

# Supplemental Appendix

## *Banking Networks and Economic Growth: From Idiosyncratic Shocks to Aggregate Fluctuations*

Shohini Kundu<sup>1</sup>    Nishant Vats<sup>2</sup>

### Contents

|                                                                         |            |
|-------------------------------------------------------------------------|------------|
| <b>A Framework</b>                                                      | <b>A3</b>  |
| A.1 Relationship between domestic growth and foreign shocks . . . . .   | A3         |
| A.2 Aggregate Effects . . . . .                                         | A4         |
| A.2.1 Variance Equation . . . . .                                       | A4         |
| A.2.2 Covariance Equation . . . . .                                     | A5         |
| A.3 Equivalence between baseline and collapsed specification . . . . .  | A5         |
| <b>B Did Banks Expand Across State lines?</b>                           | <b>A7</b>  |
| B.1 Data . . . . .                                                      | A7         |
| B.2 Results . . . . .                                                   | A7         |
| <b>C Properties of Idiosyncratic Shocks</b>                             | <b>A10</b> |
| C.1 Presence of Fat Tails . . . . .                                     | A10        |
| C.2 Idiosyncratic shocks can predict future economic growth . . . . .   | A14        |
| C.3 Persistence of Idiosyncratic shocks . . . . .                       | A15        |
| C.4 Bank Debt and Sample Firms . . . . .                                | A15        |
| C.5 Effect of Idiosyncratic Shocks on Bank Capital . . . . .            | A18        |
| <b>D Framework for IV Design</b>                                        | <b>A19</b> |
| D.1 Violation of the Exclusion Restriction . . . . .                    | A20        |
| D.2 Economic Magnitude of the Effect from 2SLS specification . . . . .  | A21        |
| <b>E Additional Results &amp; Discussion</b>                            | <b>A24</b> |
| E.1 Baseline Results . . . . .                                          | A24        |
| E.2 Weighted Regression Results . . . . .                               | A26        |
| E.3 Heterogeneous Treatment Effects . . . . .                           | A27        |
| E.4 What Explains the Heterogeneity in State-Level Estimates? . . . . . | A28        |
| <b>F Theoretical Model</b>                                              | <b>A31</b> |
| F.1 Setup . . . . .                                                     | A31        |
| F.1.1 Households . . . . .                                              | A32        |
| F.1.2 Firms . . . . .                                                   | A32        |
| F.2 Banks . . . . .                                                     | A33        |

---

<sup>1</sup>Shohini Kundu is at the Anderson School of Management, University of California Los Angeles. email: [shohini.kundu@anderson.ucla.edu](mailto:shohini.kundu@anderson.ucla.edu)

<sup>2</sup>Nishant Vats is at the Olin Business School at Washington University in St Louis, St Louis, USA. email: [vats@wustl.edu](mailto:vats@wustl.edu)

|          |                                                               |            |
|----------|---------------------------------------------------------------|------------|
| F.2.1    | Solving the model . . . . .                                   | A34        |
| F.3      | Parameterization and Theoretical Findings . . . . .           | A34        |
| <b>G</b> | <b>Robustness</b>                                             | <b>A37</b> |
| G.1      | Alternative Measures of Idiosyncratic Shocks . . . . .        | A41        |
| G.2      | Factor Structure with Heterogeneous Exposures . . . . .       | A45        |
| G.3      | Placebo Test . . . . .                                        | A49        |
| G.3.1    | Randomizing the Timing of Deregulation . . . . .              | A49        |
| G.3.2    | Randomizing idiosyncratic shocks . . . . .                    | A50        |
| G.4      | Geography-based Measurement Error . . . . .                   | A51        |
| G.4.1    | State Level Value Added Shocks . . . . .                      | A51        |
| G.4.2    | Oil Discoveries as State Level Idiosyncratic Shocks . . . . . | A53        |
| G.5      | Addressing Concerns Related to Migration . . . . .            | A57        |
| G.6      | Dropping the states of South Dakota and Delaware . . . . .    | A60        |

## Appendix A Framework

### A.1 Relationship between domestic growth and foreign shocks

This section develops a simple framework where the transmission of idiosyncratic shocks to domestic economic growth depends on banking linkages. Let there be  $i, j \in I$  states and  $k \in K$  banks. Banks can operate across states. For simplicity, we assume that there are no other linkages between states except banking linkages. Bank lending growth is defined as a sum of aggregate shock, a bank specific capital shock, local and foreign shocks. We interpret these foreign shocks as shocks to expected future returns on capital that are uncorrelated with the bank capital shocks and other fundamental shocks.

$$\frac{\Delta l_{it}^k}{l_{i,t-1}^k} = a_t + \eta_t^k + v_{it} - \sum_{j \neq i} \frac{l_{j,t-1}^k}{l_{i,t-1}^k} v_{jt} \quad (\text{A.1})$$

Equation A.1 defines the bank lending growth function where,  $l_{it}^k$  is the lending of bank  $k$  in state  $i$  at time  $t$ ,  $\frac{\Delta l_{it}^k}{l_{i,t-1}^k}$  denotes bank lending growth, and  $\frac{l_{j,t-1}^k}{l_{i,t-1}^k}$  refers to the lending depth of bank  $k$  in state  $j$ .  $a_t$  denotes aggregate shocks with variance  $\sigma_a^2$ .  $\eta_t^k$  denotes shocks to bank capital which affects banks' loan supply ability. The variance-covariance matrix of these shocks is  $\Sigma_\eta = \sigma_\eta^2 \mathbf{1}$ , where  $\mathbf{1}$  denotes the identity matrix. The bank lending policy function so far is similar to the one employed in [Landier, Sraer and Thesmar \(2017\)](#), and assumes the presence of active, within-bank internal capital markets that generate commonality in lending growth between states conditional on bank capital shocks. The innovation is the addition of domestic,  $v_{it}$ , and foreign shocks,  $v_{jt}$ , which are uncorrelated with shocks to bank capital and aggregate shocks. We make two additional assumptions. First, banks have a fixed amount of loanable funds, and, states compete for them. Therefore, local shocks enter equation A.1 with a positive sign whereas foreign shocks enter with a negative sign. This assumption is similar in spirit to [Stein \(1997\)](#) which emphasizes the critical role of internal capital markets in the transfer of funds, within conglomerates, towards the most deserving projects. Second, we assume that the impact of these shocks is proportional to the lending depth of the bank. This assumption articulates the importance of banking relations, i.e., banks respond more to these shocks when they are deep in the economy. The variance-covariance matrix of  $v_{it}$  is given by  $\sigma_v^2 \mathbf{1}$ , where  $\mathbf{1}$  denotes an identity matrix. We make additional assumptions that include  $\mathbb{E}[a_t \eta_t^k] = 0$ ;  $\mathbb{E}[a_t v_{it}] = 0 \forall i \in I$ ;  $\mathbb{E}[\eta_t^k v_{it}] = 0 \forall i \in I$  and  $\forall k \in K$ ;  $\mathbb{E}[v_{jt} v_{it}] = 0 \forall i \neq j$ .

Economic growth in state  $i$  can be described by the equation A.2, where we posit that lending shocks affect economic growth –  $\mu > 0$  and  $\frac{\Delta y_{it}}{y_{i,t-1}}$  refer to economic growth.  $\varepsilon_{it}$  are fundamental shocks to economic growth, i.e., shocks that are unrelated to credit growth shocks. The variance of these shocks

is given by  $\sigma_\varepsilon^2$  and  $\mathbb{E}[\varepsilon_{it}\varepsilon_{jt}] = 0 \forall i \neq j$ ,  $\mathbb{E}[\varepsilon_{it}a_t] = 0$  and  $\mathbb{E}[\varepsilon_{it}v_{jt}] = 0$ .

$$\frac{\Delta y_{it}}{y_{i,t-1}} = \mu \frac{\Delta l_{it}}{l_{i,t-1}} + \varepsilon_{it} \quad (\text{A.2})$$

Combining equation A.1 and A.2 with the accounting identity  $\Delta l_{it} = \sum_{k \in K} \Delta l_{it}^k$  gives the following equation:

$$\frac{\Delta y_{it}}{y_{i,t-1}} = \mu \{a_t + v_{it} + \sum_{k \in K} \eta_t^k \frac{l_{i,t-1}^k}{l_{i,t-1}} - \sum_{j \neq i} v_{jt} \sum_{k \in K} (\frac{l_{i,t-1}^k}{l_{i,t-1}} \times \frac{l_{j,t-1}^k}{l_{t-1}^k})\} + \varepsilon_{it} \quad (\text{A.3})$$

where,  $\sum_{k \in K} \frac{l_{i,t-1}^k}{l_{i,t-1}} \times \frac{l_{j,t-1}^k}{l_{t-1}^k}$  denotes the sum of the depth of each bank  $k$  in state  $j$  ( $j \neq i$ ) multiplied with the relative importance of bank  $k$  in state  $i$ , capturing the extent of banking integration between state  $i$  and  $j$ . Equation A.3 shows that economic growth in state  $i$  is positively related to the aggregate shocks, capital shocks, and domestic shocks and negatively related to foreign shocks. While the effect of bank capital shocks increases as the reliance on that bank for external funding increases, the foreign shocks negatively affect domestic economic growth depending on the banking integration between the foreign and the domestic economy. A key testable implication from equation A.3 is that foreign idiosyncratic shocks negatively affect domestic economic growth via banking linkages. This forms the basis of our empirical strategy, combining measurement of foreign shocks and exogenous shocks to banking linkages between the domestic and the foreign economy.

## A.2 Aggregate Effects

In this section, we present a simple framework wherein foreign shocks and banking integration increase aggregate volatility by increasing the variance of economic growth of each state, and decreases aggregate volatility by potentially decreasing the covariance in economic growth. We quantify the net effect of these two forces in Section 7. Aggregate volatility is the sum of volatility in economic growth of each state and their respective covariance. Hence, we derive the expressions for the variance and the covariance. We begin by re-writing the principal equation derived in Section A.1.

$$\frac{\Delta y_{it}}{y_{i,t-1}} = \mu \{a_t + v_{it} + \sum_{k \in K} \eta_t^k \frac{l_{i,t-1}^k}{l_{i,t-1}} - \sum_{j \neq i} v_{jt} \sum_{k \in K} (\frac{l_{i,t-1}^k}{l_{i,t-1}} \times \frac{l_{j,t-1}^k}{l_{t-1}^k})\} + \varepsilon_{it} \quad (\text{A.4})$$

### A.2.1 Variance Equation

The variance of economic growth in state  $i$  using equation A.4 is given by equation A.5 where,  $H_{it} \equiv \sum_{k \in K} \{\frac{l_{i,t-1}^k}{l_{i,t-1}}\}^2$ , and  $H_{kt}^{-i} \equiv \sum_{j \neq i} (\frac{l_{j,t-1}^k}{l_{t-1}^k})^2$ . Equation A.5 connects domestic economic volatility with banking integration and foreign shocks, wherein the volatility of economic activity in state  $i$  increases as

banking integration increases.

$$\text{Var}\left[\frac{\Delta y_{it}}{y_{i,t-1}}\right] = \mu^2 \sigma_a^2 + \mu^2 \sigma_\eta^2 \times H_{it} + \mu^2 \sigma_v^2 \left(1 + \sum_{k \in K} \left(\frac{l_{i,t-1}^k}{l_{i,t-1}}\right)^2 \times H_{kt}^{-i}\right) + \sigma_\varepsilon^2 \quad (\text{A.5})$$

### A.2.2 Covariance Equation

Next, we employ equation A.4 to derive the covariance equation. For simplicity in notation we present the covariance of economic growth for state 1 and 2 in equation A.6.

$$\begin{aligned} \text{Cov}\left[\frac{\Delta y_{1t}}{y_{1,t-1}}, \frac{\Delta y_{2t}}{y_{2,t-1}}\right] &= \mu^2 \{ \sigma_a^2 + \sigma_\eta^2 \left( \sum_{k \in K} \frac{l_{1,t-1}^k}{l_{1,t-1}} \times \frac{l_{2,t-1}^k}{l_{2,t-1}} \right) - \sigma_v^2 \left( \frac{l_{1,t-1} + l_{2,t-1}}{l_{2,t-1}} \right) \left( \sum_{k \in K} \frac{l_{1,t-1}^k}{l_{1,t-1}} \times \frac{l_{2,t-1}^k}{l_{2,t-1}} \right) \right. \\ &\quad \left. + \sigma_v^2 \sum_{j \neq 1,2} \left( \sum_{k \in K} \frac{l_{1,t-1}^k \times l_{j,t-1}^k}{l_{1,t-1} \times l_{j,t-1}^k} \times \sum_{k \in K} \frac{l_{2,t-1}^k \times l_{j,t-1}^k}{l_{2,t-1} \times l_{j,t-1}^k} \right) \} \end{aligned} \quad (\text{A.6})$$

The net effect of financial integration on covariance seems ambiguous. However, the negative term associated with financial integration and idiosyncratic shocks is of order 3 whereas the positive term associated with financial integration and idiosyncratic shocks is of order 4. It remains a quantitative question whether the net effect of financial integration and idiosyncratic shocks on covariance is positive or negative. The overall effect of financial integration will depend on the strength of the covariance term relative to the variance term if the covariance term is net negative. We address these quantitative issues in Section 7.

### A.3 Equivalence between baseline and collapsed specification

Our baseline specification estimates a regression at the  $(i, j, t)$  level where each observation corresponds to a state-pair  $(i, j)$  at time  $t$ .

$$\Delta GDP_{it} = \beta_0 Post_{i,j,t} \times \Gamma_{j,t-1}^{ind} + \beta_1 Post_{i,j,t} + \alpha_{i,j} + \theta_i \times t + \theta_{jt} + \varepsilon_{i,j,t}, i \neq j \quad (\text{A.7})$$

$\Delta GDP_{it}$  denotes real GDP growth for state  $i$ ,  $\Gamma_{j,t-1}^{ind}$  denotes state-level idiosyncratic shock for state  $j$ , and  $Post_{i,j,t}$  is a binary variable taking a value of 1 if banks in state  $j$  are allowed to operate in state  $i$ .  $\alpha_{i,j}$  denotes state-pairwise fixed effects, controlling for all time invariant state-pair specific heterogeneity such as distance.  $\theta_{jt}$  captures time-varying heterogeneity for state  $j$ . We do not include the level term for  $\Gamma_{j,t-1}^{ind}$  as it is absorbed within  $\theta_{jt}$ . We also control for  $\theta_i \times t$  denoting the linear trend specific to state  $i$ .<sup>23</sup>  $\varepsilon_{i,j,t}$  denotes the idiosyncratic term in the baseline specification. This regression equation is estimated at state-pair level as the variable  $Post_{i,j,t}$  exhibits variation at state-pair level. Furthermore, the state-pair

<sup>23</sup>Our results are robust to the exclusion of  $\theta_i \times t$ .

level regression enables us to account for time-varying factors specific to the origin state of idiosyncratic shocks, as well as any time-invariant heterogeneity at the state-pair level. As the regression is estimated at the state-pair level, the regression error term is likely to exhibit correlation at the state-pair level. Hence, the regression standard errors are estimated by two-way clustering at the state  $i$  and state  $j$  levels.

As a clarification, note that each state <sub>$i$</sub>  appears  $N - 1$  times in the regression sample each year using the granular residual of each state <sub>$j$</sub>  as a regressor at a time ( $i \neq j$ ), where  $N$  is the total number of states in the sample. To implement this, we estimate  $N - 1$  equations for each state  $i$  as represented by the following system of equations:

$$\begin{aligned}\Delta GDP_{i,t} &= \beta_0 Post_{i,j,t} \times \Gamma_{j,t-1}^{ind} + \beta_1 Post_{i,j,t} + \alpha_{i,j} + \theta_i \times t + \theta_{j,t} + \varepsilon_{i,j,t} \\ &= \dots \\ &= \beta_0 Post_{i,k,t} \times \Gamma_{k,t-1}^{ind} + \beta_1 Post_{i,k,t} + \alpha_i \times \alpha_k + \theta_i \times t + \theta_{k,t} + \varepsilon_{i,k,t}\end{aligned}\tag{A.8}$$

Adding all of the equations for state  $i$  yields the following:

$$\begin{aligned}(N - 1)\Delta GDP_{i,t} &= \beta_0 \left( \sum_j Post_{i,j,t} \times \Gamma_{j,t-1}^{ind} \right) + \beta_1 \sum_j Post_{i,j,t} \\ &\quad + \sum_j \alpha_{i,j} + (N - 1)\theta_i \times t + \sum_j \theta_{j,t} + \sum_j \varepsilon_{i,j,t}\end{aligned}\tag{A.9}$$

$$\begin{aligned}\Rightarrow \Delta GDP_{i,t} &= \frac{\beta_0}{N - 1} \left( \sum_j Post_{i,j,t} \times \Gamma_{j,t-1}^{ind} \right) + \frac{\beta_1}{N - 1} \sum_j Post_{i,j,t} \\ &\quad + \left( \theta_i \times t + \frac{1}{N - 1} \cdot \sum_j \alpha_{i,j} \right) + \frac{1}{N - 1} \cdot \sum_j \theta_{j,t} + \frac{1}{N - 1} \cdot \sum_j \varepsilon_{i,j,t} \\ &= \beta_0^* \left( \sum_j Post_{i,j,t} \times \Gamma_{j,t-1}^{ind} \right) + \beta_1^* \sum_j Post_{i,j,t} + \beta_2^* \sum_j \Gamma_{j,t-1}^{ind} + \theta_i^* \times t + \theta_t^* + \varepsilon_{i,t}^*\end{aligned}\tag{A.10}$$

$$\text{where, } \frac{1}{N - 1} \cdot \sum_j \theta_{j,t} = \beta_2^* \cdot \frac{1}{N - 1} \cdot \sum_j \Gamma_{j,t-1}^{ind} + \theta_t^*$$

Equation A.10 can be estimated at state-year ( $i, t$ ) level and the interpretation of the estimate of interest, i.e., the coefficient of the interaction term of Post and foreign idiosyncratic shocks, is similar to the estimate obtained from the system of equation in specification 3.

## Appendix B Did Banks Expand Across State lines?

The mechanism outlined in this paper relies on the assumption that banks did indeed expand across state lines post banking integration. While state-pairwise banking deregulation simulates the geographic expansion across state lines by diminishing regulatory frictions, the actual expansion is an equilibrium outcome which may not have been affected by the removal of regulatory barriers. In this section, we investigate if banks did expand across state lines.

### B.1 Data

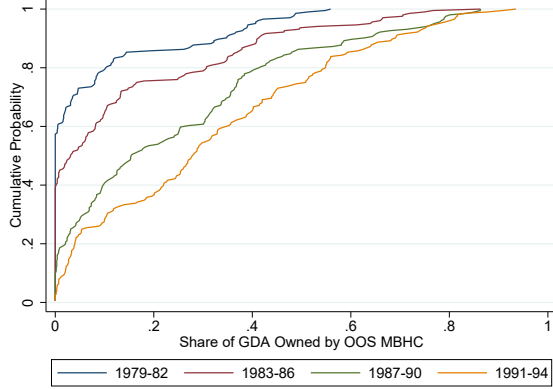
We employ state-level annual data on the share of gross domestic banking assets held by out-of-state Multi Bank Holding Companies (MBHCs). This data comes from [Berger, Kashyap and Scalise \(1995a\)](#). We use the share of gross domestic assets held by out-of-state MBHCs as a proxy for geographic expansion by out-of-state banks. A shortcoming of this measure is that it covers only a subset of all out-of-state banks, namely, out-of-state MBHCs. This suggests that our measure of geographic expansion by out-of-state banks is biased downwards. However, in light of the findings of [Berger, Kashyap and Scalise \(1995b\)](#), which notes that despite the exponential growth of assets in the banking industry between 1979 and 1994, the majority of independent banking organizations (top-tier bank holding companies and unaffiliated banks) disappeared during this time, we surmise that the error caused from mismeasurement is likely small. We use this dataset because unlike the Call Reports dataset employed in [Morgan, Rime and Strahan \(2004\)](#) and [Landier, Sraer and Thesmar \(2017\)](#) this dataset does not rely on the assumption that the lending by a bank is exclusively in the state where the bank is headquartered.

### B.2 Results

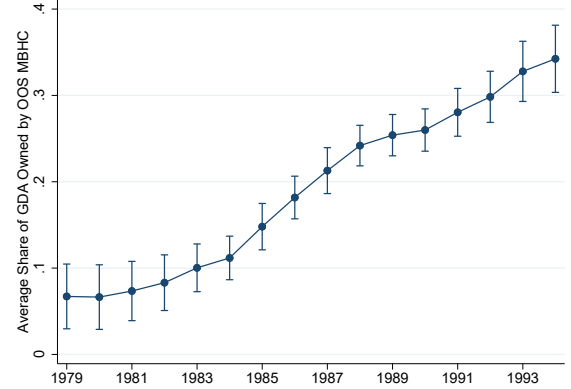
Figure [B.1a](#) reports the cumulative density function (CDF) of the share of gross domestic assets owned by out-of-state MBHCs in each state for four periods between 1979 and 1994. The period of 1979-82 refers to the four years from 1979 to 1982. This is the period before deregulation, during which, ~60% of states did not have any assets held by out-of-state MBHCs. The two periods between 1983 and 1990 (1983-86 and 1987-90), refer to the phase of active deregulation. By the end of 1990, 50% of all states had deregulated with at least 50% of all other states. The period between 1991 and 1994 is the last phase of deregulation before the passage of IBBEA in 1994. From 1979 to 1994, we see that the CDF of the share of gross domestic assets held by out-of-state MBHCs first order stochastically dominates the CDF from the previous period. This is prima facie evidence supporting the hypothesis that geographic expansion of banks occurred contemporaneously with banking deregulation. To further explore the increase in out-of-state banking presence within a given state, we run a regression of the share of gross domestic assets owned by out-of-state MBHCs on time dummies while controlling for state fixed effects. Figure [B.1b](#) plots the yearly margins and 95% confidence interval from this regression. The share of gross domestic assets owned by out-of-state MBHCs grew from ~7% in 1979 to ~35% in 1994. Growth

is relatively flat from 1979 through 1982, and picks up steadily after 1982 with a small period of low growth in the year 1990.

Figure B.1: Geographic Expansion by Out-of-State Banks Over Time



(a) CDF plots for share of GDA Owned by OOS MBHCs



(b) Within state temporal variation in share of GDA Owned by OOS MBHCs

The figure plots the temporal variation in the share of gross domestic assets (GDA) owned by out-of-state (OOS) MBHCs. Panel B.1a plots the cumulative distribution function (CDF) for the share of GDA owned by OOS MBHCs. Each line presents the CDF for a four year period between 1979 and 1994. Panel B.1b reports the average share of GDA owned by OOS MBHCs within a state. The estimate are generated by regression the share of GDA owned by OOS MBHCs on year dummies and controlling for state fixed effects. The 95% CI are generating by clustering standard errors at state level.

