

# *Digital disruption and market structure: the case of internet banking*

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Lyons, B. and Zhu, M. ORCID: <https://orcid.org/0000-0001-6687-3267> (2024) Digital disruption and market structure: the case of internet banking. *Industrial and Corporate Change*. ISSN 1464-3650 doi: <https://doi.org/10.1093/icc/dtae013>  
Available at <https://centaur.reading.ac.uk/115147/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1093/icc/dtae013>

Publisher: Oxford University Press

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

**CentAUR**

Central Archive at the University of Reading

Reading's research outputs online

# Digital disruption and market structure: the case of internet banking

Bruce Lyons<sup>1</sup> and Minyan Zhu<sup>2,\*</sup>

<sup>1</sup>School of Economics and Centre for Competition Policy, University of East Anglia, Norwich NR4 7TJ, UK.

e-mail: [b.lyons@uea.ac.uk](mailto:b.lyons@uea.ac.uk) and <sup>2</sup>Department of Economics, University of Reading, Reading RG6 6EL, UK.

e-mail: [minyan.zhu@reading.ac.uk](mailto:minyan.zhu@reading.ac.uk)

\*Main author for correspondence.

We provide an econometric study of the adoption of internet banking, a case of a potentially disruptive digital technology which could devalue/replace an incumbent legacy. Our aim is to better understand the extent to which it disrupted the market structure where incumbents start with a strong customer base. We study both regional integration and national concentration dimensions of market structure in EU member states during the period of 1997–2018. We find that internet banking was initially introduced earlier in more concentrated markets. Although consumer uptake was slower over time than in less concentrated markets, the initial higher consumer penetration in more concentrated markets was sustained until market maturity. We further find a substantial de-concentrating effect of internet banking, and evidence of integration in previously regionalized markets following uptake.

**JEL classification:** L11, O33, F15, G21, L81

## 1. Introduction

Much has been written about how digital technology is transforming the ways firms compete,<sup>1</sup> especially with the rise to dominance of a small number of global big tech firms that have entered or created markets with internet-only products and no substantial existing customer base (such as Google, Amazon, Facebook [Meta] and Netflix). However, much less is known about how the adoption of digital technology, when promoted *alongside* traditional face-to-face customer models (e.g., financial services), disrupts market structure. Internet banking (also known as online banking or web banking)—a product for which incumbent banks introduced an internet consumer interface alongside the pre-existing face-to-face networks of brick-and-mortar retail outlets (branches)—is a leading example of a potentially disruptive technology that can be examined empirically.

Internet banking is not limited to the original innovator or restricted by license agreements. In contrast with the branch banking business model (which has a very expensive cost structure requiring a heavy investment in a branch network), internet banking is not geographically tied to where the customers are. More importantly, in what is becoming a much wider digital transformation of consumer banking, internet banking has the potential to replace or devalue the legacy systems of incumbents, as it allows users to access online almost all services traditionally available from a branch. In principle, this should facilitate both de novo entry and cross-entry by previously regional banks, which should be expected to lead to long-term changes in market

<sup>1</sup> See, for example, [The Furman Review \(2019\)](#).

structure. This raises questions such as the short- and medium-term effects of initial market structure on the introduction and consumer adoption of digital services, and the consequent direction and speed of change in endogenous market structure.

One of our aims is therefore to understand the extent to which the adoption of digital technology (by way of its speed and extent) disrupts the structure of a market where incumbents start with a strong customer base, and to draw implications for market contestability as well as advantages/challenges faced by incumbents and entrants. By focusing on EU member states with vastly varying degrees of regionalization across their (pre-internet) banking markets, we are further able to separate the regional integration and national concentration dimensions of market structure.

We specify our research questions more precisely in the next section, with summarized findings and our contribution to the relevant literature. [Section 3](#) sets out our modeling approach and model specifications. [Section 4](#) describes our data and presents some stylized facts. [Section 5](#) explains our empirical strategy, and [section 6](#) details our estimation results. [Section 7](#) presents robustness checks. [Section 8](#) concludes and draws implications from our findings.

