My Epic Failures

Ben Olken

The selection problem

- You will probably read ~1,000 papers in grad school – if not more.

- But they are all highly selected!
    - They got to the point where they got written.
    - They got published. Probably in a good journal.
    - And they were sufficiently influential to make it onto a grad school syllabus.

- This is not the typical experience.
    - Instead, most projects fail. Even your professors' projects! All the time. Seriously.
    - And even most papers that succeed change a lot from conception to conclusion.

Goal for today

- Goal for today: examples of all the many, many ways my projects have died

- Why? Besides catharsis for me.

- Two goals:
    - Recognize that this is normal
    - Help you get examples of when it makes sense to kill a project and move on

- Agenda
    - My litany of woes
    - A success story.... if at first you don't succeed, sometimes you should keep trying!
    - And sometimes you should adapt.

A catalog of project death

- Here are some ways my projects have died.

- Selection rule:
    - Each of these examples has an actual project were I put in several weeks of work on it – in most cases, much, much more than that. So these are all real projects.
    - Almost all of these projects have some documents associated with them (slide decks, results documents, etc).

- Excluding therefore the most common reason projects die: Decided another project was more promising and never started it.

My litany of woe

- Found other papers that were too similar
- Model was too hard / intractable / uninteresting
- Theorem turned out not to be true
- Data fundamentally didn't really exist for what we wanted to do
- Couldn't get the data
- Realized the data wasn't good enough
- Couldn't kill the pre-trends
- Data / empirical problems we couldn't surmount
- Couldn't find the right natural experiment
- No first stage
- Results were inconclusive
- Pilots revealed project was completely misconceived
- Couldn't convince a partner to do something
- Couldn't find the right field setting
- Field visits revealed something totally unexpected
- Government didn't pass a law we expected
- Couldn't agree on NDA

My litany of woe

- Found other papers that were too similar
- Model was too hard / intractable / uninteresting
- Theorem turned out not to be true
- Data fundamentally didn't really exist for what we wanted to do
- Couldn't get the data
- Realized the data wasn't good enough
- Couldn't kill the pre-trends
- Data / empirical problems we couldn't surmount
- Couldn't find the right natural experiment
- No first stage
- Results were inconclusive
- Pilots revealed project was completely misconceived
- Couldn't convince a partner to do something
- Couldn't find the right field setting
- Field visits revealed something totally unexpected
- Government didn't pass a law we expected
- Couldn't agree on NDA

How not to get data

- Project idea: "More Donuts, Less Crime"
  - Does police presence deter / displace crime?

- Empirical idea:
  - Police tend to go to donut shops at night to get coffee late at night.
  - So if you open a donut shop, how does that change crime?
  - Control group: other fast food restaurants.

- Data
  - Had geocoded data on all openings of Dunkin Donuts and McDonalds in Massachusetts
  - Wanted to match to geocoded crime data

- Requested this from the Commonwealth of Massachusetts

How not to get data

From: eleches@fas.harvard.edu
Sent: Tuesday, March 11, 2003 2:58 PM
To: XXX@pol.state.ma.us
Cc: bolken@fas.harvard.edu
Dear Mr. YYY:
Thank you very much for your response regarding the crime data from the state of Massachusetts. My co-author (Benjamin Olken) and I have attached a short research proposal outlining the purpose of our project. If you have any questions about our project please feel free to contact us. Please be assured that we are willing to work with you to ensure confidentiality of the data. If you would like to find out more about us, you might want to access our web pages which are linked below.
Sincerely, Benjamin Olken and Christian Pop-Eleches

How not to get data

Outside of police stations, one place where police are likely to congregate, particularly at night (when one thinks that the presence of police nearby may be a more important deterrent to crime) are stores or restaurants that are open during the night, where police officers may take their breaks. If one can identify these stores, one could examine the relationship between the placement of these stores and crime rates in the immediate surrounding area....