We formally investigate the effect the deregulation timing on the share of gross domestic assets owned by out-of-state MBHCs in Table B.1. For each state, we identify the median deregulation year. Median deregulation year is defined as the year by which that state has deregulated cross-state banking activity with 50% of all other states. The variable  $Post_{i,j,t}$  ( $=1$ ) takes a value of 1 for all yearly observation for a state after the median deregulation year. The point estimate for  $Post_{i,j,t}$  ( $=1$ ) is positive and statistically significant. The  $Post_{i,j,t}$  ( $=1$ ) variable can explain  $\approx 11\%$  of variation in the heterogeneity in the share of gross domestic assets owned by out-of-state MBHCs during the sample period. Column (3)-(5) report within state estimator for the  $Post_{i,j,t}$  ( $=1$ ) variable while controlling for aggregate annual shocks. Economically, the estimate implies that the share of gross domestic assets owned by out-of-state MBHCs grew by at least 7 pp post median deregulation year.

Table B.1: Geographic Expansion by Out-of-State Banks and Deregulation Timing

|                    | (1)                | (2)                | (3)                | (4)                | (5)                |
|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| Post (=1)          | 0.1758<br>(0.0445) | 0.1966<br>(0.0262) | 0.0753<br>(0.0283) | 0.0708<br>(0.0324) | 0.0706<br>(0.0336) |
| State FE           |                    | Yes                | Yes                | Yes                |                    |
| Year FE            |                    |                    | Yes                |                    |                    |
| Region-Year FE     |                    |                    |                    | Yes                | Yes                |
| State Linear Trend |                    |                    |                    |                    | Yes                |
| # Obs              | 816                | 816                | 816                | 816                | 816                |
| $R^2$              | 0.1049             | 0.7135             | 0.7861             | 0.8257             | 0.8263             |

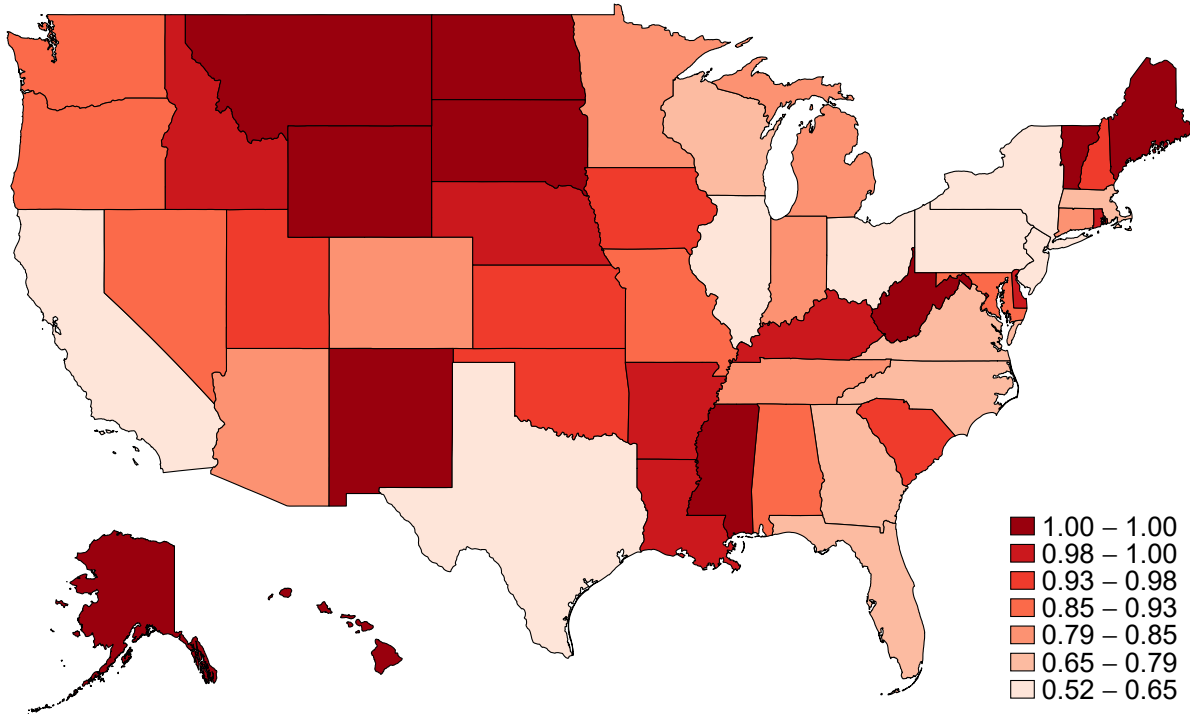
The table reports the regression of the share of gross domestic assets (GDA) owned by out-of-state (OOS) MBHCs on the Post (=1) variable. The variable Post (=1) takes a value of 1 after the median deregulation year. Median deregulation year is defined as the year by when that state deregulated with at least 50% of other states. The data on the share of GDA owned by OOS MBHCs comes from [Berger, Kashyap and Scalise \(1995a\)](#). Standard errors reported in parentheses are clustered by state.

## Appendix C Properties of Idiosyncratic Shocks

### C.1 Presence of Fat Tails

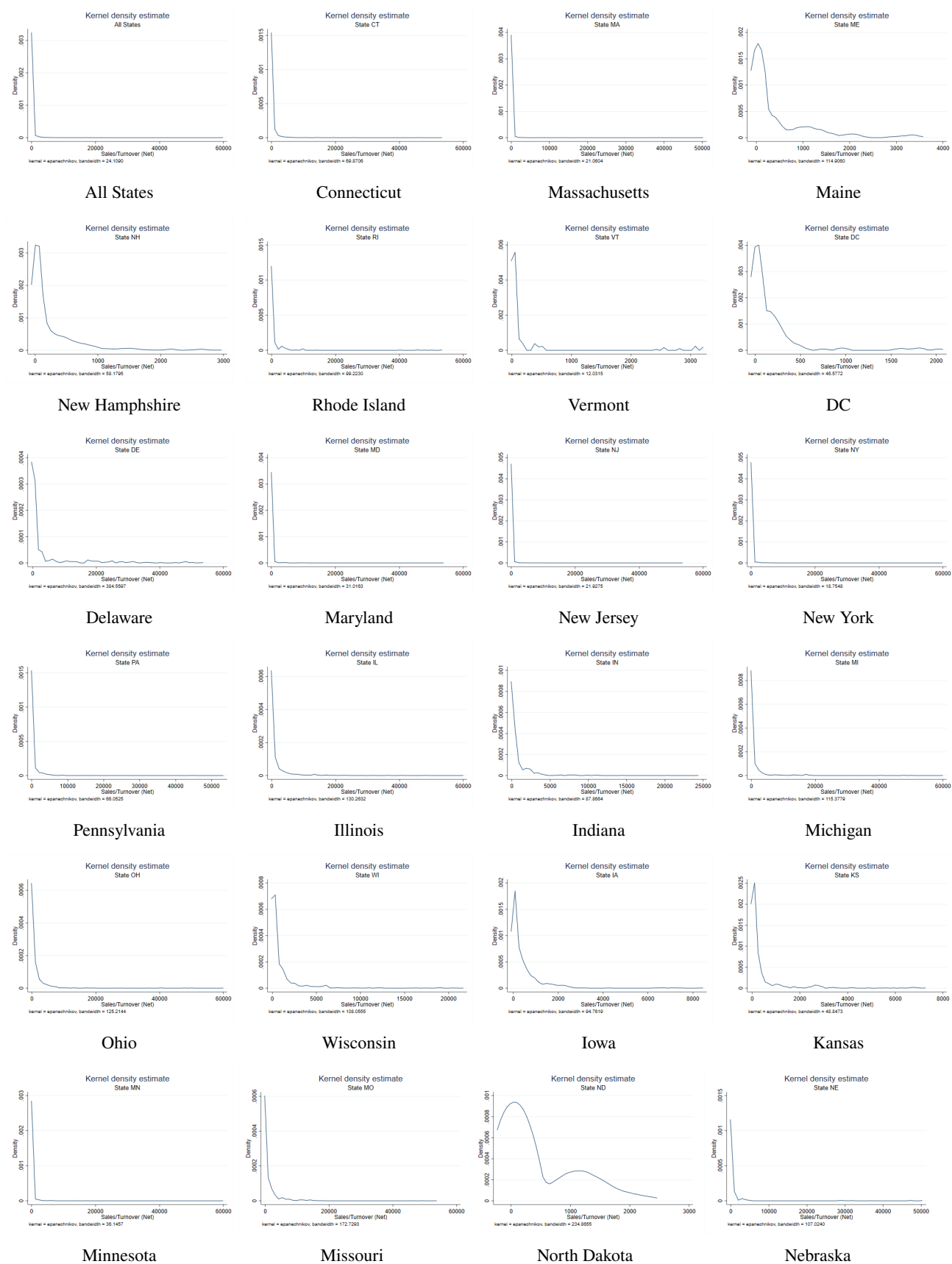
We begin our analysis by verifying that each state is dominated by large firms. We examine the ratio of sales by top 10 firms by sales to the sales of all firms for each state and find strong evidence of dominance of state-level economies by large firms. Figure C.1 shows the average proportion of sales of top 10 firms by sales relative to the total sales by all firms head-quartered in that state. The minimum value of the ratio is 0.52 indicating that top 10 firms by sales account for at least 50% of sales by all firms in that state. This is prima-facie evidence of the existence of fat tails. There is some heterogeneity in the sales share of top 10 firms by state but on average top 10 firms account for 85% of total sales. Note that in some states, such as North Dakota, South Dakota, West Virginia etc., top 10 firms account for all the sales. This is primarily because the total number of firms headquartered in that state are less or equal to 10. We supplement this analysis with a more formal description of the distribution of sales of all firms in each state. The distributions reported in Figure C.2 provide strong evidence of the sales being fat tailed in each state.

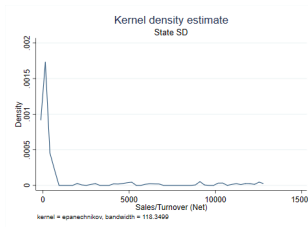
Figure C.1: Cross-Sectional Distribution of Sales Share of Top 10 Firms



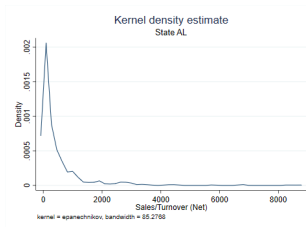
The figure plots the cross-sectional distribution of sales of top 10 firms in the state to the sales of all firms in that state between 1978 to 1995. We report the time-series average of the sales ratio of top 10 firms for each state. The legend denotes the ratio of sales of top 10 firms in the state to the sales of all firms in that state.

Figure C.2: Sales Distribution for HQ Firms in Each State

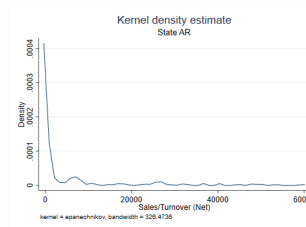




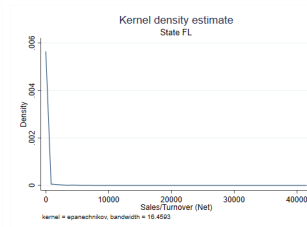
South Dakota



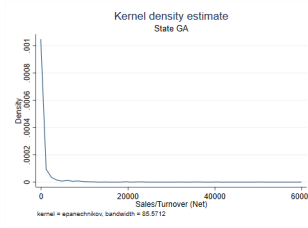
Alabama



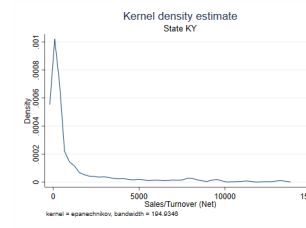
Arkansas



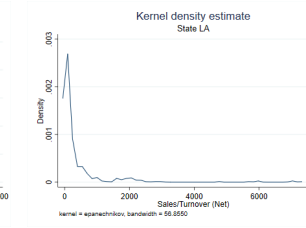
Florida



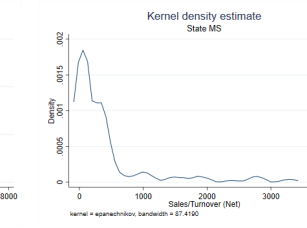
Georgia



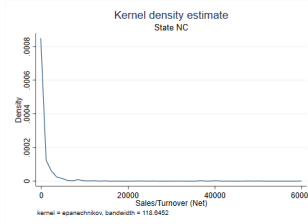
Kentucky



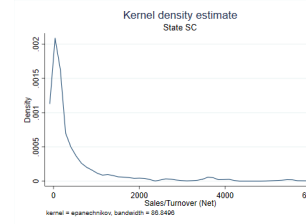
Louisiana



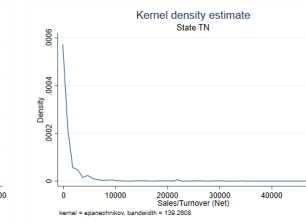
Minnesota



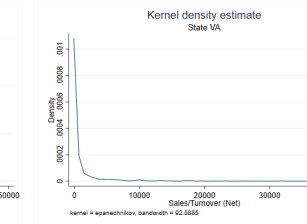
North Carolina



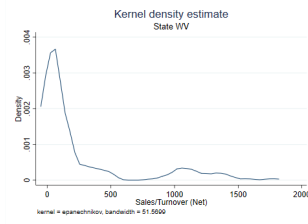
South Carolina



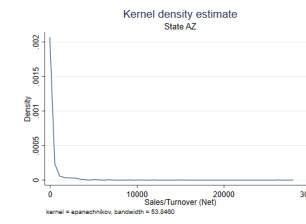
Tennessee



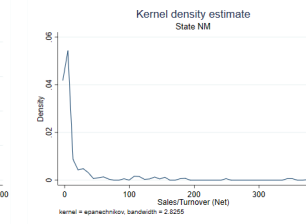
Virginia



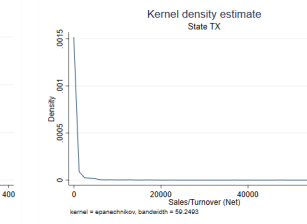
West Virginia



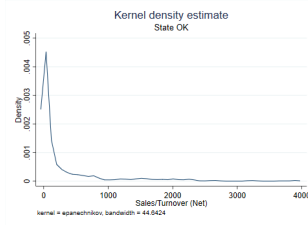
Arizona



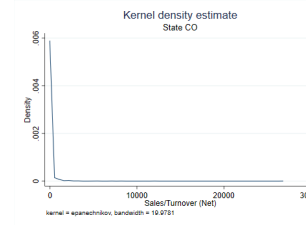
New Mexico



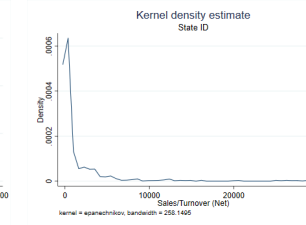
Texas



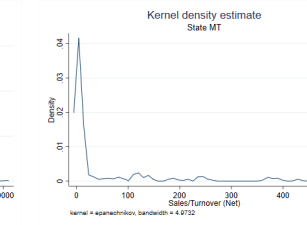
Oklahoma



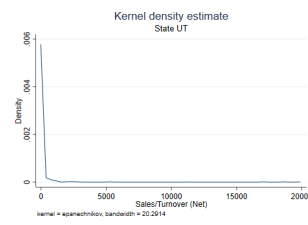
Colorado



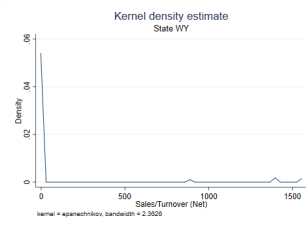
Idaho



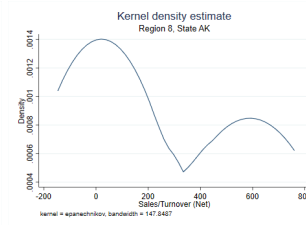
Montana



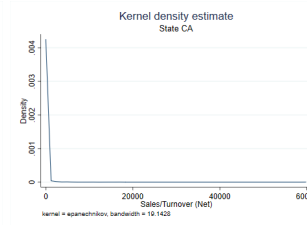
Utah



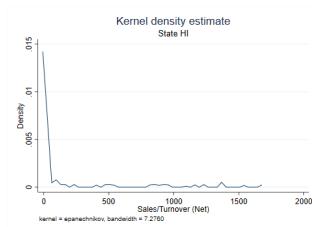
Wyoming



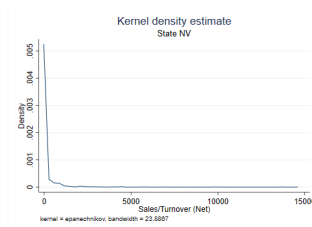
Alaska



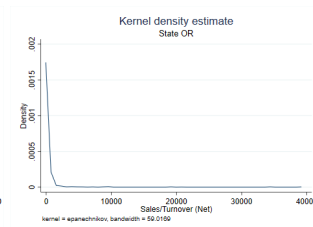
California



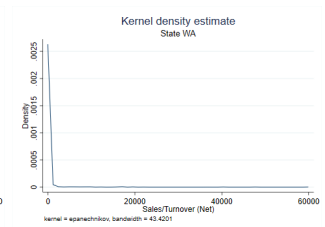
Hawaii



Nevada



Oregon



Washington

## C.2 Idiosyncratic shocks can predict future economic growth

This section reports the graphical relation between idiosyncratic shocks and subsequent annual economic growth for certain states.

Figure C.3: Relation between  $\Gamma_t$  and  $\Delta gdp_{t+1}$  for Selected States

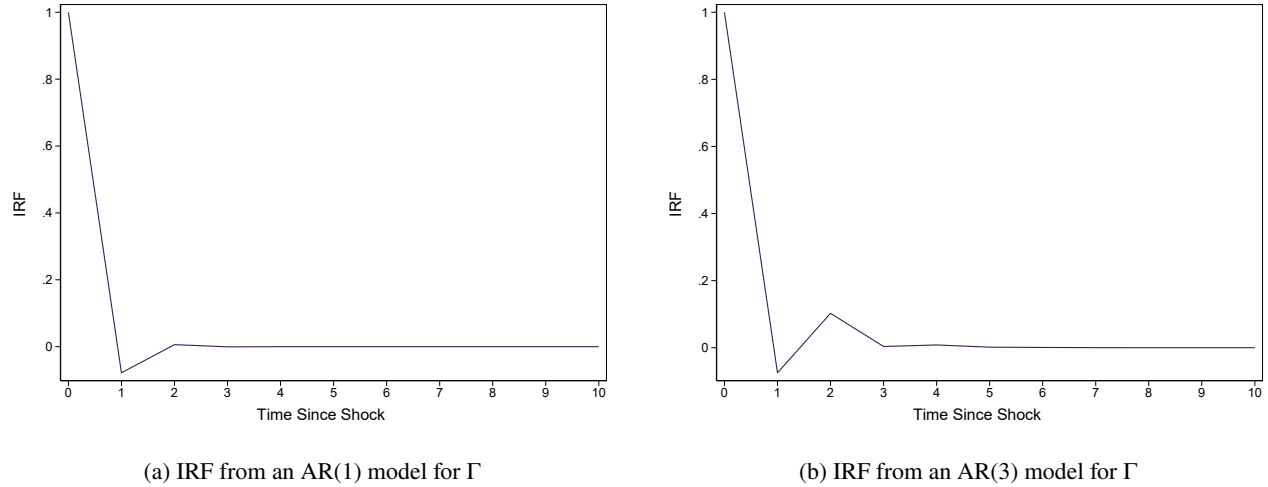


The figure plots the relation between idiosyncratic shocks  $\Gamma_t$ , not adjusting for annual industry shocks, and subsequent annual economic growth  $\Delta gdp_{t+1}$  for some selected states. All variables are standardized to mean 0 and variance 1. The sample period spans from 1977 to 2000.

### C.3 Persistence of Idiosyncratic shocks

This section reports the impulse response functions for idiosyncratic shocks obtained from a pooled AR(1) and an AR(2) model.

Figure C.4: Impulse Response Functions (IRF) for  $\Gamma$  from AR(p) models



The figure plots the impulse response functions from an AR(1) and an AR(3) model for  $\Gamma_t^{ind}$ . We estimate a panel VAR AR(p) model to estimate the impulse response functions.

### C.4 Bank Debt and Sample Firms

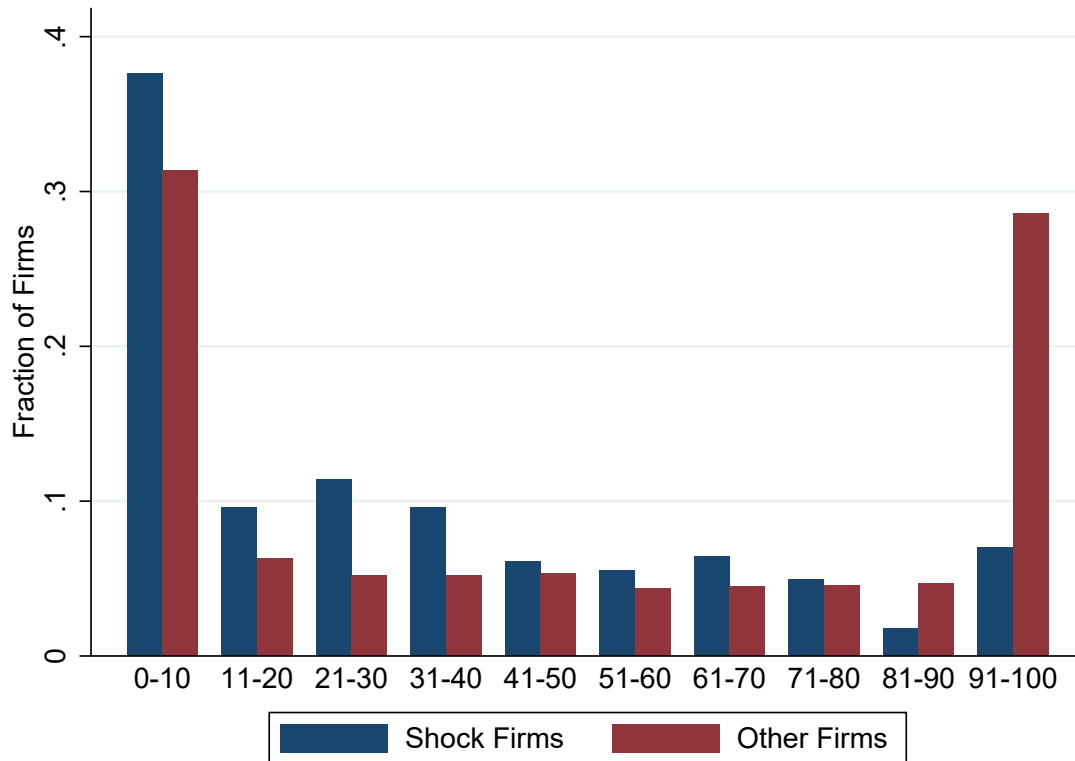
This section compares the ratio of bank debt to total debt for firms used to construct state level idiosyncratic shocks (shock firms) to all other firms in the S&P Capital IQ database. The data on bank debt and total debt comes from Capital IQ database. Due to data limitations we can only compare the bank debt to total debt ratio from 1989 onwards. Total debt is constructed by adding secured and unsecured debt for each firm. Table C.1 compares the mean and the median bank debt to total debt ratio for the shock firms and other firms. The median (mean) bank debt to total debt ratio for shock firms is 23.63% (30.35%), compared to a value of 44.63% (48.03%) for other firms. The mean and the median of bank debt to total debt ratio is lower for shocks firms by  $\approx 20$  pp relative to other firms. The t-statistic for the difference in the mean (median) bank debt to total debt for the two groups is 8.17 (4.17). This indicates that the shock firms are substantially less reliant on bank debt as source of external financing. We further validate this by examining the distribution of bank debt to total debt ratio across the two group of firms in Figure C.5.

