

ECON 594: Applied Economics

Introduction and Ideas for Ideas

Sam Norris
University of British Columbia

Plan for today

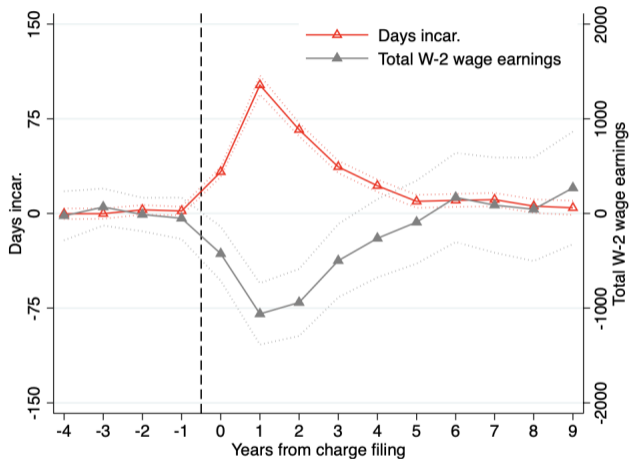
Overview

Ideas for ideas

Who am I?

- Professor, Vancouver School of Economics
- Economics of crime, labour economics, applied econometrics
 - Effects of prisons and criminal records, labour markets for justice involved people
 - Effect of immigration on wages/employment of natives
 - How does IV work with multiple instruments or treatments?
- Email: sam.norris@ubc.ca
- Office Hours: Thursdays, 8:30am–12:30pm, 10-minute slots (May 21, then weekly June 4 – August 6)
 - Iona 209 (or Zoom link if you cannot come in-person)
 - Sign-up sheet here
- Class: Mon/Tues/Wed, 10–11:30am
 - No class on May 18 (Victoria Day)

My favourite graph, currently



Who is Uyseok?

Teaching Assistant: Uyseok Lee

- Ph.D. Student in the department
- Email: leeuyseok@gmail.com
- Office Hours: Mondays 3–5pm and Wednesdays 3–5pm, in Iona 434
- Office hours signup sheet

Who are you?

What you'll do

- Write a good empirical paper: one that tells us something about the world
- Real data, a credible research design, an actual contribution
- I assume you already know the empirical methods (e.g. diff-in-diff), but we'll review the basics
- Other lectures on where ideas come from, how to make tables and figures, how to present, how to write

AI usage

- I encourage you to use AI: speed up coding, test models, check your writing
 - Claude is almost certainly better than you at coding
 - Your comparative advantage is creativity, big-picture thinking, analytical clarity
- But you still need to understand your own research
 - A one-shot Claude paper is pointless
- The successful projects come from working with AI, not delegating to it
- Oral exam on August 11: I'll ask why you made the design and specification choices you made, and how you think about the results
 - 10 minutes, in person

The lectures

- How to Use AI
- Panel Data and Difference-in-Differences
- Event Study Designs
- Tables and Figures
- Regression Discontinuity
- Instrumental Variables
- Measurement Error and Other Empirical Concerns
- How to Present
- How to Write a Paper

Course evaluation

- Progress report: 20%
- Class presentation (and attendance at presentations): 30%
- Final paper: 50%

Progress report (20%)

- Upload to Canvas by Wednesday June 3rd midnight
- Maximum three pages (excluding references) on:
 - The research question and why it matters
 - What your contribution is relative to existing work (citations needed here)
 - Your plan for data (sources, planned cleaning, etc.)
 - Your plan for analysis (what is the approach, e.g. regression discontinuity, and what is your empirical specification, will you cluster the standard errors, etc.)
- See Canvas for a few high-scoring past examples
- OK for the project to change later

Class presentation (30%)

- 20-minute presentation of preliminary version of the project
- Randomly assigned to one of six dates: June 15, 16, 17, 22, 23, 24
- By this point you should have some preliminary regressions
- “How to Present” class on June 2nd for more practical tips
- I will count attendance toward the grade (support your classmates!)

Final paper (50%)

- Following the short-article format of AER:Insights:
 - Maximum 6,000 words (see link above for how to calculate)
 - Maximum five main exhibits (figures or tables)
 - An abstract of 100 or fewer words
 - You may have an appendix with additional figures and tables (e.g. robustness checks)
- Due Sunday August 9th midnight
- Oral examinations August 11

Plan for today

Overview

Ideas for ideas

Coming up with an idea

- The process is very different to PhD idea generation
 - Very little time (3 weeks until progress report due)
 - Important to focus on feasibility
 - Less so on contribution
- The idea should be “empirical”
 - Use and analyse real (not simulated) data
 - Broadly applied microeconomics
 - Can be descriptive or causal
- Some methods more feasible (e.g. diff-in-diff better than experiments)

