

# Weekly Report about Fixed Effects, "Banking Market", and Samples

March 12, 2025

# Content

Definition of “banking market” in GP wp

Adding county fixed effects

Unbalanced obs. across categories of the fixed effect

Relation between “primarycompany” and “inst\_name” for big 4 primary companies

Definition of “banking market” in GP wp

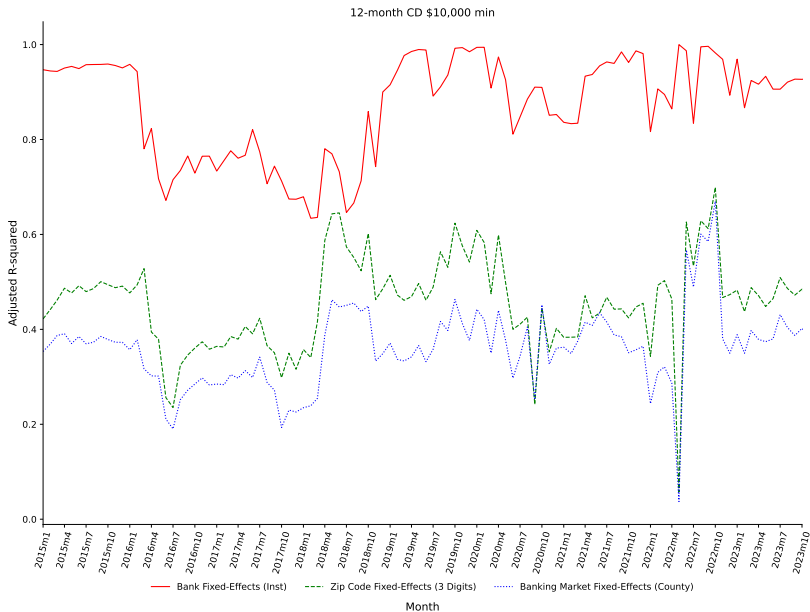
# Definition of “banking market” in GP wp

In P16 footnote, *Regional Federal Reserve Banks* define “banking markets”. The geographic delimitation of each banking market is available online: <https://cassidi.stlouisfed.org/index>.

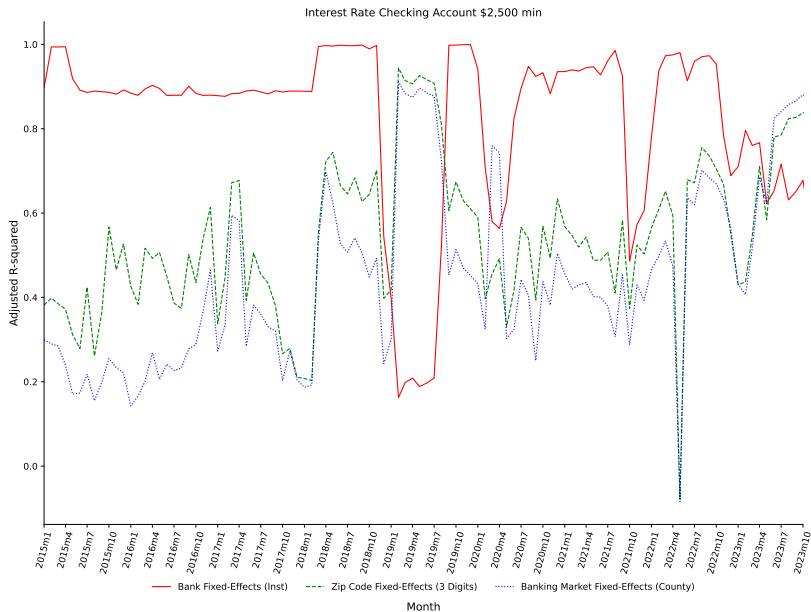
- ▶ Each “banking market” is composed of several cities, and sometimes, one banking market is a partial county.
- ▶ Ratewatch data have the “city” field, so we can match “banking market” using “city”.
- ▶ Besides, CASSIDI also has HHI data for a specific “banking market”.