Table C.1: Bank Debt to Total Debt - Shock Firms and Other Firms

|                            | Mean   | Median | St Dev |
|----------------------------|--------|--------|--------|
| Shock Firms                | 30.35  | 23.63  | 29.94  |
| Other Firms                | 48.03  | 44.63  | 40.09  |
| Difference                 | -17.68 | -20.99 |        |
| t-Statistic for Difference | 8.17   | 4.17   |        |

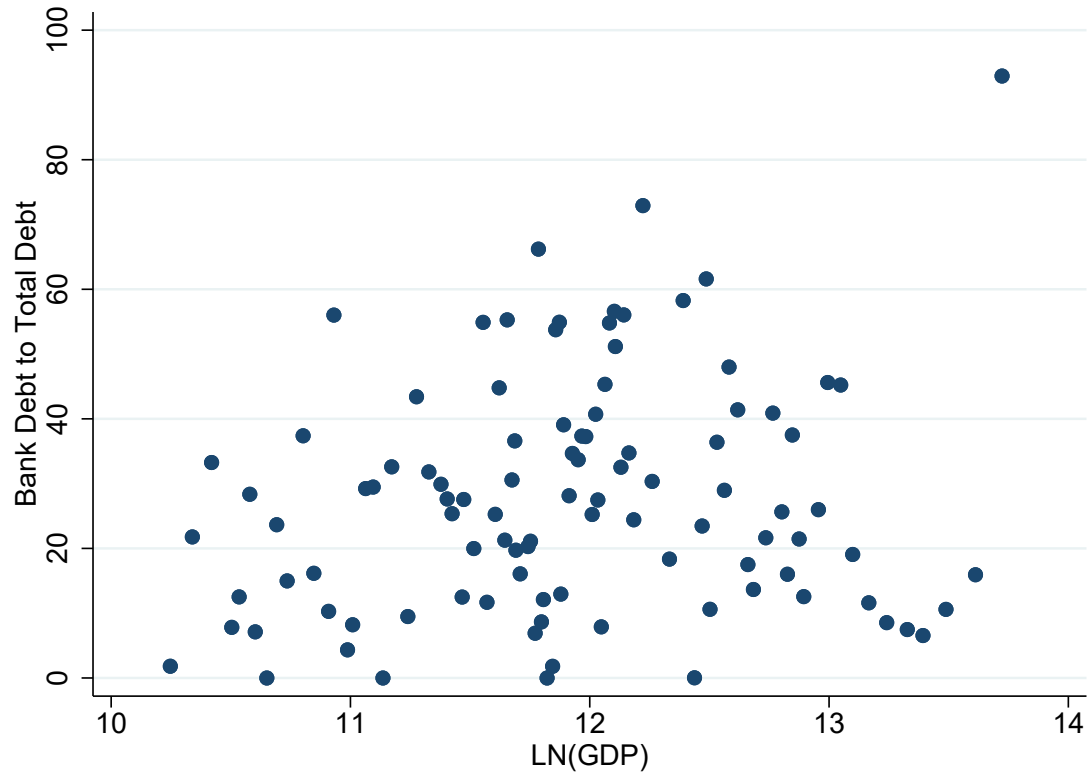
This table reports mean, median and the standard deviation for the bank debt to total debt ratio in percentage. Shock firms refer to the top 10 firms in each state used to construct state-level idiosyncratic shocks. All other firms not used to construct state-level idiosyncratic shocks are classified as other firms.

Figure C.5: Bank Debt to Total Debt - Shock Firms and Other Firms



The figure plots fraction for firms in each bin of bank debt to total debt ratio across the shock firms and other firms. The x-axis plots the bin for the total bank debt to total debt ratio. There are 10 bins, representing deciles of the ratio of bank debt to total debt. *Shock firms* refer to the top 10 firms in each state used to construct state-level idiosyncratic shocks. All other firms not used to construct state-level idiosyncratic shocks are classified as *other firms*.

Figure C.6: Bank Debt to Total Debt and Size of the State



The figure presents the scatter plot of bank debt to total ratio (y-axis) and the size of the economy (x-axis) for the shock firms. *Shock firms* refer to the top 10 firms in each state used to construct state-level idiosyncratic shocks. The size of the economy is measured using the natural logarithm of the nominal GDP of the state. The bank debt to total ratio on the X-axis is the average value of the bank debt to total ratio for the shock firms in the state.

## C.5 Effect of Idiosyncratic Shocks on Bank Capital

Table C.2: Effect of Idiosyncratic Shocks on Bank Capital

| <b>Panel A: Bank Constraint<sub><i>j,t</i></sub></b> | (1)                 | (2)                 | (3)                 |
|------------------------------------------------------|---------------------|---------------------|---------------------|
| $\Gamma_{j,t-1}^{ind}$                               | -0.0013<br>(0.0040) | -0.0005<br>(0.0032) | -0.0027<br>(0.0021) |
| Year FE                                              |                     | Yes                 | Yes                 |
| State FE                                             |                     |                     | Yes                 |
| <i>N</i>                                             | 1,154               | 1,154               | 1,154               |
| <i>R</i> <sup>2</sup>                                | 0.0000              | 0.3573              | 0.7714              |
| <b>Panel B: Bank Deposits<sub><i>j,t</i></sub></b>   | (1)                 | (2)                 | (3)                 |
| $\Gamma_{j,t-1}^{ind}$                               | -0.0115<br>(0.0165) | -0.0049<br>(0.0165) | 0.0157<br>(0.0119)  |
| Year FE                                              |                     | Yes                 | Yes                 |
| State FE                                             |                     |                     | Yes                 |
| <i>N</i>                                             | 1,154               | 1,154               | 1,154               |
| <i>R</i> <sup>2</sup>                                | 0.0002              | 0.1133              | 0.7578              |

This table reports the effect of idiosyncratic shocks on bank constraint (Panel A) and bank deposits (Panel B). The main independent variable is  $\Gamma_{j,t-1}^{ind}$  which denotes the idiosyncratic shocks in state *j* constructed by aggregating the Domar weighted labor productivity shocks of top 10 firms, by sales in state *j* after accounting for industry-year fixed effects. The unit of observation is at the state-year level. We account for the degree of bank constraint for each state by weighting a bank's level of constraint in each state by the bank's share of lending in that particular state. Bank constraint is measured as  $\frac{Liabilities}{Assets}$  in each year. We sum across these values at the state-year level to produce a measure of bank constraint at the state level for each year. Bank deposits are measured as  $\frac{Deposits}{Assets}$  in each year. We sum across these values at the state-year level to produce a measure of bank deposits at the state level for each year. All non-binary variables used in the regression are standardized to mean zero and variance one. Standard errors reported in parentheses are clustered by state *j*.

## Appendix D Framework for IV Design

This section describes the theoretical framework underlying our IV strategy for identifying the relation between bank lending and economic growth. We denote the growth rate in state  $i$  ( $j \neq i$ ) as  $g_i$ .  $g_i$  is a function of  $L_i$ , the loan supply in state  $i$ ,  $U_i$  and  $U_j$ , unobserved characteristics for each state in the state-pair  $(i, j)$ ,  $\phi_{i,j}^{NB}$ , denotes the integration of state-pair  $(i, j)$  via non-banking channels, and  $\epsilon_i$ , an idiosyncratic component in state  $i$ . The loan supply,  $L_i$  is a function of the  $g_i$  and  $g_j$ , growth rates for each state in the state-pair  $(i, j)$ ,  $\phi_{i,j}^B$ , denoting the banking integration of state-pair  $(i, j)$ ,  $V_i$  and  $V_j$  which denote unobserved characteristics for each state in the state-pair  $(i, j)$ , and  $\eta_i$ , an idiosyncratic component in state  $i$ . The growth rate,  $g_i$ , and the loan supply,  $L_i$ , are assumed to be as in equation D.1 and D.3 respectively yielding equation D.2 and D.4 under the assumption of separability.<sup>24</sup>

$$g_i = f(L_i, U_i, g_j, U_j, \phi_{i,j}^{NB}, \epsilon_i) \quad (D.1)$$

$$= f_1(L_i) + f_2(U_i, \epsilon_i) + f_3(g_j, \phi_{i,j}^{NB}, U_j) \quad (D.2)$$

$$L_i = h(g_i, g_j, \phi_{i,j}^B, V_i, V_j, \eta_i) \quad (D.3)$$

$$= h_1(g_i, \eta_i, V_i) + h_2(g_j, \phi_{i,j}^B, V_j) \quad (D.4)$$

This system of equations is plagued by a major source of endogeneity, namely, simultaneity bias, as both the growth rate and loan supply are jointly determined in equilibrium. We address this concern using an IV strategy. The loan supply is instrumented by  $\Gamma_j$ , idiosyncratic shocks to large firms in state  $j$ , and,  $\tilde{\phi}_{i,j}^B$ , exogenous shocks to the banking integration of state-pair  $(i, j)$ . Specifically, we assume that the instrument has the form:  $L_i = m[\Gamma_j, \tilde{\phi}_{i,j}^B] \equiv z_{i,j}$ . Assuming the validity of the exclusion restriction and relevance of the instrument, yields the moment condition,  $\mathbb{E}[\{g_i - f(L_i, 0, 0, 0, 0, 0)\}z_{i,j}] = 0$ . We project  $h_2(\cdot)$  using  $z_{i,j}$  onto  $f_1(\cdot)$  to identify the effect of loan supply shocks on economic growth. We instrument for bank loan supply in state  $i$ ,  $\log(l_{i,t})$ , with the interaction term of idiosyncratic shocks,  $\Gamma_{j,t-1}$ , in state  $j$ , and the timing of when state  $i$  permits banks in state  $j$  to branch within state  $i$ ,  $Post_{i,j,t}$ .<sup>25</sup> We estimate the effect of shocks to loan supply on economic growth via a two stage least square estimation (2SLS) in the following setup where equation D.5 and D.6 represent the first and the

<sup>24</sup>  $f$  is separable in,  $f_1$ ,  $f_2$ , and  $f_3$  which depend on observable characteristics in state  $i$  ( $L_i$ ), unobserved and idiosyncratic components in state  $i$  ( $U_i, \epsilon_i$ ), and state-partner ( $j$ ) components ( $g_j, \phi_{i,j}^{NB}, U_j$ ), as in equation D.2.  $h$  is separable in two functions,  $h_1$  and  $h_2$  which depend on state  $i$  characteristics ( $g_i, \eta_i, V_i$ ), and state-partner ( $j$ ) characteristics ( $g_j, \phi_{i,j}^B, V_j$ ) as in equation D.4

<sup>25</sup>  $\log(l_{i,t})$  refers to new commercial and industrial (C&I) loans given by all banks in state  $i$  during time  $t$ , capturing the flow of new loans.

second stage respectively.<sup>26</sup>

$$\log(l_{i,t}) = \alpha_2 + \beta_2 \Gamma_{j,t-1} \times Post_{i,j,t} + \beta_3 Post_{i,j,t} + \alpha_{i,j} + \theta_i \times t + \theta_{jt} + \varepsilon_{ijt} \quad (D.5)$$

$$\Delta gdp_{i,t} = \alpha_1 + \beta_1 \hat{\log}(l_{i,j,t}) + \alpha_{i,j} + \theta_i \times t + \theta_{jt} + \mu_{ijt} \quad (D.6)$$

We want to highlight that the structure of the first-stage equation mirrors our baseline specification, with equation D.5 estimated at the state-pair level. This yields  $N1$  equations per state and year ( $N$  being the total number of states). The predicted value of  $\log(l_{i,t})$  based on shocks in state  $j$  is then used in the second-stage specification. Consequently, the second stage is also estimated at the state-pair level, resulting in  $N1$  equations per state and year, with the predicted lending in state  $i$  based on shocks in state  $j$ , denoted as  $\hat{\log}(l_{i,j,t})$ , as the key variable.

## D.1 Violation of the Exclusion Restriction

Here, we discuss violations of the exclusion restriction in identifying the relation between bank lending and economic growth, and consider two counterfactual cases to assess how our point estimates may change. Our analysis suggests that the violation of the even weak identifying assumption biases our empirical strategy to estimate a magnitude of zero.

The Pre estimate reported in 2b indicates that, in aggregate, the relation between GDP growth in state  $i$  and idiosyncratic shocks in state  $j$  is weakly positive. Hence, the counterfactual cases capture incidents in which states behave as complements in the absence of banking linkages. The strong and weak forms of the exclusion restriction are as follows. The strong form of the exclusion restriction is that idiosyncratic productivity shocks in state  $j$  impact bank lending in state  $i$  strictly through loan supply, not loan demand. Even if the strong form does not hold, we can still identify the relation between bank lending and economic growth, as long as the covariance in loan demand between the two states is fixed around the deregulation shock, or that the covariance in loan demand between the two states is sticky relative to loan supply around the deregulation shock.

### Counterfactual #1:

Consider the case where states are linked by cross-state sales. If a firm in Virginia sells largely to consumers in Maryland and the state of Maryland experiences a large negative shock in a given year, consumption will fall in Maryland in that year. This means that the demand for the Virginian firm's goods will fall, which in turn, decreases total sales for that year. The decline in quantity suggests that the magnitude of our point estimates in Table 7 are downward biased.

### Counterfactual #2:

Consider the case where states are linked by input-output linkages. For illustration, suppose there is corporate law firm based in Connecticut and a corrupt firm in New York. The corrupt firm in New York

---

<sup>26</sup>Note that the estimation is run at the state-pair level. Therefore, for each pair we estimate the shocks to loan supply in state  $i$  coming from state  $j$  and use the projected loan supply from the first stage to estimate  $\beta_1$  in the second stage.

requires attorneys from the law firm in Connecticut to continue operating. If the law firm in Connecticut experiences a large negative shock, the corrupt firm in New York will suffer. In this case, the demand for the corrupt firm's goods will fall. Similar to the case above, the reduction in demand suggests that the magnitude of our point estimates in Table 7 are downward biased.

Another concern regarding linkages is the potential for positive shocks in one state to translate into negative shocks in another state, particularly if they share similar industry compositions. For instance, if oil-producing firms drive aggregate GDP growth in both Texas and Kansas, a positive shock in the oil sector in Texas (e.g., the discovery of new oil fields) could lead to a negative shock for the oil sector in Kansas. This could result in increased GDP growth in Texas, but decreased GDP growth in Kansas. This negative correlation due to industry similarity could influence our results. To address this concern, we control for industry similarity between the two states in Table 14 and find that our results remain robust despite this potential transmission of shocks.

In light of these considerations, one may question the validity of the exclusion restriction. However, our findings suggest that even if this strong identifying assumption is relaxed, the magnitude of our estimates is either downward biased or unlikely to substantially alter our primary results.

## D.2 Economic Magnitude of the Effect from 2SLS specification

This section presents the details of calculating the economic magnitude of the effect based on the 2SLS coefficients presented in Table 7.

$$\frac{\Delta y - \mathbb{E}(\Delta y)}{\sigma(\Delta y)} = \alpha + \beta \log x + \varepsilon \quad (\text{D.7})$$

Taking expectation of equation D.7 for  $x = x$  and  $x = x + dx$ , we get the following under the assumption  $\mathbb{E}[\varepsilon] = 0$ :

$$\mathbb{E} \left[ \frac{\Delta y - \mathbb{E}(\Delta y)}{\sigma(\Delta y)} \right] = \alpha + \beta \mathbb{E}[\log x] \quad (\text{D.8})$$

$$\mathbb{E} \left[ \frac{\Delta y' - \mathbb{E}(\Delta y)}{\sigma(\Delta y)} \right] = \alpha + \beta \mathbb{E}[\log(x + dx)] \quad (\text{D.9})$$

Subtracting equation D.9 from equation D.8 gives the following:

$$\begin{aligned}
\mathbb{E} \left[ \frac{\Delta y' - \mathbb{E}(\Delta y)}{\sigma(\Delta y)} \right] - \mathbb{E} \left[ \frac{\Delta y - \mathbb{E}(\Delta y)}{\sigma(\Delta y)} \right] &= \beta \{ \mathbb{E}[\log(x + dx)] - \mathbb{E}[\log x] \} \\
\mathbb{E} \left[ \frac{\Delta y' - \Delta y}{\sigma(\Delta y)} \right] &= \beta \left\{ \mathbb{E} \left[ \log \left( x \left( 1 + \frac{dx}{x} \right) \right) \right] - \mathbb{E}[\log x] \right\} \\
&= \beta \left\{ \mathbb{E}[\log(x)] + \mathbb{E} \left[ \log \left( 1 + \frac{dx}{x} \right) \right] - \mathbb{E}[\log x] \right\} \\
&= \beta \mathbb{E} \left[ \log \left( 1 + \frac{dx}{x} \right) \right] \\
&\approx \beta \mathbb{E} \left[ \frac{dx}{x} \right] \\
\mathbb{E} [\Delta y' - \Delta y] &\approx \beta \mathbb{E} \left[ \frac{dx}{x} \right] \times \sigma(\Delta y)
\end{aligned} \tag{D.10}$$

Put  $\sigma(\Delta y) = 0.03254$  (from Table 2) in equation D.10 we get the following change in  $\Delta y$  for 1 percentage point (pp) change in  $x$ :

$$\mathbb{E} [\Delta y' - \Delta y] = 0.03254 \times \beta$$

The following table provides the economic magnitude of the effect in percentage points for three values of  $\beta$  based on different specifications in Table 7:

| Coefficient Type                          | Strictest<br>Specification | Smallest<br>Magnitude | Largest<br>Magnitude |
|-------------------------------------------|----------------------------|-----------------------|----------------------|
| Source                                    | Column 8                   | Column 2              | Column 4             |
| $\sigma(\Delta y)$                        | 0.03                       | 0.03                  | 0.03                 |
| $\beta$                                   | 4.14                       | 1.88                  | 7.60                 |
| <b>Economic Magnitude of Effect in pp</b> | <b>0.13</b>                | <b>0.06</b>           | <b>0.25</b>          |

Table D.1: OLS Regression

| $\Delta gdp_{it}$                                       | (1)                 | (2)                | (3)                 |
|---------------------------------------------------------|---------------------|--------------------|---------------------|
| $\log(C\&I - Loan_{i,t})$                               | -0.0086<br>(0.0347) | 0.0741<br>(0.0759) | 0.0695<br>(0.0813)  |
| $Post_{i,j,t}$                                          | 0.2203<br>(0.0794)  | 0.0363<br>(0.0826) | 0.0489<br>(0.0709)  |
| $\Gamma_{i,t-1}^{ind}$                                  |                     |                    | 0.0025<br>(0.0171)  |
| $\Gamma_{i,t-2}^{ind}$                                  |                     |                    | -0.0048<br>(0.0145) |
| Constant                                                | 0.0463<br>(0.5895)  |                    |                     |
| Region <sub><i>i</i></sub> -Year FE                     |                     | Yes                | Yes                 |
| State <sub><i>i</i></sub> -State <sub><i>j</i></sub> FE |                     | Yes                | Yes                 |
| State <sub><i>j</i></sub> -Year FE                      |                     | Yes                | Yes                 |
| State <sub><i>i</i></sub> -Linear Trend                 |                     | Yes                | Yes                 |
| $N$                                                     | 50,950              | 50,950             | 50,563              |
| $R^2$                                                   | 0.0117              | 0.6943             | 0.7145              |

This table presents the estimates from the OLS regression of GDP growth in state  $i$  on the natural logarithm of lending in state  $i$  while controlling for banking deregulation between state  $i$  and state  $j$ .  $\Gamma_{j,t-1}^{ind}$  denotes the idiosyncratic shocks in state  $j$  constructed by aggregating the Domar weighted labor productivity shocks of top 10 firms, by sales in state  $j$  after accounting for industry-year fixed effects. The unit of observation in each regression is a state<sub>*i*</sub>-state<sub>*j*</sub>-year pair. Observations are weighted by the share of exports from state  $i$  to state  $j$ , using the 1977 Commodity Flow Survey Data. All non-binary variables except  $\log(C\&I - Loan_{i,t})$ , used in the regression are standardized to mean zero and variance 1. Standard errors reported in parentheses are two-way clustered by state<sub>*i*</sub> and state<sub>*j*</sub>.

## Appendix E Additional Results & Discussion

This section reports additional results and discussion that either support or add credibility to the main results in the paper. We refer the readers to these results in the paper wherein required. The additional results do not substantially add to the results reported in the paper but as outlined, add credibility to the results.

### E.1 Baseline Results

Table E.1: Robustness: Collapsed Version of the Specification

| $\Delta gdp_{it}$                                                                           | (1)                 | (2)                 | (3)                 | (4)                 |
|---------------------------------------------------------------------------------------------|---------------------|---------------------|---------------------|---------------------|
| $\frac{\sum_j Post_{i,j,t}}{N-1} \times \frac{\sum_{j,j \neq i} \Gamma_{j,t-1}^{ind}}{N-1}$ | -0.1928<br>(0.0505) | -0.1978<br>(0.0522) | -0.1335<br>(0.0540) | -0.1285<br>(0.0546) |
| $\frac{\sum_{j,j \neq i} \Gamma_{j,t-1}^{ind}}{N-1}$                                        | 0.0667<br>(0.0257)  | 0.0675<br>(0.0258)  | 0.0003<br>(0.1879)  | 0.0391<br>(0.2009)  |
| $\frac{\sum_j Post_{i,j,t}}{N-1}$                                                           | 0.3062<br>(0.0769)  | 0.3251<br>(0.0822)  | 0.2161<br>(0.1264)  | 0.2038<br>(0.1240)  |
| State <sub><i>i</i></sub> FE                                                                |                     | Yes                 | Yes                 | Yes                 |
| Region <sub><i>i</i></sub> -Year FE                                                         |                     |                     | Yes                 | Yes                 |
| State <sub><i>i</i></sub> -Linear Trend                                                     |                     |                     |                     | Yes                 |
| N                                                                                           | 1,173               | 1,173               | 1,173               | 1,173               |
| R <sup>2</sup>                                                                              | 0.0252              | 0.1405              | 0.6119              | 0.6586              |

This table reports the results from collapsing the baseline specification at the state<sub>*i*</sub> and year level. The dependent variable is the change in the real GDP growth rate in percentage.

