Instrumental Variables

October 27, 2010

Direct and Indirect Colonial Rule in India

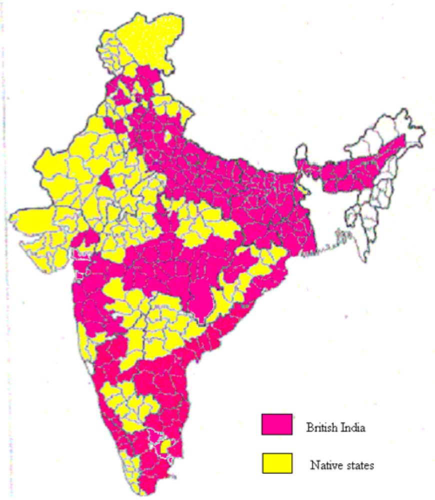


Figure 1 : British India and Native States

Colonial Rule as a "Treatment"

- 415 districts, some of which were formerly part of British India.
- Indicate treatment by a dummy variable: $D_i = 1$ if the district remained under direct colonial control until independence or $D_i = 0$ if it remained part of a "native" state.
- Focus on sample of districts (n=136) which beginning in 1848 were parts of native states.
- 36 districts (26%) were eventually *annexed* by the British Government.

Naive Estimates of the Effect of Colonial Rule

- DV is a measure of public good provision in the district (from the 1980/1990 censuses).
- Ideally, we want to know:

$$\mathbb{E}[\delta] = \mathbb{E}[Y(1) - Y(0)]$$

- We only observe the following:

$$\mathbb{E}[Y(1)|D = 1] - \mathbb{E}[Y(0)|D = 0]$$

- Works under the assumption that districts were annexed "as-if" random.

Naive Estimates

Dependent variables: Proportion of villages having public goods (mean of 1981 and 1991 data)			
Primary school	0.7720	-0.016	-0.007
		(0.032)	(0.039)
Middle school	0.2485	-0.046	-0.047
		(0.034)	(0.031)
High school	0.1260	-0.068*	-0.061*
		(0.040)	(0.033)
Primary health center	0.0415	-0.024*	-0.015*
		(0.014)	(0.008)
Primary health subcenter	0.0753	-0.002	-0.007
		(0.017)	(0.017)
Canals	0.0477	-0.010	-0.024*
		(0.014)	(0.014)
Roads	0.4344	0.043	-0.010
		(0.065)	(0.067)
Combined public goods	0.2535	-0.017	-0.026
		(0.025)	(0.021)

The Doctrine of Lapse

Lord Dalhousie, Governor-General of India from 1848-1856, enacted a new policy regarding annexation:

> *I hold that on all occasions where heirs natural shall fail, the territory should be made to lapse and adoption should not be permitted, excepting in those cases in which some strong political reason may render it expedient to depart from this general rule.*

The Doctrine of Lapse

- Now we have a new variable, $Z_i = 1$ if the district's ruler without an heir or $Z_i = 0$ with an heir. Note that $Z_i \neq D_i \forall i$, so the Doctrine of Lapse is not always followed.
- Iyer claims that while the assumption that $(Y(0), Y(1)) \perp\!\!\!\perp D$ is not plausible, the assumption $(Y(0), Y(1)) \perp\!\!\!\perp Z$ is plausible.
- Under her assumptions,

$$\begin{aligned} \hat{\delta}_Z &= \mathbb{E}[Y(1)|Z = 1] - \mathbb{E}[Y(0)|Z = 0] \\ &= \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] \end{aligned}$$

Do we care about this estimand?

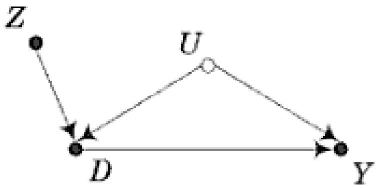
The Doctrine of Lapse as an Instrument

- Let's rewrite D_i as $D_i(Z_i)$ to reflect the fact that whether or not a district is annexed can depend on whether or not the ruler dies without an heir.
- Similarly, we can rewrite the potential outcomes as $Y_i(D_i, Z_i)$.
- The causal effect of Z on colonial rule for an individual district can be written as $D_i(1) - D_i(0)$. The causal effect of Z on socioeconomic outcomes can be written as $Y_i(1, D(1) - Y_i(0, D(0))$, which is the "intent to treat" effect.

