

Let's Call their Bluff:
The Politics of Econometric
Methodology

by

Jean-Paul Azam

Toulouse School of Economics, UT- Capitole

NEPS Lecture, The Hague, June 25, 2019

Misleading Regressions

- Non-causal regressions can put you in trouble
- The literature on “aid-ineffectiveness” is littered with examples of wrong conclusions that deeply influenced policy-makers.
- The classic one is Boone (1996), who showed that foreign aid seems to have no impact on economic growth, without controlling for endogeneity.
- Arndt et al. (2017) overturned this result by showing that foreign aid has a beneficial impact on economic growth and many other outcomes, after controlling for endogeneity in a simple two-stage approach.
- But donors have been misled for twenty years.

Citizen Oversight

- The key point in evaluating policy effectiveness from historical data is to control for the policy maker's endogenous responses.
- This requires that the econometrician seeks to uncover the policy maker's true preferences, which might be quite at variance with the proclaimed ones.
- Azam & Berlinschi (2010) have thus shown that foreign aid is in fact allocated by OECD countries with a view to reduce immigration from low and lower-middle countries.
- When these findings were first presented at Nuffield College in June 2009, Adrian Wood, who had worked a long time at DFID, came at the coffee break to tell me:
- “Jean-Paul, it is true that the Foreign Office exerted a lot of pressure on our aid allocation with a view to reduce immigration, but we don't do it anymore”.
- He obviously crossed his heart with his hand in saying that, but I am not sure he expected me to believe him.

Do we Buy Immigrants?