Adding county fixed effects

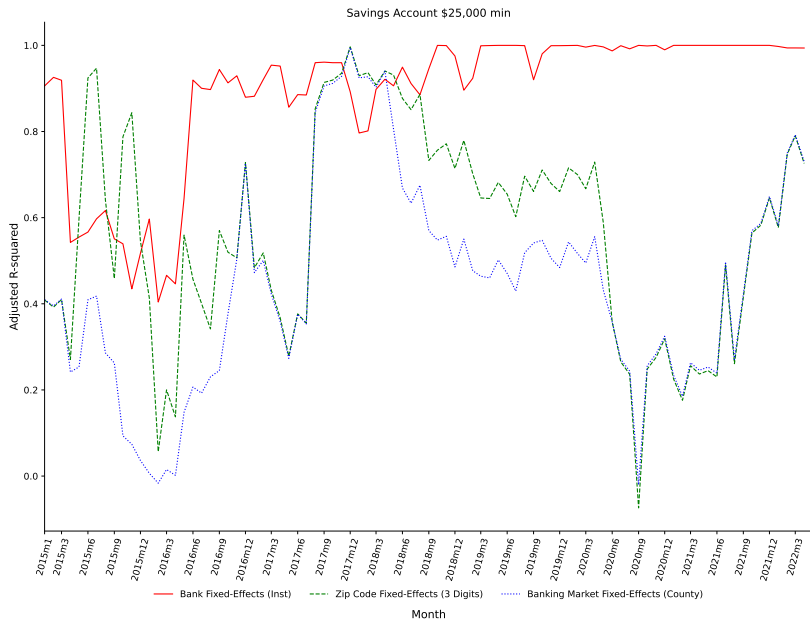
# 12MCD0K, multi-branches sample



# INTCK2.5K, multi-branches sample

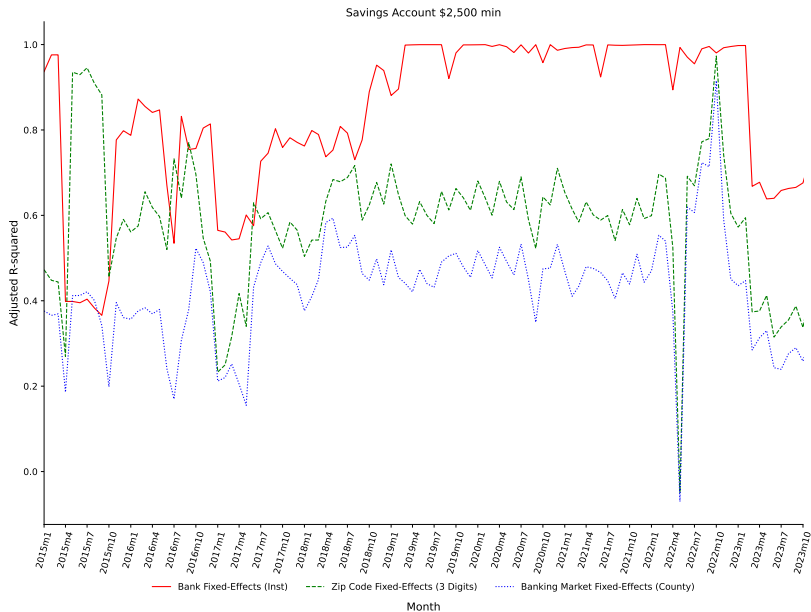


# SAV25K, multi-branches sample





# SAV2.5K, multi-branches sample



# Adding county fixed effects

- ▶ The averages across months products, for county and 3-digits fixed effects, are close.
  - ▶ (multi-branches)  $r2\_zip3$ : 0.5403,  $r2\_county$ : 0.4219
  - ▶ (all sample)  $r2\_zip3$ : 0.4409,  $r2\_county$ : 0.4225
- ▶ In line charts, the  $adj.R^2$  from regressions on county fixed effects and 3-digit fixed effects are relatively close and exhibit similar movements.

# Adding county fixed effects

Why county might be reasonable? The first 3 digits of a zip code and a county are not the same.

- ▶ Taking the reg data of multi-branch sample, product 12MCD10K, as an example, calculate:
  - ▶ The # of unique zip\_3digit values within each county.
  - ▶ The # of unique counties within each zip\_3digit.
- ▶ If either of these variables  $\neq 1$ , the observation count is 163,344, out of a total of 265,923 observations.

	zip_3digit_count_in_county	county_count_in_zip_3digit
mean	2.1858	1.5400
std	1.7953	0.9514

Unbalanced obs. across categories of the fixed effect

# Unbalanced obs. across categories of the fixed effect

- ▶ Remember that, INTCK2.5K shows unusual periods (Dec 2018 – Aug 2019),  
where  $adj.R^2$  for fixed effects, bank  $\ll$  3 digits zip code.
- ▶ However, within each bank, pricing is indeed unified. So why is the overall  $adj.R^2$  low in the regression?
- ▶ The reason is that, when applying fixed effects, one needs to **consider the balance of observations across categories.**

