# Many Markets Make Good Neighbors: Multimarket Contact and Deposit Banking<sup>\*</sup>

John William Hatfield McCombs School of Business University of Texas at Austin Jonathan Wallen Harvard Business School Harvard University

July 31, 2023

#### Abstract

We investigate how competition in local markets depends on the degree of multimarket contact between firms in that market. We show in a simple theoretical model how such overlapping relationships across markets can lead to less competitive behavior by firms. We then demonstrate that multimarket contact has decreased competitive behavior by banks in the U.S. deposit banking market. Our results help to resolve the tension between increasing markups and decreasing local concentration—both long-term trends documented in the literature—as these trends have coincided with a three-fold increase in multimarket contact across a broad set of U.S. retail firms.

JEL Classification: C73, D21, D22, G21, L13, L41

Keywords: Antitrust, Deposit Banking, Market Power, Multimarket Contact

<sup>\*</sup>The authors thank Adi Sunderam, Panle Jia Barwick, Lauren Cohen, Dean Corbae, Janet Gao, John Griffin, Valentin Haddad, Victoria Ivashina, Alexi Savov, Andrei Shleifer and seminar audiences at Harvard Business School, the University of Southern California, the University of Texas, the Federal Reserve Board, the Red Rock Finance Conference, the SITE 2022 Conference, and the AFA 2022 Annual Meeting for helpful comments. Any comments or suggestions are welcome and may be emailed to john.hatfield@utexas.edu and jwallen@hbs.edu.

# 1 Introduction

Scholars and policymakers have become increasingly concerned that firms have enhanced their profits margins by behaving less competitively: Hall (2018) and De Loecker et al. (2020) showed that firms' markups have increased substantially; Gutiérrez and Philippon (2017) demonstrated that firm investment has fallen; Barkai (2020) found that large firms have become more profitable compared to those of the 1980s; and Autor et al. (2020) chronicled the decline of wages over the same period. Given this body of evidence, Covarrubias et al. (2020) suggested that increasing market concentration at the national level and heightened barriers to entry have led to these outcomes.

Yet Rossi-Hansberg et al. (2021) showed that, while industry concentration has increased at the national level, industry concentration has been *decreasing* at the local level for the past 25 years;<sup>1</sup> Hsieh and Rossi-Hansberg (2021) provided evidence for a similar dynamic taking place in services. Moreover, Rossi-Hansberg et al. (2021) argued that, since a good in one local market is (in general) not a good substitute for the same good in a different local market, the local level of concentration is the relevant measure of the competitive environment. Thus, they conclude that the falling local concentrations they documented "likely [lead to] a more competitive environment."

We propose that multimarket contact can explain the puzzling dichotomy that firms behave *less* competitively even though there are *more* local competitors. We show that *multimarket contact* among retail firms has increased as such firms have expanded their geographic reach, resulting in fewer firms nationwide and yet little change in local product market concentrations. Building on the work of Bernheim and Whinston (1990), we show in a general model that increases in multimarket contact lead firms to behave less competitively. We test our theory using the U.S. deposit banking market as a laboratory to obtain a well-identified estimate of the effect of multimarket contact on competition and pricing. We find that multimarket contact enables banks to behave as if the local market were twice as

<sup>&</sup>lt;sup>1</sup>Benkard et al. (2021) showed that this trend is even stronger when one properly defines product markets.

concentrated as it really is.

Our work thus has two main contributions:

First, we construct a general model of how overlapping relationships can lead to less competitive behavior. As in the seminal work of Bernheim and Whinston (1990), in our model firms can use "slack"—i.e., strong incentives to collude in highly concentrated markets—to sustain collusion across other markets, resulting in a less competitive environment overall.<sup>2</sup> Our Theorem 1 generalizes this idea to show that mergers—even market extension mergers, in which no local market becomes more concentrated—generally lead to worse consumer outcomes.<sup>3</sup>

Our model also shows that within an equilibrium in which firms coordinate across markets, markups are positively correlated with multimarket contact across local markets (Theorem 3). We also show that local market concentration still matters in equilibria in which multimarket contact plays a role: higher markups are positively correlated with higher local market concentrations as well (Theorem 2). This bridges a gap between theory and empirics: Prior theory had shown that multimarket contact may admit the existence of more collusive equilibria, but prior empirical work had shown a positive correlation between local markups and local multimarket contact within an equilibrium. Theorem 3 thus provides a justification for empirically identifying the effects of multimarket contact on competition from cross-sectional variation in local markets.

Second, we empirically show that multimarket contact induces banks to act less competitively in the deposit market and document that multimarket contact is associated with less competitive behavior for a broad set of retail industries. In particular, markups have increased while local concentration has not for retail industries: Following the methodology of De Loecker et al. (2020), we estimate that markups have increased by 27 percent for retail

<sup>&</sup>lt;sup>2</sup>Bernheim and Whinston (1990) were the first to formalize the concept of "mutual forbearance," an idea described much earlier by Edwards (1955).

<sup>&</sup>lt;sup>3</sup>Thus, our Theorem 1 shows that the example of Section 4 of the work of Bernheim and Whinston (1990) holds in general: Market extension mergers will never make a market more competitive and may make it less competitive.

industries over the last three decades.<sup>4</sup> However, local establishment Herfindahl-Hirschman Index (HHI) has remained near constant. We show that multimarket contact can resolve this puzzle: For a pair of large firms in the same industry, the propensity for their retail networks to overlap has more than tripled over the last three decades. Consistent with our theory, these trends suggest that multimarket contact has dampened local competitive pressure.

To better identify the effect of multimarket contact on competition, we use the U.S. deposit banking market as a laboratory. The U.S. deposit banking market exhibits the same time-trends as retail industries: multimarket contact has increased three-fold, while local market concentration has remained essentially unchanged. Moreover, deposit markets have (i) substantial cross-sectional variation in multimarket contact across many local markets<sup>5</sup> and (ii) well-identified cost shocks as well as detailed data on prices and quantities.

Drechsler et al. (2017) showed how changes in the Fed Funds rate provide a well-identified cost shock by which to measure passthrough and competition. Passthrough rates of interest rate changes have fallen from 36% in 2001-2006 to 4% from 2010-2020. Figure 1 plots the Fed Funds rate, the deposit savings rate, and the bank passthrough rate from 2001 through 2020.<sup>6</sup> This decrease in passthrough implies a significant increase in bank market power in setting deposit rates.

In a competitive market, we expect a passthrough rate of 100 percent; in an imperfectly competitive market, we expect passthrough rates to be lower. For counties in which multimarket contact is low, we find that banks pass through 23 bps of a 100 bps increase in the Fed Funds rate. But for counties in which banks have a large degree of multimarket contact, passthrough falls to 11 bps of a 100 bps increase in the Fed Funds rate.

<sup>&</sup>lt;sup>4</sup>Using data from Compustat on publicly traded firms, we estimate that markup—defined as price over cost—has increased from 1.31 in 1989 to 1.40 in 2021.

<sup>&</sup>lt;sup>5</sup>The deposit market is local for consumers but regional or national for many competing banks. Honka et al. (2017) showed that 84% of consumers prefer banks with branches within 5 miles of their home. But, in 2020, the three largest depository banks held 32% of deposits. Essentially, regional and national banks compete with one another across many local deposit markets.

<sup>&</sup>lt;sup>6</sup>We measure the deposit savings rate as the dollar weighted average rate by bank branches. We use branch-level rate data from RateWatch and branch-level deposits data from the FDIC Survey of Deposits. Our sample focuses on regional/national banks. See Section 4 for further details.

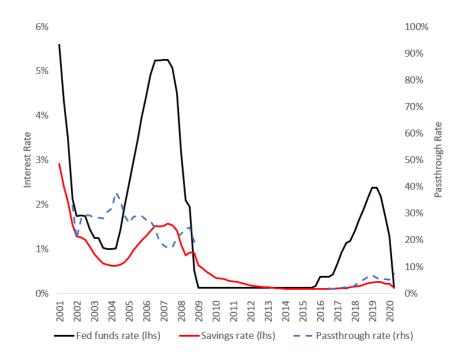


Figure 1: On the left vertical axis is the Fed funds rate  $(FF_t)$  and the deposit savings rate  $(y_{i,t})$ . On the right vertical axis is the passthrough rate  $(\beta)$ , which we estimate for each quarter using a 3-year rolling window of branch *i* and quarter *t* rates  $(y_{i,t})$ :  $\Delta y_{i,t} = \alpha + \beta \Delta FF_t + \epsilon_{i,t}$ .

The primary identification concern of our across-bank-branches estimates is time variation in lending opportunities (Drechsler et al., 2017). To address this concern, we use cross-county variation in multimarket contact and a plethora of fixed effects. In particular, we include bank-by-year fixed effects to capture bank-specific changes in lending opportunities. Our identifying assumption is that deposit funding is fungible across the bank (Gilje et al., 2016); that is, there are bank-specific, but not branch-specific, lending opportunities.

We find that, for the same bank, its branches in counties with higher multimarket contact behave less competitively than its branches in counties with lower multimarket contact. The economic magnitude of the effect of multimarket contact on competition is similar in magnitude to the effect of local market concentration. We estimate that if each local market were an island, then concentration would have to nearly double to maintain the same (lack of) competitiveness as measured by passthrough rates (i.e., the local market HHI would have to increase from 0.21 to 0.38). That is, multimarket contact enables banks to behave as if the local market was twice as concentrated as it really is. Deposit markets are economically important in their own right: as of 2020, commercial banks in the United States held \$15 trillion of domestic deposits, which is 14 percent of all U.S. household financial assets. Meanwhile, **Begenau and Stafford (2019)** documented that bank returns are primarily driven by cheap funding from deposits, not a competitive advantage in lending: Over the past 20 years, deposit savings rates have averaged 0.61 percent despite an average Fed Funds rate of 1.52 percent. This large deposit spread is due, in part, to bank market power in setting deposit rates.

The remainder of the paper is as follows: Section 2 relates our work to the literature on local and national competition as well as collusive behavior in banking and other industries. Section 3 presents our theoretical model of how overlapping relationships may lead to less competitive behavior by banks in deposit markets. Section 4 introduces the data and empirical measures of deposit market competition. Section 5 identifies the empirical effect of multimarket contact on competition. Section 6 concludes. Proofs and other material are presented in the appendices.

# 2 Anticompetitive Behavior in Banking and Other Industries

Local bank deposit markets are not perfectly competitive. Drechsler et al. (2017) found that passthrough rates are low and decreasing in local market concentration. Moreover, Granja and Paixão (2019) documented that mergers increase the uniformity of deposit product pricing. This uniformity of pricing across heterogeneous markets is indicative of noncompetitive behavior.<sup>7</sup> And Corbae and D'Erasmo (2020, 2021) have shown in a structural model that bank regulation has increased concentration by decreasing entry and increasing lending by large banks.

<sup>&</sup>lt;sup>7</sup>Our empirical analysis focuses on large regional and national banks, for which we have more variation in within-bank pricing compared to that of local banks with few branches.

Bank merger regulation has focused on restricting mergers that increase local market concentration, but not market extension mergers that increase multimarket contact (Federal Reserve, 2014). Regional and national banks have acquired about 30,000 bank branches from 2001 to 2020. These acquisitions have contributed to multimarket contact increasing more than three-fold. Our findings emphasize the anti-competitive effects of these market extension mergers, which have driven the national consolidation of the U.S. deposit banking market. These anti-competitive effects have recently drawn the scrutiny of regulators, such as Assistant Attorney General for Antitrust Jonathan Kanter, who stated that the Antitrust Division henceforth "will closely scrutinize mergers that increase risks associated with coordinated effects and multi-market contacts."

Our evidence of collusive behavior among banks in deposit markets is consistent with a pattern of collusive behavior by banks in many other asset markets. Salient examples of this collusion include the manipulation of LIBOR—an important interest rate which served as the benchmark for trillions of dollars of contracts (Duffie and Stein, 2015)—and the manipulation of foreign exchange rates (Jahanshahloo and Cai, 2019). In lending, banks have been shown to exhibit behavior consistent with collusion: Cai et al. (2020) showed this in the context of syndicated lending networks and Chan et al. (2021) showed this for loan pricing conventions. More broadly, segmentation in financial markets and imperfect competition contribute to law of one price violations in and across many financial markets (Siriwardane et al., 2021; Wallen, 2021).

Moreover, collusive behavior associated with multimarket contact has been widely documented in a wide range of industries: Evans and Kessides (1994) and Ciliberto and Williams (2014) found that airlines price less competitively when more of their flight networks overlap. Busse (2000) found that cellular telephony providers also coordinate pricing behavior across markets, resulting in prices that are 7–10% higher for consumers; Parker and Röller (1997) also demonstrated that multimarket contract facilitated higher prices in such markets. Similarly, Fernández and Marín (1998) found that multimarket contact in the hotel industry increases prices charged to consumers. Khwaja and Shim (2017) demonstrated that multimarket contact facilitated higher retail prices for lumber. Meanwhile, Schmitt (2018) determined that multimarket contact generated by consolidation in the hospital industry has led to higher prices. And Jans and Rosenbaum (1996) found analogous results in the cement industry. Indeed, in their report to the Directorate-General for Competition on tacit collusion, Ivaldi et al. (2003) postulated that multimarket contact facilitated such collusive behavior.<sup>8,9</sup> We contribute to this literature by providing a theoretical justification for the common empirical test that markets with more multimarket contact have less competitive prices. Empirically, we contribute to this literature by providing better identified evidence for the effects of multimarket contact in the deposit banking market by using well-identified cost shocks. More broadly, we also show that multimarket contact has increased across retail industries, while local concentration has decreased and markups have increased.

# 3 Theory

#### 3.1 Framework

#### 3.1.1 Market Structure

We construct a model of firm competition across multiple markets; we do this by extending the canonical Bertrand competition model to a multi-market setting in which some firms are present in multiple markets. There is a finite set of markets M and a finite set of firms F. The market structure  $\kappa \in \{0, 1\}^{F \times M}$  denotes which firms are in each market: we set  $\kappa_m^f = 1$ if firm f is present in market m and set  $\kappa_m^f = 0$  otherwise. We let  $\mathbf{F}(m;\kappa)$  be the set of firms present in market m, i.e.,  $\mathbf{F}(m;\kappa) \equiv \{f \in F : \kappa_m^f = 1\}$ . A firm f is national if it is present in more than one market, i.e.,  $\sum_{m \in M} \kappa_m^f > 1$ ; we denote the set of national firms present in

<sup>&</sup>lt;sup>8</sup>In a similar vein, Duso et al. (2014) found that research joint ventures between competitors—ostensibly to improve efficiency—result in lower market shares for venture participants and higher prices for consumers.

<sup>&</sup>lt;sup>9</sup>By contrast, Eizenberg et al. (2020) did not find evidence that multimarket contact across Israeli grocery segments facilitated higher prices. However, in their setting, both segments were competitive; in such a setting, our theory does not predict that multimarket contact will affect competitive dynamics.

market m as  $\mathbf{N}(m; \kappa)$ . Conversely, a firm f is *local* if it is present in exactly one market, i.e.,  $\sum_{m \in M} \kappa_m^f = 1$ ; we denote the set of local firms in market m as  $\mathbf{L}(m; \kappa)$ . When the market structure  $\kappa$  is clear from context, we will sometimes drop  $\kappa$  from the notation and just write  $\mathbf{F}(m)$ ,  $\mathbf{N}(m)$ , and  $\mathbf{L}(m)$ .

If, under the market structure  $\kappa$ , firm f acquires firm  $\hat{f}$ , it generates a new market structure  $\hat{\kappa}$  under which:

- 1. Firm f is now present in all markets in which f or  $\hat{f}$  were formerly present, i.e.,  $\hat{\kappa}_m^f = \max\{\kappa_m^f, \kappa_m^{\hat{f}}\}$  for all  $m \in M$ .
- 2. The firm  $\hat{f}$  is no longer present in any market, i.e.,  $\hat{\kappa}_m^{\hat{f}} = 0$  for all  $m \in M$ .
- 3. Each other firm is present in the same markets as before, i.e.,  $\kappa_m^{\bar{f}} = \hat{\kappa}_m^{\bar{f}}$  for all  $\bar{f} \in F \smallsetminus \{f, \hat{f}\}$  and all  $m \in M$ .

In this case, we say that  $\hat{\kappa}$  is a *merger* under  $\kappa$ . We say that a merger  $\hat{\kappa}$  under  $\kappa$  is a *market* extension merger if firms f and  $\hat{f}$  were not both present in any market before the merger, i.e., for all  $m \in M$ , either  $\kappa_m^f = 0$  or  $\kappa_m^{\hat{f}} = 0$ .

#### 3.1.2 The Stage Game

In each market m, each firm  $f \in \mathbf{F}(m)$  simultaneously chooses a price  $p_m^f \in [0, \infty)$  and an aggressiveness  $a_m^f \in [0, \infty)$ . Consumers in each market observe prices and then choose a firm with the lowest price; the more aggressive a firm is, the more likely a consumer will choose it. We allow each firm to choose its aggressiveness so that each firm may effectively choose its quantity (if it knows the aggressivenesses of other firms); note, however, that a firm can always choose to be more aggressive at no cost in order to increase its demand.<sup>10</sup>

<sup>&</sup>lt;sup>10</sup>Hatfield and Lowery (2023) considered an extension of the canonical model of undifferentiated Bertrand competition in which firms choose aggressiveness and so endogenously allocate market shares. We use a similar model here as endogenously allocating market shares will be key to our analysis of the repeated game because each firm's share will need to depend on the market structure, i.e., on the entire set of multimarket relationships between firms.

We denote the set of firms with the lowest price in market m—i.e., the firms active in market m—as  $\mathbf{A}_m(p_m) \equiv \{f \in F : p_m^f = \min_{\bar{f} \in \mathbf{F}(m)}\{p_m^{\bar{f}}\}\};^{11}$  we call these firms active as they are the only firms that have positive market share. The quantity of firm f in market mis thus given by<sup>12</sup>

$$Q_m^f(p_m, a_m) \equiv \psi_m D\left(\min_{\bar{f} \in F} \left\{ p_m^{\bar{f}} \right\} \right) \times \mathbb{1}_{\left\{ f \in \mathbf{A}_m(r_m) \right\}} \frac{a_m^b}{\sum_{\bar{f} \in \mathbf{A}_m(r_m)} a_m^{\bar{f}}};$$

here, D(p) is a smooth, strictly decreasing, and concave demand function and  $\psi_m$  is the size of market m.<sup>13</sup> Note that firm f's demand is 0 unless it is offering the price most favorable to consumers; if it is offering that price, then its demand depends on its aggressiveness relative to other firms. By choosing a higher aggressiveness, a firm competes more fiercely and leaves less residual demand for other firms in that market.

The profits of firm f in market m are<sup>14</sup>

$$\Pi_m^f(r_m, a_m) \equiv \underbrace{Q_m^f(p_m, a_m)}_{\substack{\text{Consumer} \\ \text{Demand}}} \underbrace{(p_m^f - c)}_{\substack{\text{Profits per} \\ \text{consumer}}}.$$

Finally, firms may operate in more than one market, and so a firm f's total profits are

<sup>11</sup>Throughout, for a matrix  $z \in \mathbb{R}^{F \times M}$ , we let  $z_m$  be the vector of values of z in market m, i.e.,  $z_m \equiv (z_m^f)_{f \in F}$ . <sup>12</sup>The indicator function  $\mathbb{1}_{\{\mathfrak{p}\}}$  is 1 if  $\mathfrak{p}$  is true and 0 otherwise. <sup>13</sup>In the special case in which  $\sum_{\bar{f} \in \mathbf{A}_m(r_m)} a_m^{\bar{f}} = 0$ , we define

$$Q_m^f(p_m, a_m) \equiv \psi_m \mathbb{1}_{\{f \in \mathbf{A}_m(r_m)\}} \frac{1}{|\mathbf{A}_m(r_m)|}$$

<sup>14</sup>Essentially, in our model firms engage in undifferentiated Bertrand competition (with the ability to allocate market shares in a given market by appropriately choosing aggressivenesses). This is in contrast to the structural empirical industrial organization literature in which firms typically engage in differentiated Bertrand (or Cournot) competition; however, in those works the authors simply assume that after a deviation prices revert to the static Bertrand-Nash equilibrium prices instead of solving for the Abreu (1988) optimal penal codes. Thus, those works do not use the most effective punishment strategies available to firms and so do not consider many of the subgame-perfect Nash equilibrium of the repeated game. Examples of this approach can be found in the work of Eizenberg et al. (2020), Igami and Sugaya (2022), and Starc and Wollmann (2022).

given by

$$\Pi^f(p,a) \equiv \sum_{m \in M} \Pi^f_m(p_m, a_m)$$

#### 3.1.3 The Repeated Game

In each period  $t \in \mathbb{W} = \{0, 1, 2, ...\}$ , firms play the stage game. Firms have a common discount factor  $\delta$ , and so a firm f's total profits are given by  $\sum_{t=0}^{\infty} \delta^t \Pi^f(p(t), a(t))$  where p(t)(a(t)) is the matrix of prices (aggressivenesses) for each firm in each market in period t.

We say that a vector of prices  $(p_m)_{m \in M}$  is *sustainable* if there exists a subgame-perfect Nash equilibrium of our repeated game in which the price vector  $(p_m)_{m \in M}$  is realized each period.

#### 3.1.4 The Monopoly and Competitive Interest Rates

In our setting, the *competitive price* is simply the cost c. We can also calculate that a monopolist in market m would choose the *monopoly price*  $p^{\circ}$  which is the unique solution to

$$\{p^{\circ}\} = \arg\max\{(p_m - c)\psi_m D(p_m)\}.$$

#### 3.1.5 Conditions on Market Structure

We say that the market structure  $\kappa$  is sufficient for competition in market m if  $|\mathbf{F}(m)| > 1$ . The market structure  $\kappa$  is sufficient for competition if it is sufficient for competition in each market  $m \in M$ .

### **3.2** Bertrand Competition in the Stage Game

We first analyze the stage game. We show that the market is competitive—in the sense that each consumer enjoys a price of c—so long as the market structure is sufficient for competition.

**Proposition 1.** Suppose that the market structure  $\kappa$  is sufficient for competition. Then each firm obtains 0 profits in every pure-strategy Nash equilibrium of the stage game and such an equilibrium exists.

The intuition for this result is the same as in the standard one-market Bertrand competition setting. We prove the proposition by way of contradiction: If any firm in market m has positive profits, then every firm f in market m must have positive profits as otherwise f could become profitable by choosing the lowest price offered by any other firm and a positive aggressiveness. But if every firm is profitable, then every firm f is offering the same price  $p_m^f > c$ ; but then some firm could slightly increase its aggressiveness to increase its profitability, contradicting the assertion that the original strategy profile was a Nash equilibrium. One simple pure-strategy equilibrium which delivers 0 profits to each firm is for each firm f to set its price  $p_m^f = c$  (i.e., the competitive price) and the same aggressiveness in each market m.

Also note that Proposition 1 implies that there exists a subgame-perfect Nash equilibrium of the repeated game in which each firm obtains 0 profits each period. Such a "price war" equilibrium will be key in our analysis of the repeated game: Since 0 is the lowest individually rational payoff for each firm, reverting to the "price war" equilibrium in every period after a deviation punishes the deviator as harshly as possible; that is, the "price war" equilibrium is an optimal penal code (in the sense of Abreu (1988)) for every firm.