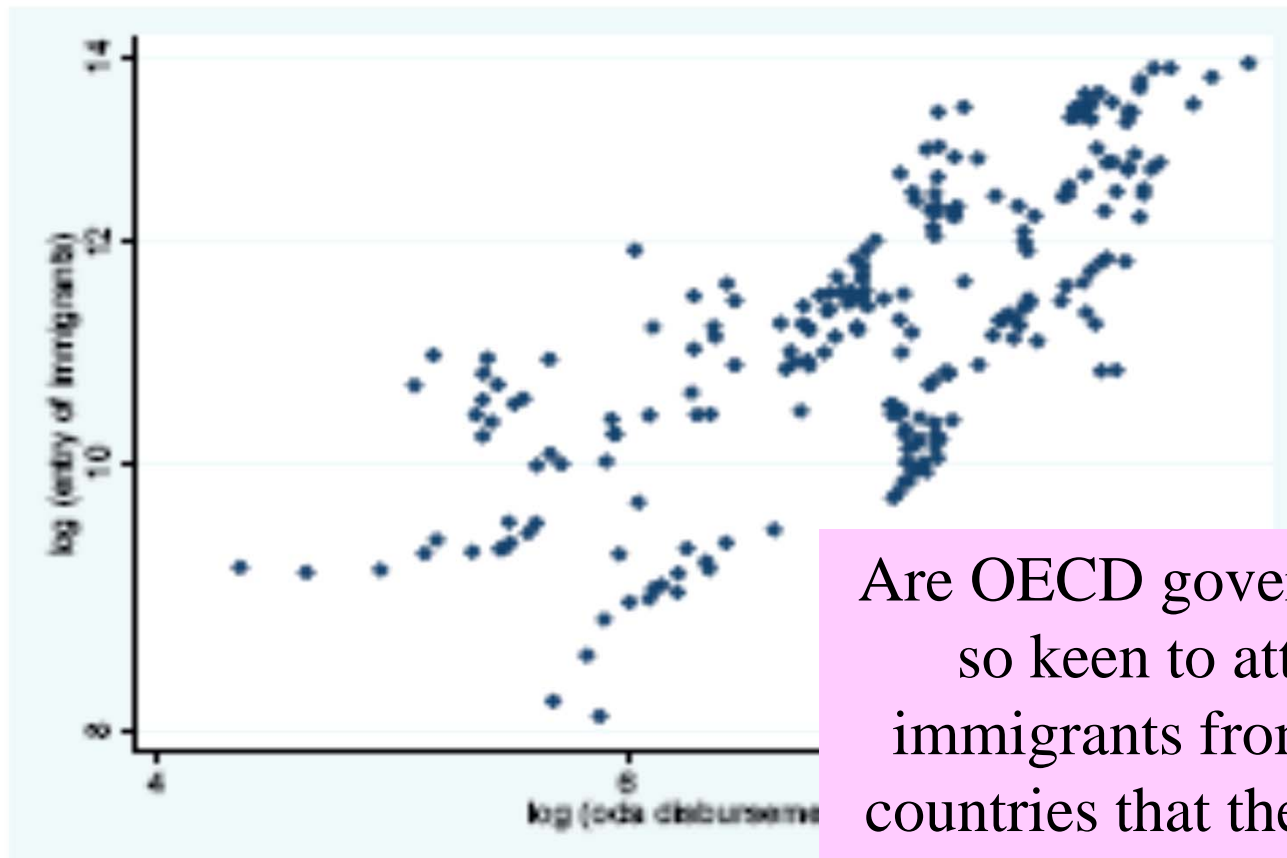


Figure 2: ODA Disbursements and E

Are OECD governments so keen to attract immigrants from poor countries that they bribe their governments?

TABLE 1. Regression Results on Flows of Legal Migrants from Low-Income and Lower-Middle-Income Countries

Variable	(1)	(2)
Unemployment rate	-0.30 ^{***} (0.09)	-0.18 ^{***} (0.08)
Social expenditures (percent of GDP)	0.32 ^{***} (0.09)	0.30 ^{***} (0.09)
Log of per capita GDP	0.54 (1.42)	9.53 ^{***} (2.82)
Log of stock of foreign population	0.19 (0.57)	0.57 (0.50)
Log of official development assistance (ODA) disbursements	0.46 (0.32)	-3.68 ^{***} (1.15)
Log of multilateral disbursements	—	—
Endogeneity bias, ODA	—	4.47 ^{***} (1.26)
Endogeneity bias, multilateral disbursements	—	—
Number of observations	118	117
F-test	9.50	9.84

Source: Authors' calculations.

The usual suspects also play their part in the plot

Foreign aid is effective for controlling immigration

And it is actually used for that

Comments

- Adding controls like the (lagged) unemployment rate and the (lagged) share of social expenditures in GDP reduces the upward bias shown by the chart to insignificance.
- However, it leaves unexplained why rich countries are giving so much aid money to migrants' source countries.
- Controlling for endogeneity solves this problem by showing that OECD countries give aid money *because* it helps to reduce immigration from low and lower middle income countries, among other things.
- Hence, foreign aid is effective at something after all, but donors are not too proud to say what it is good for.

ANNEX TABLE A.1. First-Stage Reduced-Form Equations

	Log of ODA disbursements for eq. (2)
Log of public expenditure on order and security	0.21*** (0.07)
Percentage of right-wing members in parliament	-0.003 (0.002)
Log of government revenues	
Unemployment rate	0.03 (0.02)
Social expenditures (percent of GDP)	-0.009 (0.013)
Log of per capita GDP	1.87*** (0.36)
Log of stock of foreign population	0.04 (0.08)
N	159
F-test	48.59

Source: Authors' calculations.

Right-wing governments, as revealed by expenditures on law and order, disburse more foreign aid to control immigration from poor countries than the others.

The (French) socialists prefer keeping the money for their own public expenditures and raise the minimum wage and payroll and profit taxes, and hence unemployment, maybe to deter immigration.

The Case for Political Cliometrics

- Econometric investigation of historical data is basically a game between civic-minded econometricians and policy makers whose deep motivations are potentially hidden.
- It is a basic tenet of economic theory that preferences can never be observed directly but must be inferred instead by applying revealed preference theory.
- Econometricians have devised various two-stage methods that help us to discover some determinants of policy-makers' preferences, and thus can play a key part in informing citizens' oversight.

A Basic Framework

The econometrician wants to test whether a policy tool p is effective to reduce the quantity q of an outcome that the policy maker deems "bad".

These two variables are linked by a linear relation:

$$q = \alpha + \beta x - \gamma p + \delta e + \varepsilon, \quad (1)$$

where x , e and ε are exogenous variables.

The Greek parameters, including $\{\alpha, \beta, \gamma, \delta\}$, are assumed positive unless specified otherwise, while ε is a random disturbance such that $E(\varepsilon) = 0$.

A Basic Framework

The econometrician wants to test whether a policy tool p is effective to reduce the quantity q of an outcome that the policy maker deems "bad".

These two variables are linked by a linear relation:

$$q = \alpha + \beta x - \gamma p + \delta e + \varepsilon, \quad (1)$$

where x , e and ε are exogenous variables.

