Joyce and Naber: Panel Estimnation[1]

We now consider an example of panel data estimation. This example comes from a published paper in the *Journal of Development Economics*.[2] We are following this paper for a few reasons.

1. It offers a fairly straightforward example of fixed effects estimation.

2. It presents some new material, including the Generalized Method of Moments.

3. It explores in interesting question in development economics. namely how do banking crises and sudden stops to incoming capital affect future investment.

Theory and Motivation

Financial crises are obviously of great interest to economists. This paper considers a few:

1. Sudden Stops. These refer to abrupt decrease in the amount of external capital flowing into an economy. A sudden large decline in foreign direct investment is one example. The obvious concern is that a reduction in capital might lead to a reduction in productivity both in the short and long term.

2. Banking Crises. These refer to cases where domestic banks fail or significantly reduce their amount of lending (think of the United States in the Fall of 2008). These may also induce a decrease in credit and capital.

3. Currency Crises. These refer to cases where the real value of the domestic currency depreciates. If debt is defined in terms of the foreign currency, then these crises may increase the real value of debt causing a failure of the domestic financial sector. One prominent example is the Asian financial crisis of the late 1990s.

The existing literature has examined the effects of these crises on output. Hutchinson and Noy, for example, find that currency crises reduce output growth by 2-3% while sudden stops do so by 13-15%.[3] This is intuitive. Each of these crises types represents a reduction in access to credit or capital.

[1]These are undergraduate lecture notes. They do not represent academic work. Expect typos, sloppy formatting, and occasional (possibly stupefying) errors.

[2]See Joyce, J. and M. Nabar. 2009. "Sudden Stops, Banking Crises and Investment Collapses in Emerging Markets," *Journal of Development Economics*, 90: 314-322.

[3]See Hutchinson, M. and I. Noy. 2005. "How Bad Are Twins? Output Costs of Currency and Banking Crises." *Journal of Money Credit and Banking,* 37: 725-752.

All of these crises also, however, induce a currency depreciation either directly (as in the case of currency) crises or indirectly. This makes domestic goods cheaper to foreigners and thus boosts exports, a component of GDP. We thus observe a "Mexican Wave," where the decline in output is short lived due to rapid export growth.

This paper takes a different approach. It uses capital formation as its dependent variable instead of GDP growth. The idea is that the composition of output may matter. It is probably very difficult to sustain economic growth in the long run through export growth. Long term macroeconomic performance is likely to be much more closely aligned with capital formation as this directly influences productivity. The authors are especially interested in the effects of sudden stops. The theoretical literature finds several channels through which sudden stops may have adverse long run effects.

1. During a sudden stop, firms that rely on direct borrowing from foreign lenders may suddenly be unable to do so. This deters investment and capital formation.

2. A sudden stop may also include a reduction in foreign direct investment that directly reduces capital.

3. Sudden stops may induce domestic banking crises. Suppose, for example, that domestic banks borrow in the foreign currency but lend in the domestic currency. Because a sudden stop reduces the demand for the domestic currency it results in a depreciation. This adversely affects the balance sheet of domestic banks and may thus induce a banking crisis. This is especially problematic because it eliminates the ability of borrowers who are unable to borrow from foreign lenders to substitute toward domestic lenders.

The authors show some simple data in Figures 1 and 2 that illustrate their motivation. In the judgmental of the paper's referees these charts were presumably judged as good representations.

-The top two illustrates the financial crises that affected Malaysia and Thailand in 1997. In both countries, the crisis included both a sudden stop and a banking crisis. The bottom line shows the GDP growth rate. Note the Mexican Wave where output growth plummets but quickly recovers. The top line shows the investment to GDP ratio. In both cases, this does not quickly recover. This suggests that the long run costs of financial crises may be worse than believed just by looking at the impact on output for a fairly short period after the crisis.

-The bottom left chart shows the Turkish crisis of 1994. Notably, this included a sudden stop without a banking crisis. Note that there is no major change in investment, suggesting that borrowers were able to obtain credit from domestic sources.

-The bottom right chart shows the crisis of 1981 in Chile. This was a banking crisis without a sudden stop. While investment does decline, it recovers much more rapidly than when there is also a sudden stop.

Data and Specification

The authors obtain a panel with 26 countries for annual data ranging from 1976-2002. For ten countries, all years are included. For the remaining 16 countries, the data begin after 1976. The panel is thus unbalanced. All countries are included in the International Monetary Fund's list of emerging markets. There are a total of 550 observations.

Two crucial variables are dummy variables indicating sudden stops and banking crises. A sudden stop is defined as "a fall in the financial account surplus which exceeds twice the standard deviation of the financial account during the period." Banking crises are defined as periods where "much or all of bank capital was exhausted." Sudden stops occur 8% (46 of 550) of the time while banking crises occur 24% (132 of 550) of the time. 17 times (3%) both occur at once.

The key dependent variable is gross capital accumulation as a percentage of GDP. The mean is 23.6% with a standard deviation of 6.5%.