The main independent variable is the interaction term of  $\frac{\sum_{j,j \neq i} \Gamma_{j,t-1}^{ind}}{N-1}$  and  $\frac{\sum_j Post_{i,j,t}}{N-1}$ .

$\frac{\sum_{j,j \neq i} \Gamma_{j,t-1}^{ind}}{N-1}$  denotes the average of idiosyncratic shocks in all state *j*, where *j* ≠ *i*. The state-level idiosyncratic shocks are constructed by aggregating the Domar weighted labor productivity shocks of the top 10 firms, by sales in state *j* after accounting for industry-year fixed effects.  $\frac{\sum_j Post_{i,j,t}}{N-1}$  refers to the index of deregulation for state<sub>*i*</sub> at time *t*, and is calculated as the fraction of states with which state<sub>*i*</sub> has deregulated banking. The unit of observation in each regression is at the state<sub>*i*</sub>-year level. All non-binary variables used in the regression are standardized to mean zero and variance 1. Standard errors reported in parentheses are clustered by state<sub>*i*</sub>.

Table E.2: Rolling GDP Growth Correlation and Deregulation

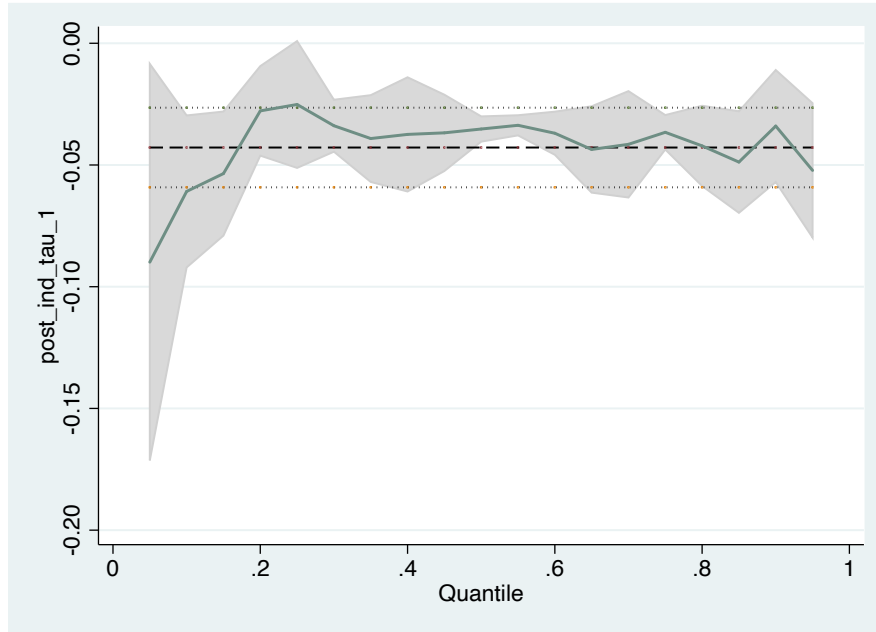
| $Corr(\Delta gdp_{it}, \Delta gdp_{jt})$  | (1)                 | (2)                 | (3)                 | (4)                 | (5)                 | (6)                 |
|-------------------------------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
|                                           | 5 years             | 6 years             | 7 years             | 8 years             | 9 years             | 10 years            |
| $Post_{i,j,t}$                            | -0.0276<br>(0.0069) | -0.0275<br>(0.0099) | -0.0253<br>(0.0103) | -0.0206<br>(0.0096) | -0.0177<br>(0.0076) | -0.0149<br>(0.0062) |
| State <sub>i</sub> -State <sub>j</sub> FE | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 |
| State <sub>j</sub> -Year FE               | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 |
| State <sub>i</sub> -Year FE               | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 |
| N                                         | 55,300              | 55,300              | 55,300              | 55,300              | 55,300              | 55,300              |
| $R^2$                                     | 0.6882              | 0.7293              | 0.7584              | 0.7818              | 0.8000              | 0.8160              |

This table reports the results from the estimation of the following regression specification:

$$Corr(\Delta gdp_{it}, \Delta gdp_{jt}) = \beta_0 Post_{i,j,t} + \alpha_{i,j} + \theta_{it} + \theta_{jt} + \varepsilon_{ijt}, i \neq j$$

The dependent variable is the rolling GDP growth correlation. The rolling correlation with GDP is calculated over a window of 5-10 years, with the specific window size indicated in columns 1-6, respectively. The unit of observation in each regression is at the state<sub>i</sub>-state<sub>j</sub>-year level. Standard errors reported in parentheses are two-way clustered by state<sub>i</sub> and state<sub>j</sub>.

Figure E.1: Point Estimate Difference between Pre & Post Period in Figure 2b: OLS & Quantile Regression Estimates



The figure plots the point estimate for the difference in the relation between GDP growth in state  $i$  and idiosyncratic shocks in state  $j$  where  $i \neq j$  in the pre and post deregulation period. *Pre* refers to a sample of all state-pairs before banking integration. *Post* refers to a sample of all state-pairs after banking integration as in Figure 2b. The dashed red line reports the OLS estimate with 95% confidence interval and the blue line reports the estimate obtained from the quantile regression for different quantile of  $\Delta GDP$  along with the 95% confidence interval in grey.

## E.2 Weighted Regression Results

Table E.3 reports the results from a weighted estimation. We compute the share of exports from state  $i$  to state  $j$ , and the share of imports coming from state  $j$  to state  $i$  using the 1977 Commodity Flow Survey Data. The share measures the magnitude and the direction of real linkages from  $i$  to  $j$ . Columns (2) and (3) weight each observation by the share of exports and imports respectively. We also report the equal-weighted regression for comparison in column (1). The estimates in column (2) and (3) are negative and statistically significant – similar to column (1). In terms of magnitude, a one standard deviation  $\Gamma_{j,t-1}$  shock increases economic growth in state  $i$  by 0.13-0.19 pp post banking integration. This estimate is larger than the baseline estimate of 0.05 pp. Hence, by accounting for the strength of banking linkages using non-banking linkages, we find a larger effect of idiosyncratic shocks in state  $i$  on economic growth than in state  $j$  post banking integration.

Table E.3: Weighted Estimation (Weighted by Exports/Imports)

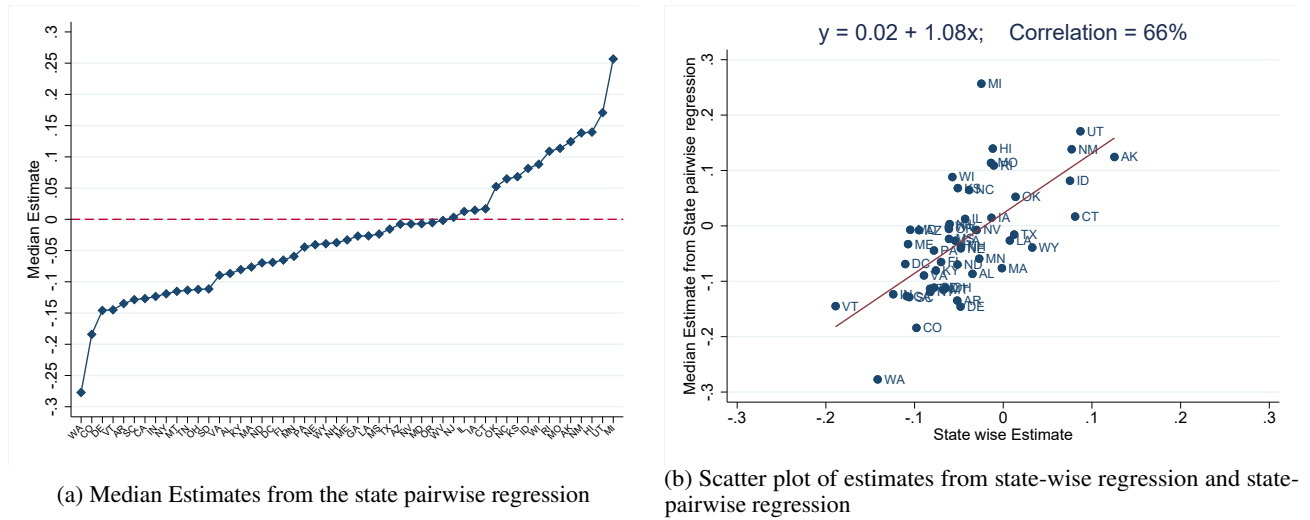
| $\Delta gdp_{it}$                          | (1)                 | (2)                 | (3)                 |
|--------------------------------------------|---------------------|---------------------|---------------------|
| $Post_{i,j,t} \times \Gamma_{j,t-1}^{ind}$ | -0.0164<br>(0.0007) | -0.0583<br>(0.0244) | -0.0397<br>(0.0156) |
| $Post_{i,j,t}$                             | 0.0783<br>(0.0491)  | 0.0452<br>(0.0767)  | 0.0816<br>(0.0603)  |
| Region $_i$ -Year FE                       | Yes                 | Yes                 | Yes                 |
| State $_i$ -State $_j$ FE                  | Yes                 | Yes                 | Yes                 |
| State $_j$ -Year FE                        | Yes                 | Yes                 | Yes                 |
| State $_i$ -Linear Trend                   | Yes                 | Yes                 | Yes                 |
| N                                          | 57,700              | 50,838              | 51,312              |
| $R^2$                                      | 0.6583              | 0.6946              | 0.6646              |
| Weights                                    | Equal               | Export ('77)        | Import ('77)        |

This table reports the results from the estimation of baseline specification where each observation is weighted by the strength of real linkages. The dependent variable is the change in the real GDP growth rate in percentage. The main independent variable is  $\Gamma_{j,t-1}^{ind}$  which denotes the idiosyncratic shocks in state  $j$  constructed by aggregating the Domar weighted labor productivity shocks of top 10 firms, by sales in state  $j$  after accounting for industry-year fixed effects. The regression is weighted by exports and imports. Column (1) presents the baseline regression result of Table 3. Column (2) presents the baseline regression weighted by exports. Column (3) presents the baseline regression weighted by imports. We compute the share of exports going from state  $i$  to state  $j$ , and the share of imports coming from state  $j$  to state  $i$  using the 1977 Commodity Flow Survey Data. The share measures the magnitude and the direction of real linkages from  $i$  to each  $j$ . Each observation in column (2) and (3) is weighted by share of exports and imports respectively. The unit of observation in each regression is a state $_i$ -state $_j$ -year pair. All non-binary variables used in the regression are standardized to mean zero and variance 1. Standard errors reported in parentheses are two-way clustered by state $_i$  and state $_j$ .

### E.3 Heterogeneous Treatment Effects

Furthermore, to ensure that the estimates are not driven by extreme values, Figure E.2a plots the state-level median estimate obtained from state-pairwise regression. We run the baseline regression at state-pair level and estimate the coefficient of the interaction term. The mean (median) value of the median estimate is -0.025 (-0.033) with ~69% of state-level estimates being strictly negative. The state-level estimates from this exercise are qualitatively similar to the ones obtained from the state-wise estimation reported in Figure 4. In fact, the correlation between these state-level estimates and the estimates described earlier is 66%, see Figure E.2b.

Figure E.2: Median Estimates from the State Pairwise Regression & its Correlation with Estimates from the State-wise Regression in Figure 4



The figure E.2a plots the median value of the estimates obtained from state-pairwise regression of economic growth in state  $i$  on the interaction term of post and idiosyncratic shocks in state  $j$ , the level terms of both post and idiosyncratic shocks in state  $j$ . For each state we take the median value of all the state-pairwise estimates and plot them in increasing order. The figure E.2b plots the relation between the two state-level estimates. The figure plots the median value of the estimates obtained from state-pairwise regression of economic growth in state  $i$  on the interaction term of post and idiosyncratic shocks in state  $j$ , the level terms of both post and idiosyncratic shocks in state  $j$ . For each state we take the median value of all the state-pairwise estimates and plot them in increasing order. This is plotted along the Y-axis. The state-level estimates obtained from the state-pairwise regression are plotted along the X-axis.

## E.4 What Explains the Heterogeneity in State-Level Estimates?

In this section we discuss reasons for heterogeneity in the state-level estimates. We attempt to explain this heterogeneity using two key variables - (1) the median timing of deregulation, i.e., early versus late-deregulation states, and (2) the degree of penetration by out-of-state banks.

We analyze the growth in the share of gross domestic assets owned by out-of-state MBHCs for early and late deregulators based on the median deregulation year for each state. We define all states with a median deregulation year before 1991 as *early deregulation states* and all other states as *late deregulation states*.<sup>27</sup> Figure E.3 shows that the average share of gross domestic assets owned by out-of-state MBHCs grew steadily from 6% in 1979 to 47% in 1994 for early deregulation states, whereas the average share of gross domestic assets owned by out-of-state MBHCs grew modestly from 7% in 1979 to 29% in 1994 for late-deregulation states. The heterogeneity in the banking response by late and early deregulators has earlier been documented by Mian, Sufi and Verner (2020) for inter- and intra-state banking deregulation. Here, we document a similar heterogeneity for state-pairwise banking deregulation.

The findings discussed in the previous paragraph suggest that the majority of out-of-state banking expansion occurred in early deregulation states. Assuming that changes in banking expansion flow from changes in banking regulation, we hypothesize that the negative and larger magnitude  $\beta$  estimates from the baseline regression are from states with earlier dates of regulation. Figure E.4a reports the scatter plot of state-level estimates and median deregulation year. Consistent with our hypothesis we find that the state-level estimate decreases and approaches zero as the median deregulation year increases. Exploring this issue further, Figure E.4b plots the scatter plot of state-level estimates with the change in share of gross domestic assets owned by out-of-state MBHCs. As expected, the best fit line is downward sloping, indicating that the large negative state-level estimates are correlated with states that experienced the largest growth in share of gross domestic assets owned by out-of-state banks.

Table E.4 reports the results from the regression of state-level coefficients on median deregulation year and a quadratic function of the change in the share of gross domestic assets owned by out-of-state MBHCs from 1979 through 1994. The median deregulation year explains around 13% of the variation in the state-level estimate. Moreover, the positive sign of the point estimate indicates a one year increase in the median deregulation year, increases the state-level point estimate by 0.008.<sup>28</sup> Hence, states that deregulated later are associated with greater state-level estimates,  $\beta$ . This estimate is statistically significant and relevant as the point estimate is 0.12 times the standard deviation of the state-level estimates discussed in Section 4.3.1. The quadratic function of the share of gross domestic assets owned by out-of-state MBHCs over the years 1979 and 1994 explains roughly 20% of the variation in the estimate. While the linear term is insignificant, the squared term is statistically significant at the 1% level and enters the regression with the expected negative sign. An increase in the change in out-of-state banking asset

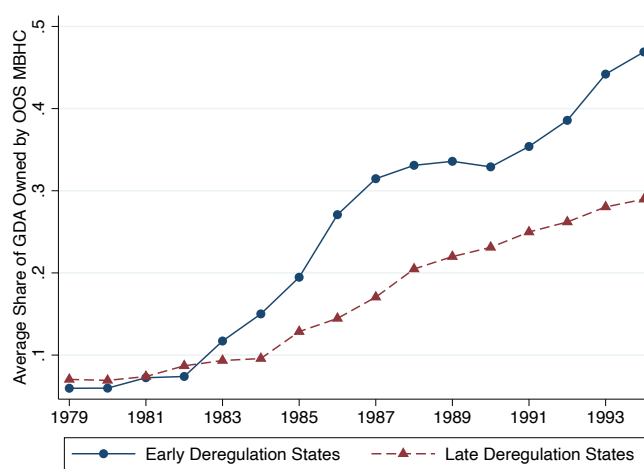
---

<sup>27</sup>1991 is the median value for all states.

<sup>28</sup> $\Delta\beta_s = 0.1237 \times \sigma_{\beta_s} = 0.1237 \times 0.061 = 0.008$ .

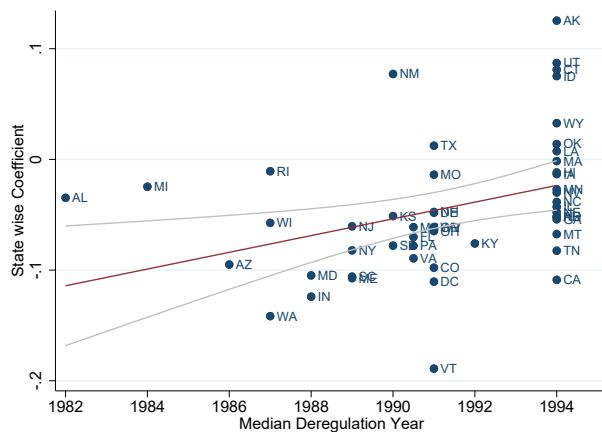
share decreases the point estimate of the coefficient. Taken together, the median deregulation year and change in out-of-state banking assets explain ~25% of variation in the state-level estimates.

Figure E.3: Out-of-State Banking Expansion in Early and Late-Deregulation States

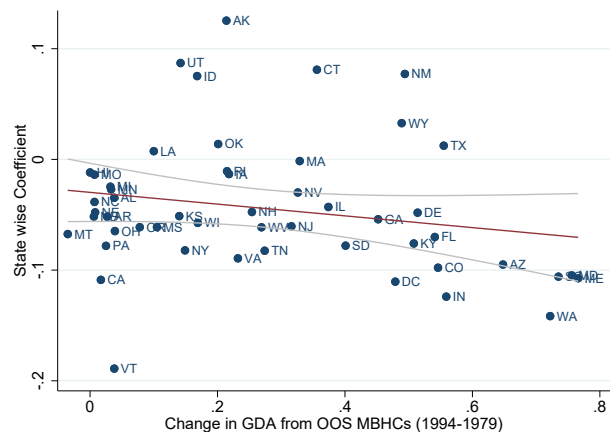


The figure plots the average share of gross domestic banking assets owned by out-of-state MBHCs across early and late-deregulation states. Data on share of gross domestic assets owned by out-of-state MBHCs comes from [Berger, Kashyap and Scalise \(1995a\)](#). Early deregulation states are defined as states that deregulated banking restrictions with at least 50% of other states before 1991, and late-deregulation states are states that deregulated with at least 50% of other states on or after 1991.

Figure E.4: State-level Estimate, Timing of Deregulation and Out-of-State Banking Penetration



(a) Median Deregulation Year



(b) Change in Gross Domestic Assets Owned by out-of-state MBHCs (1994-1979)

The figure plots the relation between the state-level estimated presented in Figure 4 and the median year of deregulation (Figure E.4a) and the change in the share of gross domestic assets owned by out-of-state MBHCs (Figure E.4b). The median year of deregulation is set equal to the year when the state has deregulated with at least 50% of other states. Data on share of gross domestic assets owned by out-of-state MBHCs comes from [Berger, Kashyap and Scalise \(1995a\)](#). The change in the share of gross domestic assets owned by out-of-state MBHCs is computed over the years 1979 and 1994.

Table E.4: State-level Estimates, Deregulation Timing and Out-of-State Banking Expansion

| Dep Var: State-level Estimate | (1)                | (2)                 | (3)                 |
|-------------------------------|--------------------|---------------------|---------------------|
| Median Deregulation Year      | 0.1237<br>(0.0507) |                     | 0.0846<br>(0.0456)  |
| $\Delta$ Asset                |                    | 0.0197<br>(0.1316)  | 0.0532<br>(0.1411)  |
| $\Delta$ Asset <sup>2</sup>   |                    | -0.4271<br>(0.1122) | -0.3639<br>(0.1019) |
| N                             | 51                 | 51                  | 51                  |
| R <sup>2</sup>                | 13.11%             | 19.39%              | 24.74%              |

The table reports the regression of state-level estimates on median deregulation year,  $\Delta$  Asset, and  $\Delta$  Asset<sup>2</sup>. The state-level estimates are constructed by running the baseline specification for each state  $i$  separately. The median year of deregulation is set equal to the year when the state has deregulated with at least 50% of other states. Data on share of gross domestic banking assets owned by out-of-state MBHCs comes from [Berger, Kashyap and Scalise \(1995a\)](#).  $\Delta$  Asset measures the change in the share of gross domestic assets owned by out-of-state MBHCs is computed over the years 1979 and 1994. All non-binary variables used in the regression are standardized to mean zero and variance 1. Robust standard errors reported in parentheses.

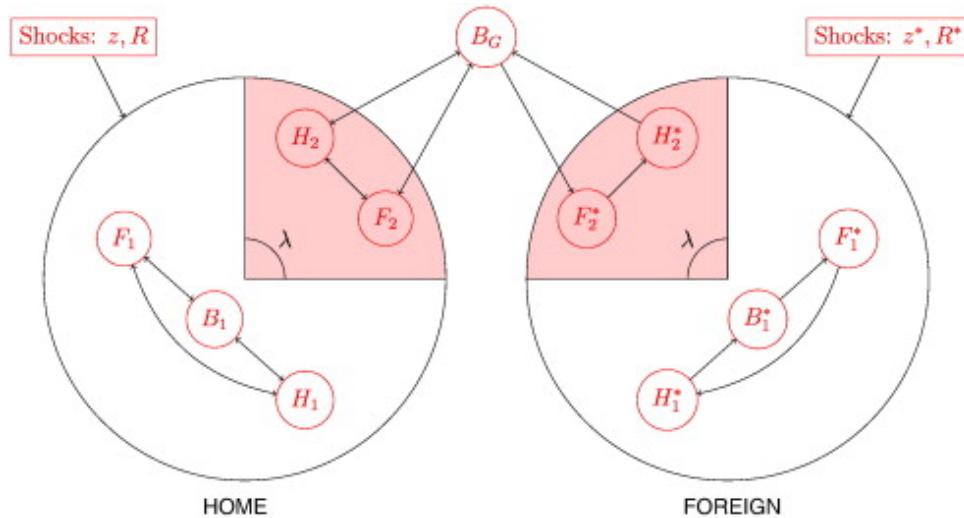
## Appendix F Theoretical Model

In this section, we outline the model of [Kalemli-Ozcan, Papaioannou and Perri \(2013b\)](#), and replicate their key theoretical finding.

