Kellogg School of Management
TECHNICAL NOTE
NORTHWESTERN UNIVERSITY

DAVID DRANOVE 7-112-005

# Practical Regression: Fixed Effects Models

*This is one in a series of notes entitled "Practical Regression." These notes supplement the theoretical content of most statistics texts with practical advice on solving real world empirical problems through regression analysis.*

## The "Experiment"

The technical note "Convincing Empirical Research in Ten Steps"[1] states that good empirical research should be convincing. Perhaps the most convincing empirical studies come from biomedical research, in which researchers use double-blind randomized controlled studies. By carefully controlling the experimental design, the researchers are reasonably sure that the only difference between the treatment group and the control group in the medical study is the presence of the treatment. (For example, one group of patients receives a drug; the other receives a placebo; and there are no other differences between the groups.) With such a clean experimental design, the results are bound to be convincing.

Such clean experimental designs are rare in social science research, however. As a result, research studies are subject to all sorts of biases. The technical note "Omitted Variable Bias"[2] describes the best-known bias. In many circumstances, it is possible to design social science studies that are reasonably immune to this bias. At a minimum, one can state the assumptions required for the results to be unbiased, and often these assumptions are plausible. The most important class of studies that overcome omitted variable bias is known as fixed effects models. Fixed effects models are designed to capture experiments that occur naturally in real world (i.e., non-experimental) data.

## Introduction to Fixed Effects Models

Suppose you want to learn the slope of the demand for back massages. **Table 1** shows your data from four cities:

[1] David Dranove, "Practical Regression: Convincing Empirical Research in Ten Steps," Case #7-112-001 (Kellogg School of Management, 2012).

[2] David Dranove, "Practical Regression: Introduction to Endogeneity: Omitted Variable Bias," Case #7-112-004 (Kellogg School of Management, 2012).

**Table 1:** A Single Cross-Section of Data

| City | Year | Price ($) | Per Capita Quantity |
|---|---|---|---|
| Chicago | 2003 | 75 | 2.0 |
| Galesburg | 2003 | 50 | 1.0 |
| Milwaukee | 2003 | 60 | 1.5 |
| Madison | 2003 | 55 | 0.8 |

This is *cross-section* data—data from several cities at a single point in time.

If you regress per capita quantity on price, you will obtain a coefficient on price of 0.45, suggesting that each $1 increase in price is associated with a 0.45 increase in per capita massages. You highly doubt that the demand curve for massages is upward sloping, so you do a little storytelling. You wonder if the quality of massages is higher in Chicago or if "sophisticated urbanites" have a higher demand for massages. Either of these scenarios could simultaneously drive up the price and the quantity, resulting in omitted variable bias.

Your storytelling convinces you that the data does not contain an "experiment" that allows you to identify the effect of price on demand. The variation in price across cities does not arise from a controlled environment in which all other factors that might affect demand are held constant. This is why your coefficient on price might be biased. You could address this problem by trying to determine the quality of massages and the "sophistication" of consumers and controlling for these in your regression, but you might not be able to gather this data—and you could probably convince yourself, though additional storytelling, that your model still suffers from omitted variable bias.

This is a classic example of endogeneity bias. There are several ways to cope with this bias. One technique, known as *instrumental variables*, involves replacing price in your regression with a proxy that satisfies certain requirements. Instrumental variables are discussed at length in the "Causality and Instrumental Variables" technical note.[3] A second technique, known as fixed effects regression, is a natural extension of basic regression methods.

There are two requirements for fixed effects regressions:

- You must have a time series.[4] Thus, if the cross-section unit of observation is the city, the unit of observation for the complete data set must be the city/period, where period could be week, year, etc.
- There must be variation over time (i.e., action) within each cross-section unit. Thus, if the unit of observation is the city/year for the massage regression, then prices must vary over time within some or all of the cities.

[3] David Dranove, "Practical Regression: Causality and Instrumental Variables," Case #7-112-010 (Kellogg School of Management, 2012).

[4] This is not technically true, as the data can be organized in some other way that facilitates the inclusion of fixed effects coefficients. This method is not easily explained without first describing fixed effects in time series data, however.

