Always tricky to defend ex-post

- **What is the external validity?**
  - Do we learn something beyond the narrow empirical setting?
  - Tension between internal & external validity

- **Is this the right framing of the paper?**
  - Which contribution should be featured?

Substitution and complementarities

- The **methodology** affects the **contribution**
  - Weak identification: risk of overpromising & underdelivering
  - Tight identification: risk to the external validity

- Sometimes, **methodology** & **contribution** are substitutes:
  - Perfect methodology can substitute for narrow contribution
  - Bigger contribution can substitute for weaker methodology
  - Papers with a new x ("we propose a new measure of…")
  - Papers with a new y ("we document a new phenomenon…")
  - Papers with a new $x \to y$ link ("we propose a novel driver of…")

(2/4) Data

- Assembling a new dataset is not officially required in empirical CF
- However, it is a huge X-factor for junior scholars
  – Easier to find & explain the contribution
  – Signals skills, creativity, and dedication
- Premium for:
  – Hard-earned datasets (skills, labor, capital)
  – Potential for multiple papers
  – Sharing the dataset, when the time is right
- Various tips:
  – Start early; it takes lots of trial and error
  – Search for low-hanging fruit (inexpensive; overlooked)
  – Learn Python, NPL, scraping techniques

(3/4) Methodology

- Most of this course will focus on methodology
- Part 2 of this guide uses a case study to illustrate the main identification strategies
- Two general tips:
  - You must develop a deep understanding of the <u>institutional setting</u> of your paper
  - You learn a lot by doing!

(4/4) Writing

Writing

- This is how you communicate parts 1-3 to your peers
  - Contribution; data structure; and identification strategy

- Neglecting this step is a sure way to kill the paper prematurely:
  - Think long and hard on the “right” framing
  - Beware of over-selling and under-selling
  - Title > Abstract > 1st paragraph > Introduction > Rest of paper
  - Adhere to general professional standards:
    - Avoid typos & grammar issues
    - Use LaTex
    - Figures & tables: aesthetic & self-explanatory

Additional factors

Other considerations

- **Timeline**:
  - Writing & publishing takes longer than you think
  - Writing & publishing is more uncertain than you think
  - Tenure path narrows faster than you think

- **Coauthors:**
  - Solo versus coauthors; peer versus senior co-author
  - Complementary skills, availability, communication

Sidenote #1: publication timeline

- In a typical journal, the steps are:
  - Desk-reject: yes/no decision, typically by the editor
  - Paper sent out to 1-3 referees (sometimes also an associate editor)
  - Decision by the editor

- Outcomes in top journals (exceptions apply):
  - Desk rejection rate: 30% or higher
  - Overall rejection rate: 95%
  - First round <u>at least</u> 70 days; <u>at least</u> two rounds total
  - Total process <u>at least</u> one year, conditional on ex-post success

- Common strategies: the earlier the better; multiple balls in the air; look for multiple shots (dual submission)…

Sidenote #2: tenure clock

- Typical timeline:
  - Midterm evaluation <= almost always after 3 years
  - Promotion to Associate <= almost always after 6 years
  - Tenure <= after 6 or 9 years
  - Possibility for tenure extensions and accelerations

- Primary criterion: top-3 publications
  - Many school will also recognize econ top-5, accounting top-3
  - Some schools expand the top-3 top top-4/5/6/7
  - Number of required publications varies a lot
  - Quality & impact of publications also matter

Sidenote #3: the process

- No “correct” process: you can…
  - Start with an idea (identify an important gap)
  - Start with an identification opportunity (methodology)
  - Start with data

Part 2:
Case study

“Incentivizing Financial Regulators”

- My first solo publication
  - First draft in my 2nd year at the PhD program
  - Ultimately became my job market paper
  - Accepted at the RFS a month after I graduated from NYU
  - Won the Rising Scholar Award in 2022

- We will use it to illustrate key steps of writing & evaluating a paper:
  - Contribution
  - Empirical strategies

I study how promotion incentives within the public sector affect financial regulation. I assemble individual data for all SEC enforcement attorneys between 2002 and 2017, including enforcement cases, salaries, and ranks. Consistent with tournament model, attorneys with stronger promotion incentives are involved in more enforcement, especially against severe misconduct, and in tougher settlement terms. For identification, I rely on cross-sectional tests within offices and ranks and on exogenous variation in salaries resulting from a conversion to a new pay system. The findings highlight a novel link between incentives and regulation and show that the regulator's organizational design affects securities markets. (*JEL* H11, J31, J45, K22, M51, M52)

- $x =$ Promotion incentives
- $y =$ SEC enforcement
- Story: promotion incentives increase SEC enforcement
- Channel: bigger incentive => bigger effort => more enforcement

Contribution

What are the contributions

- There are three listed contributions:
    1. Promotion incentives affect financial regulation
    2. Promotion incentives affect state bureaucracy
    3. Tournament incentives affect employee-level output

Critical assessment

- **Significance:** why is this an important contribution?
- **Relevance:** why is this relevant for a top-3 finance journal?
- **Framing:** which contribution is the primary one?
- **Null hypothesis:** what is the null? Is it reasonable?
- **External validity:** Do we learn something beyond the narrow empirical setting?