### F.1 Setup

This is a model of international business cycles with banks. There are two countries, e.g., *home* and *foreign* (distinguished by superscript \*), each with two segments with size  $\lambda$  and  $1 - \lambda$  respectively. The  $\lambda$  segments (segment 2) of each country are financially integrated, while the  $1 - \lambda$  segments are financially separate (segment 1), i.e., a  $1 - \lambda$  share of the domestic and foreign economies operate in autarky so that banks intermediate only between households and firms in that  $1 - \lambda$  segment, respectively. In each segment of each country, there are households which supply labor to firms and save with banks. Firms pay dividends and wages to the households, and make investment decisions. In addition, firms borrow from banks. Banks in segments 2 of each country are *global banks* as  $\lambda$  share of each economy is financially integrated. For illustration of the schema of the economy in the model, we reproduce below The figure 1 from [Kalemli-Ozcan, Papaioannou and Perri \(2013b\)](#). The model focuses on two types of shocks that

Figure F.1: The structure of the economy



Source: This figure is taken from [Kalemli-Ozcan, Papaioannou and Perri \(2013b\)](#)

drive economic fluctuations: a standard productivity shock, and banking shocks that affect the value of risky assets held by banks. In particular, we use the model to study how exogenous changes to financial integration affect output correlation, cross-border transmission of shocks, and synchronization of the business cycle.

### F.1.1 Households

In each segment  $i$  of each country, there is a continuum of identical, infinitely-lived household with preferences:

$$E_0 \sum_{t=0}^{\infty} \beta^t U(c_{it}, l_{it})$$

where  $c_{it}$  denotes consumption and  $l_{it}$  denotes labor,  $\beta \in (0, 1)$  is the discount factor,  $E_0$  denotes expectation at date 0 across time and possible states of the world. Utility is subject to the following budget constraint:

$$c_{it} + \frac{D_{it+1}}{R_{it}} = w_{it}l_{it} + d_{it} + D_{it}$$

where  $D_{it}$  denotes the amount of bank deposits that are carried over,  $w_{it}$  is the wage rate,  $d_{it}$  are firms' dividends, and  $R_{it}$  is the gross rate of return of bank deposits. The consumers' problem is to choose  $c_{it}$ ,  $l_{it}$  and  $D_{it}$ . Consumers in segment 2 can shop for banks in both countries, so by arbitrage deposit rate is the same in segment 2 of both the countries:

$$R_{2t} = R_{2t}^* \forall t$$

### F.1.2 Firms

Firms operate a technology  $F$  that uses capital,  $k_{it}$  and labor  $l_{it}$  to produce a good. Production is subject to stochastic, country specific, productivity shocks  $z_t$  and  $z_t^*$ . It is assumed that firms need to pay workers before they realize sales, hence, firms must borrow from the bank working capital that is equal to the wage bill. Firms in segment  $i$  pay gross lending rate  $R_{it}^e$  on bank loans

$$\begin{aligned} d_{it} &= e^{z_t} F(k_{it}, l_{it}) - R_{it}^e w_{it} l_{it} - x_{it} \\ k_{it+1} &= (1 - \delta)k_{it} + x_{it} - \phi k_{it} \left[ \frac{x_{it}}{k_{it}} - \delta \right]^2 \\ \begin{bmatrix} z_t \\ z_t^* \end{bmatrix} &= A_z \begin{bmatrix} z_{t-1} \\ z_{t-1}^* \end{bmatrix} + \begin{bmatrix} \epsilon_t^z \\ \epsilon_t^{z*} \end{bmatrix} \end{aligned}$$

where  $R_{it}^e$  is a gross lending rate on bank loans,  $x_{it}$  is the investment in physical capital,  $\delta$  is the depreciation rate,  $\phi$  represents capital adjustment costs. In terms of the shock process,  $A_z$  is a  $2 \times 2$  matrix and  $[\epsilon_t^z, \epsilon_t^{z*}]$  is a vector of iid innovations with mean 0, standard deviation  $\sigma_\epsilon^z$  and correlation  $\rho_\epsilon^z$ . The

firms' problem in each country and segment  $i$

$$\max_{l_{it}, k_{it}, x_{it}} E \sum_{t=0}^{\infty} d_{it} Q_{it}$$

where  $Q_{it} = \beta_t U_c(c_{it}, l_{it})$  – the MRS of domestic consumers (owners of firm) which is the stochastic discount factor. Moreover, in the financially integrated segment, firms can shop for banks, therefore:

$$R_{2t}^e = R_{2t}^{e*}$$

## F.2 Banks

Banks operating in segmented areas raise deposits  $\frac{D_{1t+1}}{R_{1t}}$  and  $\frac{D_{1t+1}^*}{R_{1t}^*}$  respectively from consumers in home and foreign areas. Global banks' deposits are given by  $\frac{D_{2t+1} + D_{2t+1}^*}{R_{2t}}$ . Further, it is assumed that deposit-raising is costly, therefore banks need to pay  $\iota$  of deposits that represents a gamut of forces (intermediation cost/term spread/net interest margin).

In this economy, banks have the option of extending loans to firms, which are considered to be *risk-free* loans, or investing in *risky* technology. Banks in segment 1 only lend to firms in that segment/country and only invest in risky tech of that country. Banks in segment 2 can lend to firms in both countries and invest in a diversified international fund with equal shares of risky tech of both countries

In addition, banks experience stochastic gross returns on risky tech in the two countries (equal mean in each country),  $R_t^m$  and  $R_t^{m*}$ .

- Credit shocks follow a bivariate auto-regressive process

$$\begin{bmatrix} R_t^m \\ R_t^{m*} \end{bmatrix} = \begin{bmatrix} \bar{R}^m \\ \bar{R}^{m*} \end{bmatrix} + A_R \begin{bmatrix} R_{t-1}^m \\ R_{t-1}^{m*} \end{bmatrix} + \begin{bmatrix} \epsilon_t^R \\ \epsilon_t^{R*} \end{bmatrix}$$

where  $A_R$  is a  $2 \times 2$  matrix and  $[\epsilon_t^R, \epsilon_t^{R*}]$  is a vector of iid innovations with mean 0, standard deviation  $\sigma_\epsilon^R$  and correlation  $\rho_\epsilon^R$ .

First, banks decide how much to invest in the risky asset without knowing the realization of returns  $R_t^m$  and  $R_t^{m*}$ . It is assumed in the model that the expected return on risky asset is high enough, so each bank invests maximum share of deposits allowed by regulation, i.e.,  $0 < \bar{m} < 1$ . After returns  $R_t^m$  and  $R_t^{m*}$  are observed but not cashed, banks offer competing loans to firms. Because firms borrow enough working capital to finance their wage bill, the equilibrium amount of loans in the economy is given by:

$$\begin{aligned} L_{1t} &= w_{1t} l_{1t}; & L_{1t}^* &= w_{1t}^* l_{1t}^* \\ L_{2t} &= w_{2t} l_{2t}; & L_{2t}^* &= w_{2t}^* l_{2t}^* \end{aligned}$$

At the end of period, banks receive proceeds from lending and risky investments, and pay back deposit and interest to consumers, as well as margin costs,  $\iota$ .

### F.2.1 Solving the model

The equilibrium conditions from solving the model are as follows:

#### Equilibrium: Consumers and firms

Consumers and firms solve problems given prices and shocks. Banks invest a share  $\bar{m}$  in risky portfolio and make zero profits in each segment  $\forall t$ :

$$\begin{aligned}\bar{m}R_{1t}^m + (1 - \bar{m})R_{1t}^e &= R_{1t} + \iota \\ \bar{m}R_{1t}^{m*} + (1 - \bar{m})R_{1t}^{e*} &= R_{1t}^* + \iota \\ \bar{m}(\frac{1}{2}R_{2t}^m + \frac{1}{2}R_{2t}^{m*}) + (1 - \bar{m})R_{2t}^e &= R_{2t} + \iota\end{aligned}$$

Revenues per unit of deposit from risky capital and lending = Cost for bank

#### Equilibrium: Goods market clearing

Investment in banking deposits, physical capital, and consumption are equal to production and resources generated by risky tech, net of margin costs  $\forall t$

$$\begin{aligned}c_{1t} + x_{1t} + (D_{1t+1} - D_{1t}) &= e^{z_t} F(k_{1t}, l_{1t}) + \frac{D_{1t+1}}{R_{1t}} (\bar{m}(R_t^m - 1) - \iota) \\ c_{1t}^* + x_{1t}^* + (D_{1t+1}^* - D_{1t}^*) &= e^{z_t^*} F(k_{1t}^*, l_{1t}^*) + \frac{D_{1t+1}^*}{R_{1t}} (\bar{m}(R_t^{m*} - 1) - \iota) \\ c_{2t} + c_{2t}^* + x_{2t} + x_{2t}^* + (D_{2t+1} - D_{2t})(D_{2t+1}^* - D_{2t}^*) &= \\ = e^{z_t} F(k_{2t}, l_{2t}) + e^{z_t^*} F(k_{2t}^*, l_{2t}^*) + \frac{D_{2t+1} + D_{2t+1}^*}{R_{2t}} (\frac{\bar{m}}{2}(R_t^m + R_t^{m*} - 2) - \iota)\end{aligned}$$

#### Equilibrium: Financial intermediation market clearing

Demand for working capital from firms in the segment equals supply of loans in that segment (fraction invested in risk-free  $\times$  total deposits)  $\forall t$ .

$$\begin{aligned}L_{1t} &= (1 - \bar{m}) \left( \frac{D_{1t+1}}{R_{1t}} \right) \\ L_{1t}^* &= (1 - \bar{m}) \left( \frac{D_{1t+1}^*}{R_{1t}} \right) \\ L_{2t} + L_{2t}^* &= (1 - \bar{m}) \frac{(D_{2t} + D_{2t}^*)}{R_{2t}}\end{aligned}$$

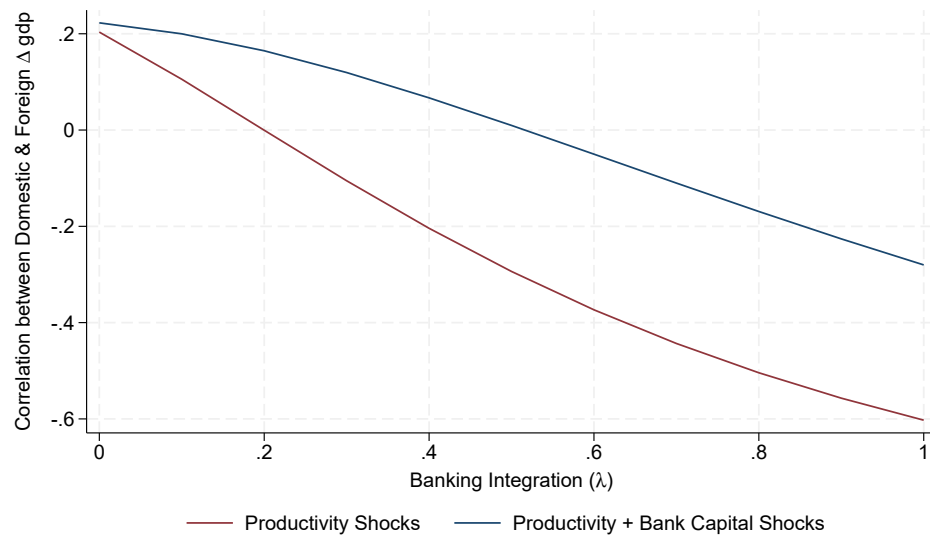
## F.3 Parameterization and Theoretical Findings

### Functional forms and baseline parameter values

- **Utility:**  $U(c, l) = \log(c) - \lambda l$
- **Production:**  $F(k, l) = k^\alpha l^{1-\alpha}$
- **Weight of labor:** 2.3
- **Capital share:**  $\alpha = 0.36$
- **Depreciation rate:**  $\delta = 0.075$
- **Productivity process:**  $A_Z = \begin{bmatrix} 0.95 & 0 \\ 0 & 0.95 \end{bmatrix}$ ;  $\rho_\epsilon^z = 0.2$ ,  $\sigma_\epsilon^z = 0.70\%$  (productivity only);  $\sigma_\epsilon^z = 0.485\%$  (productivity and credit)
- **Adjustment cost:**  $\phi = 0.43$
- **Degree of integration:**  $\lambda = [0, 1]$
- **Share of risky assets in banks portfolio:**  $\bar{m} = 0.18$
- **Credit shocks process:**  $A_R = \begin{bmatrix} 0.95 & 0 \\ 0 & 0.95 \end{bmatrix}$ ;  $\rho_\epsilon^R = 0.2$ ,  $\sigma_\epsilon^R = 3\%$ ;  $\bar{R}^m = 1.06$
- **Intermediation cost**  $\iota = 4\%$

In Figure F.2, we consider how the output correlation between home and foreign economies varies as a function of the degree of financial integration under two parameterizations: productivity shocks only, and productivity and banking shocks. The blue line represents an economy with only productivity shocks. This line indicates that a higher level of banking integration is associated with less correlated output cycles, and greater negative comovement in the output cycles. The red line represents an economy with both bank capital shocks and productivity shocks. The difference between these two lines increases with the degree of banking integration. This suggests that there is a positive marginal effect of banking integration on the comovement in output cycles between two economies in “crisis” periods with both capital and idiosyncratic shocks (Kalemli-Ozcan, Papaioannou and Perri (2013b)).

Figure F.2: Financial Integration and Output Correlation



The figures plot the output correlation between the home and foreign areas using synthetic data produced from the model for varying levels of financial integration. The red line represents an economy with both bank capital shocks and productivity shocks. The blue line represents an economy with only productivity shocks.

## Appendix G Robustness

Table G.1: Robustness - Alternative Specification

| $\Delta gdp_{it}$                                          | (1)                 | (2)                 | (3)                 | (4)                 |
|------------------------------------------------------------|---------------------|---------------------|---------------------|---------------------|
| $\sum_{j \neq i} Post_{i,j,t} \times \Gamma_{j,t-1}^{ind}$ | -0.0196<br>(0.0037) | -0.0095<br>(0.0056) | -0.0125<br>(0.0052) | -0.0125<br>(0.0054) |
| $\sum_{j \neq i} \Gamma_{j,t-1}^{ind}$                     | 0.0070<br>(0.0021)  | -0.0070<br>(0.0168) | -0.0023<br>(0.0174) | -0.0022<br>(0.0179) |
| $\sum_{j \neq i} Post_{i,j,t}$                             | 0.0063<br>(0.0016)  | 0.0039<br>(0.0032)  | 0.0041<br>(0.0025)  | 0.0041<br>(0.0026)  |
| Region <sub><i>i</i></sub> × Year FE                       |                     | Yes                 | Yes                 | Yes                 |
| State <sub><i>i</i></sub> FE                               |                     |                     | Yes                 |                     |
| State <sub><i>i</i></sub> -Linear Trend                    |                     |                     |                     | Yes                 |
| # Obs                                                      | 1,173               | 1,173               | 1,173               | 1173                |
| $R^2$                                                      | 0.0285              | 0.5171              | 0.6122              | 0.6124              |

This table presents the estimates for an alternative specification, in which we aggregate the idiosyncratic shocks across state  $j$ . The dependent variable is the change in the real GDP growth rate in percentage. The main independent variable is  $\sum_{j \neq i} Post_{i,j,t} \times \Gamma_{j,t-1}^{ind}$  which denotes the aggregated value of idiosyncratic productivity shocks to top 10 firms in state  $i$  interacted with  $Post_{i,j,t}$ . All non-binary variables used in the regression are standardized to mean zero and variance 1. The unit of observation in each regression is a state<sub>*i*</sub>-year. Standard errors reported in parentheses are clustered by state<sub>*i*</sub>.

Table G.2: Alternative Shock Specification

$$\Delta gdp_{it} = \beta_0 Post_{i,j,t} \times \Gamma_{j,t-1}^{ind,state} + \beta_1 Post_{i,j,t} + \alpha_{i,j} + \theta_i \times t + \theta_{jt} + \varepsilon_{ijt}, i \neq j$$

| $\Delta gdp_{it}$                                | (1)                 | (2)                 | (3)                 | (4)                 | (5)                 | (6)                 |
|--------------------------------------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| $Post_{i,j,t} \times \Gamma_{j,t-1}^{ind,state}$ | -0.0669<br>(0.0223) | -0.0020<br>(0.0006) | -0.0069<br>(0.0016) | -0.0088<br>(0.0006) | -0.0281<br>(0.0109) | -0.0276<br>(0.0105) |
| $\Gamma_{j,t-1}^{ind,state}$                     | 0.0746<br>(0.0165)  | -0.0001<br>(0.0003) | 0.0022<br>(0.0006)  | 0.0027<br>(0.0004)  |                     |                     |
| $Post_{i,j,t}$                                   | 0.2441<br>(0.0645)  | 0.0086<br>(0.0789)  | 0.0770<br>(0.0604)  | 0.0776<br>(0.0471)  | 0.0880<br>(0.0527)  | 0.0806<br>(0.0491)  |
| Year FE                                          | Yes                 |                     |                     |                     |                     |                     |
| Region <sub>i</sub> -Year FE                     |                     |                     | Yes                 | Yes                 | Yes                 | Yes                 |
| Region <sub>j</sub> -Year FE                     |                     |                     | Yes                 | Yes                 |                     |                     |
| State <sub>i</sub> -State <sub>j</sub> FE        |                     |                     |                     | Yes                 | Yes                 | Yes                 |
| State <sub>j</sub> -Year FE                      |                     |                     |                     |                     | Yes                 | Yes                 |
| State <sub>i</sub> -Linear Trend                 |                     |                     |                     |                     |                     | Yes                 |
| N                                                | 57,700              | 57,700              | 57,700              | 57,700              | 57,700              | 57,700              |
| $R^2$                                            | 0.0201              | 0.3094              | 0.5168              | 0.6113              | 0.6115              | 0.6583              |

This table reports the results from the estimation of baseline specification. The dependent variable is the change in the real GDP growth rate in percentage. The main independent variable is  $\Gamma_{j,t-1}^{ind,state}$  which denotes the idiosyncratic shocks in state  $j$  constructed by aggregating the Domar weighted labor productivity shocks of top 10 firms, by sales in state  $j$ , after accounting for industry-year and state-year fixed effects. The unit of observation in each regression is a state<sub>i</sub>-state<sub>j</sub>-year pair. All non-binary variables used in the regression are standardized to mean zero and variance 1. Standard errors reported in parentheses are two-way clustered by state<sub>i</sub> and state<sub>j</sub>.

Table G.3: Can Idiosyncratic Shocks to Small Firms Predict Shocks to Large Firms?

| $\Gamma_{j,t}^{ind}$         | (1)                 | (2)                | (3)                 | (4)                 | (5)                 | (6)                 | (7)                 | (8)                 |
|------------------------------|---------------------|--------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| $\Gamma_{j,t}^{small,ind}$   | -0.0379<br>(0.0660) |                    |                     | -0.0242<br>(0.0653) | -0.0136<br>(0.0609) | -0.0077<br>(0.0594) | -0.0204<br>(0.0647) | -0.0203<br>(0.0652) |
| $\Gamma_{j,t-1}^{small,ind}$ |                     | 0.0957<br>(0.0807) |                     | 0.0802<br>(0.0864)  | 0.0700<br>(0.0768)  | 0.0741<br>(0.0755)  | 0.0782<br>(0.0711)  | 0.0782<br>(0.0716)  |
| $\Gamma_{j,t-2}^{small,ind}$ |                     |                    | -0.0901<br>(0.0724) | -0.0776<br>(0.0757) | -0.0842<br>(0.0680) | -0.0803<br>(0.0676) | -0.0985<br>(0.0679) | -0.0985<br>(0.0684) |
| Year FE                      |                     |                    |                     |                     | Yes                 | Yes                 | Yes                 | Yes                 |
| State FE                     |                     |                    |                     |                     |                     | Yes                 | Yes                 |                     |
| Region X Year FE             |                     |                    |                     |                     |                     |                     | Yes                 | Yes                 |
| State-Linear Trend           |                     |                    |                     |                     |                     |                     |                     | Yes                 |
| # Obs                        | 781                 | 781                | 781                 | 781                 | 781                 | 781                 | 781                 | 781                 |
| $R^2$                        | 0.0011              | 0.0072             | 0.0063              | 0.0122              | 0.0899              | 0.1160              | 0.2872              | 0.2873              |

This table reports the results from a regression of idiosyncratic shocks to large firms at time  $t$  on lags of idiosyncratic shocks to small firms.  $\Gamma_{j,t}^{ind}$  denotes the idiosyncratic shocks in state  $j$  constructed by aggregating the Domar weighted labor productivity shocks of top 10 firms, by sales in state  $j$ , at time  $t$ .  $\Gamma_{j,t}^{ind,small}$  denotes the idiosyncratic shocks in state  $j$ , constructed by aggregating the Domar weighted labor productivity shocks of small firms – firms that are not in the top 10 firms, by sales – in state  $j$  at time  $t$ . Standard errors are robust.