The three Ps

- Procrastination is bad
 - Tempting to spend weeks excited about an idea before working on it
 - Work out quickly if idea #1 works so that you have time for a backup
 - “Fail fast” idea #2
- Prioritisation is good
 - Tempting to do the fun stuff first (e.g. reading a book about a policy)
 - But you should first check for “fatal flaws” (e.g. does the data exist?)
- Pitch early and often (to students as well as me)

Examples from previous years

- Of the 9 best papers from 2019 to 2021 (most posted to Canvas):
 - 8 are “causal” and 1 is “descriptive”
 - Methods: diff-in-diff / fixed effects (3), IV (4), RD (1), experiment (1), bunching (1)
 - Fields: Labour (5), Development (1), Political Economy (1), Behavioural (1), Macro (1)
- Idea origins (7 of 9 theses):
 - Newspaper articles
 - Disagreed with another paper / from reading a paper in class
 - Knew a policy / issue from growing up in place X

Yellen Fire in a Crowded Theatre: Central Bank Communication in Times of Crisis

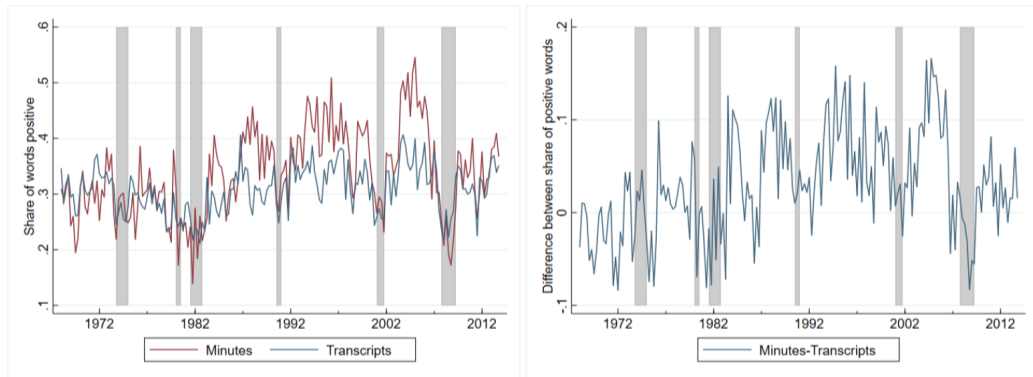
By DOGA AKAR BILGIN

August 5, 2019

In this paper, I examine how the Fed's public statements differ from its private sentiment. Using a Markov-switching regression, I find that Fed minutes are persistently optimistic relative to transcripts during expansions, while they are in line during recessions. Moreover, I find evidence of state-dependent passthrough with respect to new developments, though this result is not robust to alternative specifications. This work provides contributions to the empirical literature on central bank transparency and a new set of stylized facts for theoretical work.

The descriptive one: results

FIGURE 1. SENTIMENT IN TRANSCRIPTS VERSUS MINUTES



Note: The chart on the left shows the sentiment in minutes and FOMC transcripts based on the Loughran and McDonald (2011) dictionary. The chart on the right shows the difference between the sentiment in these two documents. As can be seen, minutes become substantially more positive than transcripts during expansions starting in the mid-1980s.

What makes this thesis compelling?

- Descriptive facts can be hard to package as a paper but:
 - Interesting research question: does the Fed communicate strategically?
 - Clever idea for empirical test: compare **private** transcripts to **public** minutes
 - Requires innovative measurement: natural language processing
 - Can't just write this paper by writing the code "line y x"
 - No empirical evidence before on strategic communication of the Fed