The Greek parameters, including $\{\alpha, \beta, \gamma, \delta\}$, are assumed positive unless specified otherwise, while ε is a random disturbance such that $E(\varepsilon) = 0$.

A Basic Framework

The econometrician wants to test whether a policy tool p is effective to reduce the quantity q of an outcome that the policy maker deems "bad".

These two variables are linked by a linear relation:

$$q = \alpha + \beta x - \gamma p + \delta e + \varepsilon, \quad (1)$$

where x , e and ε are exogenous variables.

The Greek parameters, including $\{\alpha, \beta, \gamma, \delta\}$, are assumed positive unless specified otherwise, while ε is a random disturbance such that $E(\varepsilon) = 0$.

Key Assumptions

Assumption 1: Asymmetric information:

- (i) The policy maker observes x , p , and e *ex ante* and q *ex post*.
- (ii) The econometrician observes q , x and p *ex post*.

Assumption 2: Efficient Information Processing:

The policy maker uses her information efficiently so that $E(e \varepsilon) = 0$.

Assumption 3: Quadratic Loss Function:

The policy maker seeks to minimize the following loss function:

$$\min_p L = \left(\frac{\pi}{2}\right) p^2 + \frac{1}{2} (\theta + E(q))^2, \text{ s.t. } E(q) = \alpha + \beta x - \gamma p + \delta e. \quad (2)$$

The preference parameter θ increases the policy maker's aversion to q and it is her private information, i.e., unobserved by the econometrician. 12

Key Assumptions

Assumption 1: Asymmetric information:

- (i) The policy maker observes x , p , and e *ex ante* and q *ex post*.
- (ii) The econometrician observes q , x and p *ex post*.

Assumption 2: Efficient Information Processing:

The policy maker uses her information efficiently so that $E(e \varepsilon) = 0$.

Assumption 3: Quadratic Loss Function:

The policy maker seeks to minimize the following loss function:

$$\min_p L = \left(\frac{\pi}{2}\right) p^2 + \frac{1}{2} (\theta + E(q))^2, \text{ s.t. } E(q) = \alpha + \beta x - \gamma p + \delta e. \quad (2)$$

The preference parameter θ increases the policy maker's aversion to q and it is her private information, i.e., unobserved by the econometrician. 13

Key Assumptions

Assumption 1: Asymmetric information:

- (i) The policy maker observes x , p , and e *ex ante* and q *ex post*.
- (ii) The econometrician observes q , x and p *ex post*.

Assumption 2: Efficient Information Processing:

The policy maker uses her information efficiently so that $E(e \varepsilon) = 0$.

Assumption 3: Quadratic Loss Function:

The policy maker seeks to minimize the following loss function:

$$\min_p L = \left(\frac{\pi}{2}\right) p^2 + \frac{1}{2} (\theta + E(q))^2, \text{ s.t. } E(q) = \alpha + \beta x - \gamma p + \delta e. \quad (2)$$

The preference parameter θ increases the policy maker's aversion to q and it is her private information, i.e., unobserved by the econometrician. 14

Key Assumptions

Assumption 1: Asymmetric information:

- (i) The policy maker observes x , p , and e *ex ante* and q *ex post*.
- (ii) The econometrician observes q , x and p *ex post*.

Assumption 2: Efficient Information Processing:

The policy maker uses her information efficiently so that $E(e \varepsilon) = 0$.

Assumption 3: Quadratic Loss Function:

The policy maker seeks to minimize the following loss function:

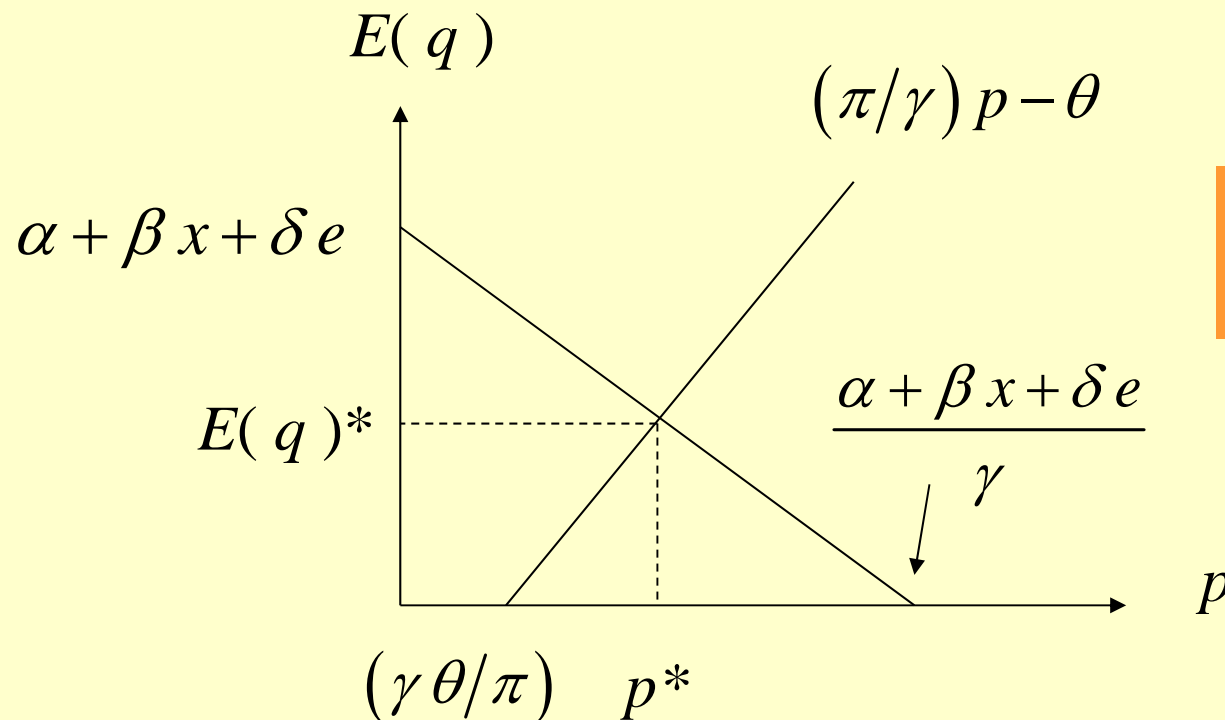
$$\min_p L = \left(\frac{\pi}{2}\right) p^2 + \frac{1}{2} (\theta + E(q))^2, \text{ s.t. } E(q) = \alpha + \beta x - \gamma p + \delta e. \quad (2)$$

The preference parameter θ increases the policy maker's aversion to q and it is her private information, i.e., unobserved by the econometrician. 15

The Policy Decision

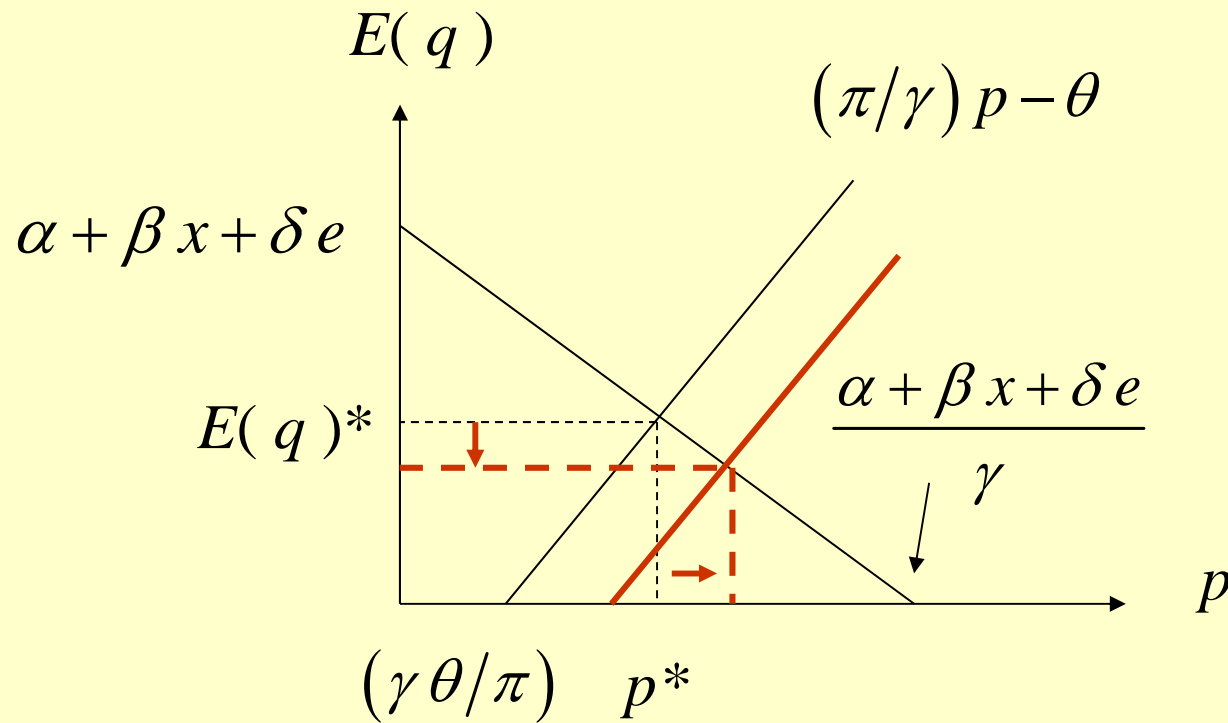
The first-order condition is:

$$\pi p^* = \gamma(\theta + E(q)^*) \Leftrightarrow E(q)^* = (\pi/\gamma)p^* - \theta$$



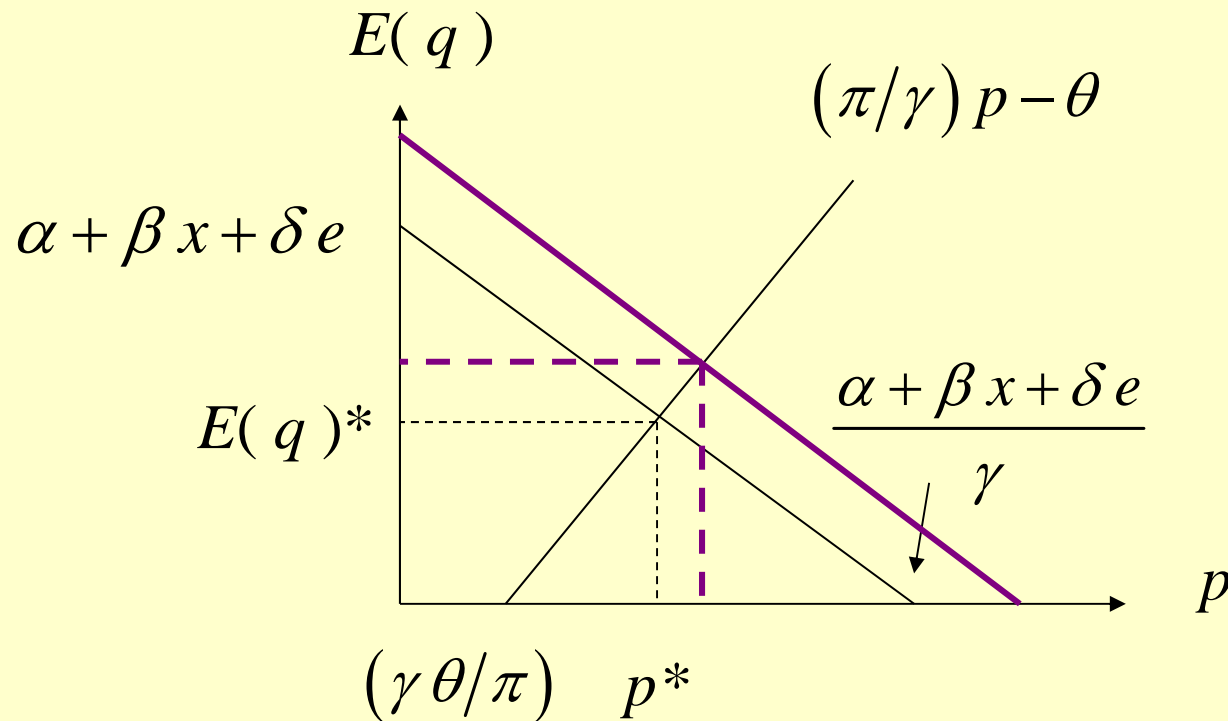
The policy choice yields $p^*(e, \theta, \dots)$

The Identification Problem (i)