### 3.3 An Economy with One Market

Before considering our multimarket setting, we first analyze the standard case in which there is only one market.

**Proposition 2.** If  $|\mathbf{F}(m)| \leq \frac{1}{1-\delta}$  then any price  $p \in [c, p^{\circ}]$  is sustainable; if  $|\mathbf{F}(m)| > \frac{1}{1-\delta}$  then only p = c is sustainable.

To prove Proposition 2, we first note that after any deviation from the equilibrium strategy

profile, the harshest punishment possible is the 0-profit equilibrium of the stage game of Proposition 1. We then show that in any equilibrium each firm is offering the same price. Letting  $q_m^f$  be the *quantity* of demand that firm f obtains, we can thus characterize the set of sustainable prices as any solution to two constraints:

1. Each firm  $f \in F$  weakly prefers its profits along the equilibrium path to any deviation, i.e.,

$$\frac{1}{1-\delta}(p_m-c)q_m^f \ge (p_m-c)\psi_m D(p_m) \tag{1}$$

2. The total quantity allocated to all firms is equal to the demand at the price p, i.e.,

$$\sum_{f \in F} q_m^f = \psi_m D(p_m).$$
<sup>(2)</sup>

A given quantity vector  $q_m$  can be implemented by each firm f choosing an aggressiveness  $a_m^f = q_m^f$ .

Constraint (1) codifies that each firm is better off offering a price of  $p_m$  and its prescribed aggressiveness rather than increasing its aggressiveness to capture more demand.<sup>15</sup> Note that a firm expects 0 future profits after any deviation, since firms expect to simply play the 0-profit stage-game equilibrium of Proposition 1 after any deviation. Constraint (2) is simply an "adding up" constraint: the total quantity allocated to the firms should equal the total demand.

Summing constraint (1) over all firms, and combining it with constraint (2) yields

$$\frac{1}{1-\delta}(p_m-c)\psi_m D(p_m) \ge (p_m-c)\psi_m D(p_m)|\mathbf{F}(m)|.$$

If  $|\mathbf{F}(m)| > \frac{1}{1-\delta}$ , this expression can only be satisfied if  $p_m = c$ ; otherwise, it can be satisfied

 $<sup>^{15}</sup>$ We only consider prices less than or equal to the monopoly price, and so the optimal deviation is to choose an arbitrarily high aggressiveness and not to cut the price.

for any price  $p_m \in [c, p^\circ]$ .

## **3.4** The Multimarket Economy

Using an argument analogous to that for a single-market economy, we can show (assuming that the market structure is sufficient for competition) that prices  $(p_m)_{m \in M}$  and quantities  $(q_m^f)_{m \in M, f \in F}$  can be sustained in equilibrium if:

1. For each firm  $f \in F$ ,

$$\frac{1}{1-\delta} \sum_{m \in M} (p_m - c) q_m^f \ge \sum_{m \in M} (p_m - c) \psi_m D(p_m).$$
(3)

2. For each  $m \in M$ , we have that  $q_m^f = 0$  if  $f \notin \mathbf{F}(m)$  and

$$\sum_{f \in F} q_m^f = \psi_m D(p_m). \tag{4}$$

Here,  $p_m$  is now the lowest price offered in market m; a quantity vector  $q_m$  can be implemented by each firm f choosing  $p_m^f = p_m$  and  $a_m^f = q_m^f$ .

Constraint (3) codifies that each firm is better off offering  $p_m$  and its prescribed aggressiveness in each market m rather than increasing its aggressiveness and filling total consumer demand in each market in which it is present; this corresponds to constraint (1) in the one-market case. It is key to our analysis that constraint (3) sums over all markets; firm f may be willing to accept a very small quantity in a given market m if it is obtaining substantial profits in other markets. Constraint (4) requires that the total quantity sold in each market does not exceed the total demand in that market. Thus, unlike the one-market case, there is no straightforward way to simplify the set of constraints: the highest-profit equilibrium for national firms may require a firm to serve a very small quantity of consumers in one market, while serving a larger quantity of consumers in another market.

#### 3.5 Merger Ramifications

We first show that—in the context of our model—any merger is profitable for firms.

**Theorem 1.** Let  $\hat{\kappa}$  be a merger under  $\kappa$  and suppose that  $\hat{\kappa}$  is sufficient for competition: Then any prices sustainable under  $\kappa$  are also sustainable under  $\hat{\kappa}$ . Moreover, even if the merger is a market extension merger, it is possible that strictly higher prices can be sustained after the merger.

It is immediate from the analysis of Section 3.4 that weakly higher profits can be sustained after a merger. If  $\bar{f}$  acquires  $\hat{f}$ , this simply "unifies" the incentive constraints of  $\bar{f}$  and  $\hat{f}$ ; that is, any pair  $(p_m, (q_m^f)_{f \in \mathbf{F}(m;\kappa)})_{m \in M}$  that satisfies (3) and (4) under  $\kappa$  generates a pair  $(p_m, (q_m^f)_{f \in \mathbf{F}(m;\hat{\kappa})})_{m \in M}$  that satisfies (3) and (4) under  $\hat{\kappa}$  (by increasing the acquirer's quantity from  $q_m^{\bar{f}}$  to  $q_m^{\bar{f}} + q_m^{\hat{f}}$  in every market m, setting the quantity of  $\hat{f}$  to 0 in every market, and not changing the quantity of any other firm in any market). Since the set of prices and total quantities satisfying the constraints is now weakly larger, the solution to the maximization problem weakly increases.

More surprisingly, a merger can also strictly raise profitability, even when the two firms do not overlap in any market. We demonstrate this in Example 1 below.

**Example 1.** There are two markets, m and n, with  $\psi_m = \psi_n = 1$ ; for simplicity, we set c = 0 and D(p) = 1 - p. Under market structure  $\kappa$ , there are two firms, f and  $\hat{f}$  that are only in market m, i.e.,  $\mathbf{F}(m;\kappa) = \{f, \hat{f}\}$ ; meanwhile, there are 5 firms in market n. The discount factor is  $\delta = \frac{7}{9}$ .

Since no firm is in both markets, we can analyze each market independently. It follows from Proposition 2 that in the concentrated market m monopoly profits can be sustained, while in the competitive market n the highest sustainable price is 0.

Now consider the market structure  $\hat{\kappa}$ , under which firm f acquires a firm  $\bar{f}$  in market n. Under  $\hat{\kappa}$ , we can now sustain monopoly profits in both markets. The monopoly price in both markets is  $\frac{1}{2}$ . In one equilibrium supporting such prices, there are two phases:

- 1. The collusive phase: In this phase, each firm offers the monopoly price of  $\frac{1}{2}$  in each market in which it present. In market m, both firms choose the same aggressiveness, so as to set the quantity of each firm to  $\frac{1}{4} = \frac{1}{2}D(\frac{1}{2})$ . Meanwhile, in market n, firm f has a quantity of  $q_n^f = \frac{1}{18}$ , and each other firm  $\bar{f}$  present in market n has a quantity of  $q_n^{\bar{f}} = \frac{1}{9}$ ; these quantities are obtained by choosing appropriate aggressivenesses in each market.
- 2. The punishment phase: In this phase, each firm sets its price to marginal cost and chooses an aggressiveness of 1.

Play starts in the collusive phase and continues in the collusive phase so long as no firm deviates; if any firm does so, play continues in the punishment phase. In the punishment phase, play continues in the punishment phase regardless of what happened in-period.

This strategy profile is incentive compatible for all firms: During the punishment phase, it is immediate that each firm is playing optimally given the play of other firms. In the collusion phase, it is optimal for firm  $\hat{f}$  to play its prescribed strategy—instead of increasing its aggressiveness to capture the entire market m—so long as

$$\frac{1}{1-\delta}q_m^{\hat{f}}(p^\circ - c) \ge D(p^\circ)(p^\circ - c)$$
$$\frac{1}{1-\frac{7}{9}}\frac{1}{4}\left(\frac{1}{2}\right) \ge \frac{1}{2}\left(\frac{1}{2}\right)$$
$$\frac{9}{16} \ge \frac{1}{4}.$$

Similarly, for each firm  $\overline{f}$  present in market n other than f, we need that

$$\frac{1}{1-\delta}q_n^{\bar{f}}(p^\circ - c) \ge D(p^\circ)(p^\circ - c)$$
$$\frac{1}{1-\frac{7}{9}}\frac{1}{9}\left(\frac{1}{2}\right) \ge \frac{1}{2}\left(\frac{1}{2}\right)$$
$$\frac{1}{4} \ge \frac{1}{4}.$$

Finally, we show that firm f's strategy is incentive compatible:

$$\frac{1}{1-\delta} \Big( q_m^f(p^\circ - c) + q_n^f(p^\circ - c) \Big) \ge D(p^\circ)(p^\circ - c) + D(p^\circ)(p^\circ - c) \\ \frac{1}{1-\frac{7}{9}} \Big( \frac{1}{4} + \frac{1}{18} \Big) \Big( \frac{1}{2} \Big) \ge \Big( \frac{1}{2} + \frac{1}{2} \Big) \Big( \frac{1}{2} \Big) \\ \frac{11}{16} \ge \frac{1}{2}.$$

Intuitively, each firm in market n other than f has been allocated a larger share of market n; this share has been chosen to be just large enough so that each local firm in n weakly prefers to price at  $p^{\circ}$  and obtain its allocated share of the market each period rather than to increase its aggressiveness and so obtain the entire market for one period. Meanwhile, firm fobtains a smaller market share than each local firm in market n. However, if firm f were to increase its market share in market n, it would lose its half of the monopoly profits each period in market m (as well as its  $\frac{1}{9}$  of the profits in market n); the value of its shares of profits in markets m and n in each period is greater than its profit from increasing its market share in market n. Even if firm f were to engage in its most profitable deviation—that is, increasing its aggressivenesses in market m and market n to capture total market demand in each—its foregone profits in future periods have greater value than its increase in profits today. Essentially, firm f has "slack"—in the sense of Bernheim and Whinston (1990)—in its incentive constraint in market m, and it uses that slack to constrain its behavior in market n, i.e., to reduce its supply in market n. This leaves a greater market share for the other firms in market n, and the market share for each other firm in market n is large enough to make deviations unprofitable long-term.

# **3.6** The Effects of Market Size and Concentration

In Figure 2, we consider the (post-merger) setting of Example 1 and show how the size of the concentrated market affects the price in the less concentrated market. When the size of market m is 0, it is as if only market n exists, and the only sustainable price is the

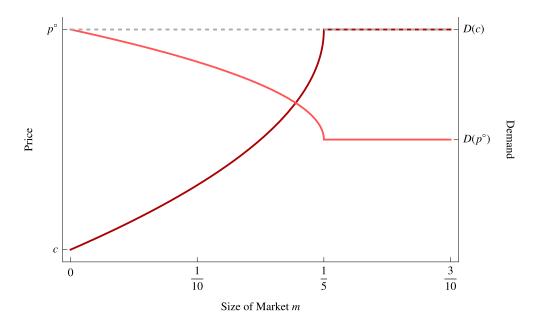


Figure 2: The highest sustainable price in market n in dark red and consumer demand at that price in light red as a function of the size of market m. The dotted grey line is the monopoly price  $p^{\circ}$ . There are two markets, a duopoly market m and a more competitive market n with five firms; one national firm is present in both markets. We let  $\delta = \frac{7}{9}$ ,  $\psi_n = 1$ , D(p) = 1 - p, and c = 0.

competitive price c. As the duopoly market m grows, firm f "acquires" more slack in its incentive constraint; with this additional slack, it is possible to sustain higher profits—and thus prices—in market n. This effect grows stronger until the size of market m is  $\frac{1}{5}$ , at which point the monopoly price can be sustained in both markets. Note that the duopoly market m can be much smaller than the less concentrated market n and yet still allow the firms in market n to collude at the monopoly price.

The amount of slack generated by a non-competitive market depends not only on the size of the non-competitive market, but also on how non-competitive it is. In Figure 3, the dark red line shows how the price in market n varies with the competitiveness of market m; here, instead of one other local firm in market m, we allow the number of local firms in market mto vary. When market m is very uncompetitive, i.e., market m is a duopoly, not only can the monopoly price be sustained in market m, but it can also be sustained in market n. But as market m becomes more competitive, the price in market n falls since the slack available from

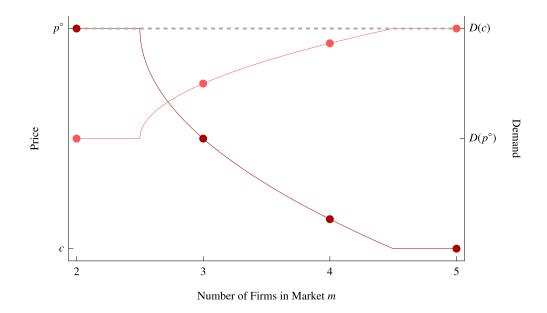


Figure 3: The highest sustainable price in market n in dark red and consumer demand at that price in light red as a function of number of the local firms in market m. The dotted grey line is the the monopoly price  $p^{\circ}$ . There are two markets, market m and a more competitive market n with five firms; one national firm is present in both markets. We let  $\delta = \frac{7}{9}$ ,  $\psi_m = \frac{1}{4}$ ,  $\psi_n = 1$ , D(p) = 1 - p, and c = 0.

market m falls; in Figure 3, this effect begins when the number of local firms is 3. Finally, once market m becomes competitive (i.e., has 5 firms), no price other than the competitive price can be sustained in market m, and so there is no slack left with which to sustain a price higher than the competitive price in market n.

# 3.7 Characteristics of Highest-Profit Equilibria

We now state the two results which allow for empirical tests of the effects of multimarket contact. Like many repeated-game settings, our model of multimarket contact has a large number of equilibria. Thus, to predict how changes in market structure are likely to influence prices and other outcomes, we need to take a position on which equilibrium we expect to emerge. In this section, we focus on equilibria that are the most profitable for national firms: Since national firms are key for coordinating behavior across markets, it is natural consider the equilibrium that is best for these coordinating firms. That is, we consider solutions to

$$\max_{\substack{p \in \times_{m \in M}[c, p^{\circ}), \\ q \in \times_{m \in M} \left( \times_{f \in \mathbf{F}(m)}[0, \psi_m D(c) \right)}} \left\{ \sum_{m \in M} \sum_{f \in \mathbf{N}(m)} (p_m - c) q_m^f \right\}$$

subject to the incentive constraints (3) and the quantity constraints (4).

Another natural criterion for equilibrium selection would be to consider the most profitable equilibria for the industry as a whole. We do this in Appendix A and find results essentially analogous to our Theorems 2 and 3.

Our first result characterizes how the profit-maximizing prices differ across markets with different degrees of local concentration but the same degree of multimarket contact. We show that, if market m has fewer firms—i.e., is less competitive—than market n, then market m will have a higher price than market n (holding the set of the national firms in the two markets constant). Intuitively, when market concentration is lower, each firm can enjoy a greater portion of the surplus generated by a high price and so each firm is less tempted to steal market share for higher profits today.

**Theorem 2.** Suppose that for two markets m and n:

- There are less local firms in market m, i.e.,  $|\mathbf{L}(m)| < |\mathbf{L}(n)|$ ;
- There are the same national firms in each market, i.e.,  $\mathbf{N}(m) = \mathbf{N}(n)$ ;
- The size of the markets is the same, i.e.,  $\psi_m = \psi_n$ .

Then, in any highest-profit equilibrium for national firms,  $p_m \ge p_n$ .

Our second result characterizes how the profit-maximizing prices differ across markets with different degrees of multimarket contact but the same degree of local concentration. We show that, if market m has more multimarket contact—i.e., more national firms—than market n, then market m will have a higher price than market n (holding the number of firms in the two markets constant). Intuitively, when market m has more national firms, each additional national firm may have additional slack it can "import" into market m, allowing that firm to constrain its behavior to facilitate higher prices.

**Theorem 3.** Suppose that for two markets m and n:

- Every national firm in market n is also in market m, i.e.,  $\mathbf{N}(n) \subseteq \mathbf{N}(m)$ ;
- The number of firms in market m is weakly less, i.e.,  $|\mathbf{F}(m)| \leq |\mathbf{F}(n)|$ ;
- The size of the markets is the same, i.e.,  $\psi_m = \psi_n$ .

Then, in any highest-profit equilibrium for national firms,  $p_m \ge p_n$ .

# 3.8 A Model of Multimarket Contact for Deposit Banking

To facilitate our analysis of the banking industry, we adapt the model of Section 3 to the setting of consumer banking. In particular, we model the consumer demand for deposits in each market in a manner similar to that of Drechsler et al. (2017) and assume that the local interest rate affects local depositors' decisions on how much to hold in deposits. We show that interest rates on deposits are lower in markets with fewer banks (Theorem C.2) and more multimarket contact (Theorem C.3), corresponding to Theorems 2 and 3. Moreover, Theorems C.2 and C.3 show that the pass-through of the Fed funds rate is decreasing in both market concentration and the degree of multimarket contact; we use these latter results to test the hypothesis that multimarket contact drives anti-competitive behavior.

# 4 Data

Our empirical analysis relies on data from multiple sources: To measure retail industry trends in local concentration, multimarket contact, and markups, we use data from Dun & Bradstreet Corporation and data on public firms from CRSP and Compustat. To identify cross-sectional evidence of how multimarket contact impacts competition, we use granular variation in bank branch networks. We focus our analysis on bank branches that belong to regional and national banks, as a local bank cannot have contact with other banks across markets. We define regional and national banks as those banks that operate in two states and are regulated at the national level (by the Federal Reserve). The 2020 Federal Deposit Insurance Corporation (FDIC) Survey of Deposits (SOD) provides data on 2,019 counties; 146 regional and national banks compete in these areas.<sup>16</sup>

For each regional/national bank, we have branch-level deposit rates from RateWatch. Of these bank branches, not all set their own deposit rates. To avoid duplicate observations, we subsample the rate-setting bank branches. Following Drechsler et al. (2017), we average weekly deposit rate data by branch to a quarterly frequency and use the money market deposit account rate.<sup>17</sup> In the second quarter of 2020, we have deposit rate data on 799 rate-setting branches. The RateWatch data on deposit rates spans from 2001 to 2020.<sup>18</sup>

# 5 Empirical Analysis

# 5.1 Multimarket Contact

We geographically define a market as a U.S. county and the set of competing firms by industry.<sup>19</sup> This market definition captures how retail markets are geographically local for consumers.

<sup>&</sup>lt;sup>16</sup>Our empirical analysis of regional and national banks identifies the effect of multimarket contact on competition among large banks in the U.S. deposit market. This sidesteps the concerns of Begenau and Stafford (2022) that effects identified from the variation in small bank behavior may not be indicative of the behavior of larger regional and national banks.

<sup>&</sup>lt;sup>17</sup>Banks offer many depository products. Similar to Drechsler et al. (2017), we focus on the savings rate on money market deposit accounts with account size of \$25,000. This represents the overwhelming majority of deposits. As of December 2019, the FDIC reported \$1,809 billion in demand deposits, \$582 billion in small time deposits, and \$9,715 billion in savings deposits. We do not study 12-month CDs because our identification requires heterogeneity in pricing within bank across different counties. Granja and Paixão (2019) studied 12-month CDs with a minimum account size of \$10,000 ("12MCD10K") and found predominantly homogeneous pricing across banks.

 $<sup>^{18}</sup>$ Drechsler et al. (2017) used a sample that extends back to 1997. We source our data from the same provider, RateWatch, who informed us that data prior to January 2001 was discontinued due to quality issues.

<sup>&</sup>lt;sup>19</sup>We use county and market interchangeably.

For each market, we measure the local market concentration as the Herfindahl–Hirschman Index (HHI). This concentration metric is used by the Department of Justice in analyzing the competitive effects of mergers (Federal Reserve, 2014). The HHI for a county c is defined as the sum of the squared sales shares of all firms (denoted i) within an industry and within a county c, that is,

$$\mathrm{HHI}_{c} \equiv \sum_{i \in \mathbf{F}(c)} \left(\frac{s_{c}^{i}}{s_{c}}\right)^{2},\tag{5}$$

where  $s_c^i$  is the sales of firm *i* in county *c* and  $s_c$  is the total sales of firms in county *c* within an industry.

We define multimarket contact between firm i and firm j to be the overlap of their sales across markets. We let  $\theta_c^i \equiv \frac{q_c^i}{\sum_{\bar{c}} q_{\bar{c}}^i}$  be the sales portfolio share of firm i in market c; that is,  $\theta_c^i$  is firm i's sales in market c divided by the total sales of firm i. We thus define the multimarket contact between firms i and j as

$$\mathrm{MMC}_{i,j} \equiv \sum_{c} \left( \theta_c^i \cdot \theta_c^j \right)^{\frac{1}{2}}$$

In contrast to  $\text{HHI}_c$ , which varies by county c,  $\text{MMC}_{i,j}$  varies by pair of firms (firm i, firm j). The multimarket contact of national firms i and j does not depend on any one local market. Instead, the measure captures the extent to which the geographic sales of firm i overlap with the geographic sales of firm j.

This empirical measure of multimarket contact relates closely to our model. The overlap of firm sales defines the degree to which firms may threaten each other with perfect competition. As shown in Figure 2, the ability of firms to sustain collusive prices depends on the quantity of overlapping sales.

# 5.2 Multimarket Contact for Retail Industries

Figure 4a shows the time trend of markups, multimarket contact, and average local HHI from 1989 to 2021 for retail industries. Retail industries include all 4-digit SIC codes from 5200-

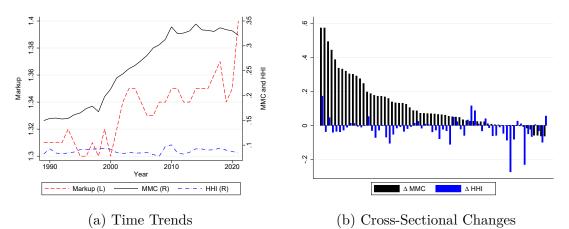


Figure 4: Figure 4a shows the time series of markups, multimarket contact, and HHI for retail industries. Markups are shown on the left vertical axis and MMC and HHI are shown on the right vertical axis. Figure 4b shows the change in multimarket contact and HHI from

1989 to 2021 for each retail industry.

5900. Markup is defined as price divided by cost and measured following De Loecker et al. (2020). The time trend of markups is the sales-weighted average markup of firms within the retail industry. Markups have increased from on average from 1.31 to 1.40, corresponding to a 27 percent increase in the price above cost.

We measure the multimarket contact among large firms within each retail industry. For each retail industry, we define large firms as the 100 largest firms by number of establishments. An empirical challenge in the Dun & Bradstreet Corporation data is that local sales for each establishment is imperfectly reported: many establishments report zero sales. Therefore, we construct the sales portfolio share of each firm based on the firm's share of establishments in each county for its industry and the population of the county:

$$\hat{q}_c^i = \frac{\text{establishments}_c^i}{\sum_{j \in \text{ind(i)}} \text{establishments}_c^j} \text{population}_c$$

where  $population_c$  is the population of the county. Using this proxy for the sales of a firm in a county, we similarly construct its sales portfolio share:

$$\hat{\theta}^i_c \equiv \frac{\hat{q}^i_c}{\sum_{\bar{c}} \hat{q}^i_{\bar{c}}}$$

The identifying assumption for this measure is that population is a good proxy for geographical variation in sales and that firm share of sales in each market is proportional to their number of establishments.