The data set in **Table 2** satisfies these requirements. The unit of observation is now the city/year, and prices vary over time within each city. (It is sufficient if prices vary over time in only some cities—but the more action the better, of course.)

**Table 2:** A Cross-Section/Time Series

| City | Year | Price ($) | Per Capita Quantity |
|---|---|---|---|
| Chicago | 2003 | 75 | 2.0 |
| Chicago | 2004 | 85 | 1.8 |
| Galesburg | 2003 | 50 | 1.0 |
| Galesburg | 2004 | 48 | 1.1 |
| Milwaukee | 2003 | 60 | 1.5 |
| Milwaukee | 2004 | 65 | 1.4 |
| Madison | 2003 | 55 | 0.8 |
| Madison | 2004 | 60 | 0.7 |

## Deconstructing the Action in Time Series Models

If you eyeball Table 2, you will see that price and quantity are inversely correlated within each of the four cities, suggesting that demand may be downward sloping after all. By focusing on patterns within each city, the effects of unobserved quality and sophistication disappear, revealing a more reasonable demand pattern. This is the intuition behind fixed effects regression.

We can get a better handle on fixed effects regression by deconstructing the action in the data (i.e., the variation from one row to the next). The variation/action comes in two "flavors":

- *Within*, or *intergroup*: variation in the average quantity from one group (i.e., city) to the next
- *Between*, or *intragroup*: variation within each group over time

The single cross-section of data (Table 1) offered only between-city variation. We have now determined that regressions relying on between-group variation are problematic due to potential endogeneity bias (the source of which is omitted variables). The solution is to focus on within-city variation—variation within each city over time. More generally, fixed effects regressions exploit within-group (e.g., city, store, etc.) variation over time. *Between-group variation is not used to estimate the regression coefficients because this variation might reflect omitted variable bias.*

## Theory Behind Fixed Effects Regressions

If you look at the data in Table 2, it is *as if* you had four "before-and-after" experiments—you have sales and price data before and after price changes 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 you believe the demand curve has the same slope in each of the four cities. To estimate this slope, you will want to find the average of the four before-and-after effects.[5] You could do this by computing the changes in price and the changes in quantity in each city—ΔP and ΔQ—and then regressing ΔQ on ΔP. This is perfectly fine, and many research papers estimate such *first difference models*. This is a good way to go if you have just two periods of data for each city.

For example, we can construct *diffq* and *diffp* from Table 2 by subtracting the 2004 values of quantity and price from the 2003 values in each city. (Thus, *diffp* for Chicago is 10, *diffq* for Chicago is -0.2, and so forth.) Here is the regression output:

```
. regress diffq diffp

      Source |       SS       df       MS              Number of obs =       4
-------------+------------------------------           F(  1,     2) =  152.11
       Model |  .046883563     1  .046883563           Prob > F      =  0.0065
    Residual |  .000616438     2  .000308219           R-squared     =  0.9870
-------------+------------------------------           Adj R-squared =  0.9805
       Total |  .047500001     3  .015833334           Root MSE      =  .01756

------------------------------------------------------------------------------
       diffq |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       diffp |  -.0253425   .0020548   -12.33   0.007    -.0341835   -.0165014
       _cons |   .0390411   .0127497     3.06   0.092    -.0158163    .0938985
------------------------------------------------------------------------------
```

Sure enough, the coefficient on *diffp* is negative; demand is downward sloping.

This before-and-after methodology seems fine, but it is limited to two time periods. You may want to increase the number of experiments by obtaining several years of data. You 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.[6]

Think of the within-group action as consisting of two components:

- Variations in prices around the mean price for the group
- Variations in quantities around the mean quantity for the group

You want to know if variations in quantities around their means are related to variations in prices around their means.

[5] In this example there are the same number of observations and the same amount of action in a city. As a result, the estimated slope is the average of the slopes in each city. If the amount of action varied by city, the regression would give more "weight" to the cities with more action.