Perhaps the most common type of store that fits this criteria is, stereotypical though this may seem, a donut store. According to Glenn Norstrem of the St. Paul police department, "The stories about cops and donuts come from the fact that police officers work odd hours when most restaurants and stores are closed. In past years, the donut shops were all that were open at night. When an officer wanted to stop for a break, coffee and donuts were convenient." Other possibilities to investigate include 24 hour convenience stores, such as Store 24 and 7-Eleven.

In Massachusetts, by far the largest chain of donut stores is Dunkin Donuts, with over 500 stores located in Massachusetts--approximately 1 store for every 12,000 people. We have obtained detailed geo-coded data (e.g., latitude and longitude) for every Dunkin Donuts in the state of Massachusetts. In order to ensure that the effect comes from police (rather than just locations that tend to attract fast food franchises), we have also obtained similar data for other types of fast-food restaurants (e.g., Burger King) that we do not believe attract as many police officers, so that we can ensure that the effect we detect is from the presence of police, rather than just the location of fast-food franchises.

Pre-trends that wouldn't go away

- Indoor air pollution: Does switching from kerosene to LPG cooking have health effects?

- Empirical idea:
  - Indonesia had the world's largest conversion campaign
  - District-by-district, phased out subsidized kerosene, and introduced subsidized LPG as well

- What we did
  - Worked with the Indonesian government planning ministry (via many meetings and a special data request) to find, and then share, the dates each district was switched
  - Created a pseudo-panel at the district level by coding up many years of household surveys
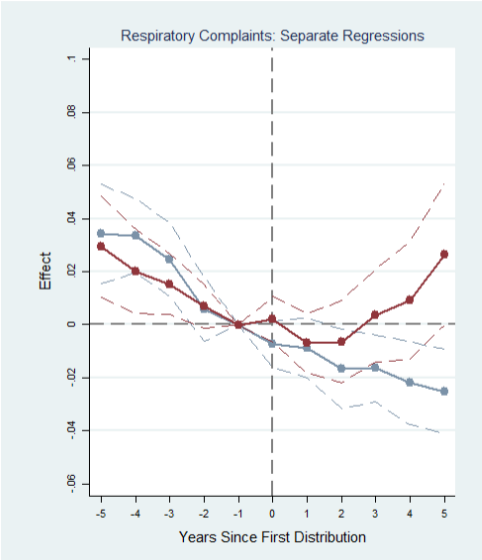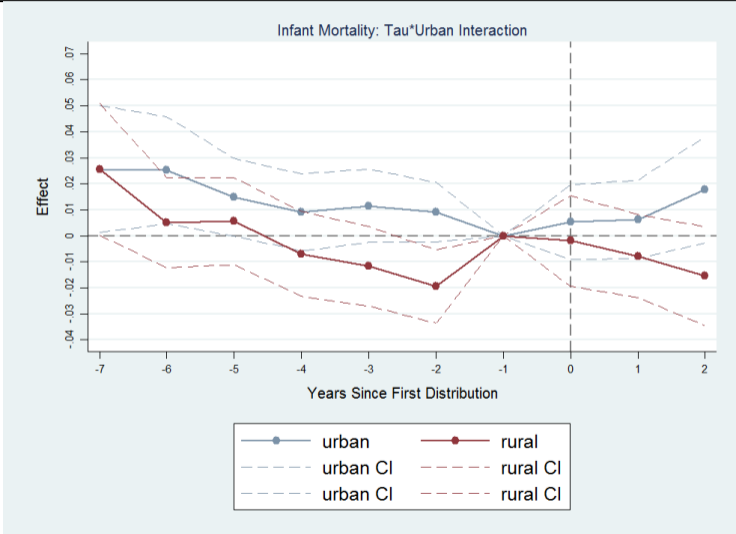 - Found a beautiful first stage