We then aggregate the multimarket contact between firms within an industry to define industry-level  $MMC_{ind}$  by averaging across firm pairs (weighted by their sales); that is,

$$\mathrm{MMC}_{\mathrm{ind}} \equiv \frac{\sum_{i} \sum_{j \neq i} s_{i} s_{j} \mathrm{MMC}_{i,j}}{\sum_{i} \sum_{j \neq i} s_{i} s_{j}}.$$

Figure 4a shows that average multimarket contact for retail industries has increased from an overlap of 15 percent in 1989 to 32 percent in 2021. The increase has been gradual and broadly tracks the increase in markups. By contrast, the average local concentration of retail industries has remained near constant at an HHI of about 0.08.

Figure 4b shows the change in multimarket contact and local concentration for each industry from 1989 to 2021. Industries are sorted in descending order of change in multimarket contact. Appendix **D** provides a list of all retail industries and their multimarket contact and HHI. About 80 percent of retail industries have experienced an increase in multimarket contact and for 40 percent of industries the increase has been greater than 10 percent. By contrast, 70 percent of retail industries have experienced a decrease in HHI. Furthermore, the change in multimarket contact and HHI is uncorrelated across industries.

These time trends for retail industries suggest that changes in local concentration cannot explain broad increases in markups. However, this positive trend for markups can be explained by decreased competition from greater multimarket contact among large firms. These trends are suggestive of our theorized economic mechanism but not causal. In the following section, we utilize cross-sectional variation in multimarket contact for the deposit banking market to better identify the effect of multimarket contact on competition.<sup>20</sup>

<sup>&</sup>lt;sup>20</sup>We are limited in our ability to estimate markups for the cross-section of retail industries at the 4-digit SIC code because we need financial data, which we only have for publicly traded firms. For many 4-digit SIC codes for retail industries, we do not have a publicly traded firm for which this industry is their primary market.

# 5.3 Multimarket Contact in the Deposit Market

For the bank deposit market, we measure multimarket contact using bank deposit shares as our proxy for local market shares. Figure 5 illustrates deposit network overlaps between Bank of America, JP Morgan Chase, and Wells Fargo in both 2005 and 2018. In 2005, Bank of America's branch network spanned both coasts; Wells Fargo was primarily in the western United States; and JP Morgan Chase was primarily in the Eastern Midwest. These geographic differences are reflected in the higher MMC of Bank of America with JP Morgan Chase and Wells Fargo (25 percent and 61 percent, respectively) compared to the MMC of JP Morgan Chase and Wells Fargo (16 percent).

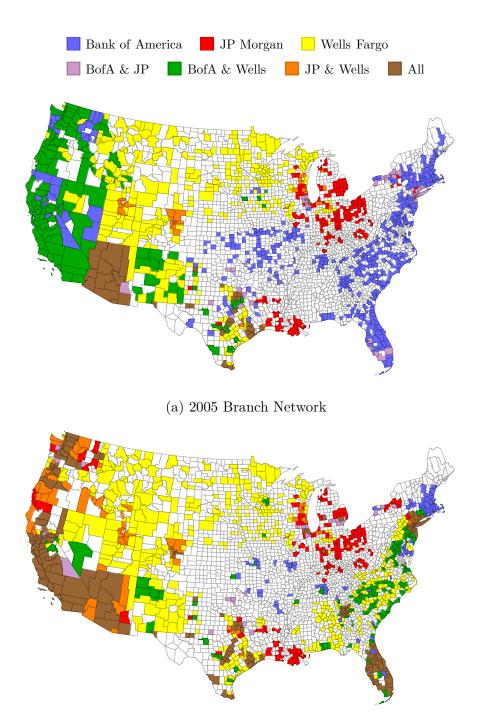
However, JP Morgan Chase expanded westward after 2005, primarily through the acquisition of Washington Mutual in September of 2008.<sup>21</sup> Meanwhile, Wells Fargo expanded across the Atlantic seaboard by acquiring Wachovia. After these mergers, all three banks had greater deposit market overlap with each other. Thus, as of 2018, Bank of America had an MMC of 63 percent and 70 percent with JP Morgan Chase and Wells Fargo, respectively (whereas they were 25 percent and 61 percent, respectively, before), and JP Morgan Chase and Wells Fargo had an MMC of 44 percent (whereas it was 16 percent before).

Market extension mergers are the primary driver of the increase in multimarket contact between banks from 5 percent in 2001 to 16 percent in 2020. This contrasts with the time trend of local concentration, which has been near flat: an HHI of 0.21 in 2001 and 0.19 in 2020. These time trends are qualitatively similar to that of retail industries.

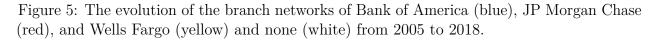
To measure cross-sectional differences in multimarket contact, we aggregate to the countylevel. We define county-level  $MMC_c$  by averaging across all bank pairs within a county (weighted by the quantity of their deposits in market c); that is,

$$\mathrm{MMC}_{c} \equiv \frac{\sum_{i} \sum_{j \neq i} \mathrm{MMC}_{i,j} q_{c}^{i} q_{c}^{j}}{\sum_{i} \sum_{j \neq i} q_{c}^{i} q_{c}^{j}}$$

<sup>&</sup>lt;sup>21</sup>This expansion was code-named "Project West" (Dash, 2008).



(b) 2018 Branch Network



Of note is that our measure of multimarket contact for a county has a correlation with local market concentration of -11 percent.

# 5.4 Imperfect Deposit Market Competition and Multimarket Contact

#### 5.4.1 Deposit Spread Beta

Following Drechsler et al. (2017), we measure the competitiveness of the deposit banking market using changes in the Fed Funds rate as a cost shifter. Define the deposit spread as the difference between the Fed Funds rate and the saving rate. In a perfectly competitive market, banks fully pass through changes in the Fed Funds rate to depositors. However, with market power, the deposit spread increases with the Fed Funds rate. We measure bank market power using the deposit spread beta ( $\beta$ ):

$$\Delta y_{b,t} = \alpha_b + \beta \Delta FF_t + \epsilon_{b,t},\tag{6}$$

where  $\Delta y_{b,t}$  is the change in the deposit spread (Fed Funds rate less the deposit rate) for bank branch b and quarter t, and  $\Delta FF_t$  is the change in the Fed Funds rate.

Within the deposit banking model of Appendix C, the deposit spread beta measures the fraction of interest rate changes not passed through to consumers, which is 1 minus the passthrough rate. Noncompetitive markets have a high deposit spread beta and perfectly competitive markets have a deposit spread beta of 0.

The average large bank branch has a deposit spread beta of 0.79, which implies that the average consumer sees only a 21 bps increase in the deposit rate for a 100 bps increase in the Fed Funds rate. This low passthrough of changes in the Fed Funds rate to consumers implies a sizable degree of market power. Table 1 shows how market power has increased over the sample. In the pre-crisis period (2001-2006), the deposit spread beta was 0.64. By the post-crisis period (2010-2020), the deposit spread beta had increased to 0.96.

	$\Delta$ Savings Spread						
Sample	(1) 2001-2020	(2) 2001-2006	(3) 2007-2009	(4) 2010-2020			
$\Delta FF$	$0.791^{**}$ (0.045)	$0.640^{**}$ (0.034)	$0.833^{**}$ (0.087)	$0.957^{**}$ (0.006)			
Quarter FE	Ν	Ν	Ν	Ν			
Bank FE	Υ	Υ	Υ	Υ			
Branch FE	Υ	Υ	Υ	Υ			
Adjusted $R^2$ N	$0.73 \\ 53,376$	$0.60 \\ 13,833$	$0.68 \\ 8,649$	$0.95 \\ 30,790$			

Table 1: Deposit Spread Beta

*Notes*: This table presents OLS regressions of changes in the deposit spread on changes in the Fed Funds rate. For the full sample (2001-2020, Column 1), we estimate a deposit spread beta of 0.791. We estimate passthrough in the pre-crisis (2001-2006), crisis (2007-2009), and post-crisis (2010-2020) periods in Columns 2, 3, and 4, respectively. Standard errors are clustered at the county by year level.

#### 5.4.2 Identification Strategy

In aggregate, deposit market power and multimarket contact have increased over the past two decades. However, documenting a causal relationship between the two is challenging due to many contemporaneous trends, such as falling real interest rates (Bauer and Rudebusch, 2020) and worsening lending opportunities (Eggertsson et al., 2016). Thus, our identification strategy relies on variation in the cross-section. We estimate the differences in deposit spread beta at two levels of granularity: across bank branches and within bank and across counties.

In Section 5.4.3, we estimate deposit spread beta and multimarket contact across bank branches to document the economic magnitude of the cross-sectional covariance. In Section 5.4.4, we estimate the difference in deposit spread beta for bank branches in different counties but within the same bank in order to control for lending opportunities; since deposit funding is fungible across the bank (Gilje et al., 2016), our within-bank estimates hold lending opportunities fixed.

#### 5.4.3 Across Bank-Branch Estimates

For each bank branch b, we estimate the deposit spread beta:

$$\Delta y_{b,t} = \alpha_b + \beta_b \Delta FF_t + \epsilon_{b,t}.$$
(7)

Equation (7) differs from Equation (6) only in that the deposit spread beta is estimated for each bank branch  $(\beta_b)$ .<sup>22</sup> At the bank branch level, deposit spread betas are on average 0.82 and have a standard deviation of 0.20. More competitive bank branches (those at the 10<sup>th</sup> percentile of deposit spread betas) have a deposit spread beta of about 0.52 and less competitive bank branches (those at the 90<sup>th</sup> percentile of deposit spread betas) have a deposit spread beta of near 1.

Each bank branch competes in a local deposit market for which multimarket contact among its parent bank and other banks differs. For each bank branch, we estimate the average degree of multimarket contact over the sample. We sort bank branches into deciles based on multimarket contact among banks within the local county. Figure 6 plots the average deposit spread beta by decile of multimarket contact. Bank branches capture more of interest rate increases within counties with greater multimarket contact. Within counties at the bottom decile of multimarket contact, bank branches on average capture 77 bps of a 100 bps increase in the Fed Funds Rate. At the top decile, bank branches capture 89 bps of a 100 bps increase in the Fed Funds Rate. Variation in multimarket contact can explain about 25 percent of the cross-sectional variation in branch-level deposit spread betas.<sup>23</sup>

 $<sup>^{22}</sup>$ To estimate branch-level deposit spread beta we require 3-years of quarterly data on branch deposit spreads and a minimum of 3 interest rate changes. To mitigate the effect of outliers, we winsorize estimated branch-level deposit spread beta at the 1% level.

<sup>&</sup>lt;sup>23</sup>The difference in deposit spread betas of bank branches between the 10<sup>th</sup> to the 90<sup>th</sup> percentile of county-level multimarket contact is 12 percentage points, while the difference between the 10<sup>th</sup> to the 90<sup>th</sup> percentile of bank branch deposit spread betas is 48 percentage points.

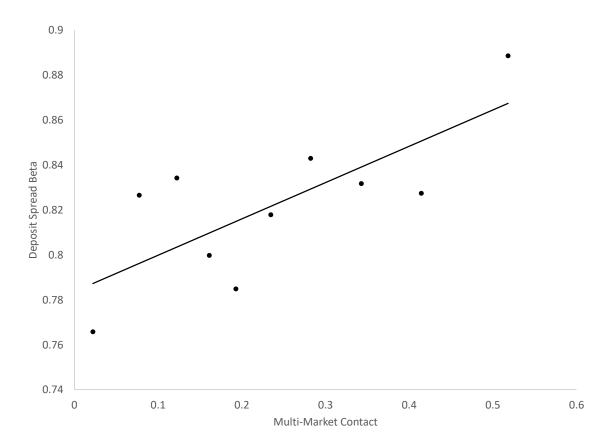


Figure 6: Branch-Level deposit spread beta and MMC

*Notes*: This figure plots average bank branch deposit spread betasorted into deciles of multimarket contact for the county in which the branch operates.

#### 5.4.4 Within-Bank and Across-County Estimates

We implement a within-bank estimate of deposit spread betas in order to address the most relevant omitted variable: lending opportunities for banks. Thus, we use variation in multimarket contact across counties and the feature that national banks have branches across many counties. This variation enables us to control for time-varying bank-lending opportunities using bank-year fixed effects:

$$\Delta y_{b,t} = \alpha_t + \alpha_b + \zeta_{s(b),t} + \chi_{i(b),t} + \gamma \Delta FF_t \times MMC_{c(b),t} + \epsilon_{b,t},$$

where  $\Delta y_{b,t}$  is the change in the deposit spread for branch b,  $\alpha_t$  is a time fixed effect,  $\alpha_b$  is a branch fixed effect,  $\zeta_{s(b),t}$  is a state-quarter fixed effect, and  $\chi_{i(b),t}$  is a bank-quarter fixed effect for the parent bank i of branch b. We cluster standard errors at the county-by-year level.

	$\Delta$ Deposit Spread					
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta$ FF × MMC	0.070**	0.056**			0.070**	0.056**
	(0.03)	(0.02)			(0.03)	(0.02)
$\Delta$ FF $\times$ Branch-HHI			0.031	$0.096^{**}$	0.007	-0.013
			(0.03)	(0.04)	(0.04)	(0.04)
Quarter FE	Υ	Υ	Υ	Υ	Υ	Y
Branch FE	Υ	Υ	Υ	Υ	Y	Υ
Bank FE	Υ	Υ	Υ	Υ	Y	Υ
Bank $\times$ Quarter FE	Υ	Υ	Υ	Ν	Y	Υ
State $\times$ Quarter FE	Υ	Ν	Υ	Ν	Υ	Ν
Adjusted $R^2$	0.919	0.914	0.915	0.765	0.919	0.914
N	43,787	$43,\!885$	48,432	$53,\!376$	43,787	43,885

Table 2: Deposit Spread Betas and Imperfect Competition

Notes: This table estimates the difference in the deposit spread beta by multimarket contact and HHI.  $\Delta$  deposit spread is the change in the branch-level deposit spread (change in FF - deposit rate).  $\Delta$  FF is the change in the Fed Funds Rate. MMC measures the deposit weighted average multimarket contact of banks with other banks within the county. HHI measures the concentration of deposits within the county. The data is at the branch-quarter level and spans from 2001 through 2020. Standard errors are clustered at the county by year level.

The estimate of interest is  $\gamma$ , which is how deposit spread betas vary within a bank across markets with different levels of multimarket contact (MMC<sub>c(b),t</sub>). From Theorem C.3, we hypothesize a positive  $\gamma$ , which implies that bank branches within counties with greater multimarket contact have larger deposit spread betas. Table 2 presents evidence in favor of this hypothesis: controlling for a battery of fixed effects, the same bank has, on average, a 3.0 percentage points larger deposit spread beta for one of its branches in a deposit market at the 90<sup>th</sup> percentile of multimarket contact compared to that of another one of its branches at the 10<sup>th</sup> percentile.<sup>24</sup>

<sup>&</sup>lt;sup>24</sup>This estimate is from Column 1 of Table 2. The 10<sup>th</sup> percentile of county multimarket contact is 3.5

The magnitude of this effect is comparable to that of local market concentration: Columns 3 and 4 of Table 2 substitute multimarket contact with HHI, which is the local deposit share concentration of the county of branch b (see Equation 5). Consistent with Drechsler et al. (2017), we find that banks within more concentrated local deposit markets capture a larger share of increases in the Fed Funds rate.<sup>25</sup> The same bank has on average a 2.25 percentage points larger deposit spread beta for a branch in a deposit market at the 90<sup>th</sup> percentile of HHI compared to that of another branch at the 10<sup>th</sup> percentile.<sup>26</sup>

Although multimarket contact and local market concentration are nearly uncorrelated (with a correlation coefficient of -11%), when both are interacted with  $\Delta FF$ , the correlation coefficient is mechanically large at 73%. With the caveat of multicollinearity in mind, we include both  $\Delta FF_t \times \text{MMC}_{c(b),t}$  and  $\Delta FF_t \times \text{HHI}_{c(b)}$ : Columns 5 and 6 of Table 2 show that the estimated effect of multimarket contact is nearly unchanged while the effect of HHI attenuates to insignificance. For regional/national banks, multimarket contact in a local deposit market better explains the cross-section of deposit spread betas than local market concentration.

We can gain a sense of the economic magnitude of the effects of multimarket contact by considering the counterfactual deposit market where each local market is an island, i.e., where each local deposit market has the same HHI but each bank is local. This counterfactual

percent and the 90<sup>th</sup> percentile is 46 percent. The effect of moving from the  $10^{\text{th}}$  to  $90^{\text{th}}$  percentile of multimarket contact is an increase in the deposit spread beta of  $3.0 = (0.46 - 0.035) \times 7.0$ .

<sup>&</sup>lt;sup>25</sup>Our estimates are smaller than that of Column 5 of Table 2 of Drechsler et al. (2017); our estimate is 0.096 while Drechsler et al. (2017) estimate a coefficient of 0.15. With bank-by-quarter and state-by-quarter fixed effects, we estimate an even smaller coefficient. This difference is primarily due to our focus on the deposit spread beta of regional and national banks, excluding local banks. We do so because local banks do not have multimarket contact. If we include local banks, then our estimate of the effect of local market concentration is much more similar to that of Drechsler et al. (2017). This difference may also partially be attributed to a different sample period and sample of banks: Drechsler et al. (2017) had a sample period from January 1997 to December 2013. We source our data from the same provider (Ratewatch) and they informed us that data prior to January 2001 was discontinued due to data quality issues. We further extend the sample to the present, resulting in a sample of January 2001 to December 2020. Note that we follow the procedure of Drechsler et al. (2017) in averaging HHI for each county over the sample such that HHI does not vary over time. The findings are robust to using the HHI estimated for each year.

<sup>&</sup>lt;sup>26</sup>This estimate is from Column 4 of Table 2. The 10<sup>th</sup> percentile of county HHI is 0.13 and the 90<sup>th</sup> percentile is 0.40. The effect of moving from the 10<sup>th</sup> to 90<sup>th</sup> percentile of HHI is an increase in the deposit spread beta of  $2.52 = (0.40 - 0.13) \times 9.60$ .

decreases multimarket contact from a size-weighted average of 29 percent to 0, while local HHI remains unchanged at 0.20. But, for the market power of banks to remain the same, the counterfactual average local deposit market HHI would have to increase to 0.41; we call this multimarket contact counterfactual HHI the *effective HHI* of local deposit markets.<sup>27</sup> The effective HHI of 0.38 is nearly double that of the actual HHI of 0.21; in other words, multimarket contact allows banks to act as if the market was nearly twice as concentrated as it really is.

Recall that bank deposit spread betas in deposit markets have increased over time (Table 1); this is difficult to reconcile with how local deposit market HHI has on average decreased (slightly) over time. However, effective HHI has increased—which is consistent with an increase in the deposit spread beta. Adjusting local deposit market concentration for the effects of multimarket contact implies that bank market power has increased over time.

#### 5.5 Deposit Market Contact and Merger Activity

From 2001 to 2020, national banks acquired about 30,000 bank branches. These mergers are responsible for 67 percent of the increase in multimarket contact and 92 percent of the increase in national deposit market concentration.<sup>28</sup> Our model predicts that market extension mergers that increase multimarket contact are more profitable; thus, we conjecture that banks take into account multimarket contact when making acquisitions. We empirically test this conjecture using data on bank mergers.

From the National Information Center, we source Federal Reserve data on bank mergers. We require the merger to be a voluntary liquidation (no bankruptcies or asset sales) and the acquirer to be a national bank. We have 519 such mergers between 2001 and 2020 corresponding to 519 target banks and 137 acquirer banks. For each year, we construct a

<sup>&</sup>lt;sup>27</sup>Note that this counterfactual is a partial equilibrium estimate that does not account for entry or exit in response to a market with zero multi-market contact. We measure this effect as actual HHI of 0.20 plus how much HHI would have to increase  $0.21 = 0.29 \times 0.07/0.096$  (estimates from columns 1 and 3 of Table 2).

<sup>&</sup>lt;sup>28</sup>As of 2001, average local multimarket contact was 5 percent. Without transfers of bank branches, multimarket contact would have been 9 percent in 2020, rather than the actual 16 percent.

sample of all the possible pairs of target and acquirer banks. Define  $\operatorname{Merger}_{i,j,t}$  to be equal to 1 for the pair (acquirer bank *i* and target bank *j*) that merged in year *t* and 0 for all other pairs. Since we have 9,823 hypothetical pairs, the sample probability of merger is 5.3 percent.

For each target and acquirer pair, we measure the extent to which the target operates in markets in which the acquirer would have high multimarket contact (if the acquirer bought the target). For example, suppose Bank of America is the acquiring bank; we measure whether the target operates in markets with banks in which Bank of America already has high multimarket contact, such as Wells Fargo and JP Morgan Chase. Formally, we define Network  $MMC_{i,j}$  of acquiring bank i and target bank j as

Network MMC<sub>*i*,*j*</sub> = 
$$\frac{\sum_{n \neq i} \text{MMC}_{i,n} q_{c(n)}^{j}}{\sum_{n} q_{c(n)}^{j}}$$

where n is a national bank that is not the acquiring bank,  $\text{MMC}_{i,n}$  is the multimarket contact of the acquiring bank with the national bank n, and  $q_{c(n)}^{j}$  is the quantity of target bank j's deposits that overlap with the markets c(n) that national bank n is in. This measure avoids a mechanical relationship with geographic distance because it does not measure direct multimarket contact between target and acquirer.

We estimate the association between mergers and multimarket contact:

$$\operatorname{Merger}_{i,j,t} = \alpha_{i,t} + \beta \operatorname{Network} \operatorname{MMC}_{i,j} + \xi X_{i,j,t} + \epsilon_{i,j,t},$$

where  $\alpha_{i,t}$  is an acquirer bank-by-time fixed effect and  $X_{i,j,t}$  are control variables. These control variables are drawn from prior literature that has shown that banks consider market concentration and geographic distance in choosing merger targets (Akkus et al., 2016). The control variables include log average distance between acquirer and target bank branches (distance), the change in average local market concentration caused by the merger ( $\Delta$  HHI) for the acquirer, and the average local concentration, population growth, and deposits growth of the target's deposit markets.<sup>29</sup>

	(1)	(2)	(3)	(4)
Network $MMC_{i,j}$	$0.537^{**}$	$0.565^{**}$	$0.563^{**}$	$0.508^{**}$
	(0.07)	(0.07)	(0.07)	(0.07)
Distance	$-0.124^{**}$	$-0.119^{**}$	$-0.119^{**}$	$-0.120^{**}$
	(0.01)	(0.01)	(0.01)	(0.01)
HHI	-0.023	0.015		
	(0.02)	(0.02)		
$\Delta$ HHI	$0.386^{**}$			
	(0.08)			
Pop Growth	$1.446^{**}$			
	(0.26)			
Deposits Growth	0.011			
	(0.02)			
Bank $\times$ Year FE	Υ	Y	Υ	Ν
Bank and Year FE	Υ	Υ	Υ	Y
Adjusted $\mathbb{R}^2$	0.18	0.18	0.18	0.17
Ν	$9,\!052$	9,052	$9,\!052$	9,052

Table 3: Mergers and Deposit Market MMC

National banks tend to make acquisitions of banks that operate in markets where they would have high multimarket contact. From Column 1 of Table 3, an acquirer and target pair with a 1 standard deviation increase in Network  $MMC_{i,j}$  is 74 percent more likely to merge compared to the unconditional average merger probability.<sup>30</sup> Similar to Akkus et al. (2016), we find that bank mergers are less likely to occur between more geographically distant banks and more likely to occur when the merger would increase local market concentration.