[6] This "easier way," described below, is known as a "fixed effects estimator." When there are only two time periods, the first difference estimator and fixed effects estimator will yield the exact same results. With more than two time periods, the estimates should be similar, though not exactly the same. If they are very different, some assumptions for linear regression are likely unsatisfied (e.g., uncorrelated errors, exogenous regressors). See Jeffrey M. Wooldridge, *Econometric Analysis of Cross Section and Panel Data* (Cambridge, MA: MIT Press, 2002), p. 284, for a long, though complex, discussion.

Let's begin by computing the mean prices and quantities of massages in each of the four cities. I let $P^*_{ct}$ and $Q^*_{ct}$ denote these means. **Table 3** reports these values:

**Table 3:** Cross-Section/Time Series with Mean Intragroup Values

| City | Year | Price | $P^*_{ct}$ | Quantity | $Q^*_{ct}$ |
|---|---|---|---|---|---|
| Chicago | 2003 | 75 | 80.0 | 2.0 | 1.90 |
| Chicago | 2004 | 85 | 80.0 | 1.8 | 1.90 |
| Galesburg | 2003 | 50 | 49.0 | 1.0 | 1.05 |
| Galesburg | 2004 | 48 | 49.0 | 1.1 | 1.05 |
| Milwaukee | 2003 | 60 | 62.5 | 1.5 | 1.45 |
| Milwaukee | 2004 | 65 | 62.5 | 1.4 | 1.45 |
| Madison | 2003 | 55 | 57.5 | 0.8 | 0.75 |
| Madison | 2004 | 60 | 57.5 | 0.7 | 0.75 |

The next step is to compute the variation in prices and quantities around their means (see **Table 4**). In this table, *dmnq* represents Quantity – $Q^*_{ct}$ and *dmnp* represents Price – $P^*_{ct}$ ("dmn" stands for "demeaned").

**Table 4:** Cross-Section/Time Series with Variation Around Their Means

| City | Year | Price | $P^*_{ct}$ | dmnp | Quantity | $Q^*_{ct}$ | dmnq |
|---|---|---|---|---|---|---|---|
| Chicago | 2003 | 75 | 80.0 | -5.0 | 2.0 | 1.90 | 0.10 |
| Chicago | 2004 | 85 | 80.0 | 5.0 | 1.8 | 1.90 | -0.10 |
| Galesburg | 2003 | 50 | 49.0 | -1.0 | 1.0 | 1.05 | 0.05 |
| Galesburg | 2004 | 48 | 49.0 | 1.0 | 1.1 | 1.05 | -0.05 |
| Milwaukee | 2003 | 60 | 62.5 | -2.5 | 1.5 | 1.45 | 0.05 |
| Milwaukee | 2004 | 65 | 62.5 | 2.5 | 1.4 | 1.45 | -0.05 |
| Madison | 2003 | 55 | 57.5 | -2.5 | 0.8 | 0.75 | 0.05 |
| Madison | 2004 | 60 | 57.5 | 2.5 | 0.7 | 0.75 | -0.05 |

The final step is to regress *dmnq* on *dmnp*. We should include a dummy variable for 2004 and the usual constant term. Note that all of the action in *dmnq* and *dmnp* is within city, that is, variation around the mean within each city. All between-city action was eliminated when we subtracted out the means. Thus, we have eliminated the key source of omitted variable bias, namely, unobservable between-city differences in quality and sophistication—and anything else. The dummy for 2004 controls for the average change in quantity between 2003 and 2004: it corresponds to the constant term in the first differences model.

Take a look at the output from this regression:

```
. regress dmnq dmnp year2004

      Source |       SS       df       MS              Number of obs =       8
-------------+------------------------------           F(  2,     5) =  281.39
       Model |  .034691797     2  .017345899           Prob > F      =  0.0000
    Residual |  .000308219     5  .000061644           R-squared     =  0.9912
-------------+------------------------------           Adj R-squared =  0.9877
       Total |  .035000017     7  .005000002           Root MSE      =  .00785

------------------------------------------------------------------------------
        dmnq |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        dmnp |  -.0253425   .0012996   -19.50   0.000    -.0286831   -.0220018
    year2004 |   .0390411   .0080636     4.84   0.005      .018313    .0597692
       _cons |  -.0195206    .004895    -3.99   0.010    -.0321035   -.0069376
------------------------------------------------------------------------------
```

Compare this to the earlier regression that involved differences; the coefficient on price is identical. (The standard error is smaller, but this is misleading, as we will see below.)