## 2. Research questions, related literature and our high-level empirical findings<sup>2</sup>

Taking account of the endogeneity of consumer uptake of internet banking and market structure, we investigate the following research questions (RQs):

### 2.1 RQ 1: How did market structure affect consumer acceptability and uptake of internet banking?

A delay in uptake means that potential consumer surplus is lost for that time period, and this loss is likely to be substantial for a major innovation.<sup>3</sup> Consumer uptake of a technology or new product/service often takes a number of years, particularly where acceptability depends, at least in part, on observing other users. This drag on diffusion increases in importance if there are network effects (e.g., mobile phones), learning effects (e.g., home computing), fear of side-effects (e.g., vaccines), or concerns over security (e.g., financial services). More specifically, trust and reputation are crucial for financial services because everyday transactions (payments), personal wealth (savings), or potentially large future claims (insurance) are at stake. These demand-side factors give incumbent providers a substantial advantage in a mature market with near-universal geographic coverage.

To empirically understand consumer acceptability and uptake of any disruptive innovation, we draw on the strand of literature relating to the speed of uptake (consumer diffusion) of new products or new service delivery methods. Building on the pioneering work of [Griliches \(1957\)](#), this econometric approach was first applied to a consumer product, mobile phones, by [Gruber and Verboven \(2001a,b\)](#). [Li and Lyons \(2012\)](#) develop this approach to examine the role of market structure and regulation over a more-or-less complete cycle of consumer adoption. We follow a similar methodology in order to understand how market context affects consumer acceptability of internet banking, as reflected in the speed of uptake.

There has been very little econometric research on the effect of market structure on user adoption of internet banking. [Takeddine and Sun \(2015\)](#) consider consumer usage across 33 European countries, but only as a cross-section in 2013 (so they cannot distinguish speed from timing

<sup>2</sup> The literature mentioned in this section is by no means complete. Our intention is to link to the literature we think most relevant to our study and indicate to what extent we have contributed to it.

<sup>3</sup> To get a very rough ballpark feel for the magnitude of this loss, consider a consumer with linear potential demand (standardized with unit slope). If  $q$  is each consumer's demand at the competitive price, delayed uptake foregoes  $0.5q^2$  of potential consumer surplus per period. By comparison, market power which increases price and reduces demand by 10% would mean a per period consumer surplus loss of  $[0.9 \cdot 0.1 + 0.5 \cdot 0.1 \cdot 0.1]q^2 = 0.095q^2$ . So, even without discounting, a 4-year delay in uptake of the new product would be equivalent in consumer surplus terms to enduring 21 years of this level of monopoly power.

effects, which we explain in Section 3.1).<sup>4</sup> Nickerson and Sullivan (2003) and Sullivan and Wang (2013) investigate the timing of the initial adoption of internet banking technology by US banks (i.e., by firms as opposed to by consumers) across US regions.<sup>5</sup>

We add to this literature by providing new evidence relating to consumer uptake of internet banking. This has two stages in that firms must first introduce a new technology, and then consumers must adopt it. We find that internet banking was introduced earlier in concentrated markets, possibly due to the capability of relatively large banks to convert existing customers on a large scale. However, our results also show that the speed of consumer uptake was slower in concentrated markets. This may be because incumbents most invested in an expensive brick-and-mortar network are less aggressive in encouraging their customers to take up internet banking, at least beyond those they were most likely to lose if they had not innovated early.

## 2.2 RQ2: What effect has the uptake of internet banking had on the evolution of national market concentration?

Disruptive innovation is seen as a core part of firms' strategic management that could change industry structure (Christensen 1997). Using data on the hard disk industry, Christensen and Rosenbloom (1995) illustrate that incumbents are usually successful in defending their position by addressing customers' needs within the value network in which the incumbents competed, but an entrant can change the context within which a firm competes (emerging value networks) and solve some customer problems. Therefore, "disruption" occurs when incumbents focus on existing customers and/or technologies, making it difficult for themselves to shift investment to disruptive innovations (see also Gans, 2016; Christensen *et al.*, 2015).<sup>6</sup> As a result, the trajectories of technological progress and industry structure in established markets change.

