

Applied Microeconometrics: assignment Dynamic Panel Data models

On Canvas you find the data sets `datadynpan2022.dta` (Stata version) and `datadynpan2022.csv` (same, but comma-separated format), which was used by Jan van Ours.¹ This (historical) data describes the so-called *opiumregie* in the Dutch East Indies (currently Indonesia) in the period 1923-1938. At this time, Indonesia was still a Dutch colony, in which the Dutch were running a state monopoly on the import, production and sale of opiates. The data provide the demand for opium in different districts. In this problem set we use this data set to estimate the price elasticity of hard drugs.

Except for the price, we allow the demand to depend on income, the amount of illegal opium intercepted by the authorities, and a time trend. Furthermore, we allow for structural differences between districts. Consider the model specification (referring to district i in year t)

$$\begin{aligned}\log(\text{consumption}_{i,t}) = & \beta_0 + \beta_1 \log(\text{prices}_{i,t}) + \beta_2 \log(\text{income}_{i,t}) \\ & + \beta_3 \log(\text{illegal opium}_{i,t}) + \beta_4 t \\ & + \beta_5 \log(\text{consumption}_{i,t-1}) + \eta_i + U_{i,t}\end{aligned}$$

1. Explain why first differencing the equation does not solve the endogeneity problem of lagged consumption.
2. Anderson & Hsiao propose a specific instrumental variable procedure for the model. Write down and perform the associated first stage regression. Comment on its outcomes.
3. Estimate the specification above using the Anderson & Hsiao approach. Comment on the underlying assumptions, tabulate the results and comment on the outcomes.
4. Describe the Arellano & Bond GMM estimator for this model.
5. Estimate the model parameters using the Arellano & Bond estimator, tabulate the results and discuss the parameter estimates.

¹Van Ours, J.C. (1995), The price elasticity of hard drugs: the case of opium in the Dutch East Indies, 1923–1938, *Journal of Political Economy* 103, 261–279.

6. What is in your estimate for the short-run and the long-run price elasticity of opium?
7. Now estimate the model parameters using the system estimator (Blundell & Bond). Tabulate results, compute the elasticities (as in 6.).
8. Which parameter estimates do you prefer? Explain why. Are there remaining problems with your preferred estimates?

Data

Variables in dataset *datadynpan2022.dta*:

Variable name		units
in model	in data	
consumption	quantity	quantity of opium consumed in kg (region)
prices	price	guilders/kg
income	income	real index of income
illegal opium	illegal	amount of illegal opium intercepted in kg
<i>t</i>	year	calendar year

In addition, *region* indicates the region (regions 1-22, regions 1-5 are Java and Madura, and regions 6-22 Outer Islands), which is the cross-sectional dimension in the problem, *users* indicates the number of users (a potential alternative for *quantity*).

xtabond2

Stata has a procedure `xtabond2`. It includes the original Arellano & Bond estimator and many variants beyond. Carefully read the instruction in Stata for its use and the various options. Option *gmmstyle* can be used to specify the endogenous regressors, and option *ivstyle* for the remaining regressors (which will be treated as strictly exogenous). Option *nolevel* indicates that the instruments are the lagged differences, as in the original Arellano & Bond estimator. Use this option for the Arellano & Bond estimator as treated in class. There is also an option for the twostep estimator.