

1 Introduction

Our ultimate aim is to estimate a production function of the form:

$$Score = f(ability, resources, teacher, family, peers, community) + \epsilon$$

Once we have done this policy questions can be answered by investigating the parameters obtained from this estimate. For example how does the effect of class size differ in different communities, what proportion of the variation in expenditure estimates can be attributed to district-level effects?

[Obviously I'll need to put the general background, data and theory here too. (Importantly I need to talk about data selection and assumptions about sampling process. That is, that the data I have dropped were dropped randomly).]

2 Model

Consider a simple linear model for academic performance over campuses $i = 1, \dots, n$; school districts $j = 1, \dots, 957$; and periods $t = 2003, \dots, 2011$ (where 2003 represents the academic year '02-'03):

$$\begin{aligned} \text{Gr5.Avg}_{ijt} = & \beta_0 + \beta_1 \text{Gr4.Avg}_{ijt-1} + \beta_2 \text{Gr3.Avg}_{ijt-2} \\ & + \beta_3 \text{Per.Pupil.Exp}_{ijt} + \beta_4 \text{Econ.Disadv.Per}_{ijt} \\ & + \beta_5 \text{T.Avg.Sal}_{ijt} + \beta_6 \text{T.Avg.Exp}_{ijt} \\ & + \beta_7 \text{Gr5.Class.Size}_{ijt} + \epsilon_{ijt} \end{aligned}$$

Before estimating this model it is instructive to restrict it to more specialised cases and clearly step through the assumptions required for OLS to

be an efficient, unbiased estimation strategy. This makes suggesting alternative estimation strategies much more intuitive.

2.1 Within district, one time period

Let the first restriction be $j = 1$ (WLOG let this represent DALLAS ISD) and $t = 2006$. These are arbitrary parameter values. The model then can be simplified to:

$$\begin{aligned} \text{Gr5.Avg}_i &= \beta_0 + \beta_1 \text{Gr4.Avg.Lag1}_i + \beta_2 \text{Gr3.Avg.Lag2}_i \\ &+ \beta_3 \text{Per.Pupil.Exp}_i + \beta_4 \text{Econ.Disadv.Per}_i \\ &+ \beta_5 \text{T.Avg.Sal}_i + \beta_6 \text{T.Avg.Exp}_i \\ &+ \beta_7 \text{Gr5.Class.Size}_i + \epsilon_i \end{aligned}$$

In making these restrictions we can automatically rule out concerns due to time-trends and some spatial autocorrelation, however we must still be confident that the GM assumptions are satisfied.

2.1.1 Multicollinearity

We require that the matrix $\mathbf{X}'\mathbf{X}$ where

$$\mathbf{X} = \begin{pmatrix} 1 & x'_1 \\ 1 & x'_2 \\ \vdots & \vdots \\ 1 & x'_n \end{pmatrix}$$

,

$$x_i = \begin{pmatrix} \text{Gr4.Avg.Lag1}_i \\ \text{Gr3.Avg.Lag2}_i \\ \vdots \\ \text{Gr5.Class.Size}_i \end{pmatrix}$$

be invertible.

2.1.2 Exogeneity

Measurement error, simultaneity.

2.1.3 Homoskedasticity

Downward bias, local spatial autocorrelation.

2.1.4 Normality

??

2.2 Within district, panel

A simple linear model for performance in Dallas ISD:

$$\begin{aligned} \text{Gr5.Avg}_{it} = & \beta_0 + \beta_1 \text{Gr4.Avg.Lag1}_{it} + \beta_2 \text{Gr3.Avg.Lag2}_{it} \\ & + \beta_3 \text{Per.Pupil.Exp}_{it} + \beta_4 \text{Econ.Disadv.Per}_{it} \\ & + \beta_5 \text{T.Avg.Sal}_{it} + \beta_6 \text{T.Avg.Exp}_{it} \\ & + \beta_7 \text{Gr5.Class.Size}_{it} + \epsilon_{it} \end{aligned}$$

2.3 Across districts, one time period

A simple linear model for performance in 2006:

$$\begin{aligned}\text{Gr5.Avg}_{ij} = & \beta_0 + \beta_1 \text{Gr4.Avg.Lag1}_{ij} + \beta_2 \text{Gr3.Avg.Lag2}_{ij} \\ & + \beta_3 \text{Per.Pupil.Exp}_{ij} + \beta_4 \text{Econ.Disadv.Per}_{ij} \\ & + \beta_5 \text{T.Avg.Sal}_{ij} + \beta_6 \text{T.Avg.Exp}_{ij} \\ & + \beta_7 \text{Gr5.Class.Size}_{ij} + \epsilon_{ij}\end{aligned}$$

2.4 Across districts, panel

This linear model utilises all our data.

$$\begin{aligned}\text{Gr5.Avg}_{ijt} = & \beta_0 + \beta_1 \text{Gr4.Avg}_{ijt-1} + \beta_2 \text{Gr3.Avg}_{ijt-2} \\ & + \beta_3 \text{Per.Pupil.Exp}_{ijt} + \beta_4 \text{Econ.Disadv.Per}_{ijt} \\ & + \beta_5 \text{T.Avg.Sal}_{ijt} + \beta_6 \text{T.Avg.Exp}_{ijt} \\ & + \beta_7 \text{Gr5.Class.Size}_{ijt} + \epsilon_{ijt}\end{aligned}$$

where $\epsilon_{ijt} = \alpha_i + \mu_j + v_{ijt}$ and

$$v_{ijt} \sim \mathcal{N}(0, \sigma^2) \text{ for all } i, j, t.$$

That is our errors are composed of time-invariant, campus-level and district-level heterogeneity and an idiosyncratic error.

2.5 Other Concerns

2.5.1 Hierarchical Error Structure

2.5.2 Parameter Heterogeneity

3 Estimation Strategy

Fixed effects is likely not to be appropriate in the presence of time trends.

3.1 Non-Linear Models

GMM estimation.

3.2 Dynamic Panel Model

Small T (9), large N (957).