Methodology

What is the causal story?

- The goal: see if promotion incentives (x) increase enforcement (y)

$$Enforcement = \beta \cdot Incentive + \epsilon$$

- The main finding: $\beta > 0$
- Main issue: endogeneity
  - Incentive (x) is not randomly assigned
  - In other words, hard to give β a causal interpretation

- Need to be specific! There are three typical challenges:
  1. Measurement error
  2. Reverse causality
  3. Omitted variable **<= the big one!**

(1) Measurement error

- Is this the right way to measure the <u>independent variable</u>?
  - Could lead to a biased estimator
  - In this paper: incentive includes promotion probability; two-rank promotions; relocations across offices

- Is this the right way to measure the <u>outcome variable</u>?
  - If it's a random noise, not a problem (larger SE)
  - If it's systematic (correlated with x), can cause bias
  - In this paper: not all enforcement actions are equal; could be sample selection (administrative vs. civil)

(2) Reverse causality

- Also called simultaneity bias
- x is determined by y, in addition / instead of causing y
- Typical fix is to use lagged x variables
- Side note: an extension of that is “bad controls”
  - If some of the regressors are affected by x, do not include them

(3) Omitted variable

- Try to think of an omitted variables (OV) which:
  - Is not subsumed under fixed effects and controls; AND
  - Is related to x; AND
  - Is related to y

- For example: an OV that –
  - Separates attorneys from the same year×office, AND
  - Increases with enforcement, AND
  - Increases with incentive

- Can you think of examples here?

- Example of observed OV: Tenure
  - This is a mis-specified model; fix by adding controls

- Example of unobserved OV: skills
  - Try to find a proxy: education, bonus
  - Otherwise, it is a challenge; need identification strategy
  - You can discuss the sign of the bias (attenuated versus inflated); however, with multiple regressors, it is difficult to determine

Identification strategy

- The paper uses the following strategies:
  1. Panel regression with fixed effects
  2. Instrumental variable

- We will sketch out alternative (unrealistic) strategies:
  3. Regression discontinuity design (RDD)
  4. Bunching
  5. Difference-in-differences

Panel regression

Intuition

- Use the panel structure of the data to remove omitted variables
- Almost all CF papers have some version of this strategy
- Tracking attorneys over time helps remove "bad variation" (confounding effects); lots of alternative stories can be ruled out:
  - Unobserved heterogeneity over hierarchy ranks => grade FE
  - Unobserved heterogeneity over time => year FE
  - Unobserved heterogeneity over local offices => year×office FE
  - Unobserved heterogeneity over local offices & ranks => year×office×rank FE

Things to watch out for

- Did we throw out some of the "good" variation?
  - With year×office×rank FE, we control for the target salary
  - Variation only comes from higher/lower salary
  - Cannot measure response to higher/lower target salary

- Measurement error:
  - With tight FE, the relative importance of the "bad" variation can cause measurement error

- Computation power is an effective constraints on FE

- Not recommended for non-linear models (Probit, Logit…)

- Old-school strategies, highly discouraged:
  - Remove mean from independent variables (within transformation)
  - Remove mean from dependent variable
  - Control for the mean of the dependent variable
  - Add a series of indicators (LSDV)

- However, some still use first-differences model
  - Instead of x_t, use $\Delta_{\mathrm{t}} x = x_t - x_{t-1}$

- Reporting R^2: within-R^2 or adjusted-R^2 (standard Stata output)

- Using the predicted values is an issue
  - Incidental parameters problem

Instrumental variable

Intuition

- Randomly assign (part of) the incentive
  - When transitioning to a new pay system, salaries were rounded up to the nearest point on the grid

- Assumption #1: relevance
  - Higher round-up => higher salary => lower incentive
  - This is testable ("first stage")

- Assumption #2: exclusion
  - Round-up affects enforcement <u>only through incentive</u>
  - That's the toughest assumption to defend
  - Untestable; try to rule out alternative stories one by one

Things to watch out for

- Avoid overidentification
  - Rule of thumb: one instrument per endogenous variable ("just identified")

- Examine evidence for the relevance assumption:
  - First-stage results need to be clearly reported in a table
  - F-statistic$<$10: weak instrument (high SE in the 2nd stage)
  - F-statistic$\gg$10: mechanical correlation?

