Empirical Banking and Finance: Tutorial № 5

Altaf Hussain and Simon Bong 22.06.2020

General

```
// Set working directory
cd "Y:\Downloads"
// install additional packages
ssc install estout, replace
ssc install reghdfe, replace
ssc install ftools, replace
// load data
use "Y:\Downloads\dataEmpBF_Tutorial5.dta"
```

1 Preliminary steps

1.1 Objectice of Jayaratne and Strahan (1996)

The authors aim to make use of the set up during the time of intra-state level branching deregulation in the US between 1972 and 1992 in order to investigate the causal relationship between finance and subsequent growth. Compared to cross sectional set ups they want to make use of the time dimension in the data to rule out a variety of sources for a potential bias. Moreover they use the background to argue against potential reverse causality, namely the possibility that finance is just a leading indicator of growth.

1.2 Improvement compared to previous literature

In the preceeding literature it was common to use a cross country setup without a time dimension. For that reason techniques like country fixed effects or difference in difference were not feasable to the researchers.
This probably causes a bias via country specific variables, the subset of those which is fixed overtime can be ruled out by the set up of the here considered paper.

1.3 Regression Equation

$$1 + g_{t,i} = \alpha_t + \beta_i + \gamma D_{t,i} + \varepsilon_{t,i} \tag{1}$$

where g is the growth rate of real percapita output (income) in year t and state i. α and β represent time and state fixed effects, the Dummy $D_{t,i}$ is equal to one for state i in all periods after[1] deregulation in the sense of lifting restrictions on branching via M&A took place. Later the time fixed effect α will be replaced by a time-region fixed effect.

1.4 Key Assumptions

By estimating equation (1) one makes the following key assumptions when considering the estimates as causal:

1. after adjusting the growth rate for time specific shocks[2] and state specific differences which are constant over time, all states are assumed to have the same adjusted growth rate before the deregulation.
2. the effect of lifting the restrictions on the growth rate is assumed to be constant measured in percent points and equal for all states.
3. moreover we make the "normal" OLS Assumptions

2 Data and Descriptives

```
// Task 2
estpost summarize
esttab using sumr.tex, cells("count mean sd min max") noobs
```

	Summary Statistics:				
	count	mean	sd	min	max
state	750	25.5	14.4405	1	50
region	750	2.38	1.165079	0	4
year	750	1984	4.323377	1977	1991
ind_dereg	750	.5533333	.4974792	0	1
ind_deregYear	750	.0426667	.2022394	0	1
deregulationInfo	0	.	.	.	.
GDPgr	700	101.8005	3.427472	81.84415	119.3942

2.1 About the Variables

The dataset includes a variable with the Statename for all considered 50 US states, these states are further assigned to 4 regions for which the identifier "region" takes account. For each observation the year is given, it starts with 1977 and ends in 1991.

[1] in the year of deregulation it self it remains equal to zero, the years are dropped from the dataset

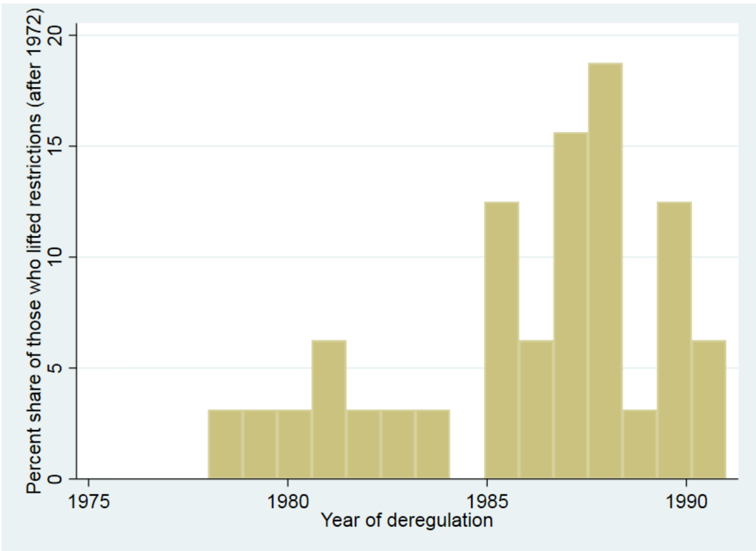
[2] which effect all states growth rates by the same amount in percent points

The field deregulationInfo contains either the information that the defegulation took place before/after the considered time horizon or the year it was implemented. This information is translated in to the two dummies ind_dereg and ind_deregYear, the first is equal to one in all years starting with the year of deregulation, the second is only equal to one in the year the regulations where lifted.

The variable GDPgr is equal to $Y_t/Y_{t-1} * 100$. In average states growed 1.8 % per year while the minimum and maximum observation reach from a 20% decrease to a 20% increase of the real per capita growth from one year to the next.

2.2 How are the years of deregulation distributed over time?

```
ge dereg_year=ind_deregYear*year
histogram dereg_year if dereg_year>0, bin(15) percent ///
ytitle(Percent share of those who lifted restrictions (after 1972))///
xtitle(Year of deregulation)
```



2.3 What does the distribution mean for the chance of finding a causal effect?

To answer this question we consider two extreme cases, first the case where in one particular year all states lift their regulations and second the case where its equally many for each considered year.