First stage
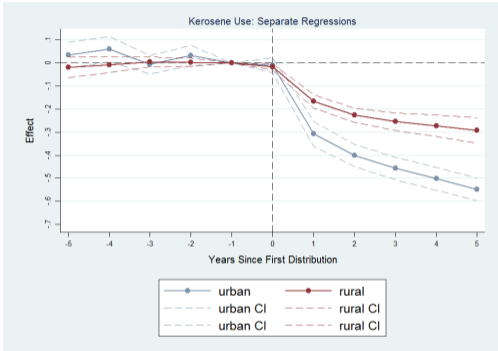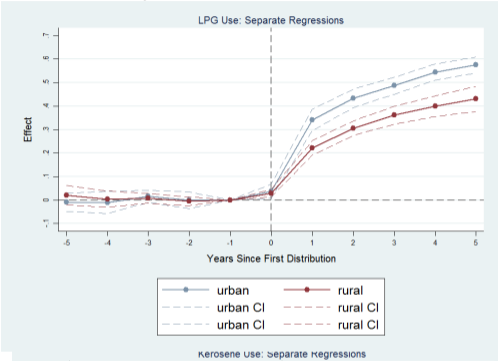
Pre-trends that wouldn't go away

- Indoor air pollution: Does switching from kerosene to LPG cooking have health effects?

- Empirical idea:
  - Indonesia had the world's largest conversion campaign
  - District-by-district, phased out subsidized kerosene, and introduced subsidized LPG as well

- What we did
  - Worked with the Indonesian government planning ministry (via many meetings and a special data request) to find, and then share, the dates each district was switched
  - Created a pseudo-panel at the district level by coding up many years of household surveys
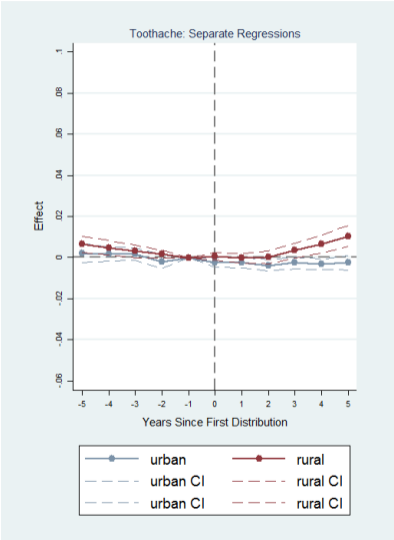 - Found a beautiful first stage

- But then...

Health outcomes

Respiratory Complaints: Separate Regressions

Health outcomes

Infant Mortality: Tau*Urban Interaction

Placebo? Toothaches!?

Toothache: Separate Regressions

Field projects that were misconceived

- Question:
  - Why are there so many failed water projects throughout the developing world?
  - Can we improve maintenance by giving communities a greater stake in maintenance? By building a revenue stream into the projects?

- What we did:
  - Extensive field visits to water projects in multiple Indonesian provinces
  - Designed to both diagnose the problem and see if our suggested interventions made sense

Sample water projects

A distributing water reservoir

A tap in one house in the sub-village

What we found

- "Sub-district XXX: From 12 villages in this sub-district, 8 have water projects, all of which are still working today."
- "In 2001, a maintenance committee was formed. Pak YY became the head of this committee as well as its technician, and he recruited a secretary and a treasurer for the 4 neighborhoods helped by PPK. Each month, the secretary went around to check individual house's water meter and the treasurer collects the payment. There are also a secretary and a treasurer for the other 8 neighborhoods. Treasurers keep all the money, and Pak YY determines the extent of damage and the amount of money to be disbursed when repairs are needed. Once every week, the maintenance committee members check for leaks, clogs, and monitor water output. Other than that the members only check the system when there are reported damages or disturbances."
- "Maintenance fee Usage fee is Rp300 (3 cents)/m3/month and maintenance fee is Rp1000 (10 cents)/month. People used to be asked to pay their fees directly to the treasurer, but there were less and less payment each month, so they changed the system in 2007. Nowadays, the heads of the neighborhood (Pak RT) collect the fees door to door and receive 10% of all collection. Currently the sanction of not paying maintenance fee is getting your water cut off after three warnings. It has happened only once, based on a decision made by the maintenance committee members.The maintenance committee has accumulated Rp17million thus far ($1,700), which will be used to increase water output."