- Strong economic story for the exclusion restriction:
  - The instrument affects the outcome only through x

- Old-school instruments, highly discouraged:
  - Instrument x with lagged x
  - Instrument x with average x (e.g., leave-one-out industry average)

- External validity: (LATE)
  - Instrument shifts x for compliers; what about non-compliers?
  - Subsample analysis of the first stage, to see if groups with small residuals (compliers) are different

- OLS and IV coefficients:
  - Need to be consistent with the OV story
  - IV$\gg$OLS: should be suspicious (although very typical…)

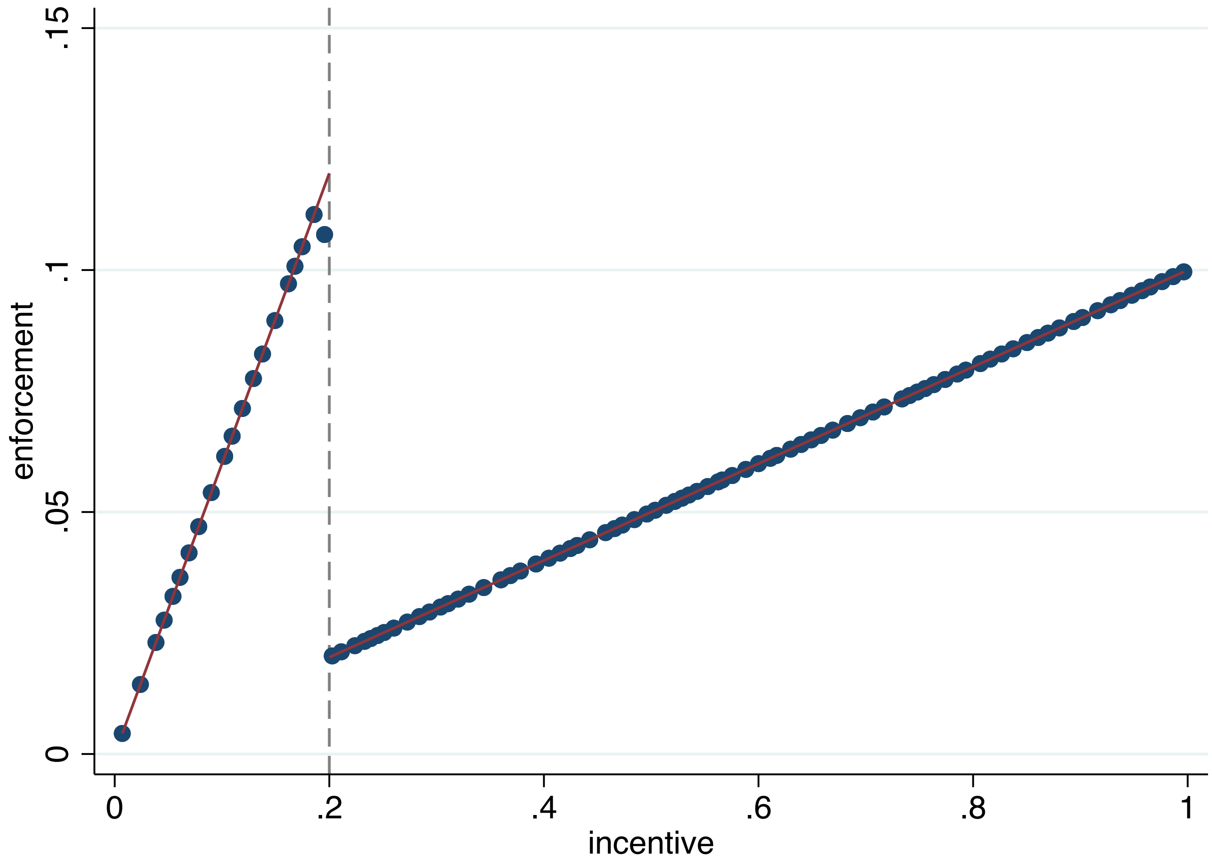
Regression discontinuity (RDD)

Imaginary scenario

- **Sharp RD:** let's assume that…
  - If incentive <0.2, probability of promotion is 100%
  - If incentive ≥ 0.2, probability of promotion drops to 0%
  - Low-ranked attorneys (high incentive) need to wait their turn