## Fixed Effects Regression in Practice

It turns out that there is a simple way to do this analysis without explicitly "demeaning" the variables. In the process, we can inform the computer that we are basing our estimates on within-city averages. This takes up degrees of freedom, ignores between-city variation, and thereby reduces the precision of the estimates.

We start with the original data. Regress *quantity* on *price*, but include in the regression dummy variables for the cities (remembering to omit one city). This effectively demeans the data by city, just as we explicitly did in Table 4. We again include a dummy for the year 2004.

```
. regress quantity price  Galesburgdummy Madisondummy Milwaukeedummy year2004

      Source |       SS       df       MS              Number of obs =       8
-------------+------------------------------           F(  5,     2) = 1983.58
       Model |  1.52844172     5  .305688344           Prob > F      =  0.0005
    Residual |  .000308219     2   .00015411           R-squared     =  0.9998
-------------+------------------------------           Adj R-squared =  0.9993
       Total |  1.52874994     7  .218392848           Root MSE      =  .01241

------------------------------------------------------------------------------
    quantity |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       price |  -.0253425   .0020548   -12.33   0.007    -.0341835   -.0165014
Galesburgdu~y |  -1.635617    .064897   -25.20   0.002    -1.914846   -1.356387
Madisondummy |  -1.720206   .0478706   -35.93   0.001    -1.926176   -1.514235
Milwaukeedu~y |  -.8934932   .0380415   -23.49   0.002    -1.057172    -.729814
    year2004 |   .0390411   .0127497     3.06   0.092    -.0158163    .0938985
       _cons |   3.907877   .1600615    24.41   0.002     3.219188    4.596566
------------------------------------------------------------------------------
```

Notice that the coefficient and standard error on price are now identical to what we obtained in the first difference regression.

**Table 5** shows how to interpret the coefficients.

**Table 5:** Interpreting the Fixed Effects Regression Coefficients

| Variable | Coefficient | What It Signifies |
|---|---|---|
| Intercept | 3.910 | This is the intercept of the demand curve in the omitted city (Chicago). |
| Price | -0.025 | Each $1 increase in price causes per capita demand to fall by 0.025. |
| Galesburgdummy | -1.640 | This is the Galesburg intercept relative to the Chicago intercept. |
| Milwaukeedummy | -1.720 | This is the Madison intercept relative to the Chicago intercept. |
| Madisondummy | -0.890 | This is the Milwaukee intercept relative to the Chicago intercept. |
| Year2004 | 0.039 | Per capita sales are higher in 2004 than in 2003 by 0.039, all else equal. |

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:

**xi:regress quantity price i.city i.year**

Stata will automatically create dummies for all but one of the city categories and 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.)

Stata offers a very powerful package, the **xt**, for dealing with cross-section/time series data. The xt package has too many options to describe in this note, but we will explain the basics.

You begin by telling Stata that you have cross-section time series data by typing:

**xtset group time**

where *group* is the name of the variable that defines the group and *time* is the name of the variable that defines the time.

You can then estimate fixed effects and a host of related regressions using **xtreg** and similar commands. You can learn more by using the Stata commands **help xt** or **help xtreg**.

## Fixed Effects Regression: An Antidote to Omitted Variable Bias

Fixed effects soak up all between-group action. In other words, they pick up the combined effects of all time-invariant predictors that differ across groups. So in the massage model, if quality differs across cities but does not vary over time within each city, the fixed effects pick up this intercity difference. They would do the same for sophistication—and for all *observable or unobservable* factors that vary across cities but do not vary over time within each city. The remaining predictors (e.g., price) in your model explain the within-group action in your data—the effects of varying these predictors over time within each group.

By controlling for all time-invariant differences in observables and unobservables, *fixed effects models greatly reduce the threat of omitted variable bias*. Unfortunately, they do not completely eliminate the threat. If unobservables vary over time within each group (e.g., the quality of massages changes over time within each city) and these changes are correlated with changes in your predictors, then your regression still suffers from omitted variable bias. In our massage example, we have to assume that unobservables such as quality and sophistication are time-invariant. Given the relatively short time frame, this assumption seems plausible. Institutional knowledge can also increase your comfort level with key assumptions.