Table G.4: Baseline Specification with Time-Varying Controls

| $\Delta gdp_{it}$                                                              | (1)                 | (2)                 | (3)                 | (4)                 | (5)                 | (6)                 |
|--------------------------------------------------------------------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| $Post_{i,j,t} \times \Gamma_{j,t-1}^{ind}$                                     | -0.0185<br>(0.0012) | -0.0173<br>(0.0034) | -0.0212<br>(0.0085) | -0.0263<br>(0.0046) | -0.0134<br>(0.0045) | -0.0195<br>(0.0108) |
| $Post_{i,j,t} \times \Gamma_{j,t-1}^{ind} \times Exports_{i,j,pre}$            |                     | -0.0024<br>(0.0084) |                     |                     |                     | -0.0048<br>(0.0081) |
| $Post_{i,j,t} \times Exports_{i,j,pre}$                                        |                     | -0.0159<br>(0.0052) |                     |                     |                     | -0.0054<br>(0.0058) |
| $\Gamma_{j,t-1}^{ind} \times Exports_{i,j,pre}$                                |                     | -0.0074<br>(0.0078) |                     |                     |                     | -0.0068<br>(0.0075) |
| $Post_{i,j,t} \times \Gamma_{j,t-1}^{ind} \times Exports_{i,j,post}$           |                     | 0.0006<br>(0.0094)  |                     |                     |                     | 0.0027<br>(0.0093)  |
| $Post_{i,j,t} \times Exports_{i,j,post}$                                       |                     | 0.0088<br>(0.0078)  |                     |                     |                     | 0.0058<br>(0.0122)  |
| $\Gamma_{j,t-1}^{ind} \times Exports_{i,j,post}$                               |                     | 0.0095<br>(0.0093)  |                     |                     |                     | 0.0096<br>(0.0090)  |
| $Post_{i,j,t} \times \Gamma_{j,t-1}^{ind} \times Imports_{i,j,pre}$            |                     | 0.0072<br>(0.0115)  |                     |                     |                     | 0.0117<br>(0.0108)  |
| $Post_{i,j,t} \times Imports_{i,j,pre}$                                        |                     | 0.0136<br>(0.0148)  |                     |                     |                     | 0.0010<br>(0.0121)  |
| $\Gamma_{j,t-1}^{ind} \times Imports_{i,j,pre}$                                |                     | -0.0198<br>(0.0068) |                     |                     |                     | -0.0199<br>(0.0075) |
| $Post_{i,j,t} \times \Gamma_{j,t-1}^{ind} \times Imports_{i,j,post}$           |                     | 0.0072<br>(0.0115)  |                     |                     |                     | 0.0117<br>(0.0108)  |
| $Post_{i,j,t} \times Imports_{i,j,post}$                                       |                     | 0.0136<br>(0.0148)  |                     |                     |                     | 0.0010<br>(0.0121)  |
| $\Gamma_{j,t-1}^{ind} \times Imports_{i,j,post}$                               |                     | -0.0198<br>(0.0068) |                     |                     |                     | -0.0199<br>(0.0075) |
| $Post_{i,j,t} \times \Gamma_{j,t-1}^{ind} \times Income_{j,t}$                 |                     |                     | 0.0103<br>(0.0165)  |                     |                     | 0.0126<br>(0.0176)  |
| $Post_{i,j,t} \times Income_{j,t}$                                             |                     |                     | -0.0062<br>(0.0640) |                     |                     | 0.0103<br>(0.0649)  |
| Income Covariance $_{i,j,t}$                                                   |                     |                     |                     | 0.2569<br>(0.0589)  |                     | 0.2718<br>(0.0631)  |
| $Post_{i,j,t} \times \Gamma_{j,t-1}^{ind} \times Income\ Covariance_{i,j,t}$   |                     |                     |                     | -0.0122<br>(0.0070) |                     | -0.0099<br>(0.0072) |
| $Post_{i,j,t} \times Income\ Covariance_{i,j,t}$                               |                     |                     |                     | 0.0135<br>(0.0922)  |                     | 0.0040<br>(0.0961)  |
| $\Gamma_{j,t-1}^{ind} \times Income\ Covariance_{i,j,t}$                       |                     |                     |                     | -0.0058<br>(0.0096) |                     | -0.0108<br>(0.0116) |
| Industry Similarity $_{i,j,t}$                                                 |                     |                     |                     |                     | 0.3700<br>(0.2067)  | 0.4431<br>(0.2130)  |
| $Post_{i,j,t} \times \Gamma_{j,t-1}^{ind} \times Industry\ Similarity_{i,j,t}$ |                     |                     |                     |                     | 0.0163<br>(0.0132)  | 0.0215<br>(0.0187)  |
| $Post_{i,j,t} \times Industry\ Similarity_{i,j,t}$                             |                     |                     |                     |                     | 0.0299<br>(0.0413)  | 0.0046<br>(0.0452)  |
| $\Gamma_{j,t-1}^{ind} \times Industry\ Similarity_{i,j,t}$                     |                     |                     |                     |                     | -0.0123<br>(0.0053) | -0.0122<br>(0.0085) |
| $Post_{i,j,t}$                                                                 | 0.0928<br>(0.0500)  | 0.0936<br>(0.0501)  | 0.0929<br>(0.0489)  | 0.0908<br>(0.0827)  | 0.1000<br>(0.0500)  | 0.0861<br>(0.0856)  |
| Region $_i$ -Year FE                                                           | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 |
| State $_i$ -State $_j$ FE                                                      | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 |
| State $_j$ -Year FE                                                            | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 |
| State $_i$ -Linear Trend                                                       | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 |
| N                                                                              | 54,048              | 54,048              | 54,048              | 54,048              | 54,048              | 54,048              |
| R <sup>2</sup>                                                                 | 0.6568              | 0.6570              | 0.6568              | 0.6624              | 0.6579              | 0.6640              |

This table reports the results from the estimation of baseline specification. The dependent variable is the change in the real GDP growth rate in percentage. The main independent variable is  $\Gamma_{j,t-1}^{ind}$  which denotes the idiosyncratic shocks in state  $j$  constructed by aggregating the Domar weighted labor productivity shocks of top 10 firms, by sales in state  $j$  after accounting for industry-year fixed effects. The unit of observation in each regression is a state $_i$ -state $_j$ -year pair. We include additional time-varying control variables including Exports from state  $i$  to state  $j$  in 1977 (*pre*-deregulation) and 1993 (*post*-deregulation), Imports from state  $i$  to state  $j$  in 1977 (*pre*-deregulation) and 1993 (*post*-deregulation), personal income per capita in state  $j$ , the similarity in industry composition between states  $i$  and  $j$ , and a 3-year forward-rolling covariance in personal income growth between states  $i$  and  $j$  which is trimmed at 1%. All non-binary variables used in the regression are standardized to mean zero and variance 1. Standard errors reported in parentheses are two-way clustered by state $_i$  and state $_j$ .

## G.1 Alternative Measures of Idiosyncratic Shocks

We begin by constructing state-level idiosyncratic shocks using only positive firm-level productivity shocks. These results are ported in Table G.5. The point estimates of the interaction term of interest are qualitatively similar to baseline results. Further, we test that our results are not driven by exceptional features in our specification of  $\Gamma_{j,t-1}$ , checking that our results are robust to alternative measures of  $\Gamma$ . These results are presented in Table G.6.  $\Gamma_{j,t-1}$  is defined as the idiosyncratic productivity shock computed using top 20 firms in state  $j$  (column 1), and top 30 firms in state  $j$  (column 2), a time-series average of idiosyncratic productivity shocks (column 3), and non-industry adjusted value (column 4).

Table G.5: Robustness - Constructing  $\Gamma_{j,t-1}^{ind}$  using only positive firm-level shocks

| $\Delta gdp_{it}$                                                   | (1)                 | (2)                 | (3)                 | (4)                 | (5)                 | (6)                 |
|---------------------------------------------------------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| $Post_{i,j,t} \times \Gamma_{j,t-1}^{ind-pos}$                      | -0.0825<br>(0.0157) | -0.0031<br>(0.0002) | -0.0052<br>(0.0014) | -0.0076<br>(0.0002) | -0.0168<br>(0.0053) | -0.0148<br>(0.0047) |
| $\Gamma_{j,t-1}^{ind-pos}$                                          | 0.0618<br>(0.0166)  | 0.0012<br>(0.0002)  | 0.0026<br>(0.0002)  | 0.0037<br>(0.0000)  |                     |                     |
| $Post_{i,j,t}$                                                      | 0.2535<br>(0.0643)  | 0.0086<br>(0.0789)  | 0.0766<br>(0.0604)  | 0.0771<br>(0.0471)  | 0.0860<br>(0.0526)  | 0.0785<br>(0.0492)  |
| Year FE                                                             | Yes                 |                     |                     |                     |                     |                     |
| Region <sub><math>i</math></sub> -Year FE                           |                     |                     | Yes                 | Yes                 | Yes                 | Yes                 |
| Region <sub><math>j</math></sub> -Year FE                           |                     |                     | Yes                 | Yes                 |                     |                     |
| State <sub><math>i</math></sub> -State <sub><math>j</math></sub> FE |                     |                     |                     | Yes                 | Yes                 | Yes                 |
| State <sub><math>j</math></sub> -Year FE                            |                     |                     |                     |                     | Yes                 | Yes                 |
| State <sub><math>i</math></sub> -Linear Trend                       |                     |                     |                     |                     |                     | Yes                 |
| N                                                                   | 57,700              | 57,700              | 57,700              | 57,700              | 57,700              | 57,700              |
| $R^2$                                                               | 0.0181              | 0.3094              | 0.5168              | 0.6113              | 0.6114              | 0.6583              |

This table reports the results from the estimation of baseline specification with state-level idiosyncratic shocks constructed using only positive firm-level labor productivity shocks. The dependent variable is the change in the real GDP growth rate in percentage. The main independent variable is  $\Gamma_{j,t-1}^{ind}$  which denotes the idiosyncratic shocks in state  $j$  constructed by aggregating the Domar weighted labor productivity shocks of top 10 firms, by sales in state  $j$ . The unit of observation in each regression is a state <sub>$i$</sub> -state <sub>$j$</sub> -year pair. All non-binary variables used in the regression are standardized to mean zero and variance 1. Standard errors reported in parentheses are two-way clustered by state <sub>$i$</sub>  and state <sub>$j$</sub> .

Table G.6: Robustness - Alternative Construction of  $\Gamma$ 

|                                           | (1)                       | (2)                       | (3)                        | (4)                     |
|-------------------------------------------|---------------------------|---------------------------|----------------------------|-------------------------|
| $\Delta gdp_{it}$                         | $\Gamma_{j,t-1}^{ind-20}$ | $\Gamma_{j,t-1}^{ind-30}$ | $\Gamma_{j,t-1}^{ind-avg}$ | $\Gamma_{j,t-1}^{norm}$ |
| $Post_{i,j,t} \times \Gamma_{j,t-1}^*$    | -0.0159<br>(0.0011)       | -0.0162<br>(0.0009)       | -0.1178<br>(0.0636)        | -0.0037<br>(0.0019)     |
| $Post_{i,j,t}$                            | 0.0782<br>(0.0491)        | 0.0782<br>(0.0491)        | 0.0777<br>(0.0491)         | 0.0778<br>(0.0492)      |
| Region <sub>i</sub> -Year FE              | Yes                       | Yes                       | Yes                        | Yes                     |
| State <sub>i</sub> -State <sub>j</sub> FE | Yes                       | Yes                       | Yes                        | Yes                     |
| State <sub>j</sub> -Year FE               | Yes                       | Yes                       | Yes                        | Yes                     |
| State <sub>i</sub> -Linear Trend          | Yes                       | Yes                       | Yes                        | Yes                     |
| N                                         | 57,700                    | 57,700                    | 57,700                     | 57,700                  |
| R <sup>2</sup>                            | 0.6583                    | 0.6583                    | 0.6583                     | 0.6583                  |

This table presents the estimates for baseline specification with alternative construction of  $\Gamma$ . The dependent variable is the change in the real GDP growth rate in percentage. The main independent variable is  $\Gamma_{j,t-1}^*$  which denotes the idiosyncratic production shocks to top 20 firms ( $\Gamma_{j,t-1}^{ind-20}$ ) in other states in column (1), to top 30 firms ( $\Gamma_{j,t-1}^{ind-30}$ ) in column (2), the a time-series average of idiosyncratic production shocks to top 10 firms ( $\Gamma_{j,t-1}^{ind-avg}$ ) in each state in column (3) and using non-industry adjusted value of  $\Gamma_{j,t-1}^*$  in column (4). The unit of observation in each regression is a state<sub>i</sub>-state<sub>j</sub>-year pair. Standard errors reported in parentheses are two-way clustered by state<sub>i</sub> and state<sub>j</sub>.

The point estimates of the interaction term in columns 1 and 2 of Table G.6 are similar to the baseline result. The estimate is larger in column 3 and smaller in column 4. The estimate is larger in column 3 because the measure  $\Gamma$  by construction incorporates future information biasing the estimate upwards. The estimate in column 4 is smaller because the local shocks are only adjusted for aggregate temporal shocks making these shocks less geographically isolated. In all specifications, the relation between idiosyncratic shocks in other states and the state-level impact on GDP growth after banking integration is statistically significant. Hence, we rule out concerns that the relation is attributed to the ad-hoc calculation of idiosyncratic shocks using top 10 firms.

Furthermore, we check whether our results are driven by outsized productivity shocks experienced by states where top 10 firms share of sales is high. We test whether our results change under alternative samples. These results are presented in Table G.7. Column (1) reports the baseline specification under complete sample, columns (2)-(5) only include a state<sub>i</sub> – state<sub>j</sub> pair if the average ratio of sales of top 10 firms to all firms between 1978 and 2000 in state *j* is less than 95%, 90%, 80%, and 70% respectively. The point estimate remains stable even after restricting the sample to varying degrees. Moreover, the relation remains statistically significant. The precision of the estimate decreases from column (1) to (3) due to the reduction in the sample size. The precision of the estimate stabilizes thereafter. Hence, the result is not driven by monopolistic states.

Table G.7: Robustness - Alternative Samples

|                                            | (1)                 | (2)                 | (3)                 | (4)                 | (5)                 |
|--------------------------------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| $\Delta gdp_{it}$                          | All                 | >95%                | >90%                | >80%                | >70%                |
| $Post_{i,j,t} \times \Gamma_{j,t-1}^{ind}$ | -0.0164<br>(0.0007) | -0.0195<br>(0.0026) | -0.0154<br>(0.0054) | -0.0145<br>(0.0054) | -0.0176<br>(0.0058) |
| $Post_{i,j,t}$                             | 0.0783<br>(0.0491)  | 0.0870<br>(0.0547)  | 0.1028<br>(0.0522)  | 0.0987<br>(0.0503)  | 0.1284<br>(0.0604)  |
| Region <sub>i</sub> -Year FE               | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 |
| State <sub>i</sub> -State <sub>j</sub> FE  | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 |
| State <sub>j</sub> -Year FE                | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 |
| State <sub>i</sub> -Linear Trend           | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 |
| $N$                                        | 57,700              | 29,900              | 25,300              | 17,250              | 8,050               |
| $R^2$                                      | 0.6583              | 0.6567              | 0.6564              | 0.6561              | 0.6569              |

This table presents the estimates for baseline specification with alternative samples. The dependent variable is the change in the real GDP growth rate in percentage. The main independent variable is  $\Gamma_{j,t-1}^{ind}$  which denotes the idiosyncratic production shocks to top 10 firms. The unit of observation in each regression is a state<sub>i</sub>-state<sub>j</sub>-year pair. Column (1) includes the entire sample, column (2), (3), (4) and (5) only includes a state<sub>i</sub>-state<sub>j</sub>-year pair if the average ratio of sales of top 10 firms to all firms between 1978 and 2000 in state<sub>j</sub> is less than 95%, 90%, 80% and 70% respectively. Standard errors reported in parentheses are two-way clustered by state<sub>i</sub> and state<sub>j</sub>.

Table G.8: Robustness - Accounting for Shocks to Small Firms

| $\Delta gdp_{it}$                                                         | (1)                 | (2)                 | (3)                 | (4)                 |
|---------------------------------------------------------------------------|---------------------|---------------------|---------------------|---------------------|
| $Post_{i,j,t} \times \Gamma_{j,t-1}^{ind}$                                | -0.0187<br>(0.0023) | -0.0189<br>(0.0032) |                     |                     |
| $Post_{i,j,t}$                                                            | 0.0821<br>(0.0500)  | 0.0822<br>(0.0500)  | 0.0820<br>(0.0501)  | 0.0818<br>(0.0500)  |
| $Post_{i,j,t} \times \Gamma_{j,t-1}^{ind,small}$                          |                     | 0.0044<br>(0.0031)  |                     |                     |
| $Post_{i,j,t} \times (\Gamma_{j,t-1}^{ind} - \Gamma_{j,t-1}^{ind,small})$ |                     |                     | -0.0153<br>(0.0036) |                     |
| $Post_{i,j,t} \times \Gamma_{j,t-1}^{Res}$                                |                     |                     |                     | -0.0186<br>(0.0023) |
| Region <sub><i>i</i></sub> -Year FE                                       | Yes                 | Yes                 | Yes                 | Yes                 |
| State <sub><i>i</i></sub> -State <sub><i>j</i></sub> FE                   | Yes                 | Yes                 | Yes                 | Yes                 |
| State <sub><i>j</i></sub> -Year FE                                        | Yes                 | Yes                 | Yes                 | Yes                 |
| State <sub><i>i</i></sub> -Linear Trend                                   | Yes                 | Yes                 | Yes                 | Yes                 |
| # Obs                                                                     | 40,400              | 40,400              | 40,400              | 40,400              |
| $R^2$                                                                     | 0.6609              | 0.6609              | 0.6609              | 0.6609              |

This table reports the results from the estimation of baseline specification. The dependent variable is the change in the real GDP growth rate in percentage. The main independent variable in columns 1 and 2 is  $\Gamma_{j,t-1}^{ind}$  which denotes the idiosyncratic shocks in state  $j$  constructed by aggregating the Domar weighted labor productivity shocks of top 10 firms, by sales in state  $j$ .  $\Gamma_{j,t-1}^{ind,small}$  ( $\Gamma_{i,t-1}^{small}$ ) denotes the idiosyncratic shocks in state  $j$  ( $i$ ) constructed by aggregating the Domar weighted labor productivity shocks of small firms – firms that are in not in the top 10 firms, by sales – in state  $j$  ( $i$ ). The unit of observation in each regression is a state<sub>*i*</sub>-state<sub>*j*</sub>-year pair. The key dependent variable in column 3 is the difference of  $\Gamma_{j,t-1}^{ind}$  and  $\Gamma_{j,t-1}^{ind,small}$  interacted with the Post variable. The key dependent variable in column 4 is the interaction of  $\Gamma_{j,t-1}^{Res}$  and Post.  $\Gamma_{j,t-1}^{Res}$  denotes the residuals from the regression of the large-firm shocks on small-firm shocks at the state-level. All non-binary variables used in the regression are standardized to mean zero and variance 1. Standard errors reported in parentheses are two-way clustered by state<sub>*i*</sub> and state<sub>*j*</sub>.

## G.2 Factor Structure with Heterogeneous Exposures

In this section, we assume that firm-level productivity shocks are heterogeneous, but have time-invariant exposure to macroeconomic shocks. We do this to investigate the claim if our measurement of idiosyncratic shocks is corrupted by the presence of a factor structure in such shocks making these shocks capture some degree of aggregate shocks and not local shocks. Under the heterogeneous but time-invariant factor structure assumption, the residuals obtained from running a firm-level regression of labor productivity shocks adjusted for industry shocks on macroeconomic variables are taken to be idiosyncratic. We define  $g_{it}^{(i)}$  as in equation 1. For each firm, we run the following regression of  $g_{it}^{(i)}$  on macroeconomic shocks for each year.

$$g_{it}^{(i)} = \alpha_i + \beta_i \Delta \Omega_t + \varepsilon_{it} \quad (\text{G.1})$$

$\Delta \Omega_t$  refers to the vector of macroeconomic shocks observed for each year. Macroeconomic shocks include change in effective Fed Funds rate, GDP growth rate, change in unemployment rate, change in inflation rate, Hamilton oil price shocks, and market risk premium. G.9 and G.10 provide a brief summary of the macroeconomic shocks employed here.

Table G.9: Summary of Data Sources for Macroeconomic Variables

| Description                             | Sources                                | Measure                              |
|-----------------------------------------|----------------------------------------|--------------------------------------|
| Change in Effective Federal Funds Rates | FRED St. Louis Fed                     | $\Delta E F F R_t$                   |
| Real Gross Domestic Product Growth      | FRED St. Louis Fed                     | $\frac{\Delta G D P_t}{G D P_{t-1}}$ |
| Consumer Price Index Growth             | FRED St. Louis                         | Annual average                       |
| Change in Unemployment Rate             | FRED St. Louis Fed                     | $\Delta \text{Unemployment Rate}_t$  |
| Hamilton Structural Oil Supply Shocks   | Christiane Baumeister Research Website | Annual average                       |
| Market Risk Premium                     | Kenneth French Data Library            | Annual average                       |

This table presents a summary of the data sources and construction methodology for the macroeconomic variables.

Table G.10: Summary Statistics of Macroeconomic Variables Across Years (Raw)

|                                        | N  | p25    | Median | p75   | Mean   | Std. Dev. |
|----------------------------------------|----|--------|--------|-------|--------|-----------|
| Change in Effective Federal Funds Rate | 24 | -1.209 | 0.025  | 1.447 | 0.050  | 1.941     |
| GDP Growth                             | 24 | 2.719  | 3.723  | 4.464 | 3.371  | 1.927     |
| CPI Growth                             | 24 | 0.666  | 0.857  | 1.326 | 1.154  | 0.759     |
| Change in Unemployment Rate            | 24 | -0.617 | -0.267 | 0.125 | -0.156 | 0.855     |
| Hamilton Structural Oil Supply Shock   | 24 | -0.237 | -0.054 | 0.269 | -0.057 | 0.415     |
| Market Risk Premium                    | 24 | -0.105 | 0.909  | 1.619 | 0.706  | 1.090     |

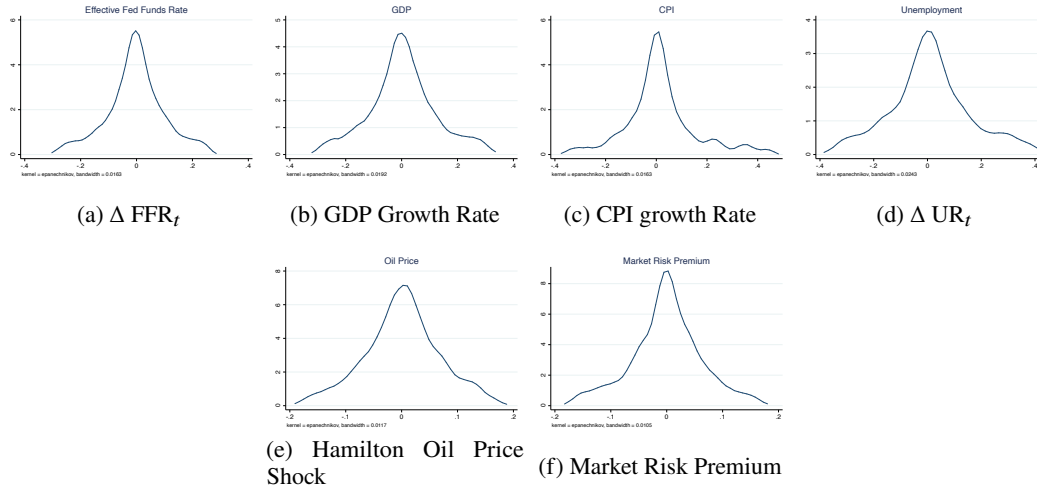
This table presents the summary statistics for the macroeconomic variables of interest from 1977-2000.

The firm-level regression allows the firms to have heterogeneous exposure to macroeconomic shocks. Figure G.1 reports the kernel density of the sensitivity of  $g_{it}^{(i)}$  to macroeconomic shocks. These sensitivities are computed at firm-level using the data between 1977 and 2000. Across all macroeconomic variables, the densities are centered around zero. This indicates that for the macroeconomic shocks considered, the average response is zero. The median and the mean estimate for sensitivity related to the monetary policy rate and unemployment rate are negative, as expected, but small in magnitude. However, the sensitivities to the monetary policy rate and unemployment rate have large variance, suggesting that firms have varied responses to these macroeconomic shocks. Sensitivities related to change in unemployment rate, inflation, GDP growth and monetary policy rate have the largest variation. Variation attributed to Hamilton shocks is rather small, as oil supply shocks have a more concentrated effect in specific industries.