Finding the right field setting

- Idea:
  - We use ordeal mechanisms to screen people for anti-poverty programs
    - E.g. poor have lower opportunity cost of time, so value time vs. money differently than the rich
  - Can we adapt similar mechanisms to labor markets?
  - Question: would requiring job applicants to complete some 'homework' as part of the application select unobservably better people?

- Version 1: Indonesian CCT
  - Conditional cash transfer program in Indonesian villages
  - Local facilitators help coach enrollees on how to achieve conditions
  - Program was recruiting 1000s of new facilitators
  - Idea: randomly introduce some additional 'homework' tasks as part of application process, i.e., health profile of village

Finding the right field setting

Finding the right field setting

## 2. Can adding hassle filter unfit applicants?

- Hassle filter can take the form of requiring applicants to draft a village health and education profile about their own village/kelurahan.
- As most applicants will have to travel to test location, MoSA can add variation on distance to test location for randomly selected applicants.

Pro	Cons
- Can allow MoSA to identify committed applicants who will perform well. - Homework may help prospective applicants get familiar with facilitators' task.	- Hassle may deter otherwise qualified applicants who have little training/experience. - Homework may overwhelm service providers who have to provide data - Service providers may refuse enquiries for homework and result may not reflect applicants' quality

Finding the right field setting

- Version 2: Indonesian Labor Insurance Program
  - National labor insurance (workers comp, disability, etc) program
  - Local staff recruit small/informal firms to pay insurance, paid on commission
  - Problem: train many people who don't actually work
  - Homework idea: require submission of list of 50 possible firms as part of application process

- Piloted in 5 locations in one province, planned expansion to 200 offices

Finding the right field setting
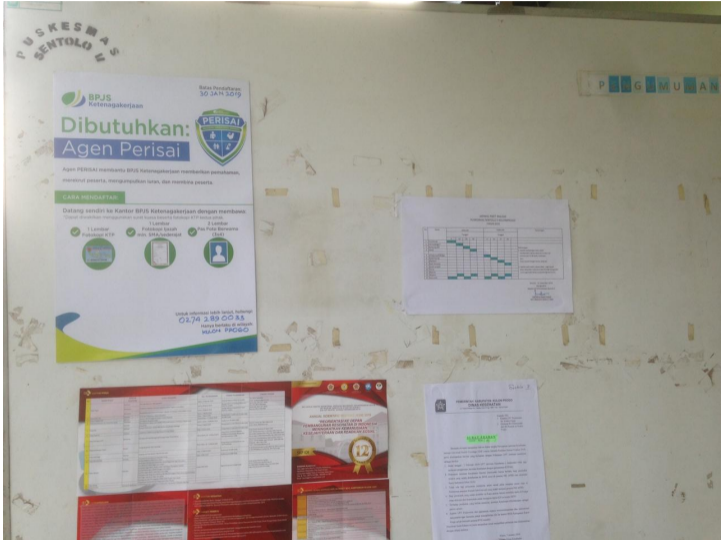
Finding the right field setting

Partners

Finding the right field setting

- Version 2: Indonesian Labor Insurance Program
  - National labor insurance (workers comp, disability, etc) program
  - Local staff recruit small/informal firms to pay insurance, paid on commission
  - Problem: train many people who don't actually work
  - Homework idea: require submission of list of 50 possible firms as part of application process

- Piloted in 5 locations in one province, planned expansion to 200 offices

- Neither project made it to full scale RCT

Sometimes you need to adapt

- Sometimes you need to adapt... for lots of reasons

- Initial design didn't work... can you find another design that does?

- Initial question or hypothesis turned out not to be that interesting or feasible.... but you learn something else instead

The Simple Economics of Extortion: Evidence from Trucking in Aceh

Benjamin A. Olken

Massachusetts Institute of Technology and National Bureau of Economic Research

Patrick Barron

World Bank

This paper tests whether the behavior of corrupt officials is consistent with standard industrial organization theory. We designed a study in which surveyors accompanied Indonesian truck drivers on 304 trips, during which they observed over 6,000 illegal payments to police, soldiers, and weigh station attendants. Using plausibly exogenous changes in the number of checkpoints, we show that market structure affects the level of illegal payments. We further show that corrupt officials use complex pricing schemes, including third-degree price discrimination and a menu of two-part tariffs. Our findings illustrate the importance of considering the market structure for bribes when designing anticorruption policy.