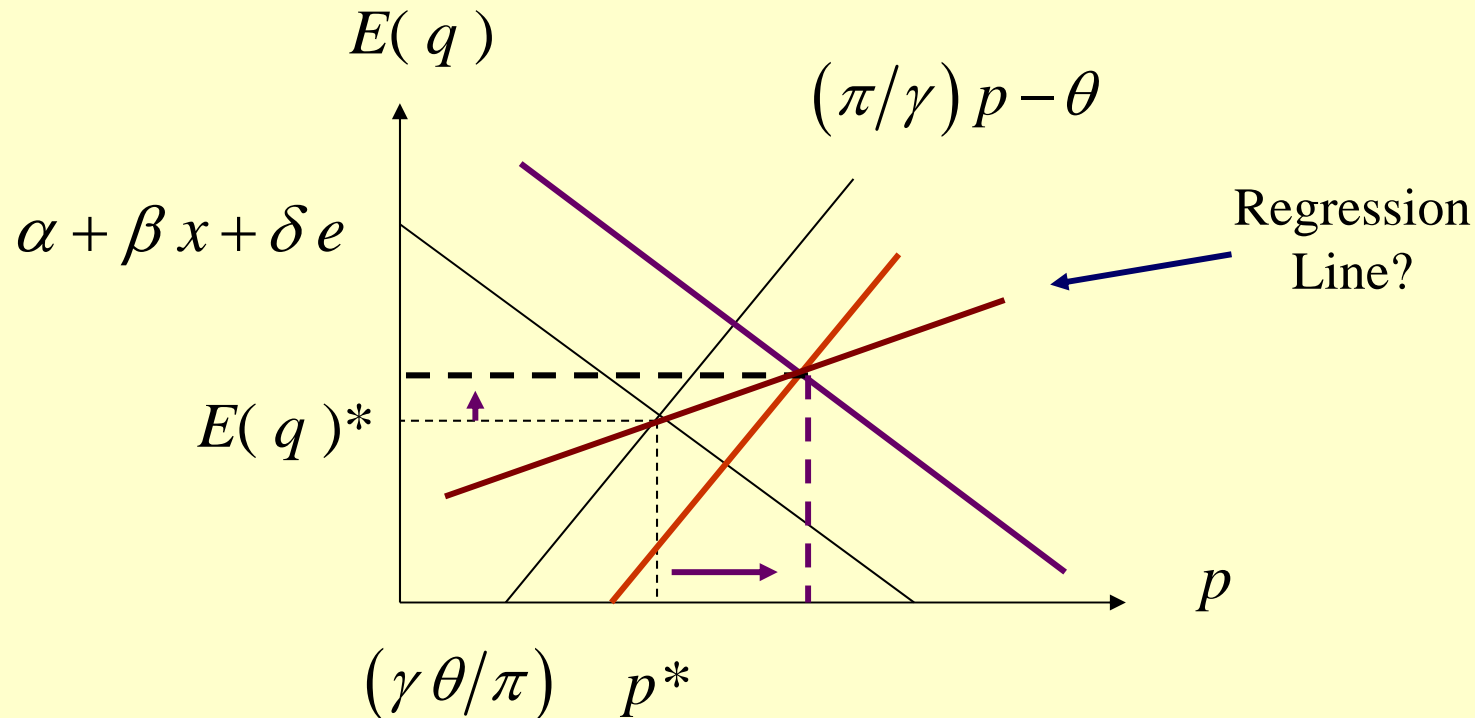
Changes in θ across observations trace out the causal relation and thus play on the side of the econometrician

The Identification Problem (ii)



Changes in e across observations trace out the first-order condition and thus play against the econometrician by contaminating the estimation of the causal relation

The Identification Problem (iii)



Then, if there is some positive correlation between θ and e over your sample, you might get a positively sloped regression line. This is a linear combination of the causal equation and the FOC.

Preference Proxies

Now, assume that the econometrician is tempted to guess that the preference proxy z such that:

$$\theta = \rho + \mu z + \zeta,$$

and $E(\zeta) = E(\zeta e) = E(\zeta \varepsilon) = E(\zeta z) = 0$, might have some empirical relevance.

This is a testable prediction using a two-stage approach to overcome the problem that θ is not directly observable.

Preference Proxies

Now, assume that the econometrician is tempted to guess that the preference proxy z such that:

$$\theta = \rho + \mu z + \zeta,$$

and $E(\zeta) = E(\zeta e) = E(\zeta \varepsilon) = E(\zeta z) = 0$, might have some empirical relevance.

This is a testable prediction using a two-stage approach to overcome the problem that θ is not directly observable.