The  $\varepsilon_{it}$  for the top 10 firms in each state are extracted from equation G.1, and aggregated at the state-level using Domar weights as in equation 2. Figure G.2 presents a binscatter plot of our standard measure of state-level idiosyncratic shock,  $\Gamma_{j,t}^{industry}$  and the idiosyncratic shock generated from the factor model,  $\Gamma_{j,t}^{factor}$ . The correlation between  $\Gamma_{j,t}^{industry}$  and  $\Gamma_{j,t}^{factor}$ , is 71.68%. Moreover, regressing  $\Gamma_{j,t}^{factor}$  on  $\Gamma_{j,t}^{industry}$  reveals that the  $R^2$  value is 51.38%, with a  $\beta$  of  $\sim 0.74$ . This indicates that the two measures of idiosyncratic shocks are highly correlated.

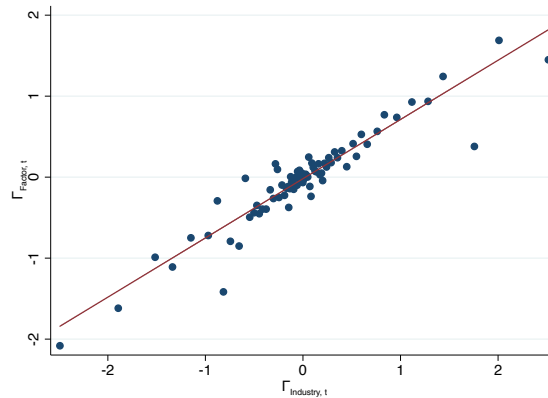
Table G.11 reports the results of the baseline estimation using the shock generated from the factor model,  $\Gamma_{j,t}^{factor}$  as the measure of state-level idiosyncratic shocks. Column (6) reports the result by constructing  $\Gamma_{j,t}^{factor}$  using all macroeconomic shocks, namely, change in effective federal funds rate, national GDP growth, oil supply shock, inflation, unemployment change, and the market risk premium.  $\Gamma_{j,t}^{factor}$  are constructed by step-wise inclusion of factors as we move from column (1) to (6). Column (1) uses a single factor, the change in the effective federal funds rate.  $\Gamma_{j,t}^{factor}$  used in columns (2)-(6) are constructed by step-wise inclusion of factors. The results in all column are quantitatively and qualitatively similar to each other and to the estimate obtained in column (6) of Table 3. The point estimates in all columns are negative, stable across different construction of  $\Gamma_{j,t}^{factor}$  and statistically significant at 1% level.

Figure G.1: Kernel Densities of Heterogeneous Exposures of firm-level shocks to Macroeconomic Variables



This figure plots the kernel density of the heterogeneous exposure of industry-year adjusted firm level labor productivity shocks to macroeconomic variables. The kernel density is plotted after trimming the variables at the 10th and 90th percentiles. Panel a, b, c, d, e and f report the kernel density for change in effective federal funds rate, GDP growth rate, CPI growth rate, change in unemployment rate, Hamilton Oil price Shocks and the market risk premium respectively. Table G.9 provides details on data sources and calculation of the macroeconomic variables employed. The table reports the summary statistics for the firm  $\beta$  values associated with the macro variables of interest.

Figure G.2: Relation between  $\Gamma_{j,t}^{factor}$  and  $\Gamma_{j,t}^{industry}$



The plot presents a binscatter plot of our standard measure of state-level idiosyncratic shock,  $\Gamma_{j,t}^{industry}$  and the idiosyncratic shock generated from the factor model,  $\Gamma_{j,t}^{factor}$ . The correlation between  $\Gamma_{j,t}^{industry}$  and  $\Gamma_{j,t}^{factor}$ , is 71.68%. Moreover, regressing  $\Gamma_{j,t}^{factor}$  on  $\Gamma_{j,t}^{industry}$  reveals that the  $R^2$  value is 51.38% between the two. The  $\beta$  value of the regression is 0.74.

Table G.11: Baseline Results with Factor Structure of Shocks

| $\Delta gdp_{it}$                             | (1)                 | (2)                 | (3)                 | (4)                 | (5)                 | (6)                 |
|-----------------------------------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| $Post_{i,j,t} \times \Gamma_{j,t-1}^{factor}$ | -0.0170<br>(0.0022) | -0.0136<br>(0.0033) | -0.0133<br>(0.0040) | -0.0145<br>(0.0037) | -0.0149<br>(0.0039) | -0.0151<br>(0.0048) |
| $Post_{i,j,t}$                                | 0.0784<br>(0.0491)  | 0.0783<br>(0.0492)  | 0.0783<br>(0.0491)  | 0.0781<br>(0.0492)  | 0.0783<br>(0.0492)  | 0.0781<br>(0.0492)  |
| Region <sub>i</sub> -Year FE                  | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 |
| State <sub>i</sub> -State <sub>j</sub> FE     | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 |
| State <sub>j</sub> -Year FE                   | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 |
| State <sub>i</sub> -Linear Trend              | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 |
| $N$                                           | 57,700              | 57,700              | 57,700              | 57,700              | 57,700              | 57,700              |
| $R^2$                                         | 0.6583              | 0.6583              | 0.6583              | 0.6583              | 0.6583              | 0.6583              |

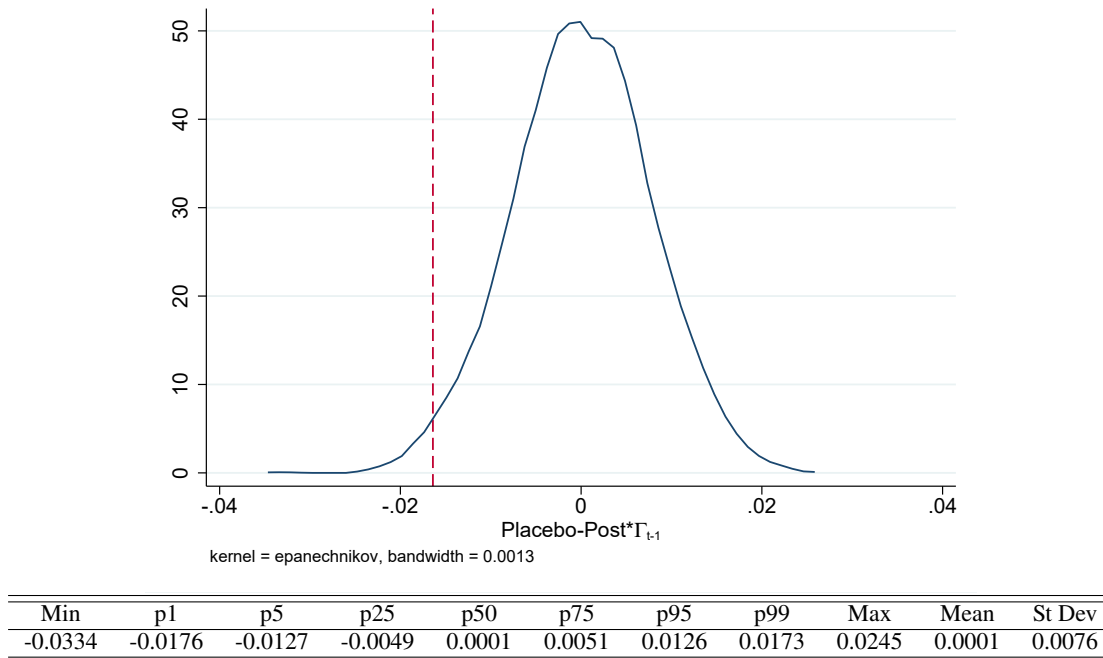
This table presents the estimates for baseline specification with alternative construction of  $\Gamma$  where the shocks are constructed using a factor structure. Column (6) reports the result, after controlling for all factors we consider, namely, change in effective federal funds rate, GDP growth, oil supply shock, inflation, unemployment change, and the market risk premium. We start in column (1) with a single factor under consideration: the change in the effective federal funds rate. As we move from column (1) to column (6), we introduce an additional factor in the model in a step-wise fashion. In column (1), the idiosyncratic shock is estimated after controlling for the change in effective federal funds rate. In column (2), the idiosyncratic shock is estimated after controlling for the change in effective federal funds rate and the GDP growth. In column (3), the shock is estimated after controlling for the change in effective federal funds rate, GDP growth, and oil supply shock. In column (4), the factors are the change in effective federal funds rate, GDP growth, oil supply shock, and inflation. In column (5), the factors are the change in effective federal funds rate, GDP growth, oil supply shock, inflation, and change in unemployment. In column (6), the factors are the change in effective federal funds rate, GDP growth, oil supply shock, inflation, change in unemployment, and market risk premium. Standard errors in parentheses are double clustered at  $state_i$  and  $state_j$  level.  $p < 0.1$ ,  $p < 0.05$ ,  $p < 0.01$

## G.3 Placebo Test

### G.3.1 Randomizing the Timing of Deregulation

We conduct a placebo test wherein we randomize the timing of banking integration. A placebo deregulation year is generated for each state-pair  $(i, j)$  from a uniform distribution between 1982 and 1994. The baseline specification is estimated using the generated placebo year. We estimate this process 3,500 times. Appendix Figure G.3 plots the kernel density of the point estimates of  $Placebo - Post_{i,j,t} \times \Gamma_{j,t-1}$  obtained from 3,500 Monte-Carlo simulations where we randomize the timing of state-pairwise banking integration. The distribution of the coefficient of the interaction term is centered around zero with a mean and standard deviation of 0.0001 and 0.0076, respectively. The dashed red line indicates the estimated point estimate from our baseline regression in Table 3 with 1.74% of the estimated coefficients of the  $Placebo - Post_{i,j,t} \times \Gamma_{j,t-1}$  lying to the left of the dashed line. Hence, we can argue that the timing of banking integration is special and results are unlikely to be driven by omitted variables as long as the structure of such variables is identical across state-pairs.

Figure G.3: Placebo Test: Randomization of the Timing of Deregulation

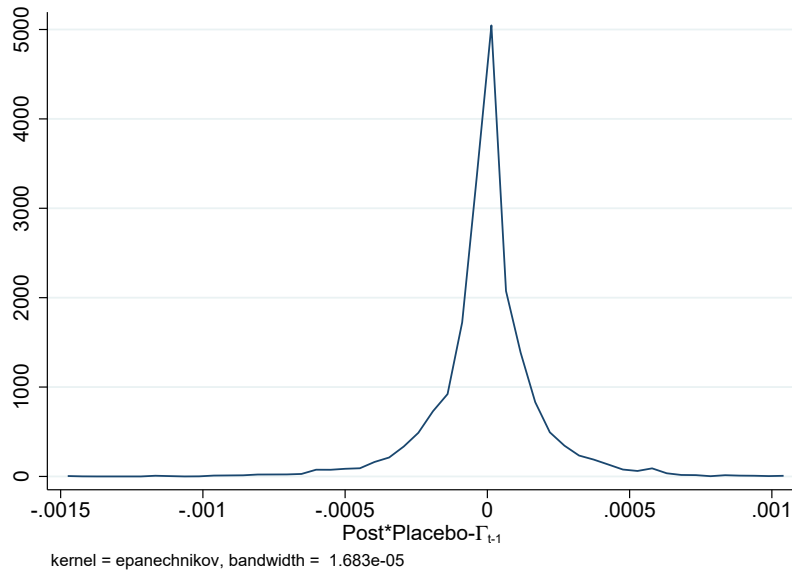


The figure plots the kernel density of the point estimates of  $Placebo - Post_{i,j,t} \times \Gamma_{j,t-1}^{ind}$  obtained from the 3,500 Monte-Carlo simulations. We generate a new date of deregulation from a uniform distribution between 1982 and 1994 for each state-pair in every simulation. We call this new deregulation year as placebo year and define the variable  $Placebo - Post_{i,j,t}$  based on the placebo year. We run our baseline specification with  $Placebo - Post_{i,j,t}$ . The table underneath The figure gives the numbers associated with the distribution of the estimates plotted in figure. The dash red line shows the point estimate from column 7 of Table 3. There are 1.74% of points to the left of the red-dashed line.

### G.3.2 Randomizing idiosyncratic shocks

We randomize the state-level idiosyncratic shocks. We generate a series of idiosyncratic shocks by randomly drawing from a Cauchy distribution with location parameter -0.0173, and scaling parameter 0.1539.<sup>29</sup> We re-run the baseline specification with the randomly generated  $Placebo - \Gamma_{j,t-1}$  and estimate the coefficient of the interaction term of  $Post_{i,j,t} \times Placebo - \Gamma_{j,t-1}$ . Figure G.4 plots the kernel density of the point estimates of  $Post_{i,j,t} \times Placebo - \Gamma_{j,t-1}$  obtained from 3,500 such Monte-Carlo simulations. The distribution of the point estimates is centred around zero with a standard deviation of 0.0002. The minimum point estimate obtained from the exercise is -0.0012 which is lower than any of the point estimates presented in Table 3. Hence, we can rule out the claim that the results are spurious in nature.

Figure G.4: Placebo Test: Randomization of  $\Gamma_{j,t-1}^{ind}$



| Min     | p1      | p5      | p25     | p50    | p75    | p95    | p99    | Max    | Mean    | St Dev |
|---------|---------|---------|---------|--------|--------|--------|--------|--------|---------|--------|
| -0.0014 | -0.0006 | -0.0003 | -0.0001 | 0.0000 | 0.0001 | 0.0003 | 0.0006 | 0.0010 | -0.0000 | 0.0002 |

The figure plots the kernel density of the point estimates of  $Post_{i,j,t} \times Placebo - \Gamma_{j,t-1}^{ind}$  obtained from the 3,500 Monte-Carlo simulations. We generate a random data for  $Placebo - \Gamma_{j,t-1}^{ind}$  using a Cauchy distribution with a location parameter of -0.0173 and scaling parameter of 0.1539. These parameters are obtained by fitting the empirical CDF to Cauchy CDF using maximum likelihood estimator (MLE). We run our baseline specification with  $Placebo - \Gamma_{j,t-1}^{ind}$ . The table underneath The figure gives the numbers associated with the distribution of the estimates plotted in figure.

<sup>29</sup>The parameters are estimated by fitting the empirical CDF of true idiosyncratic shocks to a Cauchy CDF using maximum likelihood estimator (MLE). We consider Cauchy distribution because inspection of state-level idiosyncratic shocks indicates presence of fat-tails

## G.4 Geography-based Measurement Error

### G.4.1 State Level Value Added Shocks

To validate our results, we redo our empirical exercise using value-added shocks,. These shocks are constructed as follows:

$$\Gamma_{it}^{ind} = \sum_{d \in I} \frac{VA_{d,t-1}^{(i)}}{Y_{i,t-1}} (\Delta \ln(VA_{d,t}^{(i)}) - \overline{\Delta \ln(VA_{d,t})})$$
$$\Gamma_{it}^{norm} = \sum_{d \in D} \frac{VA_{d,t-1}^{(i)}}{Y_{i,t-1}} (\Delta \ln(VA_{d,t}^{(i)}) - \overline{\Delta \ln(VA_t)})$$

where,  $I$  is the set of all industries,  $VA_{d,t}^{(i)}$  denotes the value added for a given industry,  $d$ , in a state,  $i$  at time  $t$ .  $\overline{VA_{d,t}}$  and  $\overline{VA_t}$  denote the mean growth rate in  $d$ 's industry in year  $t$  and across all industries in year  $t$  respectively. The shocks constructed using the value-added measures exhibit properties similar to our main measure of idiosyncratic productivity shocks constructed using Compustat data.  $\Gamma_{it}^{ind}$  has a median value -0.0006 and the 25th and 75th percentiles are -0.0171 and 0.0160 respectively.  $\Gamma_{it}^{norm}$  has a median value -0.0005 and the 25th and 75th percentiles are -0.0179 and 0.01564 respectively.

The results of the baseline regression are reported in Table G.12. The estimates from both regressions are statistically significant, and the point estimates are stable and within range of the previous estimates. The point estimate of the interaction term computed using this alternative measure is smaller than the baseline specification. This reduction in the point estimate can be attributed to the fact that the idiosyncratic shocks computed using value added data includes shocks to bank-dependent firms. The shocks to the bank-dependent firms can be caused by shocks to the banking sector or could result in shocks to the banking sector. Hence, these shocks are not as purely exogenous as our baseline measure of idiosyncratic shocks, hence, explains why the point estimate is smaller in magnitude.

Table G.12: Robustness - Value Added Measure of  $\Gamma$ 

| $\Delta gdp_{it}$                           | (1)                 | (2)                 |
|---------------------------------------------|---------------------|---------------------|
| $Post_{i,j,t} \times \Gamma_{j,t-1}^{ind}$  | -0.0044<br>(0.0013) |                     |
| $Post_{i,j,t} \times \Gamma_{j,t-1}^{norm}$ |                     | -0.0061<br>(0.0012) |
| $Post_{i,j,t}$                              | 0.0885<br>(0.0490)  | 0.0884<br>(0.0490)  |
| Region <sub>i</sub> -Year FE                | Yes                 | Yes                 |
| State <sub>i</sub> -State <sub>j</sub> FE   | Yes                 | Yes                 |
| State <sub>j</sub> -Year FE                 | Yes                 | Yes                 |
| State <sub>i</sub> -Linear Trend            | Yes                 | Yes                 |
| N                                           | 51,000              | 51,000              |
| R <sup>2</sup>                              | 0.6719              | 0.6719              |

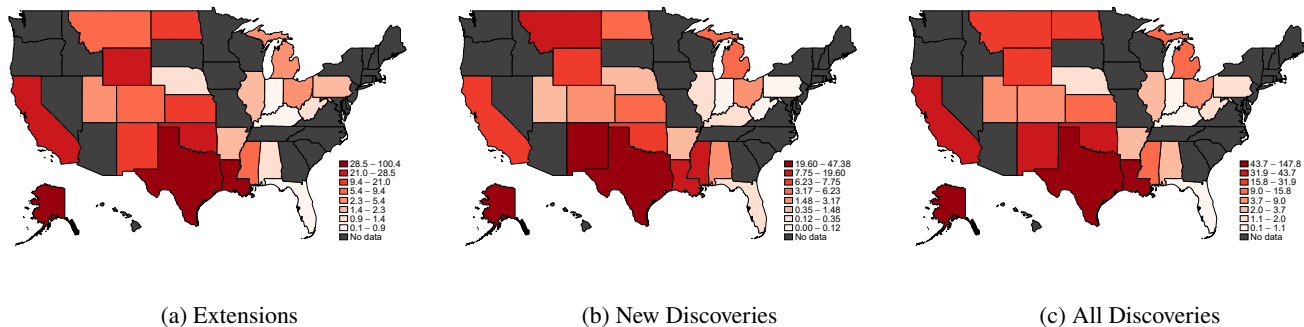
This table presents the estimates for baseline specification with alternative construction of  $\Gamma$  using the value-added measure. The dependent variable is the change in the real GDP growth rate in percentage. The main independent variable is  $\Gamma_{j,t-1}^{ind}$  and  $\Gamma_{j,t-1}^{norm}$  which denote the value-added shocks after adjusting for the mean growth rate of each industry in a given year, and for a given year, respectively. The unit of observation in each regression is a state<sub>i</sub>-state<sub>j</sub>-year pair. All non-binary variables are standardized to mean 0 and variance 1. Standard errors reported in parentheses are two-way clustered by state<sub>i</sub> and state<sub>j</sub>.

## G.4.2 Oil Discoveries as State Level Idiosyncratic Shocks

We construct another measure of state-level shocks using the discovery of new oil reserves ([U.S. Energy Information Administration \(1978-2000\)](#)). We construct three different measures of oil discovery. The first measure, *extensions*, measures the enlargement of reserves in existing reservoirs. The second measure, *new discoveries*, refers to the discovery of new reservoirs in old and new fields. The third measure, *all discoveries*, is the aggregate of the two measures – *extensions* and *new discoveries* in a state. These discoveries combine both onshore and offshore discoveries. We use the natural logarithm of one plus the magnitude of these discoveries as our measure of state-level shocks.

The magnitude of oil extensions and discoveries is measured using the number of barrels in millions. The majority of the oil discoveries occurred via *extensions* with an average discovery of 15 million barrels a year between 1978 and 2000, as compared with 8 million barrels a year of *new discoveries* during the same period. The *new discoveries* are a rare event relative to *extensions*. In terms of the geographic dispersion of these discoveries, Texas, Louisiana and New Mexico in the Southern region, experienced the largest oil discoveries during the period. The states of Illinois, Indiana, Ohio, and Michigan in the Midwest region experienced a modest degree of oil discoveries. California was the only western Pacific state to experience new oil reserves discovery during the period. See, Figure G.5 for the geographic distribution of these discoveries, and Figure G.6 for detailed summary statistics, the time series variation of oil discoveries.

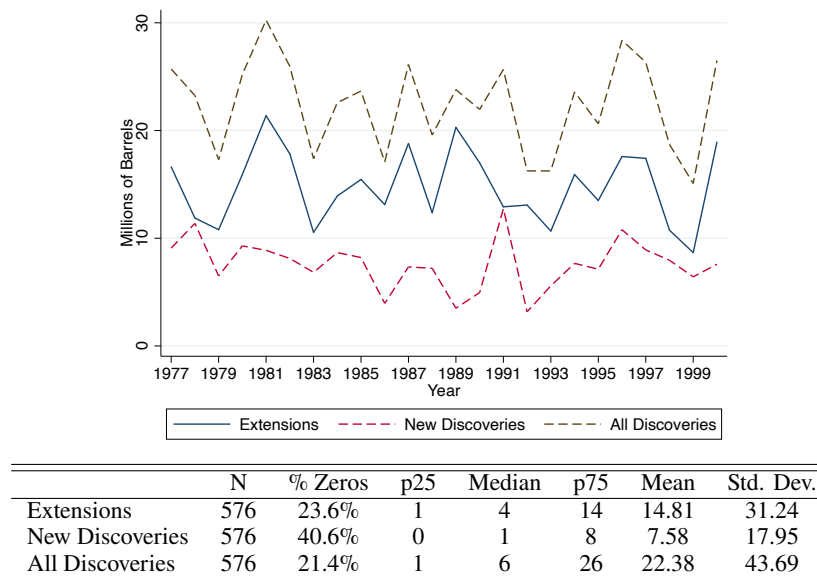
Figure G.5: Geographic Dispersion of Oil Discoveries (1977-2000)



The figure plots the geographic distribution of the average oil discoveries between 1977 and 2000 for all states that experienced at least one discovery or extension during the period. The first measure, *extensions*, measures the reserves enlargement in existing reservoirs. The second measure, *new discoveries*, refers to the discovery of new reservoirs in old and new fields. The third measure, *all discoveries*, is the aggregate of the two measures – *extensions* and *new discoveries* in a state. These measures combine both onshore and offshore discoveries. Each discovery is measured in million barrels.