- **Fuzzy RD:** similar, but less dramatic…
  - If incentive <0.2, probability of promotion is 90%
  - If incentive ≥ 0.2, probability of promotion drops to 45%

- The strategy: compare attorneys around the threshold
  - If incentive <0.2: strong effect of incentive on enforcement
  - If incentive ≥ 0.2: weak effect of incentive on enforcement



- **Forcing variable** (x) determines assignment around the **threshold**
- The **threshold** creates a **discontinuity** in the **treatment**

Things to watch out for

- Width of the band around the threshold
  - Narrow band (just above & just below): precise but noisy
  - Wide band: less noisy, but increases bias
  - Wider band requires high-order polynomials for $(x - \bar{x})$

- Placebo: are there sub-populations, or years, in which the threshold does not apply?

- Assumption #1: no manipulation of x
  - Even if they do, not all hope is lost… bunching!

- Assumption #2: local continuity
  - Outcome would have been similar around threshold, w/o treatment
  - Test for balanced covariates: are subject above/below different?

- Assumption #3: no heterogeneous effect
  - Treatment effect is locally continuous around the threshold
  - Why was the threshold chosen? Did they choose it because anticipating that subject above/below will respond differently?
  - Affects external validity, not internal validity

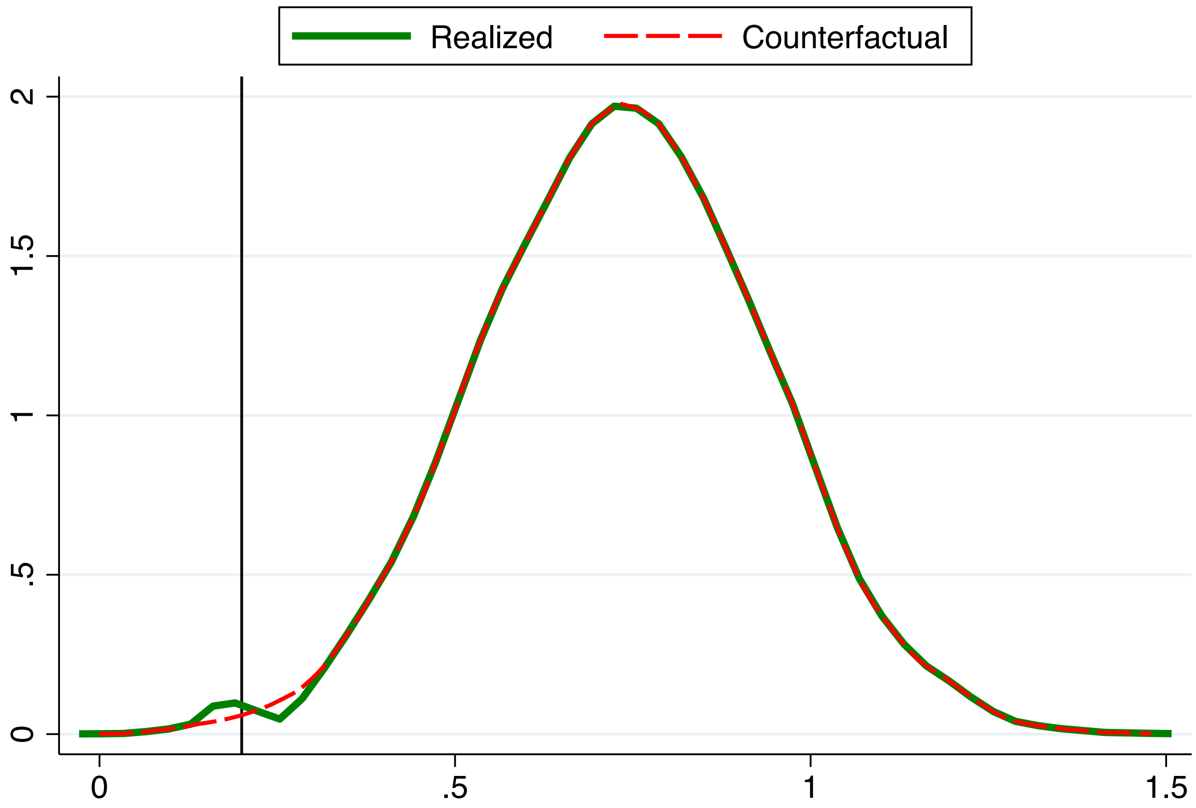
Bunching

Imaginary scenario

- Recall the fuzzy RD story…
  - If incentive <0.2, probability of promotion is 90%
  - If incentive ≥ 0.2, probability of promotion drops to 45%