Whereas disruption discussed in the aforementioned literature leads to the erosion of incumbency advantages, strategic management literature that stresses firms' "dynamic capabilities" suggests that the erosion of incumbency advantages may be limited. The dynamic capabilities of firms may favor incumbents with better access to finance and market know-how (for instance, see Teece, 2007). According to O'Reilly and Tushman (2004), appropriate internal organization of the firm can overcome managerial biases mentioned in the disruptive innovation literature. Ho and Chen (2018), by exploring the cases of Kodak and Fujifilm in the face of digital disruption, propose that incumbent firms could also succeed if they adopt disruptive innovations early and at the same time further exploit their sustaining competences.

We add to the disruptive innovation literature by providing new empirical evidence relating to the consequent market structure outcomes. Our empirical approach draws heavily on the industrial organization literature on the evolution of market structure, according to Sutton (1991, 1998). This approach is reviewed in Section 3.2. We find a large de-concentrating effect of internet banking, which suggests that new technology of such a disruptive nature has allowed the smaller challengers/entrants to erode some of the advantage of the large/established incumbents with legacy technology.

<sup>4</sup> They find that "the effects of socio-economic and technology-related factors on Internet banking diffusion are fully mediated by Internet access" [p.361]. This supports our later assumption that the maximum consumer uptake is determined by internet access.

<sup>5</sup> In terms of our consumer uptake model, their results are relevant to the timing parameter but not to the speed parameter. They find that larger banks in more concentrated markets adopt earlier, which they explain in terms of the incentive to exercise their strategic option earlier than banks with smaller market shares. We find a similar result.

<sup>6</sup> The Industrial Organisation (IO) literature on innovation also extensively analyzes the different incentives to innovate between incumbents and challengers/entrants. Studies in this area often come from a different but related perspective compared with the disruptive innovation literature mentioned earlier. For instance, using data similar to Christensen and Rosenbloom (1995), Igami (2017) finds that the "replacement effect" faced by an incumbent with market power whose new products cannibalize profits already earned from their existing range explains why entrants were successful in the adoption of 3.5-inch over 5.25-inch hard disk drives in the 1980s. Based on a comprehensive survey of relevant theoretical studies, Reinganum (1989) shows that the presence/absence of technological uncertainty in the production of innovation affects incumbents and entrants' timing of innovation differently.

### 2.3 RQ3: What effect has the uptake of internet banking had on integration across previously regional markets?

Our methodology allows us to compare countries in which banks tended to have national coverage, with those in which regional banks were common. In addition to the “*de-concentrating effect*” mentioned earlier, our results suggest a consolidation force, which we call the “*extended geographic reach effect*.” Internet banking erodes regional incumbent legacy advantages (e.g., the value of local branches), facilitating cross-regional entry, bank consolidation, and the exit of weaker regional banks. Thus, the direction of travel is towards national market integration post-internet banking.<sup>7</sup>

To the best of our knowledge, this is the first econometric study investigating the interaction between a potentially disruptive digital technology and these elements of market structure. Nevertheless, there is an interesting comparison between our work and that of Pelletier *et al.* (2020). They study the spread of mobile money, weighing the competing capabilities of banks and international telecom firms as initial innovators in emerging markets. They find that banks in developing countries had focused on urban branch networks, with little coverage for the rural population, and were slow to introduce mobile banking products. In contrast, telecom firms could provide low-risk transaction execution services at the same time as they were expanding their rural mobile network coverage.

There are two relevant differences between their work and ours. First, Pelletier *et al.* examine only the initial launch of mobile banking services by “firms,” whereas we also examine the speed of uptake of internet banking by consumers. Second, we examine European markets in which all geographic areas already had access to banking services, so there was no substantial unsupplied market for entrants to gain a foothold—internet banking was an improvement in product delivery not a completely new product like mobile banking in many emerging countries with large rural populations.

## 3. Modeling approach and model specifications

### 3.1 Approach to modeling the user uptake of internet banking

Our approach is to adapt a classic contagion model of the user uptake of a new service and then focus on how market structure may be expected—separately—to influence the timing and speed of diffusion.

We start from the standard logistic function first used by Griliches (1957), which describes how the technology diffusion process follows an S-shaped function:  $IB_{it} = \frac{M_{it}}{1 + \exp(-(a_{it} + b_{it}t))}$ , where  $IB_{it}$  is the number of users who have adopted internet banking in country  $i$  at time  $t$ ,  $M_{it}$  is the maximum number of potential users.