Education and Health: Evidence From The Cultural Revolution, 1966-1976

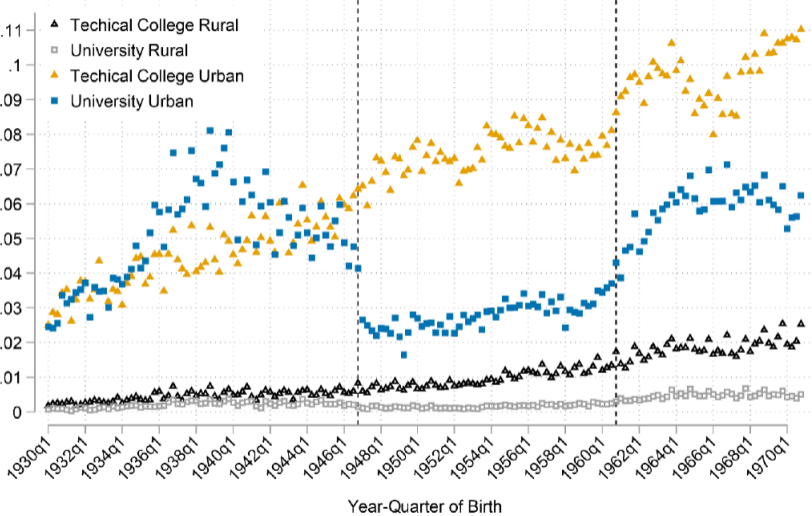
Zhaoxuan Wang

ECON 594 MA Major Essay
University of British Columbia

Abstract

This essay examines the long-term consequences on educational attainment and health due to the massive school disruptions and the suspension of the university entrance examination during China's Cultural Revolution. Using a regression discontinuity approach and an instrumental variable approach, this study quantifies a substantial loss in educational attainment among the affected cohorts born between 1947 and 1961. These impacts, however, did not lead to poor health status in their late life and were not passed down to their children. The results show no significant effects of education on physical health, mental health, chronic condition, hospitalization, and employment disability. By contrast, some evidence indicates a causal role of education in enhancing cognition, reducing drinking and smoking, and improving one's capacity to work.

An RD in year of birth



Why does this work well?

- Regression discontinuity design in birth cohort:
 - Many possible datasets to use for outcomes
 - Just need to know year of birth
 - Large sample using 1% population censuses (18 million)
- Bells and whistles (e.g. post-double-selection from Belloni et al. 2014)

The one that builds on a research design known to work well

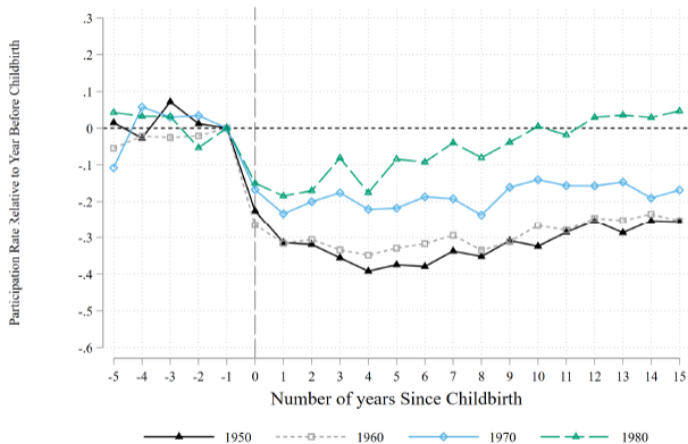
Jamie Emery's

“Who Pays the Child Penalty? Evidence From the Panel Study of Income Dynamics”

Abstract

The impact of children on the labour market outcomes of women relative to men (the “child penalty”) is well-documented, yet there is a paucity of evidence on the mechanisms behind it. I explore the importance of gendered norms and preferences and present three main findings. First, there is extensive heterogeneity in the penalty across women based on their race, marital status, and birth cohort. Second, there is a strong link between the penalty and individual-level gender-related beliefs. Third, the penalty is passed down from parents to their daughters. Taken together, my findings demonstrate that gender norms drive the penalty.

Child penalty heterogeneity



Why does this work well?

- Low-risk
 - Child penalty event study regressions are tried and tested
 - Little fear of fatal identification flaw
 - Focus on unexplored heterogeneity
- Many other possible ideas with this flavour:
 - Take an existing paper and test in country X instead of Y
 - Take an existing paper but explore new outcomes or heterogeneity
- Beautiful figures (my weakness)

The one with null results

The effects of immigration on voting in Canada

Maria teNyenhuis*

July 31, 2020

Abstract

An influx of asylum seekers has been associated with increased support for far-right anti-immigrant parties in some countries. In Canada, however, high levels of immigration do not appear to have led to a strong shift in public opinion or voting behaviour. This is somewhat exceptional in an increasingly anti-immigrant world, but there is little systematic analysis on this subject. In this paper, I employ a fixed effects model to examine the effects of immigration on votes for the three main parties in seven federal elections. To account for potential endogeneity of immigrant location, I supplement my analysis using an instrumental variable strategy that relies on the historical immigrant settlement patterns. I do not find strong evidence that voting behaviour is affected by immigration. Overall, as a number of papers have asserted before, immigration does not seem to be a key partisan issue in Canadian elections.

What matters is the ex ante research design

- Null results are fine if...
 - The question was reasonable to ask and unanswered
 - The research design has “statistical power” to answer the question
- Put another way, null results are not so great if
 - The policy you are evaluating is tiny (so no-one expected it to have effects anyway)
 - The null is imprecisely estimated (big confidence interval)

Final thought: power

Your paper will live or die on power

- (power = the likelihood we can reject a given null)
- In other words: the SEs need to be small relative to the hypothesis we are testing
- Needed precision scales with expected effect size

Think about power early and often

- SEs decrease at a $\approx \frac{1}{\sqrt{N_c}}$ rate, where N_c = number of clusters
- 4x the clusters = 2x the precision
- Use other papers to guesstimate power
- Be hard-headed

Next up

- How to Use AI (tomorrow)
- Panel Data and Difference-in-Differences (Wednesday)
- No class next Monday May 18 (Victoria Day)
- Event Study Designs (next Tuesday)
- Tables and Figures (next Wednesday)