Empirical Banking and Finance
Tutorial 4

Konrad Adler

Institute for Financial Economics and Statistics
University of Bonn

Summer 2020

Tutorial 4

- Overall, very good work!
- Sometimes confusion between clustering and fixed effects
    - Title in the method part: clustering vs fixed effects might have been confusing

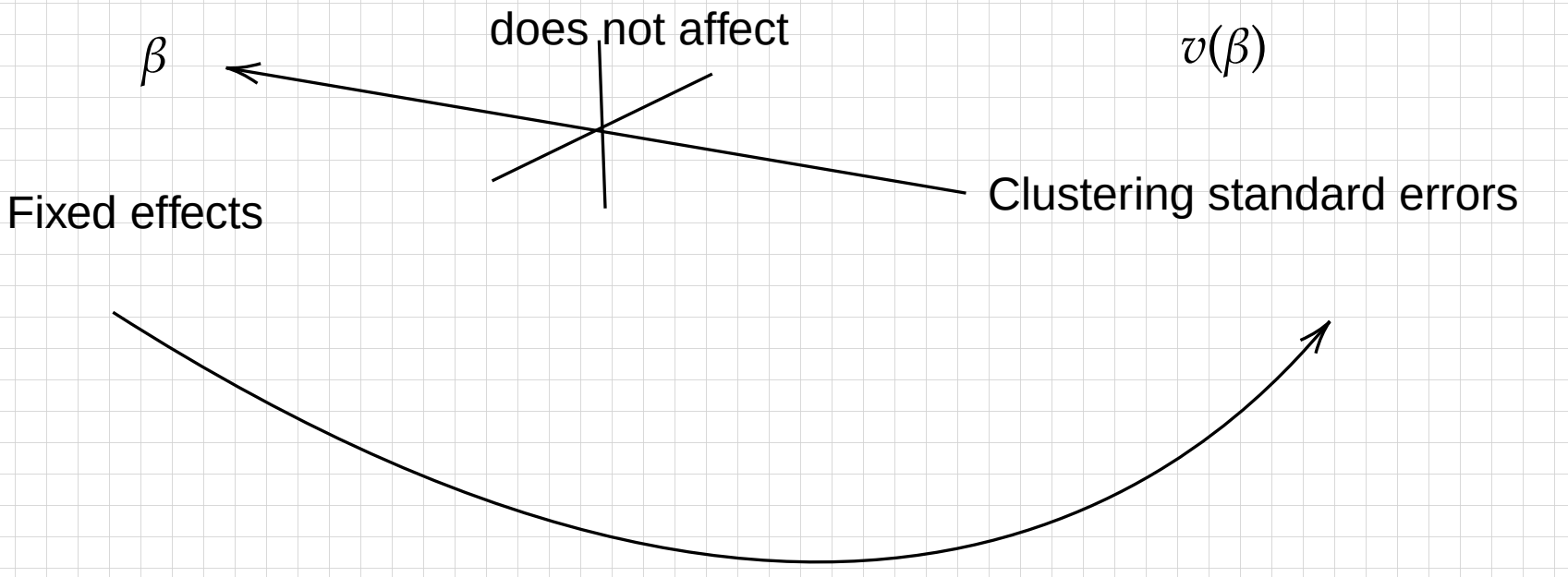
Clustering and fixed effects

- Clustering:
  - Relates to how the standard error/variance $v(\beta)$ of the coefficient β is computed
  - Clustering standard errors does not matter in the computation of β.
  - This means, among other things:
    - Clustering standard errors at the country level **does not control** for any country specific factor in the regression.
    - Clustering standard errors **does not help** against omitted variable bias
  - Clustering standard errors can help getting $v(\beta)$ right

Clustering and fixed effects

- Fixed effects:
  - Country fixed effects control for all time-invariant factors at the country level
  - This means, among other things:
    - Country fixed effects **do control** for any country specific factor in the regression.
    - Country fixed effects **do help** against omitted variable bias, as long as the omitted variable is constant over time

Clustering and fixed effects



This Lecture

Tutorial 4 Solutions

Mian (2017)

Question 1: Data & Descriptives

a) Create descriptive statistics and comment briefly.

b) Provide a scatter plot of *logGDP_future* against *private_credit_past* and compute the correlation between the two variables. Describe what the variables used measure exactly and interpret the results.