$a_{it}$  shifts the logistic curve horizontally and is known as the location or timing parameter. Its economic interpretation in our context is that if we compare two countries at the start of our observation period, for similar  $b$ , the one which introduced internet banking earlier will have a higher  $a$ . Less formally,  $a$  can be interpreted an indicator of when sufficient banks had introduced internet banking for consumer “contagion”/uptake to take off.

$b_{it}$  is the slope of the diffusion curve, which is typically referred to as the speed of diffusion. It measures how rapidly new consumers adopt internet banking once it has been introduced. The logistic functional form means that this single parameter takes account of both limited consumer learning opportunities and word-of-mouth in the early years, and fewer potential new users once uptake becomes saturated, with a maximum speed of uptake in between.

<sup>7</sup> Another market that started out with strong locally focused firms, and which has been severely disrupted by digital products, is local newspapers. Their decline across much of the world has been well documented (see, e.g., Abernathy, 2020 in the USA; Jenkins and Nielsen, 2018 in Europe). However, the two-sidedness of newspaper markets has resulted in a more complex dynamic. Revenues to fund journalism were traditionally generated both by selling print copy to local citizens and by selling advertising. Internet entry has affected both. Despite having grown their online audiences, local newspapers have had limited success in monetizing through a paywall (partly due to the consumer substitution into social media), and advertising revenues have been scooped up by Google, Facebook, and other non-journalism sites. Unlike for the online banking product, local news in one part of the country is completely different to that in another locality so there is little scope for economies of scale/regional integration. Consequently, investment in quality local journalism has had to adjust more than market structure.

We next assume that the maximum possible uptake (saturation level) is a proportion,  $\lambda$ , of current internet usage,  $IU_{it}$ , so  $M_{it} = \lambda * IU_{it}$ .

Rearranging the logistic equation and taking logs gives the following equation:

$$\log\left(\frac{IB_{it}}{\lambda * IU_{it} - IB_{it}}\right) = a_{it} + b_{it} * t.$$

There are a number of ways how market structure might affect  $a_{it}$  and  $b_{it}$ . The early introduction of internet banking  $a_{it}$  may be facilitated by the immediate scale possibilities of converting existing customers to the internet service in a concentrated market. Once introduced, the speed of consumer uptake,  $b_{it}$ , will depend on the ability and incentive for banks to market their internet services. This suggests a range of factors relating to market structure that might influence consumer uptake. These may include the original introduction of internet banking services, investments in interface quality and security, ongoing marketing, price, and the implicit price of branch banking (i.e., substitutes). Our reduced form approach bypasses these proximate influences, which anyway are almost impossible to measure at the market level.

Both  $a_{it}$  and  $b_{it}$  are also likely to be influenced by demand side factors such as income, education, and demographic factors. Previous bank investments in branch networks also affect the availability and opportunity cost of using a close substitute for internet banking.

For our estimation, both  $a_{it}$  and  $b_{it}$  are allowed to be affected by market structure, including concentration ( $C_{it}$ ) and regionalization ( $R_i$ ), branch density ( $B_{it}$ ) and a vector of controls ( $X_{it}$ ).<sup>8</sup> We therefore estimate the following equation:

$$\log\left(\frac{IB_{it}}{\lambda * IU_{it} - IB_{it}}\right) = a_0 + a_1 C_{it} + a_2 B_{it} + a_3 R_i + a_4 X_{it} + b_0 t + b_1 C_{it} * t + b_2 B_{it} * t + b_3 R_i * t + b_4 X_{it} * t + u_{it} \quad (1)$$

where  $u_{it}$  is the idiosyncratic error term. We explain our estimation strategy in Section 5.1, including how we deal with the potential endogeneity of concentration ( $C_{it}$ ) and branch density ( $B_{it}$ ).

### 3.2 Approach to modeling endogenous national concentration

Our approach to modeling endogenous market structure builds on a reduced form relationship between concentration and market size. In most markets, economies of scale lead to a negative relationship between concentration and market size in free entry equilibrium. But the slope and position of this relationship depends on factors such as the intensity of price competition,<sup>9</sup> the degree of economies of scale, and the extent of entry barriers. For instance, for a given market size, tougher price competition results in a more concentrated market as reduced margins require more customers per firm in order to cover fixed costs. Horizontal product differentiation moderates these effects by reducing the intensity of price competition, but it does not change the basic relationship.