In the first case it is really hard to distinguish between the effects a possible shock in the year of deregulation and the effect of the deregulation itself. The variance in the data makes this issue less important, since its unlikely that the a shock coincidents with the deregulation each time or we can at least better test for it.

In the second case its harder to make use of the difference in difference method in its basic version. Regarding the oberseved distribution, which is some how in between the two extreme cases, we conclude that its makes it more likely to find a causal effect compared to a distribution where all mass is at one year.

3 Regression 1

3.1 Run regression 5 from Table 2 in the paper

```
// Task 3a
drop if dereg_year>0
reghdfe GDPgr ind_dereg, a(state year) vce(r)
est sto m1
// Task 3d
reghdfe GDPgr ind_dereg, a(state year) vce(cluster state year)
est sto m2
esttab m1 m2 using reg1.tex, se obslast scalar(F) r2 ///
title("Regression 1") replace
```

Table 1: Regression 1

	(1)	(2)
	GDPgr	GDPgr
ind_dereg	1.175**	1.175
	(0.375)	(0.549)
_cons	101.1***	101.1***
	(0.226)	(0.287)
R^2	0.467	0.467
F	9.807	4.582
N	668	668

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

3.2 Comparison of the sign, size and significance of the coefficient on ind_dereg to the one in the paper.

Both in the paper and in our result we find a positive effect of the deregulation on subsequent growth. We estimate a 1.175 percent point higher annual growth after the deregulation while the figure in the paper is just 1.03 percent points. The p values are not much different for both estimations.

3.3 Interpretation of size of deregulation coefficent

We expect 1.175 percent point higher annual real state GDP per capita growth after the year deregulation. The standard deviation for the growth rate is 0.375 percent points.

3.4 Re-run the regression with standard errors clustered by state and year

The results are in Table 1 Column (2). We see that after clustering the standard error of the coefficent of interest increased significantly such that our estimate is less significant[3].
This implies that we probably underestimated the standard errors before clustering, due to correlation of errors accross states for a given year.

4 Regression 2

4.1 Granger-Causality test

```
use "Y:\Downloads\dataEmpBF_Tutorial5.dta", clear
destring deregulationInfo, g(dereg_year) force
ge lead_helper=dereg_year-2
ge lag_helper=dereg_year+2
ge lead= year==lead_helper & lead_helper>1000
ge lag= year==lag_helper & lead_helper>1000
reghdfe GDPgr ind_dereg ind_deregYear lead lag, a(state year) ///
vce(cluster state year)
est sto m3
esttab m3 using reg2.tex, se obslast scalar(F) r2 ///
title("Regression 2") replace
drop if ind_deregYear>0 // for following tasks
```

For results see Table 2

4.2 Brief comment on coefficient for lead

The coefficent for the lead variable is small and positive while the standard error is relatively large compared to the size of the coefficient. The insignificance is actually good news because otherwise it would be an indicator that our treatment, the deregulation is not the cause of the observed increase in growth rates. This way we check for example whether events precceding the deregulation are causing the change which is attributed to the deregulation.

[3] its no longer significant at the 1 percent level

Table 2: Regression 2

	(1)
	GDPgr
ind_dereg	1.158
	(0.613)
ind_deregYear	0.287
	(0.642)
lead	0.181
	(0.382)
lag	0.338
	(0.401)
_cons	101.1***
	(0.323)
R^2	0.462
F	2.142
N	700

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

5 Regression 3

5.1 Please add state-specific time trends to Regression 1

```
levelsof state, local(levels)
foreach i of local levels{
ge d_‘i’=state==‘i’
ge trend_‘i’=d_‘i’*year
}
reghdfe GDPgr ind_dereg trend*, a(state year) vce(cluster state year)
est sto m4
esttab m4 using reg3.tex, se obslast scalar(F) r2 ///
title("Regression 3") replace
```

Table 3: Regression 3

	(1)
	GDPgr
ind_dereg	1.093
	(0.781)
linear state time trends	Yes
R^2	0.531
N	668

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

5.2 Tested Assumption by adding state-specific time trend

We want to test whether the states which have not lifted their restrictions yet really form a good counterfactual for those who did.

For achieving this we would like to add state-time fixed effects but by doing we would absorb the whole variance. By adding linear state specific time trends we implement a feasable second best alternative.

When the estimate of interest becomes insignificant after adding the additional controls, we can conclude that we attibuted state specific time trends to the treatment/deregulation earlyer.

5.3 Comment on the sign, size and significance of the ind_dereg coefficient. How does it compare to Regression 1?

The coefficient still indicates a positive relation but became insignificant and smaller. This is bad news for the causal interpretation because this observation makes it likely that before we wrongly attributed state specific time trends to the treatment.