- Negative correlation between future GDP growth and past credit to GDP growth
- $logGDP\_future = (log(GDP_{t+3}) - log(GDP_t)) * 100 \rightarrow$ GDP growth
- $private\_credit\_past = (privatecredit/GDP)_{t-1} - (privatecredit/GDP)_{t-4}$
- Periods of high growth of private credit to GDP are followed by low GDP growth

c) How can you reconcile the results with the estimates on private credit to GDP on GDP growth in the two previous tutorials?

Question 1: Data & Descriptives

c) How can you reconcile the results with the estimates on private credit to GDP on GDP growth in the two previous tutorials?

- Different Measure I: here we use difference in credit to GDP
- Different Measure II: in the previous Tutorials the measures were over a much longer horizon, the interest was more on long-run growth. Here we are more interested in the impact of credit to GDP on GDP at a business cycle Frequency.
- Different Sample: Countries, this data starts only in the 1980s (for most countries)

Question 2: Regression set 1

a) Run a regression of *logGDP_future* on *private_credit_past* with country fixed effects using the Stata command *reghdfe*. For now, don't pay attention to the standard errors. Comment briefly on the sign and size of the coefficient.

- $\beta = -0.119$
- Negative sign: the negative correlation between periods of credit to GDP growth and subsequent GDP growth holds even when we account for time-invariant country charactersitics. Periods of high credit growth are followed by low GDP growth
- Size:
  - a 1 s.d. increase the past difference of credit/GDP decreases subsequent GDP growth by 0.28 s.d. of GDP growth
  - When credit/GDP increases by 16.09 percentage points (=1 s.d. of *private_credit_past*), then future GDP growth decreases by 16.09*0.119= 1.91 percentage points
  - 1.91 percentage points correspond to 0.29[1] s.d. of GDP growth
  - Not a log-log regression - I should have named the variable gdp growth!

[1] Depending on whether you use the full sample or only the regression sample

Question 2: Regression set 1

b) Run the same regression as in a) using the Stata command *reg* and including one dummy variable for each country. Compare the coefficient obtained and the R^2 measures to the previous regression.

- $\beta = -0.119$, the same as before
- R^2 in question b) corresponds to the overall R^2 in question a).
- Note that the within-R^2 in a) is much lower (8% only)

Question 2: Regression set 1

c) Compute the average for *logGDP_future* and *private_credit_past* for each country over time. Create two new variables, where you subtract the mean from each variable. Run a regression of the demeaned *logGDP_future* on the demeaned *private_credit_past*. Compare the coefficient obtained and the R^2 measures to the two previous regressions.

- $\beta = -0.119$, the same as before[2]: Fixed effects accounts to removing country-specific means of both y and x variables
- $R^2 = 8.6$ in question c) corresponds to the **within** R^2 in question a).

d) Which is the appropriate R^2 measure to report and why?

- Within R^2 is the better measure, because including country fixed effects allows to identify the coeffiient only on whithin country variation

[2] Holding the sample constant

Question 2: Regression set 1

e) If you want to give the coefficient on *private_credit_past* a causal interpretation, what is the key identifying assumption?

- All unobserved factors correlated with both *logGDP_future* and *private_credit_past* are time-invariant at the country-level.
- Confusion: Coefficients vs Standard Errors of Coefficients
  - "The key identifying assumption is that the Standard errors are independent and identically distributed"

f) In a seminar someone suggests controling for a country's legal origin. What do you answer?

- Legal origin is time-invariant and is already included/absorbed by the country FE
- Including legal origin dummies would lead to multicollinearity
- Confusion: Coefficients vs Standard Errors of Coefficients
  - "When clustering over country the clustering over legal origin is already controlled for"

Question 3: Regression set 2

a) Run the regression of *logGDP_future* on *private_credit_past* with country fixed effects using the Stata command *reghdfe* four times:

   i) Using the default standard errors
   ii) Using standard errors robust to hetereoskedasticity
   iii) Clustering standard errors at the country level
   iv) Clustering standard errors at the country level and at the year level

Question 3: Regression set 2

b) Compare the standard errors across the four types of adjustments and comment.

- Coefficient remains statistically significant whatever adjustment to s.e. we use
- Small difference when moving from no adjustment (0.015) to s.e. robust to heteroskedasticity (0.016)
- Big difference when moving from robust s.e. (0.016) to s.e. clustered by country (0.3) .
    - This can happen for different reasons
    - Possible that there are a lot of country-specific time-varying shocks
- Small difference when moving from s.e. clustered by country (0.03) to s.e. clustered by country and year (0.032)
    - This can happen for different reasons
    - There seem to be shocks affecting all countries in a given year, accounting for those common shocks increases the s.e.
    - Remember that we do not include year-FE

Question 3: Regression set 2

c) What are the "theoretical" arguments in favor of and against clustering standard errors at the country-level? What could be the reason for the additional clustering at the year level?