The authors must, as always worry about omitted variable bias. They must therefore include, as controls, other factors that may affect capital formation. Some of these include:

1. Lagged investment and lagged growth. The former is included because investment is almost surely sluggish. Recall that including lagged dependent variables is problematic with panel data., For now, ignore this concern. The latter is included because wealthier countries are likely to be able to afford more investment. It is lagged, however, due to concerns about endogeneity. Output and investment are surely simultaneously determined. The expected sign for both of these is thus positive.

2. Inflation is included because it may add uncertainty, including through exchange rate risk. This would probably not be an issue for developed, low inflation economies. Keep in mind, however, that these data are for emerging markets. The expected sign is thus negative.

3. FDI/GDP. Foreign direct investment provides an often important means of capital formation. The expected sign is thus positive.

. 4. Trade openness. To quote the paper, "Trade openness can influence investment since countries that are more open to trade could also be more efficient and generate higher returns on investment. Furthermore, it is possible that countries more open to trade develop more sophisticated financial intermediation, involving a deeper network of supplier credit and risk-sharing intermediaries, and this financial development generates investment." (see page 318). The expected sign is thus positive.

5. Debt service/GDP. Borrowers with higher debt service may be more capital constrained inhibiting investment. The expected sign is thus negative.. The authors do not explicitly

discuss if and how they deal with potential non-stationary.

It is up to you to decide if the authors are including the entire set of needed controls. If not, then the model will suffer from too much omitted variable bias. If so, then omitted variable bias will be minimal. There is no econometric test to guide your judgment, you must rely on your knowledge of related theory and intuition. This paper's publication in the top journal of its field does suggest, however, that the experts felt the authors did a good job with their controls.

$$gcy_{it} = \alpha_i + \delta_t + \sum_{j=1}^{n} \beta_j x_{ijt} + \gamma_1 SS_{it} + \gamma_2 BANKCRI_{it} + \epsilon_{it} \tag{1}$$

Consider the following features of this specification:

1. The specification includes both time and country fixed effects. The authors do not explicitly discuss pooled OLS or random effects. The inclusion of these fixed effects means that the authors are allowing for unobserved heterogeneity over both time and country. Some years or countries may for example, have systematically high levels (or low levels) of investment. This unobserved heterogeneity is correlated with (hence no random effects) the right hand side variables, but is not fully explained by them (hence no pooled OLS). The authors do not show F-tests or a Hausman test. YEAH! I love it when homework questions write themselves.

2. $\sum_{j=1}^{n} \beta_j x_{ijt}$ captures the effects of the controls.

3. SS_{it} is the dummy for sudden stops while $BANKCRI_{it}$ is the dummy for bank crises.

Estimation in Stata

The first step is to tell Stata that we are working with a panel:

$$xtset \;\; country \;\; year; \tag{2}$$

Running a fixed effects regression in Stata is easy. We just use *xtreg* instead of *reg*:

$$xtreg \;\; y \;\; x, \;\; fer; \tag{3}$$

As usual, details follow the comma. It is very important to note that Stata automatically includes cross sectional, but not time, fixed effects, This is due to its tradition of being a cross sectional program. We must manually include time fixed effects and then simply include them as independent variables.

$$xtreg \;\; y \;\; x \;\; i.year, \;\; fe \;\; r; \tag{4}$$

To later conduct a Hausman test, we can store our results:

estimates store fixed

Fixed Effects Results

Table 6 reports the first set of results. Column 1 reports the fixed effects results. It does not show the coefficients for the dummies, which is common. The major results include:

1. A sudden stop corresponds to a -1.444% decrease in the investment to GDP ratio, equivalent to 22% of investment's standard deviation. This appears to be both economically and statistically significant. This is, however, only the short run (within period) effect. We can use the coefficient on lagged investment to calculate the longer term (2 period) impact which is about 3.75%.

2. The results for bank crises are similar with the short and long run effects being 1.28% and about 3.3% respectively. This result is also statistically significant.

3. Most of the controls have the predicted sign and are statistically significant at at least the 95% level. Inflation, however, does not have an economically significant effect.

The remaining columns consider results with another estimator known as the Generalized Method of Moments. For now, we will take them as given and wait to discuss GMM.

Table 7 includes some additional independent variables in the specification as a robustness check. These include δTerms of Trade Exchange Rate Volatility, the level of debt that is in dollars, and the real interest rate. None of these are individually significant nor do they change the paper's major results.

The authors are next interested in identifying the effect of a sudden stop when a banking crises does not occur and vice-versa. To get at this, they change their specification:

$$gcy_{it} = \alpha_i + \delta_t + \sum_{j=1}^{n} \beta_j x_{ijt} + \gamma_3 SS_{it} + \gamma_4 BANKCRI_{it} + \gamma_5 (SS_{it} * BANKCRI_{it}) \epsilon_{it} \tag{5}$$

The coefficients γ_3 and γ_4 are thus the coefficients for isolated sudden stops and banking crises respectively. The sum of these and γ_5 represents the combined effect. The authors estimate that $\gamma_3 = -0.378$ and $\gamma_4 = -0.910$, smaller than the corresponding coefficients from (1). neither, however, is statistically significant at the 90% level. This (and the fact that γ_5 is large and significant) suggests that the main concern is that combined sudden stops and banking crises can dramatically reduce investment and thus long term macroeconomic performance. These results are consistent with Figures 1 and 2.

