

Topic 5 Introductory Notes: Fixed Effects

These notes are provided as a supplement to the lecture slides in BSM946 and are not expected to be used as your sole understanding of this topic. The main arguments are presented in the lecture, with extra notes here and the key reading to ensure your understanding of the topic is complete.

Introduction

Imagine we are interested in the following question: Does sleeping more before an exam have a good or bad effect on your grade? Some of your friends claim that staying up all night and drinking Redbull works, and show you their high grades as supporting evidence of the claim. We might think that the evidence supported this claim and thus we should try this revision strategy ourselves.

Assuming we have data on people's revision techniques and their exam performance we could think about regressing the hours slept on the grades students received. However, as we discussed in lectures there may be a form of unobserved heterogeneity that is omitted from our analysis and is biasing our results. Are the students that stay up all night the night before the exam different to the rest of the students on the course? All your friends that follow this strategy of staying up all night are workaholics and work harder throughout the whole term, not just that one night. There are numerous individual level characteristics that may impact on this link between sleep and grades such as ability, work ethic and people's tendency to consume alcohol. If these individual characteristics are related to both the tendency to sleep before the exam and the grade achieved in the exam then they could bias our results.

Fixed effects regressions are a simple way of removing problems like this. The basic principle can be thought of as giving each individual in the sample their own dummy. We then compare changes in the sleep for a given individual to changes in the grade for the same individual. We are thus looking at within individual changes, rather than across individual changes. In this case we are comparing the wage/hours sleep against that individual's average, instead of against the overall average of the sample. Each student thus has their own benchmark that we are comparing them to. For example, assuming that there is a genius in the class that commonly drowns their sorrows for being a genius with heavy drinking and four hours sleep the night before the exam and ends up with a score of 90%. Looking at the average student in the year, who slept eight hours and scored 64%, this may make it seem that less sleep and more alcohol leads to a better grade. This is the incorrect comparison to make. Instead we should compare the night the genius student slept for 4 hours and scored 90% against the night she slept 2 hours and scored 80%. This grade is still substantially above the average for the cohort, but for the genius' standards, it is relatively low. It should be obvious that by only running a OLS regression on an issue like this may lead to very different results than by controlling for between individual effects and only looking at within individual change. It seems important to note at this point that we can only investigate this within individual change if the following criteria are met:

1. We observe the individual more than once so that we can compare them against themselves.
2. The individual does have variation in their behaviour. If an individual that sleeps for 8 hours and never varies away from this, we cannot see how changes in their sleeping behaviour will impact upon their grade. We need to see variation in all of the variables we have an interest in, both sleep and grades.

A simple application

Suppose you want to learn the effect of price on the demand for back massages. You have the following data from UK cities:

Table 1: A single cross-section of data

Location	Year	Price	Per capita Quantity
London	2003	£75	2.0
Manchester	2003	£50	1.0
Birmingham	2003	£60	1.5
Sheffield	2003	£55	0.8

Looking at this data, you will see that **across** the four cities, price and quantity are positively correlated. In fact, if we regress per capita quantity on price, you will obtain a coefficient on price of 0.45, suggesting that a £1 increase in price is associated with a 0.45 increase in per capita massages. This should raise some alarm bells. Economic theory would suggest the demand curve for any product is not upward sloping, so we might want to think again about our estimation. Perhaps the quality of the massages is higher or lower in certain cities? Or perhaps certain cities have a higher demand for massages so each city actually faces a separate demand curve, with higher demand cities thus facing higher prices? In econometrics language, we may be worried about endogeneity, due to omitted variables, that might be biasing our results. Thus we need to control for omitted factors, such as quality and city level differences in demand, in some manner. We really cannot say much about the price and quantity relationship with this cross-section data so less move on to look at some repeated cross section data.

Table 2: Repeated Cross-Section Data

<i>Location</i>	<i>Year</i>	<i>Price</i>	<i>Per capita Quantity</i>
London	2003	£75	2.0
London	2004	£85	1.8
Manchester	2003	£50	1.0
Manchester	2004	£48	1.1
Birmingham	2003	£60	1.5
Birmingham	2004	£65	1.4

Sheffield	2003	£55	0.8
Sheffield	2004	£60	0.7

If you eyeball this data, you will see that *within* each of the four cities, price and quantity are inversely correlated, as you would expect if demand is downward sloping. By obtaining multiple observations about each city and looking at the effect of price within each city, we have removed the pernicious effect of omitted variable bias. This is the intuition behind fixed effects regression. If you examine the sales data (per capita quantity) in Table 2, you will observe that there is considerable variation (“action”) from one row to the next. This variation/action comes in two “flavors”:

1. *Intercity (across city) variation*: variation in the average quantity from one city to the next –
2. *Intracity (within city) variation*: variation within each city over time.