6 Regression 4

6.1 add region-year fixed effects to Regression 1

```
// Task 6a
drop if region==0
reghdfe GDPgr ind_dereg, a(state year region#year) vce(cluster state year)
est sto m5
esttab m5 using reg4.tex, se obslast scalar(F) r2 ///
title("Regression 4") replace
```

Table 4: Regression 4

	(1)
	GDPgr
ind_dereg	0.732
	(0.491)
_cons	101.4***
	(0.258)
R^2	0.655
F	2.223
N	641

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

6.2 What is the rationale for adding region-year FE?

We want to controll for shocks and business cycles which vary over time but not equally effect all states across the US, but are equall for all States within a particular region.

6.3 Comparison of results to regression 7. in Table II of the paper

The coefficient we have found is slightly larger but less (not) significant.
Thinking about potential sources underlying the differences we can exclude the possibility that only diffences in the estimation of the covariance matrix are causing the difference, since this would not effect the value of the coefficient. Looking on the estimated regression equation we find that in the paper region-year fixed effects were substituted for time fixed effect while we were supposed to just add them to regression 1 without undertaking further changes.

7 Summary

7.1 Do we trust the paper?

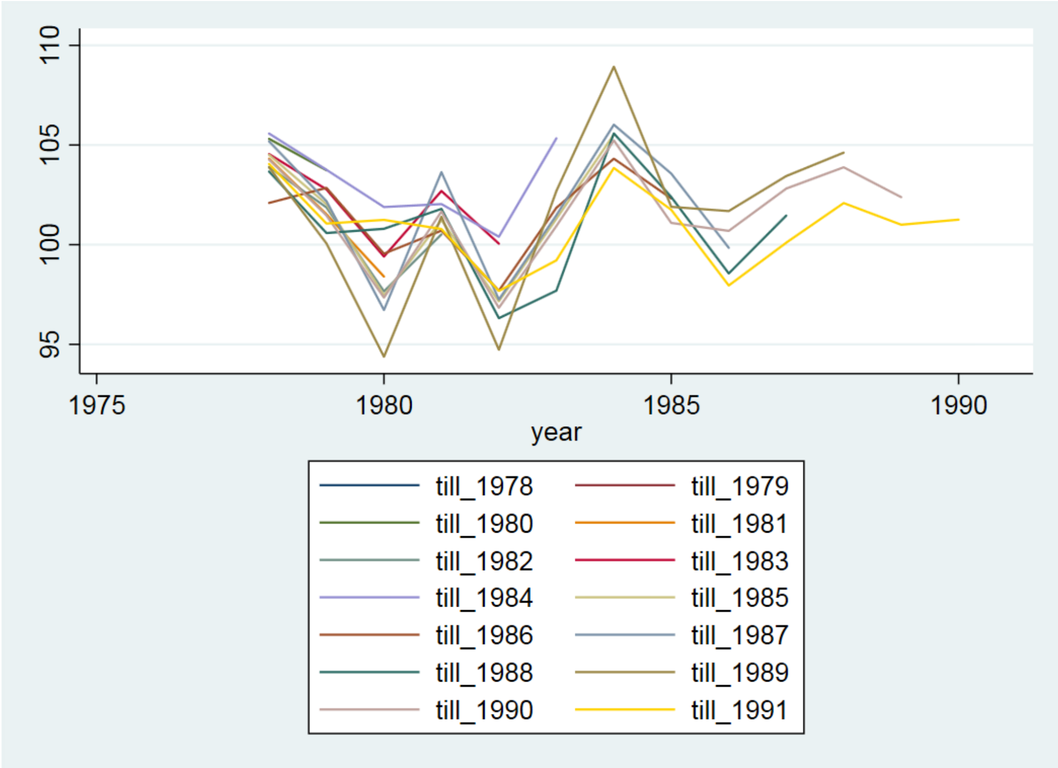
The fact that we found larger standard errors makes us a bit skeptical, while the direction of the pretictions and the argumentation makes us relatively confident.
Regarding the causality we have some doubts due to the assumption of a constant effect[4] of the deregulation after it was intruduced and also about the hypothesis that the intruduction point in time does not cause any bias related to coinciding political events. This concern became also more important due to the result of the state specific time trend test.

7.2 Possible further Robustness Checks

One example for further rubustness check which we have not considered so far woud be a common trend plot.
We created groups of states with the same ending point of their branching deregulation and computed for those the average real per capita GDP growth in all years before the ending points. To check whether we consider valid counterfactuals we plot the calculated data and observe roughly a common trend, supporting the assumptions in part 1.4.

```
// task 7
use "Y:\Downloads\dataEmpBF_Tutorial5.dta", clear
drop if ind_dereg==1
destring deregulationInfo, replace force
bysort deregulationInfo year: egen gr_avg=mean(GDPgr)
sort deregulationInfo year
quietly by deregulationInfo year:  gen dup = cond(_N==1,0,_n)
drop if dup>1
levelsof deregulationInfo, local(levels)
foreach i of local levels{
ge till_‘i’=gr_avg if deregulationInfo==‘i’
}
line till* year
```

[4] in terms of percent points on growth



Lines are ending with the year before the restrictions where lifted

Other than that we could also investigate more in the direction of confounding events as well.