- Now suppose that attorneys can manipulate their incentive
  - For example: move to an office with low target pay, resulting in smaller incentive
  - RDD estimates are biased, because subjects select their treatment

- The strategy: search for abnormal mass just below the threshold
  – Plot the realized distribution of incentives
  – Plot a counterfactual distribution **<= that's the tough part**
  – Search for unusual clustering just below the threshold



Realized
Counterfactual
0
.5
1
1.5
2
0
.5
1
1.5

Things to watch out for

- How to compute the counterfactual distribution?
  - Must develop deep understanding of how salaries are determined
  - Consider factors such as bin size & polynomial order

- Key assumption: absent the promotion policy, the distribution would have been smooth around the threshold

- Potential violations:
  - Mechanical bottlenecks, causing attorneys to be "stuck" just below the threshold
  - Pay policies that coincide with the threshold and motivating attorneys to stay just below it

Things to watch out for

- All you can show is that attorneys <u>care about</u> promotion incentives
- Essentially, a revealed preference argument
- This can be very important; but be careful with subsequent causal interpretation:
  - Suppose attorneys below the threshold (low incentive) file less enforcement that attorneys above (high incentive)
  - Does the incentive cause enforcement?
  - Not necessarily! An OV ("type") can cause attorneys to bunch & to enforce less

Natural experiment

Intuition

Imagine the following scenario

- Until 2010, promotion probability was 5%
- Since 2011, promotion probability depends on your last name:
    - A-M: promotion probability = 5%
    - N-Z: promotion probability = 50%
- We anticipate that the higher promotion rates incentivize more enforcement
- This is an example of a "natural experiment"
    - As opposed to controlled / randomized experiment
- If you see something like that, you can use the following strategies:

Simple diff (time series)

- Using only treated attorneys (N-Z), estimate:

$$\boxed{y_{i,t} = \alpha + \beta \cdot Post_t + \epsilon_i}$$

- $y_{i,t}$ = enforcement by attorney i at time t
- $Post_t = 1$ after the treatment went into effect
- β = average treatment effect (ATE)
- Assumption: absent the treatment, treated attorneys would have had the same enforcement before & after (on average)
- Violation: concurrent time-series changes
  - Suppose the SEC launched a bonus program In 2011
  - $\beta > 0$ in part because of the bonus program

Simple diff (cross section)

- Using only post-treatment period (2011 onwards), estimate:

$$\boxed{y_{i,t} = \alpha + \beta \cdot D_i + \epsilon_i}$$

- $y_{i,t}$ = enforcement by attorney i at time t
- $D_i = 1$ if attorney i is treated (last name N-Z)
- β = average treatment effect (ATE)
- Assumption: absent the treatment, treated & control attorneys would have had the same enforcement (on average)
- Violation: cross-sectional differences
  - Suppose attorneys with N-Z are highly skilled
  - $\beta > 0$ in part because of cross-sectional differences in ability

Difference-in-differences

$$\boxed{y_{i,t} = \alpha + \beta_1 Post_t + \beta_2 D_i + \beta_3 Post_t D_i + \epsilon_i}$$

- Use all the data: before & after, treated & control groups
- β_1 = average change over periods (before & after)
- β_2 = average difference between groups (treated & control)
- β_3 = difference-in-difference coefficient
- A more common approach is to use generalized diff-in-diff:

$$\boxed{y_{i,t} = \alpha + \beta_3 Post_t D_i + \lambda_i + \lambda_t + \epsilon_i}$$

Parallel trends assumption

- Absent treatment, the change in y for treated group and the change in y for control group would have been the same
- Solves some concerns in simple diff specifications:
  - OK to have time-series changes around the event, as long as they are group-invariant
  - OK to have cross-sectional differences between groups, as long as they are time-invariant
- Violation of the assumption:
  - In 2011, the SEC gave N-Z people new computers
  - In 2011, the SEC paid N-Z for continued educations
  - Changes coincide with the treatment & affect only one group

Indirect tests for PT assumption

- Pre-treatment comparison of treated & control groups
  - If they differ, raises concerns of unobservables
- Dynamic regressions
  - Compare treated & control every year separately
  - If the gaps start before the treatment: problem
- Use staggered treatment or treatment reversal
  - Harder to argue for repeated violations of PT assumption
- Placebo
  - Variables that should not clearly *not* be affected by the treatment
- Triple diff
  - Treatment should be larger for a specific group