Because fixed effects models rely on within-group action, we need repeated observations for each group and a reasonable amount of variation of key X variables *within* each group. Again, the more action the better.

The only significant limitation of fixed effects models relative to plain vanilla ordinary least squares (OLS) is that we cannot assess the effect of variables that do not vary much within their groups. For example, if we wanted to know the effect of spectator sports attendance on the demand for massages, we might not be able to use a fixed effects model because sports attendance within a city does not vary much from one year to the next. If it is crucial to learn the effect of a variable that does not show much intragroup variation, then we must forego fixed effects estimation. We would instead have to rely on intergroup variation and hope that the omitted variable bias is minimal.

Some analysts try to rule out omitted variable bias by comparing the coefficients on key predictors in models with and without fixed effects. If the coefficients are stable, they argue, the unobservables are not that important and therefore omitted variable bias is negligible. This is not foolproof, however.

## Why Fixed Effects Are So Important

When we have data that fall into categories such as industries, states, and families, we will normally want to control for characteristics that might affect the left hand side variable. Unfortunately, we can never be certain that we have all the relevant control variables, so if we estimate a plain vanilla OLS model, we may have omitted variable bias. Fixed effects regression goes a long way toward addressing these worries.

In some cases, we might believe that the set of control variables is sufficiently rich that any unobservables must be truly random. We will not be certain, however, and can often convince ourselves that the threat of omitted variable bias is real. Here again, fixed effects models allow us to minimize the bias.

## Advice on Using Fixed Effects

- If you are concerned about omitted factors that may be correlated with key predictors at the group level, then you need to estimate a fixed effects model.
- Remember this intuition: you are conducting experiments within each group and averaging the results.
- Include a dummy variable for each group, remembering to omit one of them. (Use the **xi:** feature in Stata.)
- The coefficient on each key predictor tells you the average effect of that predictor, that is, the common slope.
- You can use an F test to determine if the groups really do have different intercepts.

### *Different Slopes for Different Folks?*

In the basic fixed effects model, the effect of each predictor variable (i.e., the slope) is assumed to be identical across all the groups, and the regression merely reports the average within group effect. What if you don't believe the slopes are identical across all groups? In the extreme, you can estimate a different regression for each group. A simpler solution—and one that does not use up degrees of freedom to estimate different slopes for every predictor variable—is to include slope dummies for key predictor variables (i.e., interactions between the group dummies and the key predictor variables) in a single regression and use an F test to see if the slopes really are different.

Here is the syntax for creating slope dummies in a Stata regression:

**xi: regress quantity price i.city*price i.year**

## An Example that Cries Out for a Fixed Effects Model

There are countless examples where fixed effects modeling can minimize omitted variable bias. Most, but not all, involve time series data, in which the "experiments" involve intragroup action over time. Here is a good example. There have recently been some high-profile studies of the relationship between staffing levels in hospitals and patient outcomes. These studies use traditional OLS regression in which the unit of observation is the patient, the dependent variable is some outcome measure such as 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. Thus, the action that drives the results includes both intragroup variation in staffing over time and intergroup variation in staffing in the cross-section.

It is easy to imagine that there could be omitted variable bias in these studies. In this case, the key unobservable variable might be the severity of patients' illnesses. This is likely to be correlated in the cross-section 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 time-invariant observable and unobservable differences in severity (and 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 intrahospital variation in staffing to allow for fixed effects estimation. The lack of good data is no excuse for running deeply flawed regressions, however, and all studies to date purporting to "show" that more staffing leads to better outcomes must be viewed with caution.

Even a fixed effects model would not completely eliminate potential omitted variable bias. It may be possible that changes over time in unobservable patient severity *within each hospital* are correlated with changes over time in staffing. This is at least plausible. A hospital that experiences an increase in severity might increase staffing. If so, then unobservable severity *within* a hospital is correlated with staffing, and the omitted variable bias is still present.