What actually happened

- Idea: look at how bribes paid by truckers respond to 'IO' of corrupt officials competing with each other.

- Take 1: Collect bribes throughout island of Sumatra. Use the fact that some city pairs have multiple routes (competition), some have single route (monopoly). Estimate structurally over whole island.
  - Problem: could not separate out competition effect from differential demand.

- Take 2: Look at bribes collected at truck weigh stations. Use the fact that some weigh stations had improvements to reduce bribes paid (e.g. automatic weigh-bridges, incentives, etc). Look at spillovers to weigh stations on alternative routes.
  - Problem: intervention only occurred at 3 weighs stations in the past. Wasn't that successful at reducing corruption. And no plans to repeat it in the future.

What actually happened (continued)

- Take 3: Heard about a new road opening from Jakarta to Bandung. Look at how this increase in alternatives changes bribes paid on the old road.
  - What I did: Dropped everything and rushed to assemble a survey. Did a pilot survey!
  - Problem: No bribes in the pilot survey! Why not? Turns out corruption was already centralized. Pay a bribe at the office each month, get a sticker that says you can travel bribe-free in the province. So no bribes paid on the road, and new road would have no impact.

- Take 4: Tsunami in Aceh. Major humanitarian disaster. Also, wiped out a major road. So now only one way to get to Banda Aceh from rest of Sumatra.
  - Idea: What happens to bribes on that road when they rebuild the other road to Banda Aceh?
  - What I did: flew to Aceh, learned about situation, talked to truck drivers, redid pilot survey, etc.Lots of bribes, not centralized.
  - Problem: Road was going to take 5 years to be rebuilt. Didn't want to wait that long!

What actually happened (continued)

- Take 5: Aceh peace agreement mid-2005. Government agreed to withdraw troops from Aceh in 4 phases, starting one month hence. Exogenous reduction in number of bribe stations.
  - Idea: slightly different idea, look at double-marginalization: how do existing checkpoints on same route adapt when there are fewer other checkpoints.
  - What I did: Dropped everything, rushed to Aceh, used existing piloted survey instruments and fielded a survey ASAP. Got in the field in time, when only half of troops had already been withdrawn.

Sometimes you need to adapt

- Examples of adaptation in an RCT context

- Reducing leakage in Indonesian ID card program through transparency:
  - Our initial idea: transparency experiment, where government would send village an SMS every time rice left the warehouse to eliminate leakage along the road.
  - What the government wanted to learn about: ID cards telling beneficiaries what they were entitled to.
  - What we did: ID cards telling beneficiaries what they were entitled to.

- Many other examples. Most of my projects come from listening and finding the intersection of what is academically interesting and what partners are interested in learning.

Some general guidelines

- Figure out the mission critical steps and do them rapidly
  - Don't waste a lot time on non-mission critical parts
  - Cut your losses if the mission critical part fails

- Verify your hypothesis
  - Field visits
  - Qualitative research
  - Quick data tabs
  - Check the first stage
  - Solve a simple version of the model first

How do you know when to kill a project?

- This is probably the hardest decision to make.

- Consult your advisors.

- Can you prove the project is hopeless? If so, stop.
  - Standard errors are huge.
  - First stage is nonexistent.
  - Theorem is wrong.
  - *Exact* same paper has been written.
  - etc

- But consult your advisors since what often looks like failures are not.

One last piece of advice

- Try to smooth....

    - Good news about a project is rarely as good as it seems
        - Don't trust that first result with three stars
        - Your theorem may not actually may not be right

    - But the converse is usually true, too.
        - Results that go away often come back.
        - You and your projects can adapt