In addition to horizontal differentiation, competition can also be channeled into the escalation of *endogenous sunk costs* characterized by quality-enhancing investments that benefit all consumers without raising marginal cost (e.g., denser branch networks in the pre-internet era).<sup>10</sup> As a result, an increase in market size may then have less effect on concentration (and prices) than

<sup>8</sup> In addition to the observed heterogeneity determined by the factors (market structure, branch density, and controls) mentioned earlier, unobserved heterogeneity relating to each of the parameter which is not taken into account in our estimation may also influence the location and speed parameters, and therefore the process of diffusion. In this respect, the estimated location and speed parameters in our model should be viewed as mean estimates driven by the observed factors mentioned earlier.

<sup>9</sup> For example, Bertrand or Cournot. See, for example, Shaked and Sutton (1987), Sutton (1991), Bresnahan and Reiss (1991), and Berry (1992).

<sup>10</sup> Previous studies (see Kim and Vale 2000; Dick 2007; Cohen and Mazzeo 2010; Temesvary 2015) have provided supportive evidence that branch investment in the pre-internet era could be a quality-enhancement investment characterized as endogenous sunk costs in banking.

on enhanced quality.<sup>11</sup> While it is theoretically possible that the relationship between concentration and market size may become positive, it is empirically more typical for the relationship to remain weakly negative but less steep in the presence of such quality competition.<sup>12</sup>

We expect high user adoption of internet banking to result in a more negative relationship between market size and concentration, i.e., a “de-concentrating effect,” counteracting the influence of higher endogenous sunk costs in legacy technology such as branch network invested by established incumbents.

We adopt a well-established functional form for the relationship between concentration and market size. Following [Sutton \(1991\)](#) and followers, we specify  $y = \alpha + \frac{\beta}{\log S}$ , where  $y$  is the logistic transform of the concentration ratio,  $S$  is national market size, and  $\alpha$  and  $\beta$  are coefficients to be estimated.

We allow both  $\alpha$  and  $\beta$  to vary with regionalization,  $R$ , so we can test whether the relationship between concentration and market size differs between regionalized and non-regionalized markets.

Adding a time trend,  $t$ , gives

$$\log\left(\frac{C_{it}}{100 - C_{it}}\right) = \theta_0 + \theta_1 t + \theta_2 \frac{1}{\log S_{it}} + \theta_3 R_i + \theta_4 R_i * \frac{1}{\log S_{it}} + \theta_5 B_{it} + \varepsilon_{it} \quad (2)$$

where  $C_{it}$  is the five-firm concentration ratio,  $R_i$  is our regionalization index,  $B_{it}$  is branch density,  $S_{it}$  is total national banking assets, and  $\theta_0, \theta_1, \theta_2, \theta_3, \theta_4, \theta_5$  are coefficients to be estimated.  $\varepsilon_{it}$  is the idiosyncratic error term.

As mentioned earlier, our principal aim is to explore whether increasing user uptake of internet banking has had a de-concentrating effect. Such mechanisms are likely to operate slowly and not smoothly. Since we do not expect effects on concentration to happen either contemporaneously or with a simple time lag, we test a model that allows the relationship in [equation \(2\)](#) to shift once a threshold level of  $IB$  is reached. A dummy variable,  $D = 1$  for high  $IB$ , is interacted with all the right-hand side variables in [equation \(2\)](#), so the coefficients can be interpreted as the incremental effect of a high level of internet banking on the determination of concentration. Writing the right hand side of [equation \(2\)](#) as  $\gamma X_{it} + \varepsilon_{it}$ , we estimate [equation \(3\)](#):

$$\log\left(\frac{C_{it}}{100 - C_{it}}\right) = \gamma X_{it} + \delta D_{it} X_{it} + \varepsilon_{it} \quad (3)$$

Our estimation strategy, including how we deal with the possible endogeneity of variables  $D_{it}$  and  $B_{it}$ , is explained in [Section 5.2](#).

