Tools for Public Policy Analysis

Biswo Poudel

Kathmandu University

International Labor Organization

Two Major Strands of Analysis

- Theoretical : It involves building models, using mathematical expressions, developing theorems, and sometimes using simulations to test the predictions of the theorems. These simulations use real world parameters.

- Empirical: Many methods. But most powerful one is regression analysis.

Theoretical Analysis

- Building a model and deriving a result: A firm has a benefit function given by $f + (\alpha + \theta_t)x_t - \frac{b}{2}x_t^2$, where x is a unit of pollution. If the regulator imposes a tax of p per unit, then the benefit function is $f + (\alpha + \theta_t)x_t - p_t x_t - \frac{b}{2}x_t^2$. The firm will then set its production at $\frac{a-p_t}{b} + \frac{\theta_t}{b}$. (Karp et al.(2006), Regulation with Anticipated Learning about Environmental Damage). In the paper, they find that anticipated learning decreases the optimal level of abatement.

In general

- Theoretical analysis rely on optimization method. For example, for unconstrained models, the optimal point is derived by setting first derivative equals to zero.
- For constrained model, the method of optimization is set by Kuhn-Tucker method.
- For dynamic models, often the solutions are provided by solving optimal control methods.
- They all rely on benefit function being concave.

Regression Analysis

- What is Regression?
- Does it have causal interpretation?
- What is its interpretation (causality vs correlation)?

A Typical Regression

- A typical regression looks like the following equation:

$$y = \beta_0 + \beta_1 x_1 + \ldots \ \ldots + \beta_k x_k + u$$

- One then estimates β_0, β_1, … and provided they are significant, would like to make a statement as follows: if X_k increases by some units, by how much will y increase?
- Suppose you are regressing FoodExpenditure on Total Expenditure of different individuals and your regression result is $F_i = 94.2087 + 0.4368 T_i$
- The interpretation is if the total expenditure goes up by 1 rupees, food expenditure goes up by 43 paisa.

Interpretation cont..

- Let P be the personal expenditure and D be the personal expenditure on durable goods. Then, suppose when we run a regression of expenditure on durable goods to the total personal expenditure, and we get the following results: $\ln D_i = -9.69 + 1.90 \ln P_i$
- Interpretation: if Personal Expenditure goes up by 1%, total expenditure on durables goes up by 1.9%.

Interpretation cont..

- Let expenditure on services be S and let quarter be t. Then, suppose when expenditure on services was regressed on quarter, we get the following: $\ln S_t = 7.78 + 0.0074t$
- *Interpretation*: Expenditure on services increased by 0.74% every quarter.
- This also corresponds to the annual growth rate of about 3%.
- Suppose F is food expenditure and T is total expenditure. Then the following model $F_i = -1283.9 + 257.2 \ln T_i$ implies that an increase of total expenditure by 1% corresponds to an increase of food expense by Rs 2.57.

Endogeneity Issue

- The primary reason why OLS is an unreliable estimate of these marginal impacts is the presence of endogeneity.
- Sources of endogeneity: Omitted Variable Bias, Simultaneity, and measurement error.
- How does OVB affect the estimates? Here is a formula:

$$\hat{\rho} = \rho + \gamma' \delta_{As}$$

- Interpretation: γ is the impact of omitted variable on Y. If ability is missing from return to schooling, then the estimate that doesn't account for omitted variables overestimates true value.

Example : return to education

Controls	(1) None	(2) Age Dummies	(3) Col.(2) and Additional Controls	(4) Col.(3) and AFQT Score	(5) Col.(4), with Occupation Dummies
	.132 (0.007)	.131 (0.007)	.114 (0.007)	0.087 (0.009)	0.066 (0.010)

Methods of Natural Experiments to address endogeneity

- Instrument Variables: The equation of interest is $Y_i = \alpha + \rho s_i + A_i'\gamma + v_i$. A is unobserved.
- Suppose the instrument is Z. Then, $\rho = \frac{\text{cov}(Y_i, Z_i)}{\text{cov}(S_i, Z_i)} = \frac{\text{cov}(Y_i, Z_i)/V(Z_i)}{\text{cov}(S_i, Z_i)/V(Z_i)}$

- The quarter of birth as instrument.

	OLS		2SLS					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Years of Education	0.071	0.067	0.102	0.13	0.104	0.108	0.87	0.057
	(0.0004)	(0.0004)	(0.024)	(0.020)	(0.026)	(0.020)	(0.016)	(0.029)
Exogenous Covariates								
Age(in Quarters)								**
Age (in quarters) squared								**
9 year of birth dummies		**			**	**	**	**
50 state-of-birth dummies		**			**	**	**	**
Instruments								

Interpretation of IV

- One is LATE interpretation. It basically implies that it is an average impact on the subpopulation that is treated because of instrument.
- Suppose you are estimating return to military service on Vietnam veterans. If your IV for service is a draft lottery, then IV captures the average effect of military service on salary of those who served because they were drafted. It would exclude the volunteers.
- Bias of IV: $E[\hat{\beta}_{2SLS} - \beta] \approx \frac{\sigma_{\eta\xi}}{\sigma_{\xi}^2} \frac{1}{1+F}$

Differences-in-Differences Method

- To see the impact of an intervention, this method compares the outcomes in a agent that is exposed to the treatment to the one that is not. The regression model almost always looks like the following:

$$Y_{ist} = \alpha + \gamma NJ_s + \lambda d_t + \delta(NJ_s . d_t) + \varepsilon_{ist}$$

Does Labor Regulation Lead to Decreased Output Per Capita? Evidence from Indian states

	(1)	(2)	(3)	(4)
Labor Regulation (lagged)	-.186 (0.064)	-.185 (0.051)	-.104 (0.039)	0.0002 (0.02)
Ln Development expenditure/ capita		0.240 (.128)	.184 (.119)	.241 (.106)
Log installed electricity / capita		0.089	0.082	0.023
Log state pop		.72	.31	-1.41
Congress Majority			-0.0009	.02
Hard Left Majority			-0.05	-0.007
Janata Majority			0.008	-0.02
Regional Majority			0.006	0.026
State Specific trend	No	No	No	Yes

Regression Discontinuity Approach

- We omit the lengthy discussion of it.
- Basically, it exploits the fact that when the cut off rules for treatment are known, then comparing outcomes between two subgroups (one which narrowly qualified to participate in the treatment, and another that narrowly missed) should provide us with the causal interpretation.
- Applications: does talking in cellphone increase traffic accident?

Randomized Control Trials

- These are new methods, popularized by PAL and others.
- They design the data generating process, trying to make it as random as possible. Then they run the appropriate regression analysis. Often OLS itself may suffice.
- Kyle et al.(2017) paper of adoption of drought resistant seeds in Odessa is an example.

Question?

- Question?