Relative to our baseline shocks, oil discovery shocks are immune to geographic measurement error, and are relatively straightforward to comprehend. However, there are three limitations of these shocks. First, due to geological reasons, these shocks can be constructed for only a limited number of states. Second, these shocks are left-censored at zero and are always positive in nature. Third, the oil discovery shocks become more predictable towards the second half of the sample. We analyze the

Figure G.6: Oil Discovery: Summary Statistics & Average Over Time

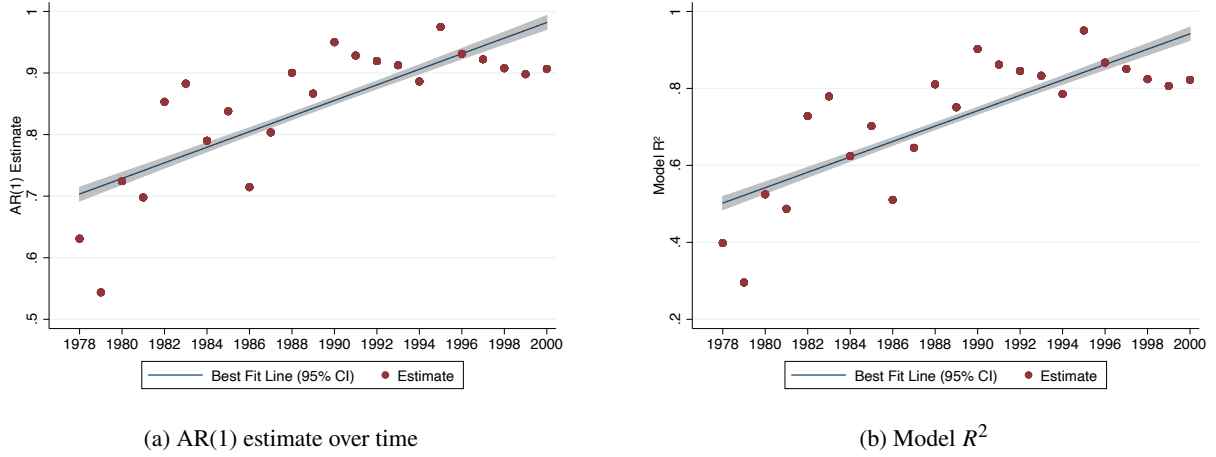


The figure plots the average oil discovery for each year between 1978 and 2000 for all states that experienced at least one discovery or extension during the period. The table reports the summary statistics - number of observations, percentage of data-points with no discoveries, first quartile, median, third quartile, mean, and standard deviation of observations for oil discoveries for the identical sample. We use three measures of oil discovery. The first measure, *extensions*, measures the reserves enlargement in existing reservoirs. The second measure, *new discoveries*, refers to the discovery of new reservoirs in old and new fields. The third measure, *all discoveries*, is the aggregate of the two measures - *extensions* and *new discoveries* in a state. These measures combine both onshore and offshore discoveries. Each discovery is measured in million barrels.

predictability of oil shocks and find that the predictability of oil shocks increases over time. We estimate the cross-sectional regression of oil discovery shock on its one period lag for each year between 1978 and 2000 and find that both the the model  $R^2$  and the AR(1) coefficient increase over time, see Figure G.7. Past oil discovery shocks provide insight into the oil endowment in that geography and facilitates learning about the geology of that area, making future discoveries more likely (Hamilton and Atkinson (2013)). However, under rational expectations, the predictability of the oil shocks only pushes the point estimate towards zero. Additionally, we control for previous period oil discoveries to account for the predictability of these shocks as in Arezki, Ramey and Sheng (2017).

Table G.13 replicates the baseline specification using oil discovery shocks. The oil discovery shocks measure banks' expectations of future economic growth in that state. The sample size is reduced as oil discovery shocks can be constructed for a selected sample of states due to natural geological reasons. Column (1), (2), (3) and (4) measure shocks in state  $j$  using our baseline idiosyncratic shocks, *extensions*, *new discoveries*, and *all discoveries*, respectively. The point estimate of the coefficient of the interaction term of oil discovery shocks and the *Post* variable is negative in all columns and comparable in magnitude to one another, as well as the baseline estimate. However, the point estimate is statistically insignificant for columns (2)-(4). The statistical insignificance of the estimates in column (2)-(4) is attributed to the loss in the power of the test due to the reduced sample size and small variation in the oil discovery shocks

Figure G.7: Predictability of Oil Shocks



The figure plots AR(1) estimates and the corresponding model  $R^2$  obtained from the cross-sectional regression of oil discovery shock on its one period lag. The cross-sectional regression is estimated for each period for a balanced sample between 1978 and 2000. The oil shock in state  $i$  at time  $t$  is defined as the natural logarithm of all discoveries plus one in state  $i$  at time  $t$ . All states that experienced at least one discovery or extension during the period 1977 and 2000 is included in the sample.

as there are a large number of zeros in the data. We provide a detailed power analysis in Figure G.8. The power analysis indicates that a sample size of  $\approx 30,000$  observations is required to have a 90% probability that we reject the null at 1% significance level when the magnitude of the effect is 0.016. By contrast, Table G.13 has  $\approx 22,000$  observations indicating a lack of power in the test given the sample size.

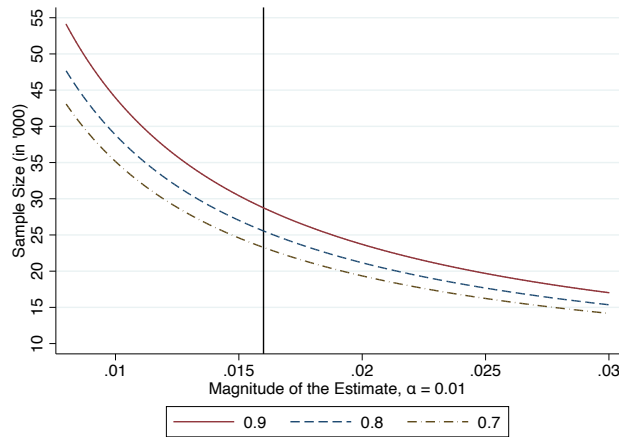
Despite the lack of power, the point estimates in column (2)-(6) are comparable to our baseline estimate of -0.016 and larger than the estimate of -0.010, estimated using baseline shocks for an identical sample. The larger magnitude of the point estimates using oil shocks relative to the baseline point estimates indicates that the geography-based measurement error attenuates the estimate in our baseline Table 3. This lends support to our argument that the geography-based measurement error is likely to bias our estimate towards finding an effect of lower magnitude.

Table G.13: Robustness - Measuring  $\Gamma$  Using Oil Discovery Shocks

|                                           | (1)                 | (2)                 | (3)                 | (4)                 |
|-------------------------------------------|---------------------|---------------------|---------------------|---------------------|
| $\Delta gdp_{it}$                         | Baseline            | Extensions          | New Disc.           | All Disc.           |
| $Post_{i,j,t} \times \Gamma_{j,t-1}$      | -0.0103<br>(0.0018) | -0.0132<br>(0.0212) | -0.0157<br>(0.0373) | -0.0364<br>(0.0382) |
| Post                                      | 0.0739<br>(0.0574)  | 0.1207<br>(0.0812)  | 0.0990<br>(0.0637)  | 0.0953<br>(0.0802)  |
| Past Exploration Control                  | No                  | Yes                 | Yes                 | Yes                 |
| Region <sub>i</sub> -Year FE              | Yes                 | Yes                 | Yes                 | Yes                 |
| State <sub>i</sub> -State <sub>j</sub> FE | Yes                 | Yes                 | Yes                 | Yes                 |
| State <sub>j</sub> -Year FE               | Yes                 | Yes                 | Yes                 | Yes                 |
| State <sub>i</sub> -Linear Trend          | Yes                 | Yes                 | Yes                 | Yes                 |
| $N$                                       | 21,850              | 21,850              | 21,850              | 21,850              |
| $R^2$                                     | 0.6688              | 0.6688              | 0.6688              | 0.6689              |

This table presents the estimates for baseline specification with alternative construction of  $\Gamma$  constructed using oil exploration shocks. The dependent variable is the change in the real GDP growth rate in percentage. The main independent variables are  $\Gamma_{j,t-1}^*$  which denotes the oil extension shocks in column (2), all discoveries including new field discoveries and new reservoirs in old fields in column (3), and, all extensions and discoveries in column (4). The baseline specification is reported in column (1) for comparison. Specifications (2-4) include a *Past Exploration Control* to control for all previous shocks in state  $j$ . This is used to control for possible serial correlation in oil discoveries (Arezki, Ramey and Sheng (2017)). The unit of observation in each regression is a state<sub>i</sub>-state<sub>j</sub>-year pair. All non-binary variables are standardized to mean 0 and variance 1. Standard errors reported in parentheses are two-way clustered by state<sub>i</sub> and state<sub>j</sub>.

Figure G.8: Oil Discovery: Power Analysis



The figure plots the iso-power curves with the required size of the sample on the Y axis and the magnitude of the effect of the X-axis. The iso-power curve gives the sample size, the required numbers of observations (in thousands), that would be required for adequately powered inference to not reject the null when the null is indeed false give the magnitude of the effect at a significance level. The iso-power curves are plotted for a significance level of 1% for power of 0.7, 0.8 and 0.9. The black line denotes the magnitude of the effect estimated from the baseline table.

## G.5 Addressing Concerns Related to Migration

This section presents two tests addressing the concern that the baseline result is not driven by interstate migration contemporaneous with the state pairwise banking deregulation. This section presents two tests to argue that the results discussed thus far are unlikely to be driven by within US migration.

In the first test we augment the baseline specification, equation 3, to include the  $\text{region}_i \times \text{region}_j \times \text{year}$  fixed effects, and  $\text{region}_i \times \text{state}_j \times \text{year}$ , where region refers to the BEA economic region of the state.<sup>30</sup> This test assumes that within US migration is likely to be smoothly distributed across space, i.e., the tendency to move between state  $i$  and state  $j$  are likely to be similar across other states in the same economic regions as state  $i$  and state  $j$ . Table G.14 reports these results. Column (1) estimate the baseline specification, equation 3, for reference. Column (2) and (3) augment the baseline specification with  $\text{region}_i \times \text{region}_j \times \text{year}$  fixed effects, and  $\text{region}_i \times \text{state}_j \times \text{year}$  respectively. The point estimate of the interaction term of  $\text{Post}_{i,j,t}$  and  $\Gamma_{j,t-1}^{ind}$  is negative and statistically significant at 1% level across all three columns indicating addition of these fixed effects have little impact on the magnitude and the significance of the estimate.

The second test, in contrast to the first test, assumes that choice set of within US migration is coarsely distributed across space. Under this setup, we randomly assign states into groups of different sizes and call these random groups as random regions and re-estimate the baseline specification with  $\text{random-region}_i \times \text{random-region}_j \times \text{year}$  fixed effects, and  $\text{random-region}_i \times \text{state}_j \times \text{year}$  fixed effects. We repeat this process of randomization of states into groups 3,500 times and estimate the distribution of the interaction term of the  $\text{Post}_{i,j,t}$  and  $\Gamma_{j,t-1}^{ind}$  while including the  $\text{random-region}_i \times \text{random-region}_j \times \text{year}$  fixed effects, and  $\text{random-region}_i \times \text{state}_j \times \text{year}$  fixed effects. Table G.16 reports the mean, median, standard deviation and t-statistic of the distribution of estimates. The mean and the median values reported in Table G.16 are negative with a small standard deviation. Moreover, a t-test of the estimates indicate that average of the distribution is less than zero. Hence, combining the results from these two tests we can rule out the results discussed in this paper are driven by within-US cross-state migration.

---

<sup>30</sup>We refer the readers to Table G.15 for the delineation of states into eight different economic regions by the Bureau of Economic Analysis.

Table G.14: Robustness - Addressing Migration Concerns Using Region Interaction Fixed Effects

| $\Delta gdp_{i,t}$                                | (1)                 | (2)                 | (3)                 |
|---------------------------------------------------|---------------------|---------------------|---------------------|
| $Post_{i,j,t} \times \Gamma_{j,t-1}^{ind}$        | -0.0164<br>(0.0007) | -0.0170<br>(0.0025) | -0.0208<br>(0.0060) |
| $Post_{i,j,t}$                                    | 0.0783<br>(0.0491)  | 0.0793<br>(0.0503)  | 0.0834<br>(0.0529)  |
| Region <sub>i</sub> -Year FE                      | Yes                 |                     |                     |
| State <sub>i</sub> -State <sub>j</sub> FE         | Yes                 | Yes                 | Yes                 |
| State <sub>j</sub> -Year FE                       | Yes                 | Yes                 |                     |
| State <sub>i</sub> -Linear Trend                  | Yes                 | Yes                 | Yes                 |
| Region <sub>i</sub> -Region <sub>j</sub> -Year FE |                     | Yes                 |                     |
| State <sub>j</sub> -Region <sub>i</sub> -Year FE  |                     |                     | Yes                 |
| $N$                                               | 57,700              | 57,700              | 57,700              |
| $R^2$                                             | 0.6583              | 0.6583              | 0.6594              |

This table reports the results from the estimation of baseline specification, in column (1), augmented to include Region<sub>i</sub>×Region<sub>j</sub>×Year fixed effects in column (2), and Region<sub>i</sub>×State<sub>j</sub>×Year fixed effects in column (3). The dependent variable is the change in the real GDP growth rate in percentage. The main independent variable is  $\Gamma_{j,t-1}^{ind}$  which denotes the idiosyncratic shocks in state  $j$  constructed by aggregating the Domar weighted labor productivity shocks of top 10 firms, by sales in state  $j$ . The unit of observation in each regression is a state<sub>i</sub>-state<sub>j</sub>-year pair. All non-binary variables used in the regression are standardized to mean zero and variance 1. Standard errors reported in parentheses are two-way clustered by state<sub>i</sub> and state<sub>j</sub>.

Table G.15: BEA Regions and their Constituents

| BEA Region     | States                                         |
|----------------|------------------------------------------------|
| New England    | CT, MA, ME, NH, RI, VT                         |
| Mideast        | NY, PA, MD, DC, DE, NJ                         |
| Great Lakes    | WI, IL, IN, OH, MI                             |
| Plains         | ND, SD, NE, KS, MO, IA, MN                     |
| Southeast      | VA, WV, KY, TN, AR, LA, MS, AL, GA, FL, SC, NC |
| Southwest      | OK, TX, NM, AZ                                 |
| Rocky Mountain | MT, ID, UT, WY, CO                             |

Table G.16: Robustness - Addressing Migration Concerns Using Random-Region Interaction Fixed Effects

| Panel A: Random-Region <sub>i</sub> ×Random-Region <sub>j</sub> ×Year FE |         |         |         |         |         |
|--------------------------------------------------------------------------|---------|---------|---------|---------|---------|
| # Groups                                                                 | 6       | 7       | 8       | 9       | 10      |
| # Simulation                                                             | 3,500   | 3,500   | 3,500   | 3,500   | 3,500   |
| Median                                                                   | -0.0121 | -0.0119 | -0.0118 | -0.0117 | -0.0115 |
| Mean                                                                     | -0.0120 | -0.0118 | -0.0118 | -0.0116 | -0.0115 |
| St Dev                                                                   | 0.0037  | 0.0040  | 0.0043  | 0.0046  | 0.0049  |
| t-statistic                                                              | 190.00  | 170.00  | 160.00  | 150.00  | 140.00  |

| Panel B: Random-Region <sub>i</sub> ×State <sub>j</sub> ×Year FE |         |         |         |         |         |
|------------------------------------------------------------------|---------|---------|---------|---------|---------|
| # Groups                                                         | 6       | 7       | 8       | 9       | 10      |
| # Simulation                                                     | 3,500   | 3,500   | 3,500   | 3,500   | 3,500   |
| Median                                                           | -0.0132 | -0.0131 | -0.0130 | -0.0132 | -0.0131 |
| Mean                                                             | -0.0131 | -0.0132 | -0.0131 | -0.0131 | -0.0131 |
| St Dev                                                           | 0.0042  | 0.0046  | 0.0050  | 0.0054  | 0.0058  |
| t-statistic                                                      | 190.00  | 170.00  | 150.00  | 140.00  | 130.00  |

This table reports the mean, median, standard deviation and t-statistic for the distribution of the interaction term of  $\text{Post}_{i,j,t}$  and  $\Gamma_{j,t-1}^{\text{ind}}$  from the estimation of baseline specification augmented to include Random-Region<sub>i</sub>×Random-Region<sub>j</sub>×Year fixed effects in panel a, and Random-Region<sub>i</sub>×State<sub>j</sub>×Year fixed effects in panel b. The dependent variable is the change in the real GDP growth rate in percentage. The main independent variable is  $\Gamma_{j,t-1}^{\text{ind}}$  which denotes the idiosyncratic shocks in state  $j$  constructed by aggregating the Domar weighted labor productivity shocks of top 10 firms, by sales in state  $j$ . The unit of observation in each regression is a state<sub>i</sub>-state<sub>j</sub>-year pair. All non-binary variables used in the regression are standardized to mean zero and variance 1. We randomly allocate states into groups and run the baseline specification with Random-Region<sub>i</sub>×Random-Region<sub>j</sub>×Year and Random-Region<sub>i</sub>×State<sub>j</sub>×Year fixed effects. We repeat this randomization 3,500 times and estimate the coefficient of the interaction term of  $\text{Post}_{i,j,t}$  and  $\Gamma_{j,t-1}^{\text{ind}}$  in each simulation. Panel a and b report the mean, median, standard deviation and t-statistic of the 3,500 values of these estimates. The columns report the number of groups into which the 50 states and DC have been grouped into.

## G.6 Dropping the states of South Dakota and Delaware

This section reports the estimation results of the baseline specification, equation 3 after dropping the states of South Dakota and Delaware from the sample. We drop these states as they had an explicit focus on attracting the credit card companies during the sample period. Table G.17 reports the results from the alternative sample. Column (1) reports the baseline regression with full sample for reference. Column (2) drops the states of South Dakota and Delaware from the set of state  $i$  while column (3) drops these states from the set of  $state_j$ . Lastly, column (4) drops the two states from both state  $i$  or state  $j$ . The results indicate the stability of the magnitude and the statistical significance of the estimate of interest across the four columns indicating the results are unlikely to be driven by the inclusion of the states of South Dakota and Delaware.

Table G.17: Robustness - Removing South Dakota & Delaware from the Sample

| $\Delta gdp_{i,t}$                         | (1)                 | (2)                          | (3)                          | (4)                             |
|--------------------------------------------|---------------------|------------------------------|------------------------------|---------------------------------|
| $Post_{i,j,t} \times \Gamma_{j,t-1}^{ind}$ | -0.0164<br>(0.0007) | -0.0153<br>(0.0034)          | -0.0180<br>(0.0003)          | -0.0167<br>(0.0019)             |
| $Post_{i,j,t}$                             | 0.0783<br>(0.0491)  | 0.0685<br>(0.0493)           | 0.0750<br>(0.0487)           | 0.0652<br>(0.0489)              |
| Region $_i$ -Year FE                       | Yes                 | Yes                          | Yes                          | Yes                             |
| State $_i$ -State $_j$ FE                  | Yes                 | Yes                          | Yes                          | Yes                             |
| State $_j$ -Year FE                        | Yes                 | Yes                          | Yes                          | Yes                             |
| State $_i$ -Linear Trend                   | Yes                 | Yes                          | Yes                          | Yes                             |
| $N$                                        | 57,700              | 55,438                       | 55,400                       | 53,184                          |
| Sample                                     | Full Sample         | -{SD & DE}<br>from state $i$ | -{SD & DE}<br>from state $j$ | -{SD & DE}<br>from state $i, j$ |
| $R^2$                                      | 0.6583              | 0.6618                       | 0.6583                       | 0.6618                          |

This table reports the results from the estimation of baseline specification after dropping the states of South Dakota and Delaware from the sample. Column (1) uses the full sample, column (2), and (3) drop the states of South Dakota (SD) and Delaware (DE) from state  $i$  and  $j$  respectively, and column (4) drops the two states from both state  $i$  and  $j$ . The dependent variable is the change in the real GDP growth rate in percentage. The main independent variable is  $\Gamma_{j,t-1}^{ind}$  which denotes the idiosyncratic shocks in state  $j$  constructed by aggregating the Domar weighted labor productivity shocks of top 10 firms, by sales in state  $j$ . The unit of observation in each regression is a state $_i$ -state $_j$ -year pair. All non-binary variables used in the regression are standardized to mean zero and variance 1. Standard errors reported in parentheses are two-way clustered by state $_i$  and state $_j$ .

## Appendix References

- Arezki, Rabah, Valerie A Ramey, and Liugang Sheng.** 2017. “News shocks in open economies: Evidence from giant oil discoveries.” *The Quarterly Journal of Economics*, 132(1): 103–155.
- Berger, Allen N, Anil K Kashyap, and Joseph M Scalise.** 1995a. “Data for: The transformation of the US banking industry: What a long, strange trip it’s been.” *Brookings Papers on Economic Activity*. [https://www.brookings.edu/wp-content/uploads/1995/06/1995b\\_bpea\\_berger\\_kashyap\\_scalise\\_gertler\\_friedman.pdf](https://www.brookings.edu/wp-content/uploads/1995/06/1995b_bpea_berger_kashyap_scalise_gertler_friedman.pdf), 1995(2): 55–218.
- Berger, Allen N, Anil K Kashyap, and Joseph M Scalise.** 1995b. “The transformation of the US banking industry: What a long, strange trip it’s been.” *Brookings Papers on Economic Activity*, 1995(2): 55–218.
- Hamilton, Kirk, and Giles Atkinson.** 2013. *Resource discoveries, learning, and national income accounting*. The World Bank.
- Kalemli-Ozcan, Sebnem, Elias Papaioannou, and Fabrizio Perri.** 2013. “Global banks and crisis transmission.” *The Journal of International Economics*, 89(2): 495–510.
- Landier, Augustin, David Sraer, and David Thesmar.** 2017. “Banking integration and house price co-movement.” *The Journal of Financial Economics*, 125(1): 1–25.
- Mian, Atif, Amir Sufi, and Emil Verner.** 2020. “How does credit supply expansion affect the real economy? the productive capacity and household demand channels.” *The Journal of Finance*, 75(2): 949–994.
- Morgan, Donald P, Bertrand Rime, and Philip E Strahan.** 2004. “Bank integration and state business cycles.” *The Quarterly Journal of Economics*, 119(4): 1555–1584.
- Stein, Jeremy C.** 1997. “Internal capital markets and the competition for corporate resources.” *The Journal of Finance*, 52(1): 111–133.
- U.S. Energy Information Administration.** 1978-2000. “Data for: U.S. Crude Oil and Natural Gas Proved Reserves.” *U.S. Department of Energy, Energy Information Administration*. <https://www.eia.gov/naturalgas/crudeoilreserves/>, Annual data on extensions and new discoveries.