

















































































































































## B.1. Staggered Difference-in-Differences Framework

### B.1.1. Lender-County-Year Level Analysis

In this subsection, we implement a staggered difference-in-differences approach to estimate the effect of the litigation wave on FHA lenders. We assume only litigated lenders were affected and that they exited the FHA market in a staggered fashion. Litigated lenders are defined as lenders that have been subjected to litigation, audits, or investigations by the DOJ for FHA-related False Claims Act cases and have settled with the DOJ. We estimate the effect on FHA lending using a lender-county-year panel and the following specification:

$$Y_{i,c,t} = \beta \textit{Litigated}_{i,t} + \delta_i + \gamma_{c,t} + \varepsilon_{i,c,t}. \quad (1)$$

where subscripts  $i$ ,  $c$ , and  $t$  represent the lender, county, and year, respectively. The dependent variable,  $Y_{i,c,t}$ , is the number of FHA home purchase mortgage originations in a given lender-county-year. The key independent variable,  $\textit{Litigated}_{i,t}$ , is determined by a combination of whether a lender was ever litigated and the year in which the lender had settled. In particular,  $\textit{Litigated}_{i,t}$  turns one when lender  $i$  was litigated and year  $t$  is in or after the year when lender  $i$  settled with the DOJ.<sup>2</sup> The specification also includes lender and county-year fixed effects denoted by  $\delta_i$  and  $\gamma_{c,t}$ , respectively. Standard errors are double-clustered at the lender and county levels. We report the results in Panel A of Table B.1, which are overall consistent with the results in Panel A of Table 4.

[Insert Table B.1 Here]

Similar to the triple differences specifications in the paper, we also conduct staggered triple differences tests that account for lenders' origination activity in the conventional mortgage market. We construct a lender-county-year-loan type panel that includes two observations for each lender-county-year, one for FHA loan volume and one for conventional loan volume. We then estimate specifications of the form:

$$Y_{i,c,t,m} = \beta_1 \textit{Litigated}_{i,t} \times \textit{FHA}_m + \beta_2 \textit{Litigated}_{i,t} + \delta_{m,t} + \psi_{m,i} + \gamma_{c,t} + \varepsilon_{i,c,t,m}. \quad (2)$$

where the new subscript  $m$  denotes the loan type (FHA versus conventional) and  $\textit{FHA}_m$  is an indicator for the observation corresponding to FHA mortgage lending.  $Y_{i,c,t,m}$  is the number of home purchase mortgage originations of the given type. We include FHA-year, FHA-lender, and county-year fixed effects denoted by  $\delta_{m,t}$ ,  $\psi_{m,i}$ , and  $\gamma_{c,t}$ , respectively. Standard errors are double-clustered at the lender and county levels. We report the results in Panel B of Table B.1, which are overall consistent with the results in Panel B of Table 4.

### B.1.2. County-Year Level Analysis

Next, we examine the aggregate effect of the DOJ litigation wave on FHA mortgage lending using county-year level data. In these tests, counties exhibit variation in their pre-period exposure

---

<sup>2</sup>We use June 30th as a cutoff for determining the settlement year. For example, a lender that settled between July 1, 2012, and June 30, 2013, is considered to have settled in 2013.

to litigated banks that exited the FHA market in a staggered fashion. We construct the key independent variable, *Exposure to Litigated Banks*, as the current fraction of lenders in the county that have already been litigated (weighted by pre-period market shares). Formally, this variable is defined as:

$$Exposure\ to\ Litigated\ Banks_{c,t} = \sum_i Market\ Share\ in\ 2010_{i,c} \times Litigated\ Bank_{i,t}, \quad (3)$$

where *Market Share in 2010*<sub>*i,c*</sub> is the fraction of the dollar volume of FHA lending in 2010 by lender *i* in county *c* relative to all FHA lending in the county. *Litigated Bank*<sub>*i,t*</sub> turns one when lender *i* is a bank and year *t* is in or after the year when lender *i* settled with the DOJ. We estimate a staggered difference-in-differences specification of the form:

$$Y_{c,t} = \beta_1 Exposure\ to\ Litigated\ Banks_{c,t} + \beta_2 Controls_{c,t-1} + \delta_c + \gamma_t + \varepsilon_{c,t}. \quad (4)$$

where  $Y_{c,t}$  represents various dependent variables we use to measure FHA lending volume at the county-year level.  $Controls_{c,t-1}$  represents one period lagged county-year level measures of: county population, median household income, poverty rate, unemployment rate, education levels, minority population share, change in house prices, and average credit scores.  $\delta_c$  and  $\gamma_t$  are county and year fixed effects, respectively. Standard errors are clustered at the county level. We report the results in Panel A of Table B.2, which are overall consistent with the results in Panel A of Table 5.