- Clustering at the country-level
  - In favor - technical answer: the residuals might be correlated, i.e. taking away the mean of the variable by including the country FE does not remove all serial correlation at the country level or cross-section correlation at the annual level.
  - In favor - economic answer: there are probably time-varying shocks at the country level that create serial correlation in the residuals, i.e. some countries might benefit/suffer from a temporary oil-price boom
  - Against: less than 50 clusters
- Clustering at the year level:
  - there are "global" shocks that create comovement in GDP across countries, i.e. the Great Depression
  - Note: with two-way clustering the "few cluster problem" does not get worse/better because there are more years ($46 > 42$) than countries (30) in the sample

Question 3: Regression set 2

d) Does the coefficient change?

   - Of course not!

e) In a seminar someone suggest clustering standard errors at the country-year level. What do you answer? What happens to the standard errors if you do that?

   - Does not make sense- this amounts to not clustering at all

Question 4: Regression set 3

a) Please replicate columns (2), (3) and (4) of Table III in Mian (2017)

- Past increase in household credit $\rightarrow$ decreases future GDP
- Past increase in firm credit $\rightarrow$ decreases future GDP
- Past increase in firm credit vs household credit: only household credit remains statistically significant as a predictor of future (negative) GDP
- Points to household debt as the culprit of credit busts
- Note: they cluster at the year and at the country level

b) Test formally whether the coefficients on household credit and firm credit in the regression of column (4) are equal. Provide the distribution, H_0, H_A and the result of the test.

- F-distribution (1,29)
- H_0 household_credit_past $=$ firm_credit_past
- H_A household_credit_past $\neq$ firm_credit_past
- p-value below 0.01: We reject H_0 that the two coefficients are equal

Question 4: Regression set 3

c) Re-run the regression of column (4) and add year fixed effects. Briefly comment on the results. How does the interpretation of the coefficients change compared to the regression without year fixed effects?

- Household credit increases are still significant, even though a bit less and the size is also smaller
- Interpretation: the coefficient is now the effect of household/firm credit on GDP relative to the country average and to the annual worldwide average. The year FE control for worldwide events affecting all countries equally.
- Mian (2017) do not include year FE because they argue that **global** household debt cycles are an important factor

This Lecture

Tutorial 4 Solutions

Mian (2017)

Growth and Credit

So far

- Credit as an indicator of financial development
- More credit helps GDP growth in the long run

Mian (2017)

- Household credit
- More household credit can lower GDP growth in the medium run

Mian (2017)

Related Literature

Series of papers by Schularick & Taylor: credit $\rightarrow$ crises

Mian & Sufi: US-county level study

Data

Unbalanced panel of countries

1960 - 2012

Annual frequency

Debt of households and nonfinancial firms

Macro variables

GDP forecasts

Mian (2017)

Findings

Household debt booms

Consumption to GDP increases
More consumption goods importet
GDP increases

Followed by household debt bust

GDP falls

Magnitude

a 1 s.d. increase in household debt to GDP over the last three years...
... decreases GDP by 2.1 percentage points over the next three years

What are the implications of their findings?

Mian (2017)

Implications of Findings

Use the empirical results to think about models

Are booms coming from demand or supply of credit?

> "the empirical evidence is more supportive of models in which a positive credit supply shock driven in part by flawed expectations of lenders explains the rise in debt during the boom."

How do they get there?

Mian (2017)

Implications of Findings

Shocks to credit demand

Rational expectations: positive correlation between change in debt today and future economic growth
Data: correlation is negative
Behavioral models: A change in borrower beliefs should lead to higher interest rates
Data: Household debt increases when interest rates are *low*

Mian (2017)

Implications of Findings

Shocks to credit supply

Behavioral models: A change in lender beliefs should lead to lower interest rates, higher credit and lower subsequent GDP growth
Data: ✓

Issue: cannot observe beliefs

Proxy for beliefs: GDP forecasts of IMF and OECD

Overestimate GDP growth at the end of household debt booms
$\rightarrow$ Market participants do not understand when they are in a credit boom

Mian (2017)

Other Findings

Which factors affect the end of the household credit boom?

Nominal rigidites
Constraints on monetary policy: fixed exchange rate regime

A decline in household debt does not lead to higher GDP growth!

It is more difficult to "downgrade" after a boom than "upgrade"