# Unbalanced obs. across categories of the fixed effect

*INTCK2.5K, Mar 2019, multi-branches*

	<b>With Wells Fargo</b>		<b>No Wells Fargo</b>	
	Bank FE	3 Zip Code FE	Bank FE	3 Zip Code FE
Obs.	2732	2732	2488	2488
$R^2$	0.204	0.926	0.997	0.641
$adj.R^2$	0.198	0.914	0.997	0.580

# Unbalanced obs. across categories of the fixed effect

Take Mar 2019 of INTCK2.5K as an example,

1. Notice that the total variance for the rate is very small—almost all values are 0.01.
2. The bank fixed effect generates several dummy variables (categorical variables). However, the number of observations within each category is highly unbalanced—Wells Fargo has 244 out of 2,732 observations, which is disproportionately high.
3. As a result, for the rate 0.01, since Wells Fargo has the largest number of observations, this dummy variable dominates the regression results. Other banks are treated as noise, even though pricing within each bank is unified at 0.01.

# Unbalanced obs. across categories of the fixed effect

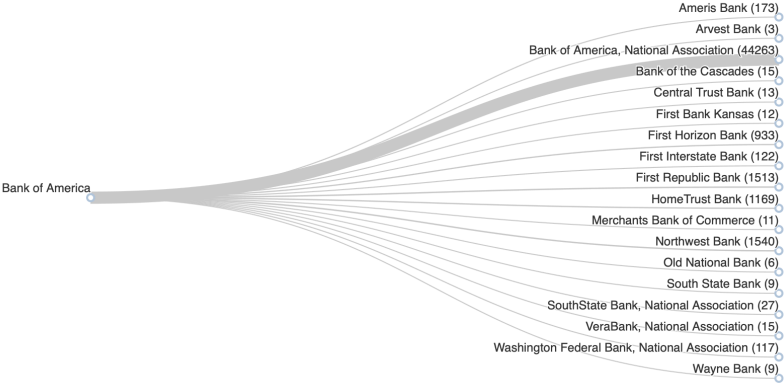
- ▶ In other words, because Wells Fargo has too many observations and the rate 0.01 has little variance, Wells Fargo dominates the model. Consequently, even though other banks (e.g., BoA, U.S. Bank, etc.) have the same pricing at 0.01, they are still considered noise, which lowers the  $adj.R^2$ .
- ▶ As for zip codes, the numbers of obs within each unique 3 digits zip code is relatively balanced, resulting in a higher  $adj.R^2$  compared to the bank fixed effect (when Wells Fargo is included in the sample).



Relation between "primarycompany" and "inst\_name" for big 4  
primary companies

# Bank of America

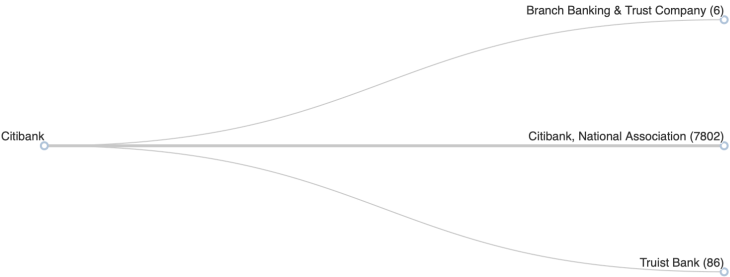
## Bank of America - Institution Ownership Structure



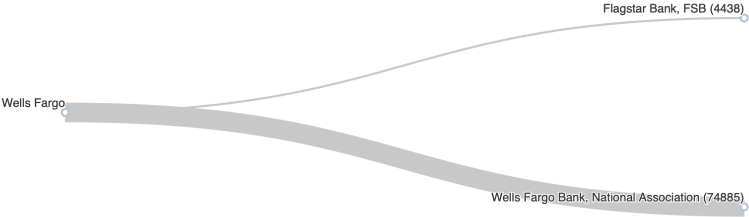
## Chase - Institution Ownership Structure



## Citibank - Institution Ownership Structure



## Wells Fargo - Institution Ownership Structure



## Relation between "primarycompany" and "inst\_name" for big 4 primary companies

- ▶ The one that truly belongs to the primary company is the entity that ends with "National Association"
  - ▶ Bank of America ← Bank of America, National Association
  - ▶ Chase ← JPMorgan Chase Bank, National Association
  - ▶ Citibank ← Citibank, National Association
  - ▶ Wells Fargo ← Wells Fargo Bank, National Association
- ▶ Other institutions have acquired branches from these primary companies. Detailed branch data can be found in `bank_institution_ownership_NotNationalAssociation_detail.xlsx`
- ▶ When conducting analysis, filter out samples where `inst_nm` does not belong to the primary company due to early-year acquisitions?