Intuition

- The idea is to capture some of the beneficial identifying property of the changes in the policy maker's preferences over the sample, in order to let most of the unobservable disturbance reveal the changes in e rather than in θ .
- Then, that remaining unobservable disturbance can in fact be estimated via a first-stage equation.
- This is done by estimating a reduced-form equation capturing the chosen policy tool, which thus reveals the hidden information used by the policy maker, among other things.
- This works as follows:

The First-Stage Regression

Substituting the preference proxy for θ , the FOC becomes:

$$E(q) = (\pi/\gamma) p - \rho - \mu z - \zeta.$$

We can substitute for $E(q) = \alpha + \beta x - \gamma p + \delta e$ and rearrange the terms to get the reduced-form policy equation as:

$$p^* = \underbrace{\frac{\gamma(\alpha + \rho)}{\pi + \gamma^2}}_{\hat{a}} + \underbrace{\frac{\gamma\beta}{\pi + \gamma^2}}_{\hat{b}} x + \underbrace{\frac{\gamma\mu}{\pi + \gamma^2}}_{\hat{m}} z + \underbrace{\frac{\gamma}{\pi + \gamma^2}(\delta e + \zeta)}_{\hat{r}},$$

where \hat{r} are the residuals of the regression of p on x and z .

Notice that: (i) \hat{r} is correlated with e , (ii) but it contains some noise produced by ζ , whose variance is lower the better z does proxy θ .₂₃

The First-Stage Regression

Substituting the preference proxy for θ , the FOC becomes:

$$E(q) = (\pi/\gamma) p - \rho - \mu z - \zeta.$$

We can substitute for $E(q) = \alpha + \beta x - \gamma p + \delta e$ and rearrange the terms to get the reduced-form policy equation as:

$$p^* = \underbrace{\frac{\gamma(\alpha + \rho)}{\pi + \gamma^2}}_{\hat{a}} + \underbrace{\frac{\gamma\beta}{\pi + \gamma^2}}_{\hat{b}} x + \underbrace{\frac{\gamma\mu}{\pi + \gamma^2}}_{\hat{m}} z + \underbrace{\frac{\gamma}{\pi + \gamma^2}(\delta e + \zeta)}_{\hat{r}},$$

where \hat{r} are the residuals of the regression of p on x and z .

Notice that: (i) \hat{r} is correlated with e , (ii) but it contains some noise produced by ζ , whose variance is lower the better z does proxy θ .₂₄

The First-Stage Regression

Substituting the preference proxy for θ , the FOC becomes:

$$E(q) = (\pi/\gamma) p - \rho - \mu z - \zeta.$$

We can substitute for $E(q) = \alpha + \beta x - \gamma p + \delta e$ and rearrange the terms to get the reduced-form policy equation as:

$$p^* = \underbrace{\frac{\gamma(\alpha + \rho)}{\pi + \gamma^2}}_{\hat{a}} + \underbrace{\frac{\gamma\beta}{\pi + \gamma^2}}_{\hat{b}} x + \underbrace{\frac{\gamma\mu}{\pi + \gamma^2}}_{\hat{m}} z + \underbrace{\frac{\gamma}{\pi + \gamma^2}(\delta e + \zeta)}_{\hat{r}},$$

where \hat{r} are the residuals of the regression of p on x and z .

Notice that: (i) \hat{r} is correlated with e , (ii) but it contains some noise produced by ζ , whose variance is lower the better z does proxy θ .₂₅

Second-Stage Equation

Then the second-stage regression equation:

$$q = \hat{\alpha} + \hat{\beta} x - \hat{\gamma} p + \hat{\phi} \hat{r} + \hat{\varepsilon}$$

is exactly equivalent to 2SLS in linear models.

This is the control function approach, which can be used in other contexts where 2SLS cannot.

As \hat{r} is a noisy proxy for e , there is a measurement issue that biases $\hat{\phi}$ toward zero, i.e., against the endogeneity diagnosis.

Second-Stage Equation

Then the second-stage regression equation:

$$q = \hat{\alpha} + \hat{\beta} x - \hat{\gamma} p + \hat{\phi} \hat{r} + \hat{\varepsilon}$$

is exactly equivalent to 2SLS in linear models.

This is the control function approach, which can be used in other contexts where 2SLS cannot.

As \hat{r} is a noisy proxy for e , there is a measurement issue that biases $\hat{\phi}$ toward zero, i.e., against the endogeneity diagnosis.