[Insert Table B.2 Here]

To conduct staggered triple differences tests, we construct a county-year-loan type panel. This panel contains two observations for each county-year, one for FHA loan volume and one for conventional loan volume. We then estimate specifications of the form:

$$\begin{aligned} Y_{c,t,m} = & \beta_1 Exposure\ to\ Litigated\ Banks_{c,t} \times FHA_m \\ & + \beta_2 Exposure\ to\ Litigated\ Banks_{c,t} \\ & + \beta_3 Controls_{c,t-1} + \delta_{m,c} + \psi_{m,t} + \varepsilon_{c,t,m}. \end{aligned} \quad (5)$$

where the subscript *m* denotes the loan type (FHA versus conventional) and  $FHA_m$  is an indicator for the observation corresponding to FHA mortgage lending.  $Y_{c,t,m}$  represents three dependent variables we use to measure lending activity. We include FHA-county and FHA-year fixed effects denoted by  $\delta_{m,c}$  and  $\psi_{m,t}$ , respectively. Standard errors are again clustered at the county level. We report the results in Panel B of Table B.2, which are overall consistent with the results in Panel B of Table 5.



**Table B.2:** Aggregate Effect of the Litigation Wave: County-Year Level Evidence, Staggered Difference-in-Differences

This table presents regressions examining the effects of the 2012 increase in DOJ litigation risk and large banks' exit from the FHA market on total FHA lending volumes. We examine mortgage lending activity from 2009 to 2017 using a county-year panel constructed from HMDA data on first-lien home purchase mortgages for owner-occupied one-to-four family dwellings. Panel A presents staggered difference-in-differences tests using the county-year panel. The dependent variables in columns 1, 2, and 3, are  $\ln(\text{Volume})$ , the natural logarithm of one plus the dollar volume of FHA loan originations,  $\text{Volume per Capita}$ , the total dollar volume of FHA originations per capita, and  $\text{Loans per 1,000}$ , the total number of FHA originations per 1,000 residents in the county. The key independent variable is  $\text{Exposure to Litigated Banks}$ , which is the current fraction of lenders in the county that have already been litigated (weighted by pre-period market shares). The specifications also include the county-year level controls listed and defined in Appendix Table A.1, as well as county fixed effects and year fixed effects. Panel B presents staggered triple differences tests examining the same effects. These tests use a county-year-loan type panel, which is expanded to include an observation for both FHA volume and conventional volume in the county-year. The key independent variable is the interaction between  $\text{Exposure to Litigated Banks}$  and the  $\text{FHA}$  indicator for the observation corresponding to FHA lending. The specifications also include county-loan type fixed effects and year-loan type fixed effects. Standard errors are clustered at the county level and are reported in parentheses. \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% level, respectively.

*Panel A: FHA Loan Originations*

	ln(Volume) (1)	Volume per Capita (2)	Loans per 1,000 (3)
Exposure to Litigated Banks	-0.098** (0.044)	-36.164*** (6.687)	-0.118*** (0.040)
Controls	Yes	Yes	Yes
County FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
2010 Mean	8.784	246.385	1.709
Adjusted R <sup>2</sup>	0.960	0.881	0.861
# Observations	23,820	23,820	23,820

*Panel B: FHA and Conventional Loan Originations*

	ln(Volume) (1)	Volume per Capita (2)	Loans per 1,000 (3)
Exposure to Litigated Banks × FHA	-0.161*** (0.049)	-114.095*** (31.850)	-0.693*** (0.117)
Exposure to Litigated Banks	0.064** (0.027)	67.052** (29.578)	0.546*** (0.108)
Controls	Yes	Yes	Yes
FHA × County FE	Yes	Yes	Yes
FHA × Year FE	Yes	Yes	Yes
2010 Mean	9.040	317.151	2.021
Adjusted R <sup>2</sup>	0.972	0.883	0.889
# Observations	47,640	47,640	47,640