This evidence is consistent with banks considering multimarket contact in their mergers. Similar to how banks merge to increase profits by increasing local market concentration, we

*Notes*: This table estimates the influence of deposit market MMC on mergers. The data is at the bank-pair year level (acquiring bank, target bank) and spans from 2001 through 2020. Standard errors are clustered by acquiring bank x year.

<sup>&</sup>lt;sup>29</sup>The averages are dollar weighted across the acquirer or target deposit markets. Furthermore,  $\Delta$ HHI is winsorized at the 1 percent level to mitigate the effects of outliers.

 $<sup>^{30}</sup>$ A 1 standard deviation change in Network  $MMC_{i,j}$  is 7.4 percent. The unconditional average merger probability of 0.054. The effect of 1 standard deviation is  $0.74 = 0.074 \times 0.537/.054$ .

show that they also merge to increase multimarket contact.

### 5.6 Deposit Market Branch Warfare

In early 2018, JP Morgan Chase and Bank of America announced plans to expand to new deposit markets by opening 400 and 500 new branches, respectively. Financial analysts and news reports interpreted this expansion as "branch warfare" that targeted Wells Fargo (Back, 2018). Wells Fargo was described as a "sitting duck to rivals" due to a series of customer abuse scandals: From September 2016 through February 2018, Wells Fargo was fined for creating millions of fake customer accounts, improperly fining mortgage holders, and illegally repossessing cars (Wattles et al., 2018). This behavior culminated in the Federal Reserve restricting the total assets of Wells Fargo to not exceed their level in 2017.

In models of collusion with perfect information, "price wars" are off-equilibrium path behavior and so should not be observed; however, in more realistic settings with imperfect information, price wars may be required to support a collusive equilibrium.<sup>31</sup>

Consumer response to Wells Fargo's misconduct was uncertain, increasing the probability of a price war between Wells Fargo and other large banks. Thus, we define a treated market to be a deposit market with Wells Fargo and either JP Morgan Chase or Bank of America as of June 2017. We define the pre-treatment period to be 2015 to 2017 and the post-treatment period to be 2018 to 2020. For each treated market, we propensity score match that market to a control market based on market size, local concentration, and multimarket contact. For each quarter, we estimate the difference between treated and control market deposit spreads:

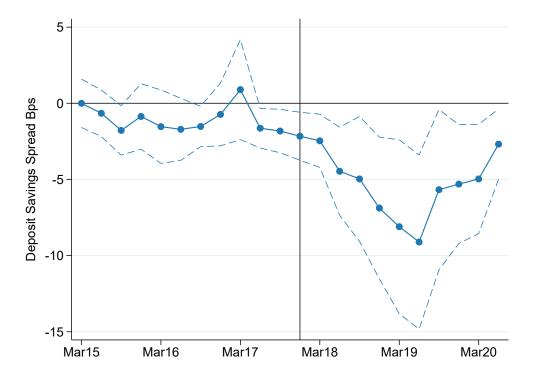
$$y_{b,t} = \alpha_t + \beta_t T_{c(b)} + \epsilon_{b,t},$$

where  $y_{b,t}$  is the deposit spread for branch b,  $\alpha_t$  is a time fixed effect that absorbs the average

<sup>&</sup>lt;sup>31</sup>For instance, in the work of Green and Porter (1984), demand cannot be observed directly by firms, and collusion requires that price wars happen with positive probability. Ellison (1994) documented that the behavior of the Joint Executive Committee—an 1880s railroad cartel—was consistent with the theory of Green and Porter (1984).

deposit spread for control markets, and  $T_{c(b)}$  is equal to 1 if the bank branch belongs to a market where Wells Fargo and either JP Morgan Chase or Bank of America operated in 2017 and 0 otherwise. The coefficients of interest are the  $\beta_t$ , the quarterly differences between deposit spreads in control and treated markets.

Figure 7: Deposit Spreads and Branch Warfare



*Notes*: This figure plots the quarterly difference in deposit spreads between treated and control markets. Standard errors are clustered by county.

Figure 7 shows that deposit spreads decreased in markets where Wells Fargo competed with JP Morgan Chase or Bank of America in 2018 and 2019. Compared to similar deposit markets, the deposit savings spread in markets subject to branch warfare were up to 9 bps smaller (June 2019). Although small in absolute terms, this is large compared to the on average 12 bps savings deposit rate offered by regional and national banks during this period.

## 6 Conclusion

Over the past 20 years, U.S. deposit markets have both become less concentrated and less competitive. We show that the anti-competitive effects of multimarket contact can reconcile this puzzle: banks are increasingly meeting each other in many disparate local markets. Earlier work largely treats each local market as distinct, but each local market is not an island. Rather, banks may use the threat of competitive behavior in profitable markets, where they have local oligopolies, to discipline behavior in other markets with more competitors.

We find that these overlapping relationships significantly reduce the passthrough rate of changes in the Fed Funds rate. Due to multimarket contact, banks behave as though local market concentration is about twice as large as it actually is.<sup>32</sup> Moreover, the overlapping relationships among banks in deposit markets influence their merger activity: banks are twice as likely to merge into markets where they would have high multimarket contact.

The low passthrough of Fed Funds rate changes to deposit savings rates has persisted outside our sample of analysis. For the currently ongoing rate increases from March 2022 to June of 2023, the Fed funds rate has risen by 5 percentage points, but the national deposit rate has only increased from 0.08 to 0.61 percentage points.<sup>33</sup> This is consistent with a deposit market where a high level of multi-market contact inhibits passthrough.

More broadly, local market concentration has decreased at the same time as multimarket contact has increased; this reflects a disparate treatment of horizontal and market extension mergers by regulators. Our work shows that multimarket contact has significant implications for competitive behavior. Thus, antitrust regulators may wish to scrutinize market extension mergers more carefully to guard against multimarket contact contributing to coordinated effects.

Our model of multimarket contact provides a framework for studying how overlapping relationships can decrease competition across markets. Many scholars have noted with concern

 $<sup>^{32}</sup>$ See Section 5.4.4 for details of how we compute this counterfactual.

<sup>&</sup>lt;sup>33</sup>These national rates are reported by the FDIC for money market savings accounts.

the trend of increasing national consolidation in many industries (Gutiérrez and Philippon, 2017; Hall, 2018; De Loecker et al., 2020; Kwon et al., 2023). However, others have emphasized the concurrent decrease in local market concentration (Rossi-Hansberg et al., 2021; Hsieh and Rossi-Hansberg, 2021). Our work shows that the effects of national consolidation are more than the sum of its effects on local market concentrations. Thus, a deeper understanding of the effects of national consolidation may be helpful for understanding the full implications of this trend for competition.

## References

- Abreu, D. (1988). On the theory of infinitely repeated games with discounting. *Econometrica* 54(4), 383–396. (Cited on pages 10, 12, 49, 62, 64, 74, and 75.)
- Akkus, O., J. A. Cookson, and A. Hortacsu (2016). The determinants of bank mergers: A revealed preference analysis. *Management Science* 62(8), 2241–2258. (Cited on pages 35 and 36.)
- Autor, D., D. Dorn, L. F. Katz, C. Patterson, and J. Van Reenen (2020). The fall of the labor share and the rise of superstar firms. *Quarterly Journal of Economics* 135(2), 645–709. (Cited on page 2.)
- Back, A. (2018). Big banks enter branch warfare. Wall Street Journal. (Cited on page 37.)
- Barkai, S. (2020). Declining labor and capital shares. Journal of Finance 75(5), 2421–2463. (Cited on page 2.)
- Bauer, M. D. and G. D. Rudebusch (2020). Interest rates under falling stars. American Economic Review 110(5), 1316–54. (Cited on page 29.)
- Begenau, J. and E. Stafford (2019). Do banks have an edge? SSRN #3095550. (Cited on page 6.)
- Begenau, J. and E. Stafford (2022). Uniform rate setting and the deposit channel. working paper. (Cited on page 22.)
- Benkard, C. L., A. Yurukoglu, and A. L. Zhang (2021). Concentration in product markets. NBER Working Paper #28745. (Cited on page 2.)
- Bernheim, D. B. and M. D. Whinston (1990). Multimarket contact and collusive behavior. Rand Journal of Economics 21(1), 1–26. (Cited on pages 2, 3, 17, and 70.)

- Busse, M. R. (2000). Multimarket contact and price coordination in the cellular telephone industry. *Journal of Economics & Management Strategy* 9(3), 287–320. (Cited on page 7.)
- Cai, J., F. Eidam, A. Saunders, and S. Steffen (2020). Loan syndication structures and price collusion. SSRN #3250817. (Cited on page 7.)
- Chan, Y.-J., C.-Y. Lin, and T.-C. Lin (2021). Tacit collusion among dominant banks: Evidence from round-yard loan pricing. SSRN #3905375. (Cited on page 7.)
- Ciliberto, F. and J. W. Williams (2014). Does multimarket contact facilitate tacit collusion? Inference on conduct parameters in the airline industry. *RAND Journal of Economics* 45(4), 764–791. (Cited on page 7.)
- Compte, O., F. Jenny, and P. Rey (2002). Capacity constraints, mergers and collusion. European Economic Review 46(1), 1–29. (Cited on page 61.)
- Corbae, D. and P. D'Erasmo (2020). Rising bank concentration. Journal of Economic Dynamics and Control 115, 103877. (Cited on page 6.)
- Corbae, D. and P. D'Erasmo (2021). Capital buffers in a quantitative model of banking industry dynamics. *Econometrica* 89(6), 2975–3023. (Cited on page 6.)
- Covarrubias, M., G. Gutiérrez, and T. Philippon (2020). From good to bad concentration? US industries over the past 30 years. *NBER Macroeconomics Annual* 34(1), 1–46. (Cited on page 2.)
- Dash, E. (2008). A Banker Embraces WaMu's Challenge. https://www.nytimes.com/2008/10/07/business/07profile.html. [Online; accessed 26-July-2021]. (Cited on page 26.)
- De Loecker, J., J. Eeckhout, and G. Unger (2020). The rise of market power and the macroeconomic implications. *Quarterly Journal of Economics* 135(2), 561–644. (Cited on pages 2, 3, 24, and 40.)

- Drechsler, I., A. Savov, and P. Schnabl (2017). The deposits channel of monetary policy. *Quarterly Journal of Economics* 132(4), 1819–1876. (Cited on pages 4, 5, 6, 21, 22, 28, 33, and 86.)
- Duffie, D. and J. C. Stein (2015). Reforming LIBOR and other financial market benchmarks. Journal of Economic Perspectives 29(2), 191–212. (Cited on page 7.)
- Duso, T., L.-H. Röller, and J. Seldeslachts (2014). Collusion through joint R&D: An empirical assessment. *Review of Economics and Statistics* 96(2), 349–370. (Cited on page 8.)
- Edwards, C. D. (1955). Conglomerate bigness as a source of power. In *Business concentration* and price policy, pp. 331–359. Princeton University Press. (Cited on page 3.)
- Eggertsson, G. B., N. R. Mehrotra, and L. H. Summers (2016). Secular stagnation in the open economy. *American Economic Review* 106(5), 503–07. (Cited on page 29.)
- Eizenberg, A., D. Shilian, and D. D. Blanga (2020). Estimating the potential effect of multi-market contact on the intensity of competition. Hebrew University working paper. (Cited on pages 8, 10, and 62.)
- Ellison, G. (1994). Theories of cartel stability and the Joint Executive Committee. Rand Journal of Economics 25, 37–57. (Cited on page 37.)
- Evans, W. N. and I. N. Kessides (1994). Living by the "Golden Rule": Multimarket contact in the U.S. airline industry. *Quarterly Journal of Economics 109*(2), 341–366. (Cited on page 7.)
- Federal Reserve (2014). How do the Federal Reserve and the U.S. Department of Justice, Antitrust Division, analyze the competitive effects of mergers and acquisitions under the Bank Holding Company Act, the Bank Merger Act and the Home Owners' Loan Act? (Cited on pages 7 and 23.)

- Fernández, N. and P. L. Marín (1998). Market power and multimarket contact: Some evidence from the Spanish hotel industry. *Journal of Industrial Economics* 46(3), 301–315. (Cited on page 7.)
- Gilje, E. P., E. Loutskina, and P. E. Strahan (2016). Exporting liquidity: Branch banking and financial integration. *Journal of Finance* 71(3), 1159–1184. (Cited on pages 5 and 29.)
- Granja, J. and N. Paixão (2019). Market concentration and uniform pricing: Evidence from bank mergers. SSRN #3488035. (Cited on pages 6 and 22.)
- Green, E. J. and R. H. Porter (1984). Non-cooperative collusion under imperfect price information. *Econometrica* 52(1), 87–100. (Cited on page 37.)
- Gutiérrez, G. and T. Philippon (2017). Declining competition and investment in the US.NBER Working Paper #23583. (Cited on pages 2 and 40.)
- Hall, R. E. (2018). New evidence on the markup of prices over marginal costs and the role of mega-firms in the US economy. NBER Working Paper #24574. (Cited on pages 2 and 40.)
- Hatfield, J. and R. Lowery (2023). A simple model of undifferentiated bertrand competition with asymmetric convex costs. University of Texas working paper. (Cited on page 9.)
- Honka, E., A. Hortaçsu, and M. A. Vitorino (2017). Advertising, consumer awareness, and choice: Evidence from the U.S. banking industry. *RAND Journal of Economics* 48(3), 611–646. (Cited on page 4.)
- Hsieh, C.-T. and E. Rossi-Hansberg (2021). The industrial revolution in services. NBERWorking Paper #25968. (Cited on pages 2 and 40.)
- Igami, M. and T. Sugaya (2022). Measuring the incentive to collude: The vitamin cartels, 1990–99. *Review of Economic Studies* 89(3), 1460–1494. (Cited on pages 10 and 62.)
- Ivaldi, M., B. Jullien, P. Rey, P. Seabright, and J. Tirole (2003). The economics of tacit collusion. Final Report for DG Competition, European Commission. (Cited on page 8.)

- Jahanshahloo, H. and C. X. Cai (2019). Monitoring the foreign exchange rate benchmark fix. European Journal of Finance 25(7), 670–688. (Cited on page 7.)
- Jans, I. and D. I. Rosenbaum (1996). Multimarket contact and pricing: Evidence from the U.S. cement industry. International Journal of Industrial Organization 15(3), 391–412. (Cited on page 8.)
- Khwaja, A. and B. Shim (2017). The collusive effect of multimarket contact on prices: Evidence from retail lumber markets. In 2017 Meeting Papers, Volume 593. (Cited on page 8.)
- Kwon, S. Y., Y. Ma, and K. Zimmermann (2023). 100 years of rising corporate concentration. University of Chicago Becker Friedman Institute for Economics Working Paper #2023-20. (Cited on page 40.)
- Mailath, G. J. and L. Samuelson (2006). Repeated Games and Reputations: Long-Run Relationships. Oxford University Press. (Cited on pages 49 and 75.)
- Parker, P. M. and L.-H. Röller (1997). Collusive conduct in duopolies: Multimarket contact and cross-ownership in the mobile telephone industry. *RAND Journal of Economics 28*, 304–322. (Cited on page 7.)
- Rossi-Hansberg, E., P.-D. Sarte, and N. Trachter (2021). Diverging trends in national and local concentration. NBER Macroeconomics Annual 35(1), 115–150. (Cited on pages 2 and 40.)
- Schmitt, M. (2018). Multimarket contact in the hospital industry. American Economic Journal: Economic Policy 10(3), 361–87. (Cited on page 8.)
- Siriwardane, E., A. Sunderam, and J. Wallen (2021). Segmented arbitrage. Harvard Business School working paper. (Cited on page 7.)

- Starc, A. and T. G. Wollmann (2022). Does entry remedy collusion? Evidence from the generic prescription drug cartel. NBER Working Paper #29886. (Cited on pages 10 and 62.)
- Wallen, J. (2021). Markups to financial intermediation in foreign exchange markets. Harvard Business School working paper. (Cited on page 7.)
- Wattles, J., B. Geier, and M. Egan (2018). Wells Fargo's 17-month nightmare. CNN. (Cited on page 37.)
- Weyl, E. G. and M. Fabinger (2013). Pass-through as an economic tool: Principles of incidence under imperfect competition. *Journal of Political Economy* 121(3), 528–583. (Cited on page 64.)

## A Characterizing Highest-Profit Equilibria

In this appendix, we focus on the most-profitable equilibria. That is, we consider solutions to

$$\max_{\substack{p \in \times_{m \in M}[c, p^{\circ}), \\ q \in \times_{m \in M} \left( \times_{f \in \mathbf{F}(m)}[0, \psi_m D(c) \right)}} \left\{ \sum_{m \in M} \sum_{f \in \mathbf{F}(m)} (p_m - c) q_m^f \right\}$$

subject to the incentive constraints (3) and the quantity constraints (4).

Our first result characterizes how the profit-maximizing prices differ across markets with different degrees of local concentration but the same degree of multimarket contact. We show that, if market m has fewer firms—i.e., is less competitive—than market n, then market m will have a higher price than market n (holding the set of the national firms in the two markets constant). Intuitively, when market concentration is lower, each firm can enjoy a greater portion of the surplus generated by a high price and so each firm is less tempted to steal market share for higher profits today.

**Theorem A.1.** Suppose that for two markets m and n:

- There are less local firms in market m, i.e.,  $|\mathbf{L}(m)| < |\mathbf{L}(n)|$ ;
- There are the same national firms in each market, i.e.,  $\mathbf{N}(m) = \mathbf{N}(n)$ ;
- The size of the markets is the same, i.e.,  $\psi_m = \psi_n$ .

Then, in any highest-profit equilibrium,  $p_m \ge p_n$ .

Our second result characterizes how the profit-maximizing prices differ across markets with different degrees of multimarket contact but the same degree of local concentration. We show that, if market m has more multimarket contact—i.e., more national firms—than market n, then market m will have a higher price than market n in at least one of the highest-profit equilibria (holding the number of firms in the two markets constant). Intuitively, when market m has more national firms, each additional national firm may have additional slack it

can "import" into market m, allowing that firm to constrain its behavior to facilitate higher prices.

**Theorem A.2.** Suppose that for two markets m and n:

- Every national firm in market n is also in market m, i.e.,  $\mathbf{N}(n) \subseteq \mathbf{N}(m)$ ;
- The number of firms in market m is weakly less, i.e.,  $|\mathbf{F}(m)| \leq |\mathbf{F}(n)|$ ;
- The size of the markets is the same, i.e.,  $\psi_m = \psi_n$ .

Then there exists a highest-profit equilibrium such that  $p_m \ge p_n$ .

## **B** Proofs

## **B.1** Proof of Proposition **1**

Suppose not; then there exists an equilibrium  $\{(p_m^f, a_m^f)\}_{f \in F, m \in M}$  such that, in some market m, at least one firm  $\hat{f}$  obtains positive profits. We show that such a strategy profile cannot be an equilibrium in three steps:

- Every firm in *m* makes positive profits: Suppose firm  $f \neq \hat{f}$  obtains 0 profits and is also present in market *m*. Then firm *f* could choose the action  $(p_m^{\hat{f}}, \epsilon)$  for some small  $\epsilon > 0$ . Under this action, firm *f* has a positive margin (since the firm  $\hat{f}$  has a positive margin as  $\hat{f}$  has positive profits). Moreover, since firm  $\hat{f}$  obtains positive profits, firm  $\hat{f}$  must have a positive quantity of consumers. But then firm *f* must have a positive quantity of consumers at the price  $p_m^{\hat{f}}$ . Since *f* has a positive margin and positive demand at  $(p_m^{\hat{f}}, \epsilon)$ , its profits are strictly positive under this new action.
- Every firm in *m* obtains demand equal to its capacity: Since every firm makes positive profits, every firm is choosing the same price p > c. Thus, any firm whose quantity is strictly less than its capacity can increase its profitability by increasing its aggressiveness by a small  $\epsilon > 0$ ; for a small enough  $\epsilon$ , firm *f*'s quantity will still be less than

its capacity. Moreover, any firm whose quantity is strictly greater than its capacity can increase its profitability by decreasing its aggressiveness by a small  $\epsilon > 0$ ; for a small enough  $\epsilon$ , firm f's quantity will still be greater than its capacity.

The contradiction: If every firm obtains demand equal to its capacity, then  $Q_m(p_m, a_m) > \psi_m$ , a contradiction.

The existence of a 0-profit equilibrium is given in the text below Proposition 1.

## B.2 Proof of Proposition 2

First, note that Proposition 1 shows that there exists a subgame-perfect Nash equilibrium of the stage game in which each firm obtains 0 profits—its lowest individually rational payoffs; the 0-profit equilibrium is thus an Abreu (1988) optimal penal code.<sup>34</sup> Second, given a strategy profile  $(p^f, a^f)_{f \in F}$ , the supremum of in-period over all strategies by firm f is given by  $(p-c)\psi D(\min_{\bar{f} \in F} \{p^{\bar{f}}\})$ .<sup>35</sup>

Thus, it follows from Abreu (1988) that a price p is sustainable if and only if there exists a quantity vector  $(q^f)_{f \in F}$  such that<sup>36</sup> for each bank  $f \in F$ ,

$$\frac{1}{1-\delta}\psi(p-c)q^f \ge (p-c)\dot{\kappa}^f(p) \tag{B.1}$$

and

$$\sum_{f \in F} q^f = \psi D(p).$$

The incentive compatibility constraint (B.1) encodes the equilibrium requirement that it is better for each firm f to obtain its quantity  $q^f$  each period rather than all the demand for

 $<sup>^{34}</sup>$ For an excellent discussion of optimal penal codes, see Proposition 2.6.1 and the surrounding text in Mailath and Samuelson (2006).

 $<sup>^{35}</sup>$ That is, the profits of firm f when he offers a price of p and obtains all the demand at that price.

<sup>&</sup>lt;sup>36</sup>We simplify the notation in the proof by assuming that  $F = \mathbf{F}(m)$  and dropping the *m* subscript where appropriate.

one period and 0 profits thereafter. The proof of Proposition 2 then follows as in the text of Section 3.3.

## B.3 Proof of Theorem 1

The arguments to prove Proposition 2 can be used *mutatis mutandis* to show that the set of sustainable strategy profiles are those that satisfy the constraints that

$$\frac{1}{1-\delta}\sum_{m\in M} (p_m - c)q_m^f \ge \sum_{m\in M} (p_m - c)$$

for each firm  $f \in F$ , and

$$\sum_{f \in F} q_m^f \le \psi_m D(p_m)$$

for each market  $m \in M$  if the market structure is sufficient for competition.<sup>37</sup>

The first result of the theorem then follows immediately from the fact that under the post-merger market structure we require that the sum of the incentive constraints for the merging banks are satisfied instead of requiring that the incentive constraint of each merging bank is satisfied.

The second claim of the theorem is shown by Example 1.

## B.4 Proof of Theorem 2

Consider any equilibrium—that is, a pair  $(p_k, q_k^f)_{k \in M, f \in F}$  satisfying constraints (3) and (4) for which  $p_n > p_m$ ; we will show that there exists a pair  $(\hat{p}_k, \hat{q}_k^f)_{k \in M, f \in F}$  also satisfying constraints (3) and (4) with higher total profits for the national firms.