Second-Stage Equation

Then the second-stage regression equation:

$$q = \hat{\alpha} + \hat{\beta} x - \hat{\gamma} p + \hat{\phi} \hat{r} + \hat{\varepsilon}$$

is exactly equivalent to 2SLS in linear models.

This is the control function approach, which can be used in other contexts where 2SLS cannot.

As \hat{r} is a noisy proxy for e , there is a measurement issue that biases $\hat{\phi}$ toward zero, i.e., against the endogeneity diagnosis.

Political Interpretation

- The residuals from the first-stage equation are a proxy for some information that (*i*) is observed by the policy maker when she makes her decision, (*ii*) is entering the causal relation, and (*iii*) is not directly observed by the econometrician.
- Hence, although it is biased towards zero, the control-function approach is providing a crucial information on the information used by the policy maker to choose her policy intervention, if it is statistically significant.
- Jean-Paul Azam & Kartika Bhatia (2017: “Provoking insurgency in a federal state: theory and application to India”, *Public Choice*, 170, 183-210, have used this argument to deny the claim that the rebels are the aggressors.

Table 4 Two-stage estimation of *rebellion* on *state_forces*

Variables	(1) State_forces	(2) Rebellion
State_forces		0.075** (0.033)
Tribal	-0.492*** (0.142)	
Iron	3.239 (3.409)	1.285*** (0.404)
Coal	9.453** (4.470)	0.877** (0.361)
Forest cover rate	-0.137* (0.081)	
Tribal_iron	0.314** (0.129)	
Tribal_coal	-0.105 (0.192)	
Tribal_forest	0.020*** (0.003)	
Bus service	0.015 (0.039)	-0.028*** (0.006)
Rural	0.084 (0.059)	0.042*** (0.009)
Hunger	-0.306* (0.176)	0.038** (0.018)
Prop_no_educ	-0.700 (0.891)	0.154 (0.112)
Logpop	1.968 (1.573)	0.159 (0.192)
Central_forces		
Residual1		0.025* (0.015)
Residual2		
Lalpha		1.087*** (0.135)
Observations	191	191
Adjusted R^2	0.3848	
F statistic (prob > F)	10.90 (0.000)	

Standard errors in parentheses

Hence, police killing of civilians is a major cause of the Naxalite uprising

This disproves the federal government's claim that the Adivasis are animated by the Maoist ideology with a view to topple the Indian democracy

The rebels know how much violence has been inflicted on the tribal people when they attack the policemen

The dependent variables are respectively the number of civilians killed by the police (1) and the number of people killed by the rebels (2)

Table 4 Two-stage estimation of *rebellion* on *state_forces*

Variables	(1) State_forces	(2) Rebellion
State_forces		0.075** (0.033)
Tribal	-0.492*** (0.142)	
Iron	3.239 (3.409)	1.285*** (0.404)
Coal	9.453** (4.470)	0.877** (0.361)
Forest cover rate	-0.137* (0.081)	
Tribal_iron	0.314** (0.129)	
Tribal_coal	-0.105 (0.192)	
Tribal_forest	0.020*** (0.003)	
Bus service	0.015 (0.039)	-0.028*** (0.006)
Rural	0.084 (0.059)	0.042*** (0.009)
Hunger	-0.306* (0.176)	0.038** (0.018)
Prop_no_educ	-0.700 (0.891)	0.154 (0.112)
Logpop	1.968 (1.573)	0.159 (0.192)
Central_forces		
Residual1		0.025* (0.015)
Residual2		
Lalpha		1.087*** (0.135)
Observations	191	191
Adjusted R^2	0.3848	
F statistic (prob > F)	10.90 (0.000)	

Standard errors in parentheses

Most of the killing of civilians performed by the state police occurs in areas where a large share of the people are tribal ones and there are significant deposits of iron ore and a significant forest cover. The latter mainly exists where bauxite deposits retain the mountains' moisture in these semi-arid areas

Hence, it is the interaction between the presence of mining potential riches and the largely tribal population of the area that entails significant violence against civilians by the police

Conclusion

- Carefully chosen preference proxies give the civic-minded econometrician a valuable tool to uncover the policy-makers deep motivations that determine their decisions.
- Using two-stage econometric analysis of real-world data can thus play a key part in informing citizens' oversight in democracies and authoritarian regimes.
- The objective pursued is to estimate simultaneously whether the policy tool has a significant impact on the outcome of interest and is actually used for that by the policy maker.