The single cross-section of data (Table 1) offered only intercity (across) variation. We now know that regressions relying on intercity variation are problematic due to potential omitted variable bias. The solution is to focus on intracity (within) variation.

We must make one key assumption if we are to believe we have really removed omitted variable bias by focusing on within variation. We must assume that there are no changes in quality or demand over time within each city that we cannot control for. For example, suppose prices went up in Chicago between 2003 and 2004 because quality increased over that time period. If we cannot adequately control for quality in our regression, then we have an omitted variable bias, because omitted quality is correlated with both price and quantity. Using econometrics language, we must assume that:

Identifying assumption: Unobservable factors that might simultaneously affect the left hand side and right hand side of the regression are time-invariant.

If we accept this assumption, which normally is not too concerning, then we have a very powerful tool to remove omitted variable bias. Within-group variation is considered over time, with across-group variables not being used to estimate the regression coefficients.

Examining the data in Table 2, it is *as if* there were four “before and after” experiments. In other words, there are sales and price data before and after prices change in each of four cities. This concept of “before and after” offers some insight into the estimation of fixed effects models.

Suppose for now that we believe the demand curve has the same shape in each of the four cities. This is a parsimonious assumption and requires that we estimate the average of the four “before-and-after” effects. One way to do this is to compute the changes in prices and the changes in quantities in each city (ΔP and ΔQ) and then regress ΔQ on ΔP . (The appropriate data from our massage

example are shown in the last two columns of Table 3.) Many research papers estimate such *difference equations*. In fact, we saw this type of regression in the beginning part of the lecture on panel data topics.

Table 3: Data for Difference Equation Estimation

<i>Location</i>	<i>Year</i>	<i>Price</i>	<i>Per capita Quantity</i>	ΔP	ΔQ
London	2003	£75	2.0		
London	2004	£85	1.8	10	-0.2
Manchester	2003	£50	1.0		
Manchester	2004	£48	1.1	-2	0.1
Birmingham	2003	£60	1.5		
Birmingham	2004	£65	1.4	5	-0.1
Sheffield	2003	£55	0.8		
Sheffield	2004	£60	0.7	5	-0.1

If we have lots of years of data, things get messier. We could, in principle, compute all of the differences (i.e., 2004 versus 2003, 2005 versus 2004, etc.) and then run a single regression, but there is an easier way. Instead of thinking of each year's observation in terms of how much it differs from the prior year for the same city, let's think about how much each observation differs from the average for that city. With two key variables, price and quantity, we will be concerned with the following:

1. Variations in prices around the mean price for each city.
2. Variations in quantities around the mean quantity for each city.

We want to know whether variations in the quantities (around their means) are related to variations in prices (around their means). This is the general idea of fixed effects regressions. With our relatively small massage data set, we can do this by hand. Table 4 reports the mean prices and quantities of massages in each of the four cities, where P^*_{ct} and Q^*_{ct} denote these means.

Table 4: Repeated cross-section with Mean within-group values

Location	Year	Price	P^*_{ct}	Quantity	Q^*_{ct}
London	2003	75	80	2.0	1.9
London	2004	85	80	1.8	1.9
Manchester	2003	50	49	1.0	1.05
Manchester	2004	48	49	1.1	1.05
Birmingham	2003	60	62.5	1.5	1.45
Birmingham	2004	65	62.5	1.4	1.45
Sheffield	2003	55	57.5	0.8	0.75
Sheffield	2004	60	57.5	0.7	0.75

Table 5 reports the variation in prices and quantities around their means.

Table 5: Repeated cross-section with variation around the mean

Location	Year	Price	P^*_{ct}	Price - P^*_{ct}	Quantity	Q^*_{ct}	Quantity - Q^*_{ct}
London	2003	75	80	-5.0	2.0	1.9	0.1
London	2004	85	80	5.0	1.8	1.9	-0.1
Manchester	2003	50	49	-1.0	1.0	1.05	.05
Manchester	2004	48	49	1.0	1.1	1.05	-.05
Birmingham	2003	60	62.5	-2.5	1.5	1.45	.05
Birmingham	2004	65	62.5	2.5	1.4	1.45	-.05
Sheffield	2003	55	57.5	-2.5	0.8	0.75	.05
Sheffield	2004	60	57.5	2.5	0.7	0.75	-.05

The final step is to regress **Quantity - Q^*_{ct}** on **Price- P^*_{ct}** . Note that by subtracting the means, we have restricted all of the action in the regression to within-city action. Thus, we have eliminated the key source of omitted variable bias, namely, unobservable across-city differences in quality, sophistication, or whatever!

It turns out that there is a simple way to do this in practice:

- 1) Start with your original price and quantity data.
- 2) Regress quantity on price, but include dummy variables for the cities (remembering to omit one city).
- 3) Add a control for the time trend if you think such a trend might be important. We would typically always include a time trend in analysis like this just to be sure!