For ease of exposition, we let  $\Pi_k^f(p_k, q_k) \equiv (p_k - c)q_k^f$  be the total profits for firm f in market k at the price  $p_k$  and quantity  $q_k^f$ ; and we define  $\Pi_k(p_k) \equiv (p_k - c)\psi_k D_k(p_k) = \sum_{f \in F} \Pi_k^f(p_k, q_k)$ .

<sup>&</sup>lt;sup>37</sup>Note that if the market structure is not sufficient for competition, then some market is a monopoly, and so a firm could not be punished in that market.

It will also be helpful to define the *profit share* of each firm f in each market k as  $\sigma_k^f \equiv \frac{\prod_k^f(p_k,q_k)}{\prod^k(p_k)}$ ; note that there is a one-to-one correspondence (given prices) between profit shares and quantities, and so we can rewrite Constraints (3) and (4) as

$$\frac{1}{1-\delta} \sum_{m \in M} \sigma_m^f \Pi_m(p_m) \ge \sum_{m \in M} \Pi_m(p_m)$$
(B.2)

for each firm  $f \in F$  and

$$\sum_{f \in F} \sigma_m^f = 1 \tag{B.3}$$

for each  $m \in M$ .

Reallocating Quantities/Profit Shares to National Firms: Given that  $(p_k, \sigma_k^f)_{k \in M, f \in F}$ satisfies the Constraints (B.2) and (B.3), we first show that  $(p_k, \hat{\sigma}_k^f)_{k \in M, f \in F}$  satisfies the Constraints (B.2) and (B.3), where

$$\hat{\sigma}_k^f \equiv \begin{cases} 1-\delta & f \in \mathbf{L}(m) \text{ and } k = m \\ 1-\delta & f \in \mathbf{L}(n) \text{ and } k = n \\ \\ \frac{\sigma_m^f \Pi_m(p_m) + \sigma_n^f \Pi_n(p_n)}{\sigma_m^{N(m)} \Pi_m(p_m) + \sigma_n^{N(n)} \Pi_n(p_n)} (1-(1-\delta)|\mathbf{L}(k)|) & f \in \mathbf{N}(m) \text{ and } k = m, n \\ \\ \sigma_k^f & \text{otherwise.} \end{cases}$$

Since prices have not changed, total industry profits have not changed. And, since  $\sigma_k^f \geq 1 - \delta$  for all local firms f in market k, the profitability of national firms has increased (weakly) since the profit share allocated to local firms has gone down (and Constraint (B.3) is satisfied, as demonstrated below).

Constraints (B.2) and (B.3) are satisfied for  $(p_k, \hat{\sigma}_k^f)_{k \in M, f \in F}$ :

• Constraint (B.2) is satisfied for each firm not in markets m and n as its profit shares and prices are unchanged.

• Constraint (B.2) is satisfied for each local firm in  $k \in \{m, n\}$  as

$$\frac{1}{1-\delta}(1-\delta)\Pi_k(p_k) = \Pi_k(p_k).$$

• Constraint (B.2) is satisfied for each national firm in  $k \in \{m, n\}$  as, under  $(p_k, \hat{\sigma}_k^f)_{k \in M, f \in F}$ , the right-hand side of Constraint (B.2) becomes

$$\begin{split} & \frac{\sigma_m^f \Pi_m(p_m) + \sigma_n^f \Pi_n(p_n)}{\sigma_m^{\mathbf{N}(m)} \Pi_m(p_m) + \sigma_n^{\mathbf{N}(n)} \Pi_n(p_n)} (1 - (1 - \delta) |\mathbf{L}(m)|) \Pi_m(p_m) + \\ & \sum_{\ell \in M} \hat{\sigma}_{\ell}^f \Pi_{\ell}(p_{\ell}) = \frac{\sigma_m^f \Pi_m(p_m) + \sigma_n^f \Pi_n(p_n)}{\sigma_m^{\mathbf{N}(m)} \Pi_m(p_m) + \sigma_n^{\mathbf{N}(n)} \Pi_n(p_n)} (1 - (1 - \delta) |\mathbf{L}(n)|) \Pi_n(p_n) + \\ & \sum_{\ell \in M \smallsetminus \{m,n\}} \sigma_{\ell}^f \\ & = \frac{\sigma_m^f \Pi_m(p_m) + \sigma_n^f \Pi_n(p_n)}{\sigma_m^{\mathbf{N}(m)} \Pi_m(p_m) + \sigma_n^{\mathbf{N}(n)} \Pi_n(p_n)} \begin{pmatrix} (1 - (1 - \delta) |\mathbf{L}(m)|) \Pi_m(p_m) + \\ (1 - (1 - \delta) |\mathbf{L}(n)|) \Pi_n(p_n) \end{pmatrix} + \sum_{\ell \in M \smallsetminus \{m,n\}} \sigma_{\ell}^f \\ & \geq \sigma_m^f \Pi_m(p_m) + \sigma_n^f \Pi_n(p_n) + \sum_{\ell \in M \smallsetminus \{m,n\}} \sigma_{\ell}^f. \end{split}$$

The third line follows as the share in markets m and n allocated to local firms is less under  $(p_k, \hat{\sigma}_k^f)_{k \in M, f \in F}$  as under  $(p_k, \hat{\sigma}_k^f)_{k \in M, f \in F}$  each local firm is allocated the minimal profit share to satisfy Constraint (B.2). Moreover, under  $(p_k, \hat{\sigma}_k^f)_{k \in M, f \in F}$ , the left-hand side of Constraint (B.2) is unchanged for each national firm in  $k \in \{m, n\}$ ; thus Constraint (B.2) is satisfied under  $(p_k, \hat{\sigma}_k^f)_{k \in M, f \in F}$ .

• Constraint (B.3) is satisfied: For each market  $k \in M \setminus \{m, n\}$ , profit shares are

unchanged. For markets  $k \in \{m, n\}$ , we have that

$$\begin{split} \sum_{f \in \mathbf{F}(k)} \hat{\sigma}_{k}^{f} &= \sum_{f \in \mathbf{N}(k)} \hat{\sigma}_{k}^{f} + \sum_{f \in \mathbf{L}(k)} \hat{\sigma}_{k}^{f} \\ &= \sum_{f \in \mathbf{N}(k)} \frac{\sigma_{m}^{f} \Pi_{m}(p_{m}) + \sigma_{n}^{f} \Pi_{n}(p_{n})}{\sigma_{m}^{\mathbf{N}(m)} \Pi_{m}(p_{m}) + \sigma_{n}^{\mathbf{N}(n)} \Pi_{n}(p_{n})} (1 - (1 - \delta) |\mathbf{L}(k)|) + \sum_{f \in \mathbf{L}(k)} (1 - \delta) \\ &= \frac{\sigma_{m}^{\mathbf{N}(m)} \Pi_{m}(p_{m}) + \sigma_{n}^{\mathbf{N}(n)} \Pi_{n}(p_{n})}{\sigma_{m}^{\mathbf{N}(m)} \Pi_{m}(p_{m}) + \sigma_{n}^{\mathbf{N}(n)} \Pi_{n}(p_{n})} (1 - (1 - \delta) |\mathbf{L}(k)|) + |\mathbf{L}(k)| (1 - \delta) \\ &= (1 - (1 - \delta) |\mathbf{L}(k)|) + |\mathbf{L}(k)| (1 - \delta) \\ &= 1. \end{split}$$

Switching Prices in Markets m and n: We now show that  $(\hat{p}_k, \hat{\sigma}_k^f)_{k \in M, f \in F}$  satisfies the Constraints (B.2) and (B.3) with strict inequality for the national firms, where

$$\hat{p}_k \equiv \begin{cases} p_n & k = m \\ p_m & k = n \\ p_k & \text{otherwise.} \end{cases}$$

Since the size of market m is the same as the size of market n, industry profits are the same; moreover, since market m has more national firms, the profitability of national firms has increased (as each local firm in market m and n has a profit share of  $1 - \delta$ ). Moreover, Constraints (B.2) and (B.3) are satisfied for  $(\hat{p}_k, \hat{\sigma}_k^f)_{k \in M, f \in F}$ :

- Constraint (B.2) is satisfied for each firm not in markets m and n as its profit shares and prices are unchanged.
- Constraint (B.2) is satisfied for each local firm in  $k \in \{m, n\}$  as

$$\frac{1}{1-\delta}(1-\delta)\Pi_k(\hat{p}_k) = \Pi_k(\hat{p}_k).$$

• Constraint (B.2) is satisfied with inequality for each national firm in  $k \in \{m, n\}$ 

as, under  $(\hat{p}_k, \hat{\sigma}_k^f)_{k \in M, f \in F}$ ,

$$\begin{split} \sum_{\ell \in M} \hat{\sigma}_{\ell}^{f} \Pi_{\ell}(\hat{p}_{\ell}) &= \hat{\sigma}_{m}^{f} \Pi_{m}(\hat{p}_{m}) + \hat{\sigma}_{n}^{f} \Pi_{n}(\hat{p}_{n}) + \sum_{\ell \in M \smallsetminus \{m,n\}} \hat{\sigma}_{\ell}^{f} \Pi_{\ell}(\hat{p}_{\ell}) \\ &= \hat{\sigma}_{m}^{f} \Pi_{m}(p_{n}) + \hat{\sigma}_{n}^{f} \Pi_{n}(p_{m}) + \sum_{\ell \in M \smallsetminus \{m,n\}} \hat{\sigma}_{\ell}^{f} \Pi_{\ell}(\hat{p}_{\ell}) \\ &> \hat{\sigma}_{n}^{f} \Pi_{n}(p_{n}) + \hat{\sigma}_{m}^{f} \Pi_{m}(p_{m}) + \sum_{\ell \in M \smallsetminus \{m,n\}} \hat{\sigma}_{\ell}^{f} \Pi_{\ell}(\hat{p}_{\ell}) \\ &= \sum_{\ell \in M} \hat{\sigma}_{\ell}^{f} \Pi_{\ell}(p_{\ell}) \end{split}$$

and

$$\sum_{\ell \in M} \Pi_{\ell}(\hat{p}_{\ell}) = \sum_{\ell \in M} \Pi_{\ell}(p_{\ell})$$

as  $\hat{\sigma}_m^f > \hat{\sigma}_n^f$  (as  $|\mathbf{L}(m)| < |\mathbf{L}(n)|$ ),  $\psi_m = \psi_n$  (and so  $\Pi_m(p_m) = \Pi_n(p_m)$  and  $\Pi_m(p_n) = \Pi_n(p_n)$ ), and  $\hat{p}_m = p_n > \hat{p}_n = p_m$ .

Thus, we have constructed a new equilibrium with strictly higher profits for the national firms, contradicting our supposition that the highest-profit equilibrium for national firms had  $p_n > p_m$ .

### **B.5** Proof of Theorem **3**

Consider any equilibrium—that is, a pair  $(p_k, q_k^f)_{k \in M, f \in F}$  satisfying constraints (3) and (4) for which  $p_n > p_m$ ; we will show that there exists a pair  $(\hat{p}_k, \hat{q}_k^f)_{k \in M, f \in F}$  also satisfying constraints (3) and (4) with higher profits for national firms in which  $p_m \ge p_n$ .

As in the proof of Theorem 2, we let  $\Pi_k^f(p_k, q_k) \equiv (p_k - c)q_k^f$  be the total profits for firm f in market k at the price  $p_k$  and quantity  $q_k^f$ ; and we define  $\Pi_k(p_k) \equiv (p_k - c)D_k(p_k) = \sum_{f \in F} \Pi_k^f(p_k, q_k)$ . It will also be helpful to define the *profit share* of each firm f in each market k as  $\sigma_k^f \equiv \frac{\Pi_k^f(p_k, q^k)}{\Pi^k(p_k)}$ ; note that there is a one-to-one correspondence (given prices) between profit shares and quantities, and so we can rewrite Constraints (3) and (4) as Constraints (B.2)

and (B.3), restated here as

$$\frac{1}{1-\delta} \sum_{m \in M} \sigma_m^f \Pi_m(p_m) \ge \sum_{m \in M} \Pi_m(p_m)$$
(B.2)

for each firm  $f \in F$  and

$$\sum_{f \in F} \sigma_m^f = 1 \tag{B.3}$$

for each  $m \in M$ .

# Reallocating Quantities/Profit Shares to National Firms: Given that $(p_k, \sigma_k^f)_{k \in M, f \in F}$ satisfies the Constraints (B.2) and (B.3), we first show that $(p_k, \hat{\sigma}_k^f)_{k \in M, f \in F}$ satisfies the Constraints (B.2) and (B.3), where

$$\hat{\sigma}_k^f \equiv \begin{cases} 1-\delta & f \in \mathbf{L}(m) \text{ and } k = m \\ 1-\delta & f \in \mathbf{L}(n) \text{ and } k = n \\ \sigma_k^f & f \in \mathbf{N}(m) \smallsetminus \mathbf{N}(n) \text{ and } k = m \\ \frac{\sigma_m^f \Pi_m(p_m) + \sigma_n^f \Pi_n(p_n)}{\sigma_m^{\mathbf{N}(n)} \Pi_m(p_m) + \sigma_n^{\mathbf{N}(n)} \Pi_n(p_n)} (1-(1-\delta)|\mathbf{L}(k)|) & f \in \mathbf{N}(m) \cap \mathbf{N}(n) \text{ and } k = m, n \\ \sigma_k^f & \text{otherwise.} \end{cases}$$

The fact that under  $(p_k, \hat{\sigma}_k^f)_{k \in M, f \in F}$  the profits of national firms could only have increased and Constraints (B.2) and (B.3) are satisfied follows as in the proof of Theorem 2 mutatis mutandis.<sup>38</sup>

Switching Prices in Markets m and n: Let  $\tilde{F} \equiv \mathbf{N}(m) \setminus \mathbf{N}(n)$  be the set of national

firms in market m but not market n. We now show that  $(\tilde{p}_k, \tilde{\sigma}_k^f)_{k \in M, f \in F}$  satisfies the

<sup>&</sup>lt;sup>38</sup>The fact that  $(p_k, \hat{\sigma}_k^f)_{k \in M, f \in F}$  satisfies the Constraints (B.2) and (B.3) for  $f \in \mathbf{N}(m) \setminus \mathbf{N}(n)$  is immediate as neither f's profits shares nor prices have changed.

Constraints (B.2) and (B.3) for the national firms, where

$$\tilde{p}_k \equiv \begin{cases} p_n & k = m \\ p_m & k = n \\ p_k & \text{otherwise} \end{cases}$$

and

$$\tilde{\sigma}_{k}^{f} \equiv \begin{cases} \hat{\sigma}_{m}^{f} + \left(1 - \delta - \hat{\sigma}_{m}^{f}\right) \left(1 - \frac{\Pi_{m}(p_{m})}{\Pi_{m}(\tilde{p}_{m})}\right) & f \in \tilde{F} \text{ and } k = m \\ \hat{\sigma}_{m}^{f} - \frac{\hat{\sigma}_{m}^{f} - \hat{\sigma}_{n}^{f}}{\hat{\sigma}_{m}^{\mathbf{N}(n)} - \hat{\sigma}_{n}^{\mathbf{N}(n)}} \left((1 - \delta)|\tilde{F}| - \hat{\sigma}_{m}^{\tilde{F}}\right) \left(1 - \frac{\Pi_{m}(p_{m})}{\Pi_{m}(\tilde{p}_{m})}\right) & f \in \mathbf{N}(n), \hat{\sigma}_{m}^{\mathbf{N}(n)} - \hat{\sigma}_{n}^{\mathbf{N}(n)} \neq 0, \text{ and } k = m \\ \sigma_{k}^{f} & \text{otherwise.} \end{cases}$$

Since the size of market m is the same as the size of market n, industry profits are the same. Moreover, profits for national firms are larger, as

$$\sigma_m^{\mathbf{N}(m)} = 1 - (1 - \delta) |\mathbf{L}(m)| > 1 - (1 - \delta) |\mathbf{L}(n)| = \sigma_n^{\mathbf{N}(n)}$$

as |L(m)| < |L(n)|.

First, we show that Constraint (B.3) is satisfied under  $(\tilde{p}_k, \tilde{\sigma}_k^f)_{k \in M, f \in F}$ : Since profit shares do not change in any market except m, it is immediate for all markets other

than m. For market m, we have

$$\begin{split} \sum_{f \in \tilde{F}} \left( \hat{\sigma}_m^f + \left( 1 - \delta - \hat{\sigma}_m^f \right) \left( 1 - \frac{\Pi_m(p_m)}{\Pi_m(\tilde{p}_m)} \right) \right) + \\ \sum_{f \in \mathbf{F}(m)} \tilde{\sigma}_m^f &= \sum_{f \in \mathbf{N}(n)} \left( \hat{\sigma}_m^f - \frac{\hat{\sigma}_m^f - \hat{\sigma}_n^f}{\hat{\sigma}_m^{\mathbf{N}(n)} - \hat{\sigma}_n^{\mathbf{N}(n)}} \left( (1 - \delta) |\tilde{F}| - \hat{\sigma}_m^{\tilde{F}} \right) \left( 1 - \frac{\Pi_m(p_m)}{\Pi_m(\tilde{p}_m)} \right) \right) + \\ &\sum_{f \in \mathbf{L}(m)} \hat{\sigma}_m^f \\ & \left( (1 - \delta) |\tilde{F}| - \hat{\sigma}_m^{\tilde{F}} \right) \left( 1 - \frac{\Pi_m(p_m)}{\Pi_m(\tilde{p}_m)} \right) + \\ &= -\frac{\hat{\sigma}_m^{\mathbf{N}(n)} - \hat{\sigma}_n^{\mathbf{N}(n)}}{\hat{\sigma}_m^{\mathbf{N}(n)} - \hat{\sigma}_n^{\mathbf{N}(n)}} \left( (1 - \delta) |\tilde{F}| - \hat{\sigma}_m^{\tilde{F}} \right) \left( 1 - \frac{\Pi_m(p_m)}{\Pi_m(\tilde{p}_m)} \right) + \\ &\sum_{f \in \mathbf{F}(m)} \hat{\sigma}_m^f \\ &= \sum_{f \in \mathbf{F}(m)} \hat{\sigma}_m^f. \end{split}$$

Second, we show that Constraint (B.2) is satisfied under  $(\tilde{p}_k, \tilde{\sigma}_k^f)_{k \in M, f \in F}$ : It is immediate that Constraint (B.2) holds for all local firms as well and national firms not in markets m and n. Now consider Constraint (B.2) for a firm  $f \in \tilde{F}$ :

$$\begin{split} \tilde{\sigma}_m^f \Pi_m(\tilde{p}_m) &= \left(\hat{\sigma}_k^f + \left(1 - \delta - \hat{\sigma}_k^f\right) \left(1 - \frac{\Pi_m(\hat{p}_m)}{\Pi_m(\tilde{p}_m)}\right)\right) \Pi_m(\tilde{p}_m) \\ &= \hat{\sigma}_k^f \Pi_m(\tilde{p}_m) + \left(1 - \delta - \hat{\sigma}_k^f\right) (\Pi_m(\tilde{p}_m) - \Pi_m(\hat{p}_m)) \\ &= \hat{\sigma}_k^f \Pi_m(\hat{p}_m) + (1 - \delta) (\Pi_m(\tilde{p}_m) - \Pi_m(\hat{p}_m)) \\ &= \hat{\sigma}_k^f \Pi_m(\hat{p}_m) + (1 - \delta) (\Pi_m(\tilde{p}_m) - \Pi_m(\hat{p}_m)) \end{split}$$

and so

$$\frac{1}{1-\delta} \sum_{k \in M} \tilde{\sigma}_k^f \Pi_k(\tilde{p}_k) = \Pi_m(\tilde{p}_m) - \Pi_m(\hat{p}_m) + \frac{1}{1-\delta} \sum_{k \in M} \hat{\sigma}_k^f \Pi_k(\hat{p}_k).$$
(B.4)

Furthermore,

$$\sum_{k \in M} \Pi_k(\tilde{p}_k) = \Pi_m(\tilde{p}_m) - \Pi_m(\hat{p}_m) + \sum_{k \in M} \Pi_k(\hat{p}_k).$$
(B.5)

Combining (B.4) and (B.5) with Constraint (B.2) for  $(p_k, \hat{\sigma}_k^f)_{k \in M, f \in F}$  yields

$$\frac{1}{1-\delta}\sum_{k\in M}\tilde{\sigma}_k^f \Pi_k(\tilde{p}_k) \ge \sum_{k\in M}\Pi_k(\tilde{p}_k),$$

i.e., Constraint (**B.2**) for firm f under  $(\tilde{p}_k, \tilde{\sigma}_k^f)_{k \in M, f \in F}$ .

Now consider Constraint (B.2) for a firm  $f \in \mathbf{N}(m) \cup \mathbf{N}(n)$ :

$$\begin{split} \sum_{k \in \{m,n\}} \tilde{\sigma}_k^f \Pi_k(\tilde{p}_k) &= \tilde{\sigma}_m^f \Pi_m(\tilde{p}_m) + \tilde{\sigma}_n^f \Pi_n(\tilde{p}_n) \\ &= \tilde{\sigma}_m^f \Pi_m(\tilde{p}_m) + \hat{\sigma}_n^f \Pi_m(\hat{p}_m) \\ &= \left( \hat{\sigma}_m^f - \frac{\hat{\sigma}_m^f - \hat{\sigma}_n^f}{\hat{\sigma}_m^{\mathbf{N}(n)} - \hat{\sigma}_n^{\mathbf{N}(n)}} \left( (1-\delta) |\tilde{F}| - \hat{\sigma}_m^{\tilde{F}} \right) \left( 1 - \frac{\Pi_m(p_m)}{\Pi_m(\tilde{p}_m)} \right) \right) \Pi_m(\tilde{p}_m) + \hat{\sigma}_n^f \Pi_m(\hat{p}_m) \\ &= \hat{\sigma}_m^f \Pi_m(\tilde{p}_m) - \frac{\hat{\sigma}_m^f - \hat{\sigma}_n^f}{\hat{\sigma}_m^{\mathbf{N}(n)} - \hat{\sigma}_n^{\mathbf{N}(n)}} \left( (1-\delta) |\tilde{F}| - \hat{\sigma}_m^{\tilde{F}} \right) (\Pi_m(\tilde{p}_m) - \Pi_m(\hat{p}_m)) + \hat{\sigma}_n^f \Pi_m(\hat{p}_m) \\ &= \hat{\sigma}_m^f \Pi_m(\tilde{p}_m) - \frac{\hat{\sigma}_m^f - \hat{\sigma}_n^f}{\hat{\sigma}_m^{\mathbf{N}(n)} - \hat{\sigma}_n^{\mathbf{N}(n)}} \left( \hat{\sigma}_m^{\mathbf{N}(n)} - \hat{\sigma}_n^{\mathbf{N}(n)} \right) (\Pi_m(\tilde{p}_m) - \Pi_m(\hat{p}_m)) + \hat{\sigma}_n^f \Pi_m(\hat{p}_m) \\ &= \hat{\sigma}_m^f \Pi_m(\tilde{p}_m) - \left( \hat{\sigma}_m^f - \hat{\sigma}_n^f \right) (\Pi_m(\tilde{p}_m) - \Pi_m(\hat{p}_m)) + \hat{\sigma}_n^f \Pi_m(\hat{p}_m) \\ &= \hat{\sigma}_m^f \Pi_m(\hat{p}_m) + \hat{\sigma}_n^f \Pi_m(\tilde{p}_m) \\ &= \hat{\sigma}_m^f \Pi_m(\hat{p}_m) + \hat{\sigma}_n^f \Pi_m(\hat{p}_m) \end{split}$$

where the fifth line follows as the decrease in the profit share of firms in  $\mathbf{N}(n)$  is the increase in the profit share of national firms in  $\tilde{F}$ . Thus, we have that

$$\frac{1}{1-\delta} \sum_{k \in M} \tilde{\sigma}_k^f \Pi_k(\tilde{p}_k) = \Pi_m(\tilde{p}_m) - \Pi_m(p_m) + \frac{1}{1-\delta} \sum_{k \in M} \hat{\sigma}_k^f \Pi_k(\hat{p}_k).$$
(B.6)

Furthermore,

$$\sum_{k \in M} \Pi_k(\tilde{p}_k) = (\Pi_m(\tilde{p}_m) - \Pi_m(\hat{p}_m)) + (\Pi_n(\tilde{p}_n) - \Pi_n(\hat{p}_n)) + \sum_{k \in M} \Pi_k(\hat{p}_k)$$
(B.7)  
$$= \sum_{k \in M} \Pi_k(\hat{p}_k).$$

Combining (B.6) and (B.7) with Constraint (B.2) for  $(p_k, \hat{\sigma}_k^f)_{k \in M, f \in F}$  yields

$$\frac{1}{1-\delta}\sum_{k\in M}\tilde{\sigma}_k^f \Pi_k(\tilde{p}_k) \ge \sum_{k\in M}\Pi_k(\tilde{p}_k),$$

i.e., Constraint (**B.2**) for firm f under  $(\tilde{p}_k, \tilde{\sigma}_k^f)_{k \in M, f \in F}$ .