Now we need assumptions:

1. SUTVA
2. ("as if") Random Assignment of Z_i.
3. Exclusion restriction: $Y_i(1, D_i) = Y_i(0, D_i)$.
4. Non-zero average causal effect of Z on D.
5. Montonicity: $D_i(1) \geq D_1(0)$ for all i.

The Exclusion Restriction



: A DAG with an unblocked back-door path and a valid IV.

The First Stage

	Post-1847 sample			
	no controls	geography	soils	main effects
	(1)	(2)	(3)	(4)
Ruler died without natural heir in 1848-1856 (Instrument)	$\boxed{0.682^{***}}$	0.673***	0.669***	0.953***
	$\boxed{(0.159)}$	(0.155)	(0.162)	(0.176)

Deriving the IV Estimator

- First, let's redefine the individual causal effect using the exclusion restriction:

$$
\begin{aligned}
Y_i(1, D(1) - Y_i(0, D(0)) &= Y_i(D(1) - Y_i(D(0)) \\
&= (Y_i(1) - Y_i(0)) \cdot (D_i(1) - D_i(0))
\end{aligned}
$$

- Now we need to focus on average effects. We've assumed that $\mathbb{E}[D_i(1) - D_i(0)] \neq 0$. So we can decompose the average causal effect as:

$$
\begin{aligned}
& \mathbb{E}[Y_i(D_i(1)) - Y_i(D_i(0))] \\
= \; & \mathbb{E}[(Y_i(1) - Y_i(0))|(D_i(1) - D_i(0) = 1)] \cdot P[D_i(1) - D_i(0) = 1] \\
& -\mathbb{E}[(Y_i(1) - Y_i(0))|(D_i(1) - D_i(0) = -1)] \\
& \cdot P[D_i(1) - D_i(0) = -1]
\end{aligned}
$$

Deriving the IV Estimator

- Ok, now we have 5 unknowns, of which we can calculate only 1 of them from the data. To get rid of some unknowns, we invoke the monotinicity assumption, which leaves us with:

$$\begin{aligned} & \mathbb{E}[Y_i(D_i(1)) - Y_i(D_i(0))] \\ = \; & \mathbb{E}[(Y_i(1) - Y_i(0))|(D_i(1) - D_i(0) = 1)] \cdot P[D_i(1) - D_i(0) = 1] \end{aligned}$$

- Now we have three unknowns, 2 which of can be estimated with the data. Under the monotonicity assumption $P[D_i(1) - D_i(0) = 1] = \mathbb{E}[D_i(1) - D_i(0)]$, which is simply the the effect of the instrument on treatment assignment status.
- Solving for $\mathbb{E}[(Y_i(1) - Y_i(0))|(D_i(1) - D_i(0) = 1)]$, we get the following expression for our IV estimator:

$$\frac{\mathbb{E}[Y_i(D_i(1)) - Y_i(D_i(0))]}{\mathbb{E}[D_i(1) - D_i(0)]}$$

What is the IV estimand?

		$D_i(0)$	
		0	1
$D_i(1)$	0	Never-taker	Defier
	1	Complier	Always-taker

Table: Causal Types

- $\mathbb{E}[(Y_i(1) - Y_i(0))|(D_i(1) - D_i(0) = 1)]$ is the causal effect of the treatment on compliers, a kind of LATE.
- In our example, it's the causal effect of British direct rule on those districts induced by the Doctrine of Lapse to lose their independent status.

IV Results

Dependent variables: Proportion of villages having public goods				
(mean of 1981 and 1991 data)				
Primary school	0.7720	-0.016	-0.007	-0.011
		(0.032)	(0.039)	(0.041)
Middle school	0.2485	-0.046	-0.047	-0.091**
		(0.034)	(0.031)	(0.037)
High school	0.1260	-0.068*	-0.061*	-0.065
		(0.040)	(0.033)	(0.042)
Primary health center	0.0415	-0.024*	-0.015*	-0.031**
		(0.014)	(0.008)	(0.013)
Primary health subcenter	0.0753	-0.002	-0.007	-0.053**
		(0.017)	(0.017)	(0.021)
Canals	0.0477	-0.010	-0.024*	-0.043
		(0.014)	(0.014)	(0.028)
Roads	0.4344	0.043	-0.010	-0.198***
		(0.065)	(0.067)	(0.066)
Combined public goods	0.2535	-0.017	-0.026	-0.075***
		(0.025)	(0.021)	(0.023)