In STATA, you could run:

regress quantity price peoriadummy milwaukee dummy madison dummy 2004 dummy

In fact, doing this on the above data would give the following:

Variable	Coefficient	What it signifies
Intercept	3.91	This is the intercept of the demand curve in the omitted city (Chicago).
price	-.025	Each \$1 increase in price causes per capita demand to fall by 0.025.
peoriadummy	-1.64	This is the Peoria intercept relative to the Chicago intercept.
milwaukee dummy	-1.72	This is the Milwaukee intercept relative to the Chicago intercept.
madison dummy	-0.89	This is the Madison intercept relative to the Chicago intercept.
2004 dum	.039	Per capita sales are higher in 2004 than in 2003 by 0.039

This is known as a “*fixed effects*” regression because it holds constant (fixes) the average effects of each city. There is a shortcut in Stata that eliminates the need to create all the dummy variables. Suppose that our variable names are **quantity**, **price**, **city** and **year**.

If we type: **xtreg quantity price i.city i.year**

STATA will automatically create dummies for all but one of the city categories as well as for the year category and then run the fixed effects regression. You can do this procedure with any number of years of data, provided there are at least two observations per city. (Cities with only one observation will drop out of the regression.)

Behind the scenes of fixed effect regressions

By including fixed effects (group dummies), you are controlling for the average differences across cities in any observable *or unobservable* predictors, such as differences in quality, inherent demand, etc. The fixed effect coefficients soak up all the across-group action. What is left over is the within-group action, which is what you want. You have greatly reduced the threat of omitted variable bias. Because fixed effects models rely on within-group action, you need repeated observations for each group, and a reasonable amount of variation of your key X variables within each group.

One potentially significant limitation of fixed effects models is that you cannot assess the effect of variables that have little within-group variation. For example, if you want to know the effect of spectator sports attendance on the demand for massages, you might not be able to use a fixed effects model, because sports attendance within a city does not vary very much from one year to the next. If it is crucial that you learn the effect of a variable that does not show much within-group variation, then you will have to forego fixed effects estimation. But this exposes you to potential omitted variable bias. Unfortunately, there is no easy solution to this dilemma.

The importance of fixed effects regression

Fixed effects regression`ns are very important because data often fall into categories such as industries, states, families, etc. When you have data that fall into such categories, you will normally want to control for characteristics of those categories that might affect the LHS variable. Unfortunately, you can never be certain that you have all the relevant control variables, so if you estimate a standard OLS model, you will have to worry about unobservable factors that are correlated with the variables that you included in the regression. Omitted variable bias would result. If you believe that these unobservable factors are time-invariant, then fixed effects regression will eliminate omitted variable bias. In some cases, you might believe that your set of control variables is sufficiently rich that any unobservables are part of the regression noise, and therefore omitted variable bias is nonexistent. But you can never be certain about unobservables because, well, they are unobservable! So fixed effects models are

a nice precaution even if you think you might not have a problem with omitted variable bias. Of course, if the unobservables are not time-invariant – if they move up and down over time within categories in a way that is correlated with the variables included in the regression – then you still have omitted variable bias. You may never be able to fully rule out this possibility.

A real world application

There have recently been high profile studies of the relationship between staffing in hospitals and patient outcomes. These studies use traditional OLS regression where the unit of observation is the patient, the dependent variable is some outcome measure like mortality (a dummy variable that equals 1 if the patient died in the hospital) and the key predictor is staffing (e.g., nurses per patient). These studies do not use fixed effects, and invariably show that hospitals with more staff have better patient outcomes. These results have had enormous policy implications.

Unfortunately, these studies may suffer from omitted variable bias. For example, the key unobservable variable might be the severity of patients' illnesses, which is notoriously difficult to control for with available data. Severity is likely to be correlated with both mortality and staffing, so that the coefficient on staffing will be biased. (What is the direction of the bias?) If you ran a *hospital fixed-effects* model, you would include hospital dummies in the regression that would control for observable and unobservable differences in severity (and all other factors) across hospitals. This would greatly reduce potential omitted variable bias.

None of the current research in this field has done so, perhaps because there is not enough intra-hospital variation in staffing to allow for fixed-effects estimation. Even a fixed effects model would not completely eliminate potential omitted variable bias. You must also hope that any changes over time in unobservable patient severity *within each hospital* are uncorrelated with changes over time in staffing. This might not be such a good assumption. As hospitals experience increases in severity, they may increase staffing. If so, then unobservable severity within a hospital is correlated with staffing and the omitted variable bias is still present. We saw this idea in the IV topic when we looked at police offices and crime rates being simultaneously determined.

That's is from me, have a look at the simulation notes on panel data and causality as well!

David Morris