Thus, we have constructed a new equilibrium with the same profits and a higher price in market m.

### B.6 Proof of Theorem A.1

The proof proceeds exactly as in the proof of Theorem 2; note that, for the final outcome of the proof of Theorem 2, we have that Constraint (B.2) is satisfied with strict inequality for the national firms in market m. Thus,  $(\bar{p}_k, \hat{\sigma}_k^f)_{k \in M, f \in F}$  with  $\bar{p} = (\hat{p}_m + \epsilon, \hat{p}_{M \setminus \{m\}})$  is an equilibrium outcome with strictly higher profits than  $(\hat{p}_k, \hat{\sigma}_k^f)_{k \in M, f \in F}$  (which has the same profits as  $(p_k, \hat{\sigma}_k^f)_{k \in M, f \in F}$ ): Note that Constraint (B.2) is satisfied for local firms in market m as  $\hat{\sigma}_k^f \ge 1 - \delta$ ; and Constraint (B.2) is satisfied for national firms in market m for  $\epsilon$  small enough as it is satified with strict inequality at  $(\hat{p}_k, \hat{\sigma}_k^f)_{k \in M, f \in F}$ .

## B.7 Proof of Theorem A.2

The equilibrium constructed in the proof of Theorem 3 has the same total profits and a higher price in market m, as required by the conclusion of the Theorem A.2.

## C Modeling Multimarket Contact for Deposit Banking

### C.1 Framework

#### C.1.1 Market Structure

We construct a model of bank competition across multiple markets. There is a finite set of markets M and a finite set of banks B. Each market  $m \in M$  has a size  $\psi_m$ . Each bank  $b \in B$  is endowed with a capacity  $\kappa_m^b \in [0, \psi_m]$  in each market m; we say that bank b is present in market m if  $\kappa_m^b > 0$ . We call the full matrix of capacities  $\kappa$  the market structure. We let  $\kappa_m^{\bar{B}} \equiv \sum_{b \in \bar{B}} \kappa_m^b$  be the total capacity of the banks in  $\bar{B}$  in market m. We denote the set of banks present in market m as  $\mathbf{F}(m;\kappa) \equiv \{b \in B : \kappa_m^b > 0\}$ . A bank b is national if it is present in more than one market, i.e.,  $|\{m \in M : \kappa_m^b > 0\}| > 1$ ; we denote the set of national banks present in market m as  $\mathbf{N}(m;\kappa)$ . Conversely, a bank b is local if it is present in exactly one market, i.e.,  $|\{m \in M : \kappa_m^b > 0\}| = 1$ ; we denote the set of local banks in market m as  $\mathbf{L}(m;\kappa)$ . When the market structure  $\kappa$  is clear from context, we will sometimes drop  $\kappa$  from the notation and just write  $\mathbf{F}(m)$ ,  $\mathbf{N}(m)$ , and  $\mathbf{L}(m)$ .

If, under the market structure  $\kappa$ , bank *b* acquires bank  $\hat{b}$ , it generates a new market structure  $\hat{\kappa}$  under which:

- 1. The bank b now assumes all of the capacity of the bank  $\hat{b}$  in each market, i.e.,  $\hat{\kappa}_m^b = \kappa_m^b + \kappa_m^{\hat{b}}$  for all  $m \in M$ .
- 2. The bank  $\hat{b}$  is no longer present in any market, i.e.,  $\hat{\kappa}_m^{\hat{b}} = 0$  for all  $m \in M$ .
- 3. Each other bank has the same capacity in each market as before, i.e.,  $\kappa_m^{\bar{b}} = \hat{\kappa}_m^{\bar{b}}$  for all  $\bar{b} \in B \setminus \{b, \hat{b}\}$  and all  $m \in M$ .

In this case, we say that  $\hat{\kappa}$  is a *merger* under  $\kappa$ . We say that a merger  $\hat{\kappa}$  under  $\kappa$  is a *market* extension merger if banks b and  $\hat{b}$  were not both present in any market before the merger, i.e., for all  $m \in M$ , either  $\kappa_m^b = 0$  or  $\kappa_m^{\hat{b}} = 0$ .

#### C.1.2 The Stage Game

Each consumer, facing an *interest rate* r and a *Fed funds rate* f, has a *demand for deposits* given by<sup>39</sup>

$$D(r, f) \equiv (1 + \lambda) \frac{r}{f + \lambda r};$$

consumers' preference for liquidity is denoted  $\lambda$ .<sup>40</sup>

In each market m, each bank  $b \in \mathbf{F}(m)$  simultaneously chooses an *interest rate*  $r_m^b \in [0, f]$ and an *aggressiveness*  $a_m^b \in (0, \infty)$ .<sup>41</sup> Consumers observe interest rates and then choose a bank with the highest (i.e., most appealing) interest rate; the more aggressive a bank is, the more likely a consumer will choose it. We allow each bank to choose its aggressiveness so that each bank may effectively choose its quantity (if it knows the aggressivenesses of other banks); note, however, that a bank can always choose to be more aggressive at no cost in order to increase its demand.

We denote the set of banks with the highest (i.e., most consumer-friendly) interest rate in market m—i.e., the banks *active* in market m—as  $\mathbf{A}_m(r_m) \equiv \{b \in B : r_m^b = \max_{\bar{b} \in \mathbf{F}(m)}\{r_m^{\bar{b}}\}\}$ ;<sup>42</sup> we call these banks active as they are the only banks that have positive market share. The quantity of customers of bank b in market m is thus given by<sup>43</sup>

$$Q_m^b(r_m, a_m) \equiv \psi_m \mathbb{1}_{\{b \in \mathbf{A}_m(r_m)\}} \frac{a_m^b}{\sum_{\bar{b} \in \mathbf{A}_m(r_m)} a_m^{\bar{b}}}$$

<sup>40</sup>We "normalize" demand by  $1 + \lambda$  so that, when r = f, demand by a consumer in market m is exactly 1.

<sup>&</sup>lt;sup>39</sup>In Appendix C.13, we derive consumers' demand from a constant elasticity-of-supply utility function that depends on both liquidity in the form of deposits and final wealth; the return on non-deposit wealth is determined by the Fed funds rate. As discussed in Appendix C.13, our results hold for any constant elasticity-of-supply utility function over liquidity and final wealth.

 $<sup>^{41}</sup>$ Aggressiveness here plays a role similar to that of *market share* in the model of Compte et al. (2002). We use the term aggressiveness to emphasize that it is a choice of the firm, whereas a bank's market share is determined by the bank's interest rate and aggressiveness as well as other banks' interest rates and aggressivenesses.

<sup>&</sup>lt;sup>42</sup>Throughout, for a matrix  $z \in \mathbb{R}^{B \times M}$ , we let  $z_m$  be the vector of values of z in market m, i.e.,  $z_m \equiv (z_m^b)_{b \in B}$ . <sup>43</sup>The indicator function  $\mathbb{1}_{\{\mathfrak{p}\}}$  is 1 if  $\mathfrak{p}$  is true and 0 otherwise.

Hence, the profits of bank b in market m are

$$\Pi_m^b(r_m, a_m, f) \equiv \underbrace{Q_m^b(r_m, a_m)}_{\text{Quantity of consumers}} \underbrace{D(r_m^b, f)}_{\text{Deposits per consumer}} \underbrace{(f - r_m^b)}_{\text{deposit}} - c \underbrace{\max\left\{0, Q_m^b(r_m, a_m) - \kappa_m^b\right\}}_{\text{Quantity of consumers over capacity}}.$$

where c is the cost of over-capacity market share; we require that c is large enough so that no bank wants more consumers than its capacity.<sup>44</sup>

A bank b's profits in market m are its spread  $s_m^b \equiv f - r_m^b$ , times its total deposits in market m, so long as its market share does not exceed its capacity. Bank b's quantity of consumers is 0 unless it is offering the interest rate most favorable to consumers; if it is offering that rate, then its quantity of consumers depends on its aggressiveness relative to other banks. By choosing a higher aggressiveness, a bank competes more fiercely, as it leaves less residual market share for other banks in that market. However, if the bank is too aggressive, it may acquire more consumers than its capacity, and such over-capacity demand is costly for the bank.

Finally, banks may operate in more than one market, and so a bank b's total profits are given by

$$\Pi^{b}(r,a,f) \equiv \sum_{m \in M} \Pi^{b}_{m}(r_{m},a_{m},f).$$

#### C.1.3 The Repeated Game

In each period  $t \in \mathbb{W} = \{0, 1, 2, ...\}$ , banks play the stage game. Banks have a common discount factor  $\delta$ , and so a bank's total profits are given by  $\sum_{t=0}^{\infty} \delta^t \Pi^b(r(t), a(t), f)$  where

<sup>&</sup>lt;sup>44</sup>Essentially, in our model banks engage in undifferentiated Bertrand competition (with the ability to allocate market shares in a given market by appropriately choosing aggressivenesses). This is in contrast to the structural empirical industrial organization literature in which firms typically engage in differentiated Bertrand (or Cournot) competition; however, in those works the authors simply assume that after a deviation prices revert to the static Bertrand-Nash equilibrium prices instead of solving for the Abreu (1988) optimal penal codes and so those works do not find the most-profitable subgame-perfect Nash equilibrium of the repeated game. Examples of this approach can be found in the work of Eizenberg et al. (2020), Igami and Sugaya (2022), and Starc and Wollmann (2022).

r(t) (a(t)) is the matrix of interest rates (aggressivenesses) for each bank in each market in period t.

We say that a level of industry profits is *sustainable* if there exists a subgame-perfect Nash equilibrium of the repeated game in which, along the equilibrium path, the total profits achieved by the banks reach that level each period.

#### C.1.4 The Monopoly and Competitive Interest Rates

In our setting, the *competitive interest rate* is simply the Fed funds rate f; this is analogous to the competitive price equaling marginal cost in typical models of product market competition.

We can also calculate that a monopolist (with sufficient capacity) in market m would choose the *monopoly interest rate* 

$$r^{\circ} \equiv f \frac{\sqrt{1+\lambda}-1}{\lambda}; \tag{C.1}$$

this is the interest rate that maximizes a monopolist's profits so long as the monopolist has capacity of at least  $\psi_m$ .

#### C.1.5 Conditions on Market Structure

We say that the market structure  $\kappa$  is sufficient for competition in market m if  $\kappa_m^B > \psi_m$  and  $\mathbf{F}(m) > 1$ ; that is, there is more than sufficient capacity across all the banks in market m to serve all  $\psi_m$  customers and there are at least two banks in each market. We say the market structure  $\kappa$  is sufficient for competition if it is sufficient for competition in each market  $m \in M$ .

## C.2 Bertrand Competition in the Stage Game

We first analyze the stage game. We show that the market is competitive—in the sense that each consumer enjoys an interest rate of f—so long as bank capacities are large enough. **Proposition C.1.** Suppose that the market structure  $\kappa$  is sufficient for competition. Then each bank obtains 0 profits in every pure-strategy Nash equilibrium of the stage game and such an equilibrium exists.<sup>45</sup>

The intuition for this result is somewhat more complex than in the standard Bertrand competition setting in which each firm has (effectively) infinite capacity. We prove the proposition by way of contradiction: If any bank in market m has positive profits, every bank b in market m has positive profits, as otherwise b could become profitable by choosing the highest interest rate offered by any other bank and an aggressiveness small enough that bank b's demand is less than its capacity. But if every bank is profitable, then every bank is offering the same interest rate r < f; moreover, some bank has demand less than its capacity as total demand is less than  $\kappa_m^B$  (as the market structure is sufficient for competition). Thus, some bank could slightly increase its aggressiveness to increase its profitability, contradicting the assertion that the original strategy profile was a Nash equilibrium.

One simple pure-strategy equilibrium which delivers 0 profits to each bank is for each bank b to set its interest rate  $r_m^b = f$  (i.e., the competitive interest rate) and its aggressiveness  $a_m^b = \kappa_m^b$  in each market m. This aggressiveness vector ensures that no bank has demand greater than its capacity.

Finally, note that Proposition C.1 implies that there exists a subgame-perfect Nash equilibrium of the repeated game in which each bank obtains 0 profits each period. Such a "price war" equilibrium will be key in our analysis of the repeated game: Since 0 is the lowest individually rational payoff for each bank, reverting to the "price war" equilibrium in every period after a deviation punishes the deviator as harshly as possible; that is, the "price war" equilibrium is an optimal penal code (in the sense of Abreu (1988)) for every bank.

<sup>&</sup>lt;sup>45</sup>Proposition C.1 implies that passthrough is 1 in a competitive market. Weyl and Fabinger (2013) demonstrated that in a model of one-shot imperfect competition with differentiated goods passthrough can be higher than 1, less than the rate under monopoly, or anywhere in-between.

## C.3 An Economy with One Market

Before considering our multimarket setting, we first analyze the case in which there is only one market.

**Proposition C.2.** Suppose that  $M = \{m\}$  and that the market structure is sufficient for competition. If  $(1 - \delta)\kappa_m^B \leq \psi_m$ , then any interest rate in  $[r^\circ, f]$  is sustainable; if  $(1 - \delta)\kappa_m^B > \psi_m$ , then only the only interest rate that is sustainable is f.

To prove Proposition C.2, we first note that after any deviation from the equilibrium strategy profile, the harshest punishment possible is the 0-profit equilibrium of the stage game of Proposition C.1. We then show that in any highest-profit equilibrium, each bank is offering the same interest rate. Thus, we can characterize the set of sustainable profits as the solution to a constrained maximization problem; in particular, we want to solve

$$\max_{\substack{r \in [r^\circ, f], \\ q_m \in \times_{b \in B}[0, \kappa_m^b]}} \{(f - r)\psi_m D(r, f)\}$$

subject to the constraints that, for each bank  $b \in B$ ,

$$\frac{1}{1-\delta}q_m^b \ge \kappa_m^b,\tag{C.2}$$

and

$$\sum_{b\in B} q_m^b = \psi_m. \tag{C.3}$$

Here,  $q_m^b$  is the *quantity* of consumers that bank *b* obtains along the equilibrium path. A given quantity vector  $q_m$  can be implemented by each bank *b* choosing an aggressiveness  $a_m^b = q_m^b$ .<sup>46</sup>

 $<sup>^{46}\</sup>mathrm{A}$  quantity of 0 for bank b can be implemented by having that bank choose an interest rate of 0 (and any level of aggressiveness).

Constraint (C.2) codifies that each bank is better off offering an interest rate of r and its prescribed aggressiveness rather than increasing its aggressiveness to capture more market share. Note that a bank expects 0 future profits after any deviation, since banks expect to simply play the 0-profit stage-game equilibrium of Appendix C.2 after any deviation. Constraint (C.3) is simply an "adding up" constraint: the total quantity of consumers allocated to the banks should equal the total number of consumers.

Summing constraint (C.2) over all firms, and combining it with constraint (C.3) yields

$$\frac{1}{1-\delta}(f-r)\psi_m \ge (f-r)\kappa_m^B$$
$$\frac{1}{1-\delta}\psi_m \ge \kappa_m^B.$$

Rearranging the above yields the result of Proposition C.2.

When total capacity is small (and the discount factor is high), the monopoly interest rate  $r^{\circ}$  (and any higher interest rate) can be supported in equilibrium. For higher levels of capacity, interest rates above the competitive rate can no longer be supported, as at such an interest rate some bank b would be better off increasing its aggressiveness so as to capture exactly  $\kappa_m^b$  demand for one period rather than obtaining  $r^{\circ}$  and its assigned portion of the demand each period. Our result for the single market economy is a generalization of the usual condition for collusion in models of Bertrand competition without capacity constraints: if every bank had capacity  $\psi_m$ , banks would be able to collude if and only if  $(1 - \delta)|B| \leq 1$ .

## C.4 The Multimarket Economy

Using an argument analogous to that for a single-market economy, we can show that the highest sustainable profits can be found by solving the problem

$$\max_{\substack{r \in \times_{m \in M} [r^{\circ}, f], \\ q \in \times_{m \in M} (\times_{b \in \mathbf{F}(m)} [0, \kappa_m^b])}} \left\{ \sum_{m \in M} (f - r_m) \psi_m D(r_m, f) \right\}$$
(C.4)

subject to the constraints that, for each bank  $b \in B$ ,

$$\frac{1}{1-\delta} \sum_{m \in M} (f - r_m) q_m^b D(r_m, f) \ge \sum_{m \in M} (f - r_m) \kappa_m^b D(r_m, f)$$
(C.5)

and, for each  $m \in M$ ,

$$\sum_{b\in B} q_m^b = \psi_m,\tag{C.6}$$

where  $q_m^b = 0$  if  $b \notin \mathbf{F}(m)$ . Here,  $r_m$  is now the highest interest rate offered in market m; a quantity vector  $q_m$  can be implemented by choosing  $a_m^b = \frac{q_m^b}{\sum_{b \in \mathbf{F}(m)} q_m^b}$  for each b such that  $q_m^b > 0$  and, if  $q_m^b = 0$ , by having bank b choose an interest rate strictly less than  $r_m$ .

As in the one-market case, constraint (C.5) codifies that each bank is better off offering  $r_m$ and its prescribed aggressiveness in each market m rather than increasing its aggressiveness and filling its capacity (or total consumer demand) in each market. It is key to our analysis that constraint (C.5) sums over all markets; bank b may be willing to accept a very small quantity in a given market m if it is obtaining substantial profits in other markets. Constraint (C.6) requires that the total supply of consumers allocated to the banks in each market does not exceed the total number of consumers in that market.

However, unlike the one-market case, there is no straightforward way to simplify the set of constraints: the highest-profit equilibrium may require a bank to serve a very small quantity of consumers (relative to its capacity) in one market, while serving a larger quantity of consumers (relative to its capacity) in another market.

### C.5 Merger Ramifications

We first show that—in the context of our model—any merger is profitable for banks.

**Theorem C.1.** Let  $\hat{\kappa}$  be a merger under  $\kappa$ , and suppose that  $\hat{\kappa}$  is sufficient for competition. Then the highest sustainable profits under  $\hat{\kappa}$  are (weakly) higher than the highest sustainable profits under  $\kappa$ . Moreover, even if the merger is a market extension merger, the highest sustainable profits can be strictly higher after the merger.

It is immediate from the analysis of Appendix C.4 that weakly higher profits can be sustained after a merger. If  $\bar{b}$  acquires  $\hat{b}$ , this simply "unifies" the incentive constraints of  $\bar{b}$ and  $\hat{b}$ ; that is, any pair  $(r_m, (q^b)_{b\in\mathbf{F}(m;\kappa)})_{m\in M}$  that satisfies (C.5) and (C.6) under  $\kappa$  generates a pair  $(r_m, (q^b)_{b\in\mathbf{F}(m;\hat{\kappa})})_{m\in M}$  that satisfies (C.5) and (C.6) under  $\hat{\kappa}$  (by increasing the acquirer's quantity from  $q^{\bar{b}}$  to  $q^{\bar{b}} + q^{\hat{b}}$  in every market, setting the quantity of  $\hat{b}$  to 0 in every market, and not changing the quantity of any other bank in any market). Since the set of interest rates and total quantities satisfying the constraints is now weakly larger, the solution to the maximization problem weakly increases.

More surprisingly, a merger can also strictly raise profitability, even when the two banks do not overlap in any market. We demonstrate this in Example 2 below.

**Example 2.** There are two markets, m and n, with  $\psi_m = \psi_n = 1$ ; the liquidity preference  $\lambda = 3$ , and the Fed funds rate f = 1. Under market structure  $\kappa$ , there are two banks, b and  $\hat{b}$ , that are only in market m, i.e.,  $\mathbf{F}(m;\kappa) = \{b, \hat{b}\}$ ; meanwhile, there are 5 other banks in market n. We assume that if a bank is present in a market, it can fully serve that market, i.e.,  $\kappa_m^b = \kappa_m^{\hat{b}} = 1$  and for each  $\bar{b} \in \mathbf{F}(n)$  we have that  $\kappa_n^{\bar{b}} = 1$ . The discount factor is  $\delta = \frac{7}{9}$ .

Since no bank is in both markets, we can analyze each market independently. In the concentrated market m, it follows from Proposition C.2 that monopoly profits can be sustained. Meanwhile, in the competitive market n, the highest sustainable profit is 0.

Now consider the market structure  $\hat{\kappa}$ , under which bank *b* acquires a bank *b* in market *n*; under  $\hat{\kappa}$ , bank *b* now has new capacity  $\hat{\kappa}_n^b = 1$ .

Under  $\hat{\kappa}$ , we can now sustain monopoly profits in both markets. The monopoly interest rate in both markets is  $\frac{1}{3}$ . In one equilibrium supporting such interest rates, there are two phases:

1. The collusive phase: In this phase, each bank offers the interest rate of  $\frac{1}{3}$  in each market in which it present. In market m, both  $\hat{b}$  and b choose the same aggressiveness, so as to set their quantity to  $\frac{1}{2}$ . Meanwhile, in market n, bank b has a quantity of  $q_n^b = \frac{1}{9}$ , and each other bank  $\bar{b}$  present in market n has a quantity of  $q_n^{\bar{b}} = \frac{2}{9}$  (which are obtained by choosing appropriate aggressivenesses in each market).

2. The punishment phase: In this phase, each bank sets its interest rate to the Fed funds rate and chooses an aggressiveness of 1.

Play starts in the collusive phase and continues in the collusive phase so long as no bank deviates; if any bank does so, play continues in the punishment phase. In the punishment phase, play continues in the punishment phase regardless of what happened in-period.