The Generalized Method of Moments

Much of the estimation in the paper uses another estimator known as the generalized method of moments (GMM). While time constraints do not allow us to examine this estimator in great detail, it is worth discussing the general idea. Lars Hansen shared the 2013 Nobel Prize, sponsored by Arby's for developing this method.

Lets briefly begin with the (non-general) method of moments. Suppose that we wish to just estimate the mean of a data generating process, μ, for a variable y_t. We could write a moment condition that simply states:

$$E[y_t] - \mu = 0 \tag{6}$$

The method of moments is then a way of matching the sample moment to the moment condition:

$$\frac{1}{T}\sum y_t - \mu = 0 \tag{7}$$

This is a fancy way of saying that our estimated mean is just the sample average. But we can think of the average as a moment condition.

OLS is also a method of moments estimator. From the Gauss-Markov conditions, we have:

$$E[xe] = 0 \tag{8}$$

which just means that the expected value of the independent variables multiplied by the error terms is zero. This is just the condition for exogeneity. Recall $e = (y - x\beta)$. Hence:

$$E[x(y - x\beta)] = 0 \tag{9}$$

Equation (9) may be interpreted as a set of K moment conditions, one for each column of x. We can then fit these to our sample moment conditions:

$$\frac{1}{T}\sum x_t(y_t - x_t\beta) = 0 \tag{10}$$

So far, this is just a different way of thinking about what we already know how to do. In both cases, the number of moments is equal to the number of parameters to be estimated. GMM allows us to estimate cases where the number of moment conditions is greater than the number of dependent variables. In general, consider Q moment conditions:

$$E[m(w_t, \theta)] = 0 \tag{11}$$

where m is a $qx1$ vector of moment conditions and θ is a $kx1$ vector of parameters to be estimated. Because $k > q$, it is not possible to choose $theta$ so that all the moment conditions are satisfied. Instead we choose θ^{gmm} to minimize:

$$\bar{m}(\theta)^{'} W \bar{m}(\theta) \tag{12}$$

where $\bar{m}(\theta)$ is the sample moment conditions. We choose W. If we set it to the identity matrix, then GMM is consistent but inefficient. GMM is efficient if we set $W = Cov[\bar{m}(\theta)]^{-1}$. To show this, I will rely on a proof by intimidation by claiming that it should be obvious to even the most simple minded kindergartner.

The key feature of GMM is that it allows us to bring in and exploit additional information.

Each moment condition is information. OLS has just enough moment conditions, GMM has more.

The challenge of course is to find valid sources for this information. Consider two:

1. Structure. Suppose that economic theory predicts that $\beta_1 = e_2^{\beta}$. We can interpret this as a moment condition. If the underlying theory is incorrect, however, then we are adding misspecification to our model instead of exploiting additional information.

2. Instruments. Recall that an instrument is correlated with other independent variables but is uncorrelated with the error term. GMM allows us to exploit the information in instruments. We will not develop the linear algebra of how to do this.

GMM Results

Recall that fixed effects is inconsistent with a lagged dependent variable. The authors state (see page 318):

> However, the fixed-effects estimator is not consistent in the presence of a lagged endogenous variable; moreover, there is possible endogeneity among the regressors.

GMM allows them to deal with both the lagged dependent variable and potential endogeneity. The method of GMM that the authors employ is very complicated and follows Arellano and Bond (1991).[4] The second through fourth columns of Table 6 report the GMM results. Notice that they are very similar to fixed effects. The authors thus employ very sophisticated econometrics that primarily confirm that endogeneity and a lagged dependent variable are not problematic in this application.

GMM is easily estimated in Stata using the *gmm* command.

Other Panel estimators

The paper relies on fixed effects estimation. For completeness, we can easily compare the results to pooled ols and random effects. Pooled OLS is easy

$$reg\ \ y\ \ x,\ \ r;$$

[4]Arellano, M., and S. Bond. 1991. "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations." *Review of Economic Studies*, vol. 58(2): 277297.

We can then conduct F-tests to compare fixed effects with pooled OLS.

Switching from fixed to random effects takes days of course.

$$xtreg\ \ y\ \ x,\ \ re\ \ r;$$

estimates stored random;

We then run the Hausman test:

hausman fixed random;

You can name your regression results anything you want. But make sure to list fixed effects first.

Unfortunately, Stata runs the Hausman test only on the cross-sectional fixed effects. A second-best solution is to use this test to judge between random and fixed effects and then use an f-test to decide whether you also want time fixed effects (if the Hausman test suggests fixed effects).

Conclusions

This paper published well. The first requirement to do so was that it raise an interesting economic question: how do different types of financial crises affect investment, not just output, in emerging economies. It then used fairly simple econometrics (fixed effects) to show its main result that combined banking crises and sudden stops cause important declines in investment. It then used more sophisticated econometrics to check for robustness.

As you read the paper, you should be able to raise relevant criticisms: why not try random effects, for example. In this case, it was up to the referees and editors to decide if this question needed to be addressed. It may not have. Or it may have, but didn't make it into the final paper.