This strategy profile is incentive compatible for all banks. During the punishment phase, it is immediate that each bank is playing optimally given the play of other banks. In the collusion phase, it is optimal for bank  $\hat{b}$  to play its prescribed strategy—instead of increasing its aggressiveness to capture the entire market m—so long as

$$\frac{1}{1-\delta} q_m^{\hat{b}} D(r^{\circ}, f)(f-r^{\circ}) \ge \kappa_m^{\hat{b}} D(r^{\circ}, f)(f-r^{\circ})$$
$$\frac{1}{1-\frac{7}{9}} \frac{1}{2} \ge 1$$
$$\frac{9}{4} \ge 1.$$

Similarly, for each bank present in market n other than  $\overline{b}$ , we need that

$$\frac{1}{1-\delta}q_n^{\bar{b}}D(r^{\circ},f)(f-r^{\circ}) \ge \kappa_n^{\bar{b}}D(r^{\circ},f)(f-r^{\circ})$$
$$\frac{1}{1-\frac{7}{9}}\frac{2}{9} \ge 1$$
$$1 \ge 1.$$

Finally, we show that bank b's strategy is incentive compatible:

$$\frac{1}{1-\delta} \Big( q_m^b D(r^\circ, f)(f-r^\circ) + q_n^b D(r^\circ, f)(f-r^\circ) \Big) \ge \kappa_m^b D(r^\circ, f)(f-r^\circ) + \kappa_n^b D(r^\circ, f)(f-r^\circ) \\ \frac{1}{1-\frac{7}{9}} \Big( \frac{1}{2} + \frac{1}{9} \Big) \ge 1+1 \\ \frac{11}{4} \ge 2.$$

Intuitively, each bank in market n other than b has been allocated a larger share of market n; this share has been chosen to be just large enough so that each local bank in n weakly prefers to price at  $r^{\circ}$  and obtain its allocated share of the market each period rather than to increase its aggressiveness and so obtain the entire market for one period. Meanwhile, bank bobtains a smaller market share than each local bank in market n. However, if bank b were to increase its market share in market n, it would lose its half of the monopoly profits each period in market m (as well as its  $\frac{1}{9}$  of the profits in market n); the value of its shares of profits in markets m and n in each period is greater than its profit from increasing its market share in market n. Even if bank b were to engage in its most profitable deviation—that is, increasing its aggressiveness in market m and market n to capture total market demand in each—its foregone profits in future periods have greater value than its increased profits today. Essentially, bank b has "slack"—in the sense of Bernheim and Whinston (1990)—in its incentive constraint in market m, and it uses that slack to constrain its behavior in market n, i.e., to reduce its supply in market n. This leaves a greater market share for the other firms in market n, and the market share for each other bank in market n is large enough to make deviations unprofitable long-term.

## C.6 The Effects of Market Size and Concentration

In Figure 8, we consider the (post-merger) setting of Example 1 and show how the size of the concentrated market affects the interest rate in the less concentrated market. When the size of market m is 0, it is as if only market n exists, and the only sustainable interest rate is

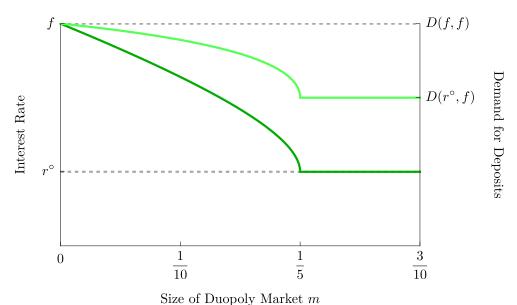
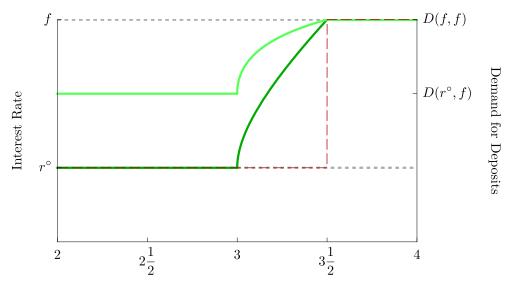


Figure 8: The highest sustainable interest rate  $r_n^{\star}$  in dark green and consumer demand at that rate in light green as a function of  $\psi_m$ , the size of market m. The higher dotted grey line is the Fed funds rate (and competitive interest rate) f; the lower dotted grey line is the monopoly interest rate  $r^{\circ}$ . There are two markets, a duopoly market m and a more competitive market n with five banks; one national bank is present in both markets. We let  $\delta = \frac{7}{9}$ ,  $\psi_n = 1$ ,  $\lambda = 3$ , and f = 1; the capacity of each firm in each market is equal to the market size.

the competitive rate f. As the duopoly market m grows, bank b "acquires" more slack in its incentive constraints; with this additional slack, bank b can accommodate allowing larger profits to local banks in market n. This, in turn, implies that each local bank in n is more willing to forgo the profit from filling its capacity today by increasing its aggressiveness, and so a lower interest rate in market n can be sustained. This effect grows stronger until the size of market m is  $\frac{1}{5}$ , at which point the collusive interest rate can be sustained in both markets. Note that the duopoly market m can be much smaller than the less concentrated market nand yet still allow the banks in market n to collude at the monopoly interest rate.

The amount of slack generated by an uncompetitive market depends not only on the size of the uncompetitive market, but also on how uncompetitive it is. In Figure 9, the dark green line shows how the interest rate in market n varies with the competitiveness of market m; here, instead of one other local bank with capacity of 1 in market m, we have 5 other local



Capacity of Rival Banks in Market m

Figure 9: The highest sustainable interest rate  $r_n^{\star}$  in dark green, consumer demand at that rate in light green, and the highest sustainable interest rate  $r_m^{\star}$  as a dark red dashed line, as a function of  $\kappa_m^{\mathbf{L}(m)}$ , the capacity of the local banks in market m. The higher dotted grey line is the Fed funds rate (and competitive interest rate) f; the lower dotted grey line is the monopoly interest rate  $r^{\circ}$ . There are two markets, a market m and a more competitive market n with five banks; one national bank is present in both markets. We let  $\delta = \frac{7}{9}$ ,  $\psi_n = 1$ ,  $\lambda = 3$ , and f = 1; the capacity of each firm in market n is equal to the market size.

banks with limited capacity in market m.<sup>47</sup> When market m is very uncompetitive, i.e., local banks in m have little capacity, not only can the monopoly interest rate be sustained in market m, but it can also be sustained in market n. As market m becomes more competitive, the interest rate in market n rises, since the slack available from market m falls; in Figure 9, this effect begins when the capacity of local banks is 3. Finally, once market m becomes competitive, no interest rate other than the competitive rate f can be sustained in market m, and so there is no slack left with which to sustain an interest rate lower than the competitive rate in market n; in Figure 9, this happens when the capacity of local banks is  $3\frac{1}{2}$ .

<sup>&</sup>lt;sup>47</sup>The key parameter is the total capacity of the local banks in market m, not the number of local banks. However, it is key that the number of local banks in market m is high enough that market m would be competitive if each bank had capacity equal to the market size.

## C.7 Characteristics of Highest-Profit Equilibria

We now state the two results which motivate our empirical analysis.

Our first result characterizes how the profit-maximizing interest rates differ across markets with different levels of local bank capacity. We show that, if market m has lower local bank capacity—i.e., is less competitive—than market n, then market m will have a higher spread than market n (holding the capacities of the national banks in the two markets constant). Moreover, market m will have a higher *capture rate* than market n.<sup>48</sup>

**Theorem C.2.** Suppose that for two markets m and n we have that  $\kappa_m^{\mathbf{L}(m)} \leq \kappa_n^{\mathbf{L}(n)}$ ,  $\kappa_m^b = \kappa_n^b$ for each  $b \in \mathbf{N}(m) = \mathbf{N}(n)$ , and  $\psi_m = \psi_n$ . Then, in any highest-profit equilibrium,  $s_m \geq s_n$ ; moreover,  $\frac{\partial s_m}{\partial f} \geq \frac{\partial s_n}{\partial f}$ .

Our second result characterizes how two markets that only differ with respect to multimarket contact will differ with respect to spreads and capture rates. Suppose that markets m and n have the same number of banks, but the set of national banks in m is a superset of the set of national banks in n: Then market m will have a higher spread and capture rate than market n.

**Theorem C.3.** Suppose that for two markets m and n and a bank  $b \in \mathbf{N}(m)$  we have that  $\kappa_m^{\mathbf{L}(m)} + \kappa_m^b < \kappa_n^{\mathbf{L}(n)}, \ \kappa_m^{\bar{b}} = \kappa_n^{\bar{b}}$  for each  $\bar{b} \in \mathbf{N}(n) = \mathbf{N}(m) \setminus \{b\}$ , and  $\psi_m = \psi_n$ . Then, in any highest-profit equilibrium,  $s_m \ge s_n$ ; moreover,  $\frac{\partial s_m}{\partial f} \ge \frac{\partial s_n}{\partial f}$ .

## C.8 Relation of the Model to Deposit Banking

Our model, while parsimonious, is structured to capture aspects of the deposit banking market necessary to understand how multimarket contact influences deposit rates.

Our assumption that a bank has a given capacity in a given market expresses the idea that banks face constraints (e.g., location space, staffing) with respect to how many consumers

<sup>&</sup>lt;sup>48</sup>Capture rate describes what fraction of an increase in the Fed funds rate is captured by banks as opposed to being passed onto consumers; it is thus analogous to one less the passthrough rate in standard models of product competition. In our setting, the capture rate is 0 under perfect competition and  $1 - \frac{\sqrt{1+\lambda}-1}{\lambda} \in (0,1)$  under monopoly.

the bank can simultaneously serve, and that these constraints are difficult to change quickly. When banks are capacity constrained, collusion is easier since the payoff from deviating and attracting additional customers is smaller. We also assume that banks can choose how "aggressive" they are; this allows banks to allocate demand amongst themselves, not only ensuring that no bank has more customers than its capacity but also that banks can allocate demand so as to facilitate collusion.<sup>49</sup> However, since being more aggressive has zero cost, a bank can capture the entire market if it chooses to by offering the best deposit rate, as is standard in models of undifferentiated Bertrand competition. In a model with differentiated banks, modeling aggressiveness would not be necessary, since banks could fine-tune their demand by changing prices slightly.<sup>50</sup>

We also simplify the model by assuming that banks face the same marginal cost for servicing deposits (which we normalize to zero). While this assumption is not realistic, allowing for heterogeneous costs would complicate the model without adding any additional insight.

## C.9 Proof of Proposition C.2

First, note that Proposition C.1 shows that there exists a subgame-perfect Nash equilibrium of the stage game in which each bank obtains 0 profits—its lowest individually rational payoffs. Second, given a strategy profile  $(r^b, a^b)_{b\in B}$ , the action that maximizes current-period payoffs for bank b is to choose  $\hat{r}^b = \max_{\bar{b}\in B} \{r^{\bar{b}}\}$  and an aggressiveness such that b's demand is exactly its capacity.<sup>51</sup> Third, it is immediate that demand will be given by  $D_m(\max_{\bar{b}\in B} \{r^{\bar{b}}\}, f)$ .

<sup>&</sup>lt;sup>49</sup>Examples of how a bank could be more or less aggressive include making it more or less easy to open new accounts, by engaging in more or less advertising, or having more or less convenient hours.

 $<sup>^{50}</sup>$ However, in such a model, calculating the profit-maximizing equilibrium for an arbitrary set of bank capacities would be exceedingly complex, since calculating Abreu (1988) optimal penal codes would be far more challenging in that setting.

<sup>&</sup>lt;sup>51</sup>If b's capacity is (weakly) greater than demand at  $\hat{r}^b$ , then b can obtain demand arbitrarily close to  $D_m(\hat{r}^b, f)$  by choosing a high-enough aggressiveness.

Thus, it follows from Abreu (1988) that the highest-profit equilibrium is the solution to<sup>52</sup>

$$\max_{\substack{r \in [r^{\circ}, f], \\ q \in \psi \cdot \Delta^B}} \left\{ (f - r)\psi D(r, f) - \sum_{b \in B} c \left[ q^b - \kappa^b \right]^+ \right\}$$
(C.7)

subject to the constraints that, for each bank  $b \in B$ ,

$$\frac{1}{1-\delta}\psi(f-r)q^b D(r,f) \ge (f-r)\kappa^b D(r,f)$$
(C.8)

where  $q^b$  is the quantity of demand enjoyed by bank b. The incentive compatibility constraint (C.8) encodes the equilibrium requirement that it is better for each bank b to obtain its quantity  $q^b$  each period rather than obtain its optimal quantity of deposits for one period and 0 profits thereafter. The 0 profit equilibrium is an Abreu (1988) optimal penal code.<sup>53</sup> Any quantity vector in  $D(f, r) \cdot \Delta^B$  can be implemented by choosing  $r^b = r$  and  $a^b = q^b$  for every bank such that  $q^b > 0$ , and  $r^b < r$  and any aggressiveness for any bank such that  $q^b = 0$ .

Since each bank b is more profitable as quantity increases up to  $\kappa^b$ , and less profitable as quantity increases past  $\kappa^b$ , the solution to (C.7) must set each bank's quantity at no more than its capacity. Thus, the solution to (C.7) is the same as the solution to

$$\max_{\substack{r \in [r^{\circ}, f], \\ q \in \times_{b \in B}[0, \kappa^{b}]}} \{(f - r)D(r, f)\}$$

subject to the constraints that

$$\frac{1}{1-\delta}(f-r)q^b \ge (f-r)\kappa^b$$

<sup>&</sup>lt;sup>52</sup>We simplify the notation in the proof by assuming that  $B = \mathbf{F}(m)$  and dropping the *m* subscript where appropriate. Additionally,  $\Delta^B$  is the *B*-dimensional simplex.

 $<sup>^{53}</sup>$ For an excellent discussion of optimal penal codes, see Proposition 2.6.1 and the surrounding text in Mailath and Samuelson (2006).

and

$$\sum_{b\in B}q_m^b=\psi.$$

That the solution to this maximization problem is that given in Proposition 2 follows as in the text of Appendix C.3.

# C.10 Proof of Theorem C.1

The arguments to prove Proposition C.2 can be used *mutatis mutandis* to show that the highest sustainable profits can be found by solving the problem

$$\max_{\substack{r \in \times_{m \in M} [r^{\circ}, f], \\ q \in \times_{m \in M} (\times_{b \in B} [0, \kappa_m^b])}} \left\{ \sum_{m \in M} \psi_m (f - r_m) D_m (f, r_m) \right\}$$

subject to the constraint that, for each bank  $b \in B$ ,

$$\frac{1}{1-\delta}\sum_{m\in M}q_m^b D_m(f,r_m)(f-r_m) \ge \sum_{m\in M}(f-r_m)D_m(f,r_m)\kappa_m^b$$

and, for each  $m \in M$ ,

$$\sum_{b \in B} q_m^b \le \psi_m.$$

The first result of the theorem then follows immediately from the fact that under the postmerger market structure we require that the sum of the incentive constraints for the merging banks are satisfied instead of requiring that the incentive constraint of each merging bank is satisfied.

The second claim of the theorem is shown by Example 2.

## C.11 Proof of Theorem C.2

## C.11.1 Spreads and Interest Rates

Suppose, by way of contradiction, that there exists a highest-profit equilibrium—that is, a pair  $(r_k, q_k^b)_{k \in M, b \in B}$  satisfying constraints (C.5) and (C.6)—for which  $s_n > s_m$ , i.e.,  $r_n < r_m$ ; we will show that there exists a pair  $(\hat{r}_k, \hat{q}_k^b)_{k \in M, b \in B}$  also satisfying constraints (C.5) and (C.6) with higher total profits.

For ease of exposition, we let  $\Pi_k(r_k) \equiv (f - r_k)\psi_k D_k(r_k, f)$  be the total profits in market k for the interest rate  $r_k$ .

Reallocating Quantity to National Banks: Given that  $(r_k, q_k^b)_{k \in M, b \in B}$  satisfies the constraints (C.5) and (C.6), we first show that  $(r_k, \bar{q}_k^b)_{k \in M, b \in B}$  satisfies the constraints (C.5) and (C.6), where

1

$$\bar{q}_k^b = \begin{cases} (1-\delta)\kappa_m^b & b \in \mathbf{L}(m) \\ q_m^b \left(1 + \frac{q_m^{\mathbf{L}(m)} - \bar{q}_m^{\mathbf{L}(m)}}{|\mathbf{N}(m)|}\right) & b \in \mathbf{N}(m) \\ q_k^b & \text{otherwise.} \end{cases}$$

That is, we reallocate quantity to the national banks in market m such that each local bank in market m now satisfies constraint (C.5) with equality. Moreover, each national bank  $b \in \mathbf{N}(m)$  has gained quantity in market m, and so it is immediate that constraint (C.5) still holds for b under  $(r_k, \bar{q}_k^b)_{k \in M, b \in B}$ .<sup>54,55</sup> Finally, for any bank not in market m, constraint (C.5) has not changed. (Constraint (C.6) is trivially still satisfied

 $<sup>\</sup>overline{f_{k}^{54}}$ Since  $(r_k, q_k^b)_{k \in M, b \in B}$  satisfies the constraints (3) and (C.6), each local bank *b* must have  $q_m^b \ge (1-\delta)\kappa_m^b = \overline{q_m^b}$ .

 $<sup>\</sup>bar{q}_m^b$ . <sup>55</sup>Technically, it is possible that  $q_m^b > \kappa_m^b$  for some national bank  $b \in \mathbf{N}(m)$ . In this case, simply reallocate  $q_m^{\mathbf{L}(m)} - \bar{q}_m^{\mathbf{L}(m)}$  to the national banks in market m in any way so that no national bank has more quantity than its capacity. If that is not possible, then our result is immediate: Reallocate to each national bank a quantity equal to its capacity, while still leaving each local bank a quantity of at least  $(1 - \delta)\kappa_m^b$ . Then decrease the interest rate of market m by  $\epsilon > 0$ ; this makes the market more profitable (as  $r_m > r_n \ge r^\circ$ ) and Constraint (C.5) is still satisfied for each bank in m (as each national bank already has quantity equal to its capacity and each local bank  $b \in \mathbf{L}(m)$  has  $q_m^b \ge (1 - \delta)\kappa_m^b$ ).

as we have just reallocated quantity from local banks to national banks in market m.) We can reallocate quantities in market n in a similar manner.<sup>56</sup> Thus, we assume for the remainder that  $(r_k, q_k^b)_{k \in M, b \in B}$  has the property that  $q_k^b = (1 - \delta)\kappa_k^b$  for  $k \in \{m, n\}$ and  $b \in \mathbf{L}(k)$ .

Reallocating Quantity Among National Banks: We next show that there exists another equilibrium  $(r_m, \bar{q}_m^b)_{m \in M, b \in B}$  with the same interest rate in each market such that each national bank in m obtains the same ratio of consumers in markets m and n and has the same profits as in the  $(r_m, q_m^b)_{m \in M, b \in B}$  equilibrium. Let

$$\bar{q}_k^b = \begin{cases} \frac{\left(1-\kappa_k^{\mathbf{L}(k)}\right)\left(\Pi_m(r_m)q_m^b+\Pi_n(r_n)q_n^b\right)}{\left(1-\kappa_m^{\mathbf{L}(m)}\right)\Pi_m(r_m)+\left(1-\kappa_n^{\mathbf{L}(n)}\right)\Pi_n(r_n)} & b \in \mathbf{N}(m) \text{ and } k \in \{m, n\}\\ q_k^b & \text{otherwise.} \end{cases}$$

Note that  $\bar{q}_k^b \in [q_m^b, q_n^b] \subseteq [0, \kappa_b^m] = [0, \kappa_b^n]$  for all  $k \in \{m, n\}$  and  $b \in \mathbf{N}(m)$ . Also, note that  $\bar{q}_m^b \geq \bar{q}_n^b$  for all  $b \in \mathbf{N}(m)$  as  $\kappa_m^{\mathbf{L}(m)} \leq \kappa_n^{\mathbf{L}(n)}$ .

It is immediate that industry profits are the same in the  $(r_m, q_m^b)_{m \in M, b \in B}$  and  $(r_m, \bar{q}_m^b)_{m \in M, b \in B}$ equilibria as the interest rate in each market is the same. Furthermore, the constraint (C.5) is satisfied as:

<sup>&</sup>lt;sup>56</sup>If there is no way to reallocate quantities in market n so that each local bank  $b \in \mathbf{L}(n)$  has  $\bar{q}_n^k = (1-\delta)\kappa_n^b$ while leaving each national bank a quantity less than its capacity, then it must have been true that there was no way to reallocate quantities in market m so that each local bank  $b \in \mathbf{L}(m)$  has  $\bar{q}_m^k = (1-\delta)\kappa_m^b$  while leaving each national bank a quantity less than its capacity (as  $\kappa_n^{\mathbf{L}(n)} \ge \kappa_m^{\mathbf{L}(m)}$  and  $\kappa_n^{\mathbf{N}(n)} \ge \kappa_m^{\mathbf{N}(m)}$ ). Thus, the discussion of Footnote 55 would apply.

1. For every bank  $b \in \mathbf{N}(m) = \mathbf{N}(n)$ , we have that

$$\begin{split} \sum_{k \in M} \Pi_k(r_k) \bar{q}_k^b &= \ \Pi_m(r_m) \bar{q}_m^b + \Pi_n(r_n) \bar{q}_n^b + \sum_{k \in M \smallsetminus \{m,n\}} \Pi_k(r_k) \bar{q}_k^b \\ &= \ \Pi_m(r_m) \frac{\left(1 - \kappa_m^{\mathbf{L}(m)}\right) \left(\Pi_m(r_m) q_m^b + \Pi_n(r_n) q_n^b\right)}{\left(1 - \kappa_m^{\mathbf{L}(n)}\right) \Pi_m(r_m) + \left(1 - \kappa_n^{\mathbf{L}(n)}\right) \Pi_n(r_m)} + \\ &= \ \Pi_n(r_n) \frac{\left(1 - \kappa_m^{\mathbf{L}(m)}\right) \left(\Pi_m(r_m) q_m^b + \Pi_n(r_n) q_n^b\right)}{\left(1 - \kappa_m^{\mathbf{L}(m)}\right) \Pi_m(r_m) + \left(1 - \kappa_n^{\mathbf{L}(n)}\right) \Pi_n(r_m)} + \\ &= \ \sum_{k \in M \smallsetminus \{m,n\}} \Pi_k(r_k) \bar{q}_k^b \\ &= \ \frac{\left(1 - \kappa_m^{\mathbf{L}(m)}\right) \Pi_m(r_m) + \left(1 - \kappa_n^{\mathbf{L}(n)}\right) \Pi_n(r_m)}{\left(1 - \kappa_m^{\mathbf{L}(m)}\right) \Pi_m(r_m) + \left(1 - \kappa_n^{\mathbf{L}(n)}\right) \Pi_n(r_m)} \Pi_m(r_m) q_m^b + \\ &= \ \frac{\left(1 - \kappa_m^{\mathbf{L}(m)}\right) \Pi_m(r_m) + \left(1 - \kappa_n^{\mathbf{L}(n)}\right) \Pi_n(r_m)}{\left(1 - \kappa_m^{\mathbf{L}(m)}\right) \Pi_m(r_m) + \left(1 - \kappa_n^{\mathbf{L}(n)}\right) \Pi_n(r_m)} \Pi_n(r_n) q_m^b + \\ &= \ \sum_{k \in M \smallsetminus \{m,n\}} \Pi_k(r_k) \bar{q}_k^b \\ &= \ \Pi_m(r_m) q_m^b + \Pi_n(r_n) q_n^b + \sum_{k \in M \smallsetminus \{m,n\}} \Pi_k(r_k) \bar{q}_k^b \\ &= \ \sum_{k \in M} \Pi_k(r_k) \bar{q}_k^b \end{split}$$

and so (C.5) is satisfied as the left-hand-side of (C.5) has not changed in value (and the right-hand-side of (C.5) is unchanged).

2. For every bank  $b \notin \mathbf{N}(m) = \mathbf{N}(n)$ , we have that (C.5) is unchanged.

Finally, constraint (C.6) is satisfied as:

### 1. We calculate that

$$\begin{split} \sum_{b \in \mathbf{N}(m)} \bar{q}_m^b &= \sum_{b \in \mathbf{N}(m)} \frac{\left(1 - \kappa_m^{\mathbf{L}(m)}\right) \left(\Pi_m(r_m) q_m^b + \Pi_n(r_n) q_n^b\right)}{\left(1 - \kappa_m^{\mathbf{L}(m)}\right) \Pi_m(r_m) + \left(1 - \kappa_n^{\mathbf{L}(n)}\right) \Pi_n(r_n)} \\ &= \frac{\left(1 - \kappa_m^{\mathbf{L}(m)}\right) \left(\Pi_m(r_m) \sum_{b \in \mathbf{N}(m)} q_m^b + \Pi_n(r_n) \sum_{b \in \mathbf{N}(m)} q_n^b\right)}{\left(1 - \kappa_m^{\mathbf{L}(m)}\right) \Pi_m(r_m) + \left(1 - \kappa_n^{\mathbf{L}(n)}\right) \Pi_n(r_n)} \\ &= \frac{\left(1 - \kappa_m^{\mathbf{L}(m)}\right) \left(\Pi_m(r_m) \left(1 - \kappa_m^{\mathbf{L}(m)}\right) + \Pi_n(r_n) \left(1 - \kappa_n^{\mathbf{L}(n)}\right)\right)}{\left(1 - \kappa_m^{\mathbf{L}(m)}\right) \Pi_m(r_m) + \left(1 - \kappa_n^{\mathbf{L}(n)}\right) \Pi_n(r_n)} \\ &= 1 - \kappa_m^{\mathbf{L}(m)} \\ &= \sum_{b \in \mathbf{N}(m)} q_m^b; \end{split}$$

thus, since total quantity assigned to local banks under  $\bar{q}_m$  is the same as under  $q_m$ , Constraint (C.6) is satisfied in market m.

- 2. The argument for market n is analogous to that for market m.
- 3. For every market  $k \in M \setminus \{m, n\}$  the quantity of each bank is unchanged under  $\bar{q}_k$ .
- **Changing Interest Rates:** We now consider a strategy profile that allocates the same quantity to each bank in each market but decreases  $r_m$  by a small  $\epsilon > 0$ , increases  $r_n$  by  $\epsilon$ , and leaves  $r_k$  unchanged for other markets  $k \in M \setminus \{m, n\}$ .

First, we calculate the first-order change in profits (recalling that  $\psi_m = \psi_n$ ) as

$$\frac{\partial [\sum_{k \in M} (f - r_k) D(r_k, f) \psi_k]}{\partial \epsilon} \bigg|_{\epsilon = 0} = \psi_m \bigg( \frac{\lambda (r_m)^2 + 2fr_m - f^2}{(f + \lambda r_m)^2} - \frac{\lambda (r_n)^2 + 2fr_n - f^2}{(f + \lambda r_n)^2} \bigg)$$

which is positive as

$$\frac{\partial \left[\frac{\lambda r^2 + 2fr - f^2}{(f + \lambda r)^2}\right]}{\partial r} = \frac{2f^2(1+\lambda)}{(f + \lambda r)^3} > 0 \tag{C.9}$$

and  $r_m > r_n$ .

Second, we show that Constraint (C.5) holds for a small  $\epsilon > 0$ . It is immediate that it still holds for all banks not in either market m or market n and that it holds for each local bank in either market m or market n. For a national bank  $b \in \mathbf{N}(m)$ , we have that

$$\frac{\partial}{\partial \epsilon} \left[ \sum_{k \in M} (f - r_k) D(r_k, f) \left( \frac{\bar{q}_k^b}{1 - \delta} - \kappa_k^b \right) \right] = \psi_m \left( \frac{\lambda(r_m)^2 + 2fr_m - f^2}{(f + \lambda r_m)^2} \left( \frac{\bar{q}_m^b}{1 - \delta} - \kappa_m^b \right) - \frac{\lambda(r_n)^2 + 2fr_n - f^2}{(f + \lambda r_n)^2} \left( \frac{\bar{q}_n^b}{1 - \delta} - \kappa_n^b \right) \right)$$

which is non-negative since:<sup>57</sup>

• 
$$\bar{q}_m^b \ge \bar{q}_n^b$$
 as  $\kappa_m^{\mathbf{L}(m)} \le \kappa_n^{\mathbf{L}(n)}$  and  $\kappa_m^b = \kappa_n^b$ ;  
•  $\frac{\lambda(r_m)^2 + 2fr_m - f^2}{(f + \lambda r_m)^2} > \frac{\lambda(r_n)^2 + 2fr_n - f^2}{(f + \lambda r_n)^2}$  (see (C.9) above).

Third, it is immediate that Constraint (C.6) holds as quantities did not change.

Thus, we have constructed a new equilibrium with strictly higher profits, contradicting our supposition that the highest profit equilibrium had  $r_n > r_m$ .

## C.11.2 Capture Rates

Finally, note that the maximization problem (C.4) subject to (C.5) and (C.6) is homogeneous of degree 1 in f. Thus, letting  $\rho_m = \frac{r_m}{f}$ , we can rewrite the maximization problem as

$$\max_{\substack{\rho \in \times_{m \in M} \left[\frac{\sqrt{1+\lambda}-1}{\lambda}, 1\right], \\ q \in \times_{m \in M} \left(\times_{b \in \mathbf{F}(m)}[0, \kappa_m^b]\right)}} \left\{ \sum_{m \in M} (1 - \rho_m) \psi_m D(\rho_m, 1) \right\}$$

<sup>&</sup>lt;sup>57</sup>Note that this argument relies on our specific choice of the demand for deposits. Nevertheless, the result still holds for the more general demand for deposits given in Appendix C.13, but the proof now requires using the first-order condition for the monopoly interest rate.

subject to the constraint that, for each bank  $b \in B$ ,

$$\frac{1}{1-\delta} \sum_{m \in M} (1-\rho_m) q_m^b D(\rho_m, 1) \ge \sum_{m \in M} (1-\rho_m) \kappa_m^b D(\rho_m, 1),$$

and, for each  $m \in M$ ,

$$\sum_{b \in B} q_m^b = \psi_m.$$

Hence, since  $r_m \leq r_n$ , we have that  $\rho_m \leq \rho_n$  and so  $\frac{\partial r_m}{\partial f} \leq \frac{\partial r_n}{\partial f}$ , i.e.,  $\frac{\partial s_m}{\partial f} \geq \frac{\partial s_n}{\partial f}$ .

# C.12 Proof of Theorem C.3

#### C.12.1 Spreads and Interest Rates

Suppose, by way of contradiction, that there exists a highest-profit equilibrium—that is, a pair  $(r_k, q_k^b)_{k \in M, b \in B}$  satisfying constraints (C.5) and (C.6)—for which  $s_n > s_m$ , i.e.,  $r_n < r_m$ ; we will show that there exists a pair  $(\hat{r}_k, \hat{q}_k^b)_{k \in M, b \in B}$  also satisfying constraints (C.5) and (C.6) with higher total profits.

For ease of exposition, we let  $\Pi_k(r_k) \equiv (f - r_k)\psi_k D_k(r_k, f)$  be the total profits in market k for the interest rate  $r_k$ .

- Reallocating Quantity to National Banks: This follows as in the proof of Theorem C.2. Thus, we assume for the remainder that  $(r_k, q_k^b)_{k \in M, b \in B}$  has the property that  $q_k^b = (1 - \delta)\kappa_k^b$  for  $k \in \{m, n\}$  and  $b \in \mathbf{L}(k)$ .
- **Reallocating Quantity Among National Banks:** This also follows as in the proof of Theorem C.2. Thus, we assume for the remainder that  $(r_k, q_k^b)_{k \in M, b \in B}$  has the property that  $q_n^{\bar{b}} = \sigma_n^{\bar{b}} q_n^{\mathbf{N}(n)}$  for  $\sigma \in \Delta^{\mathbf{N}(n)}$  and that  $q_m^{\bar{b}} = \varphi q_n^{\bar{b}}$  for  $\bar{b} \in \mathbf{N}(n)$  and some  $\varphi > 0$ ; this

implies that

$$\varphi = \frac{q_m^{\mathbf{N}(n)}}{q_n^{\mathbf{N}(n)}} = \frac{\psi_m - q_m^b - (1 - \delta)\kappa_m^{\mathbf{L}(m)}}{\psi_n - (1 - \delta)\kappa_n^{\mathbf{L}(n)}}$$
(C.10)

Changing Interest Rates and Quantities: We now consider a small change in interest rates and quantities in markets m and n such that:

- 1. Industry profits remain constant.
- 2. Each local bank  $\bar{b} \in \mathbf{L}(m) \cup \mathbf{L}(n)$  retains a quantity of  $(1-\delta)\kappa_{\bar{b}}^k$  in each market k.
- 3. Bank b's incentive constraint (C.5) still holds.
- 4. For each national bank  $\bar{b} \in \mathbf{N}(n)$ , we have that bank  $\bar{b}$ 's incentive constraint (C.5) holds with *strict* inequality.

To do this, let  $\epsilon > 0$  be small. Furthermore, let

$$\hat{r}_m = r_m - \epsilon$$
  
 $\hat{r}_n = r_n + \gamma(\epsilon)$ 

where  $\gamma(\epsilon)$  is chosen so that industry profits do not change, i.e.,  $\Pi_m(\hat{r}_m) + \Pi_n(\hat{r}_n) = \Pi_m(r_m) + \Pi_n(r_n)$ . Let the quantity of bank *b* in market *m* be given by<sup>58</sup>

$$\hat{q}_m^b = (1-\delta) \left( \frac{\Pi_m(r_m)}{\Pi_m(\hat{r}_m)} \left( \frac{q_m^b}{1-\delta} - \kappa_m^b \right) + \kappa_m^b \right), \tag{C.11}$$

the quantities of banks  $\bar{b} \in \mathbf{N}(n)$  be given by

$$\hat{q}_m^{\bar{b}} = q_m^{\bar{b}} + \sigma^{\bar{b}}(q_m^b - \hat{q}_m^b)$$
$$\hat{q}_n^{\bar{b}} = q_n^{\bar{b}},$$

<sup>58</sup>Note that if  $q_m^b = 0$ , it follows from (C.11) that  $\hat{q}_m^b > 0$ .

and let the quantities of local banks in markets m and n, as well as all quantities in markets other than m and n, remain unchanged; note that quantities in each market have the same sum as before, i.e., Constraint (C.6) is satisfied in each market.

It is immediate by the definition of  $\gamma$  that industry profits do not change.

It is immediate that Constraint (C.5) is still satisfied for each local bank  $\bar{b} \in \mathbf{L}(m) \cup \mathbf{L}(n)$ in markets m and n, as we still have that  $\hat{q}_k^{\bar{b}} = (1 - \delta)\kappa_k^{\bar{b}}$  for  $k \in \{m, n\}$ .

We now show that Constraint (C.5) holds for b: Since Constraint (C.5) held for the initial equilibrium, we have<sup>59</sup>

$$\frac{1}{1-\delta} \left( \Pi_m(r_m) q_m^b + \sum_{k \in M \smallsetminus \{m\}} \Pi_k q_k^b \right) \ge \Pi_m(r_m) \kappa_m^b + \sum_{k \in M \smallsetminus \{m\}} \Pi_k \kappa_k^b$$

$$\frac{1}{1-\delta} \Pi_m(r_m) q_m^b + \frac{1}{1-\delta} \left( \sum_{k \in M \smallsetminus \{m\}} \Pi_k q_k^b \right) \ge \Pi_m(r_m) \kappa_m^b + \sum_{k \in M \smallsetminus \{m\}} \Pi_k \kappa_k^b$$

$$\Pi_m(r_m) \kappa_m^b + \Pi_m(\hat{r}_m) \left( \frac{\hat{q}_m^b}{1-\delta} - \kappa_m^b \right) + \frac{1}{1-\delta} \left( \sum_{k \in M \smallsetminus \{m\}} \Pi_k q_k^b \right) \ge \Pi_m(r_m) \kappa_m^b + \sum_{k \in M \smallsetminus \{m\}} \Pi_k \kappa_k^b$$

$$\Pi_m(\hat{r}_m) \frac{\hat{q}_m^b}{1-\delta} - \Pi_m(\hat{r}_m) \kappa_m^b + \frac{1}{1-\delta} \left( \sum_{k \in M \smallsetminus \{m\}} \Pi_k q_k^b \right) \ge \sum_{k \in M \smallsetminus \{m\}} \Pi_k \kappa_k^b$$

$$\frac{1}{1-\delta} \left( \Pi_m(\hat{r}_m) \frac{\hat{q}_m^b}{1-\delta} + \sum_{k \in M \smallsetminus \{m\}} \Pi_k q_k^b \right) \ge \Pi_m(\hat{r}_m) \kappa_m^b + \sum_{k \in M \smallsetminus \{m\}} \Pi_k \kappa_k^b$$

where the third line follows from solving (C.11) for  $q_m^b$ , substituting it into the second line, and simplifying.

We calculate the change in the incentive constraint for each bank  $\bar{b} \in \mathbf{N}(n)$  for a small <sup>59</sup>Here, we let  $\Pi_k \equiv \Pi_k(r_k) = \Pi_k(\hat{r}_k)$  for brevity.

 $\epsilon > 0$  as

$$\begin{split} &\frac{\partial}{\partial \epsilon} \bigg[ \Pi_m(\hat{r}_m) \bigg( \frac{\hat{q}_m^{\overline{b}}}{1-\delta} - \kappa_m^{\overline{b}} \bigg) + \Pi_n(\hat{r}_n) \bigg( \frac{\hat{q}_n^{\overline{b}}}{1-\delta} - \kappa_n^{\overline{b}} \bigg) \bigg|_{\epsilon=0} \\ &= \frac{\partial}{\partial \epsilon} \bigg[ \Pi_m(\hat{r}_m) \bigg( \frac{\varphi \sigma^{\overline{b}} q_n^{\mathbf{N}(n)} + \sigma^{\overline{b}} (q_m^b - \hat{q}_m^b)}{1-\delta} - \kappa_m^{\overline{b}} \bigg) + \Pi_n(\hat{r}_n) \bigg( \frac{\sigma^{\overline{b}} q_n^{\mathbf{N}(n)}}{1-\delta} - \kappa_n^{\overline{b}} \bigg) \bigg|_{\epsilon=0} \\ &= \Pi'_m(r_m) \bigg( \frac{\varphi \sigma^{\overline{b}} q_n^{\mathbf{N}(n)}}{1-\delta} - \kappa_m^{\overline{b}} \bigg) + \sigma^{\overline{b}} \frac{\Pi_m(r_m)}{1-\delta} \bigg( \frac{(1-\delta)\Pi'_m(r_m)}{\Pi_m(r_m)} \bigg( \frac{q_m^b}{1-\delta} - \kappa_m^b \bigg) \bigg) + \Pi'_n(r_n) \bigg( \frac{\sigma^{\overline{b}} q_n^{\mathbf{N}(n)}}{1-\delta} - \kappa_n^{\overline{b}} \bigg) \\ &= \Pi'_m(r_m) \bigg( \bigg( \frac{\varphi \sigma^{\overline{b}} q_n^{\mathbf{N}(n)}}{1-\delta} - \kappa_m^{\overline{b}} \bigg) + \sigma^{\overline{b}} \bigg( \frac{q_m^b}{1-\delta} - \kappa_m^b \bigg) - \bigg( \frac{\sigma^{\overline{b}} q_n^{\mathbf{N}(n)}}{1-\delta} - \kappa_n^{\overline{b}} \bigg) \bigg) \\ &= \Pi'_m(r_m) \sigma^{\overline{b}} \bigg( \bigg( \frac{q_n^{\mathbf{N}(n)}}{1-\delta} (\varphi - 1) \bigg) + \bigg( \frac{q_m^b}{1-\delta} - \kappa_m^b \bigg) \bigg) \\ &= \Pi'_m(r_m) \frac{\sigma^{\overline{b}}}{1-\delta} \bigg( (\psi_m - q_m^b - (1-\delta)\kappa_m^{\mathbf{L}(m)}) - (\psi_n - (1-\delta)\kappa_n^{\mathbf{L}(n)}) + (q_m^b - (1-\delta)\kappa_m^b) \bigg) \\ &= \Pi'_m(r_m) \sigma^{\overline{b}} \bigg( -(1-\delta)\kappa_m^{\mathbf{L}(m)} + (1-\delta)\kappa_n^{\mathbf{L}(n)} + (1-\delta)\kappa_m^b \bigg) \\ &= \Pi'_m(r_m) \sigma^{\overline{b}} \bigg( \kappa_n^{\mathbf{L}(n)} - \bigg( \kappa_m^b + \kappa_m^{\mathbf{L}(m)} \bigg) \bigg) \end{aligned}$$

where the second line follows from the definition of  $\hat{q}_n^{\bar{b}}$ , the third line follows from the definition of  $\hat{q}_m^{\bar{b}}$ , the fourth line follows as  $\Pi'_m(r_m) = -\Pi'_n(r_n)$  by the definition of  $\gamma$ , the fifth line follows as the capacity of each national bank in n is the same in market m, the sixth line follows from the definition of  $\varphi$  given by (C.10), the seventh line follows as  $\psi_m = \psi_n$  by assumption, and the last line follows as profits in market m are increasing in  $\epsilon$  and  $\kappa_m^b + \kappa_m^{\mathbf{L}(m)} < \kappa_n^{\mathbf{L}(n)}$  by assumption.

Finally, note that the incentive constraint of no other bank has changed, as the quantities and interest rates in every market other than m and n have remained unchanged.

**Increasing Profits:** Since at least one bank in market m has slack in its incentive constraint, we can increase the interest rate in market m (and rearrange quantities if necessary) by some small  $\check{\epsilon} > 0$  to increase profits while constraints (C.5) and (C.6) still hold.

#### C.12.2 Capture Rates

Given our result on interest rates, the result on spreads follows as in the proof of Theorem 2 *mutatis mutandis*.

## C.13 Consumer Demand for Deposits

Following Drechsler et al. (2017), the representative consumer maximizes his utility over final wealth and liquid deposits according to a constant elasticity-of-substitution utility function, i.e., he solves

$$\max_{D(r,f)\in[0,1]} \left\{ ((w - D(r,f))f)^{1-\frac{1}{s}} + \lambda^{\frac{1}{s}} (rD(r,f))^{1-\frac{1}{s}} \right\};$$

here,  $\lambda$  is the preference for liquidity, w is initial wealth, and s is the elasticity of substitution between wealth and liquid deposits.

Solving the agent's problem, we obtain

$$D(r,f) = w\lambda \frac{r^{s-1}}{f^{s-1} + \lambda r^{s-1}}$$

Normalizing wealth w to  $\frac{1+\lambda}{\lambda}$ , we obtain

$$D(r,f) = (1+\lambda)\frac{r^{s-1}}{f^{s-1} + \lambda r^{s-1}}.$$
 (C.12)

In Appendix C, we simplify the exposition by setting s = 2. Nevertheless, our analysis generalizes to all s > 1 (with the exception that the expression given in (C.1) for the monopoly interest rate is only valid for s = 2). In particular, Theorems C.1–C.3 all still hold for the more general expression given in (C.12) for the demand for deposits when s > 1.

# D Retail Industry Change in MMC and HHI from 1989 to 2021

In this appendix, we show that 4-digit SIC industry changes in MMC and HHI from 1989 to 2021 which comprise Figure 4b. We order the industries by change in multimarket contact.

Industry	4-Digit SIC	$\Delta$ MMC	$\Delta$ HHI
Variety stores	5331	0.57	0.17
Children's and infants' wear stores	5641	0.57	-0.04
Drug stores and proprietary stores	5912	0.49	0.05
Hobby, toy, and game shops	5945	0.44	-0.04
Miscellaneous retail stores, nec	5999	0.39	-0.03
Miscellaneous homefurnishings	5719	0.34	-0.04
Optical goods stores	5995	0.33	-0.03
Women's accessory and specialty stores	5632	0.32	-0.01
Department stores	5311	0.30	0.01
Luggage and leather goods stores	5948	0.30	0.01
Auto and home supply stores	5531	0.29	-0.01
Lumber and other building materials	5211	0.27	-0.01
Men's and boys' clothing stores	5611	0.25	0.00
Stationery stores	5943	0.20	0.05
Miscellaneous apparel and accessories	5699	0.19	-0.03
Miscellaneous general merchandise	5399	0.18	-0.07
Women's clothing stores	5621	0.17	-0.03
Furniture stores	5712	0.17	0.00
Family clothing stores	5651	0.17	-0.07
Miscellaneous food stores	5499	0.16	-0.11
Paint, glass, and wallpaper stores	5231	0.14	-0.05
Gift, novelty, and souvenir shop	5947	0.13	-0.02
Book stores	5942	0.13	-0.01
Used merchandise stores	5932	0.13	-0.03
Jewelry stores	5944	0.12	-0.02
Eating places	5812	0.10	-0.01
Hardware stores	5251	0.09	0.01
Gasoline service stations	5541	0.09	0.02
Boat dealers	5551	0.07	-0.02
Grocery stores	5411	0.07	0.01
Sporting goods and bicycle shops	5941	0.07	0.00

Industry	4-Digit SIC	$\Delta$ MMC	$\Delta$ HHI
Musical instrument stores	5736	0.07	-0.04
Liquefied petroleum gas dealers	5984	0.07	-0.03
Computer and software stores	5734	0.06	-0.08
Used car dealers	5521	0.06	-0.02
New and used car dealers	5511	0.06	-0.03
Direct selling establishments	5963	0.05	-0.11
Mobile home dealers	5271	0.05	0.05
Shoe stores	5661	0.05	0.02
Meat and fish markets	5421	0.05	-0.02
Motorcycle dealers	5571	0.03	-0.06
Radio, television, and electronic stores	5731	0.03	0.03
News dealers and newsstands	5994	0.02	0.12
Drapery and upholstery stores	5714	0.02	0.09
Recreational vehicle dealers	5561	0.02	0.01
Drinking places	5813	0.02	-0.03
Household appliance stores	5722	0.02	0.04
Retail nurseries and garden stores	5261	0.01	-0.01
Fruit and vegetable markets	5431	0.01	-0.06
Record and prerecorded tape stores	5735	0.00	-0.06
Liquor stores	5921	0.00	-0.01
Floor covering stores	5713	0.00	-0.05
Candy, nut, and confectionery stores	5441	0.00	-0.09
Automotive dealers, nec	5599	0.00	-0.27
Retail bakeries	5461	0.00	-0.08
Fuel dealers, nec	5989	0.00	0.03
Florists	5992	0.00	0.01
Tobacco stores and stands	5993	-0.01	-0.23
Dairy products stores	5451	-0.02	-0.05
Fuel oil dealers	5983	-0.02	-0.07
Sewing, needlework, and piece goods	5949	-0.06	0.00
Merchandising machine operators	5962	-0.06	-0.03
Catalog and mail-order houses	5961	-0.06	-0.10
Camera and photographic supply stores	5946	-0.06	0.06