### 3.3 Construction of the regionalization index

Some European countries have much more regionalized banking markets than others. The roots of these differences are historic, for example, where there were proud histories of independent states prior to nineteenth-century unification (e.g., Germany and Italy). These countries tend to have lower national concentration than the countries of a similar size but with nationally integrated banks. For example, in 2009, the combined market shares of the five largest banks in Estonia and the Netherlands were 93% and 85%, respectively, while in the much more regionalized jurisdictions of Germany and Italy, this concentration ratio ( $C$ ) was only 25% and 34%, respectively. [Equations \(1\)](#) and [\(3\)](#) should ideally be estimated at the level of the competitively relevant market, which may be regional in some countries and national in others. In the absence of consistent data at the level of relevant geographic markets, we use a measure of pre-internet banking regionalization ( $R_i$ ), based on the Herfindahl index of regional concentration of bank headquarters (HQs), explained as follows.

We began with the idea that banks tend to locate their central operations close to their main demand base. Thus, a strongly regional bank (in terms of its branch and customer base) is likely

<sup>11</sup> See [Berry and Waldfogel \(2010\)](#) for an empirical test of the difference between markets where quality is enhanced by endogenous sunk costs as compared with quality enhancements that increase marginal cost.

<sup>12</sup> See [Sutton \(1991, 1998, Sutton \(2007\)\)](#) reviews the literature.

to be headquartered in the region where it is strong, whereas a bank that considers the whole country as its natural market is more likely to be headquartered in the national financial capital.

Consider a country with  $K$  regions. We require an index of regionalization,  $R$ , with the following desirable properties.

1. Minimum  $R = 0$  if all HQs are in a single region. This should apply for both a multi-region country and a small country, which forms a single region.
2.  $R$  should increase if HQs are distributed more equally between a given number of regions ( $K \geq 2$ ). Maximum  $R$  (given  $K$ ) should result from a uniform distribution of HQs (i.e., a share  $K^{-1}$  in each region).
3.  $R$  should increase if, for a given distribution of HQs, the number of regions with HQs increases.

To develop our index, we aggregated the assets of all banks headquartered in region  $k$  to create the scale of banking in that region,  $S_k$ . The region's share of national banking assets is  $\frac{S_k}{S}$  where  $S = \sum_{k=1}^K S_k$ . We propose the following index:

$$R = \left[ 1 - \sum_{k=1}^K \left( \frac{S_k}{S} \right)^2 \right].$$

The summation term is a Herfindahl index of regional concentration of bank HQs, and the "one minus" converts this to an index of regionalization.  $R$  ranges between zero (when all HQs are in one region), and  $1 - K^{-1}$  (when there is an equal number of HQs in each region). Two empirically interesting examples are where: (i) there are two equal sized regions and the remaining  $K - 2$  regions have no HQs, in which case  $R = 0.5$ , and (ii) there are four regions containing 40%, 30%, 20%, and 10% shares of HQs, in which case  $R = 0.7$ .<sup>13</sup> It is straightforward that  $R$  satisfies the first desirable property. The second and third follow from a standard property of the Herfindahl that  $\sum_{k=1}^K \left( \frac{S_k}{S} \right)^2 = \frac{1+\nu^2}{K}$ , where  $\nu$  is the coefficient of variation of regional shares.

It should be noted that even in regionalized markets, some competition might exist beyond the region. Therefore, the above measure of pre-internet banking regionalization could be viewed as a control of the national concentration, which is associated with the competition beyond regions within a nation. Whereas it is possible that competition may even exist beyond national boundary, in this paper, we consider that the relevant geographic market is not wider than national (i.e., we have not considered market structure measures beyond national level). This is consistent with the European Commission's report regarding a Sector Inquiry under Article 17 Regulation 1/2003 on retail banking<sup>14</sup> during which retail banking concentration was investigated both at national and regional levels (Section 4 in the report).

#### 4. Data description

Our dataset consists of a panel of relevant variables 1997–2018 for the 15 EU member states at the start of the period, increasing to 27 countries from 2001 (i.e., including the new members who acceded in 2004).<sup>15</sup> In this section, we describe our data on key variables including internet banking, national concentration, and regionalization.

An annual Eurostat survey since 2003 has reported the percentage of surveyed individuals by EU member states who have used internet banking in the past 3 months.<sup>16</sup> We use this as

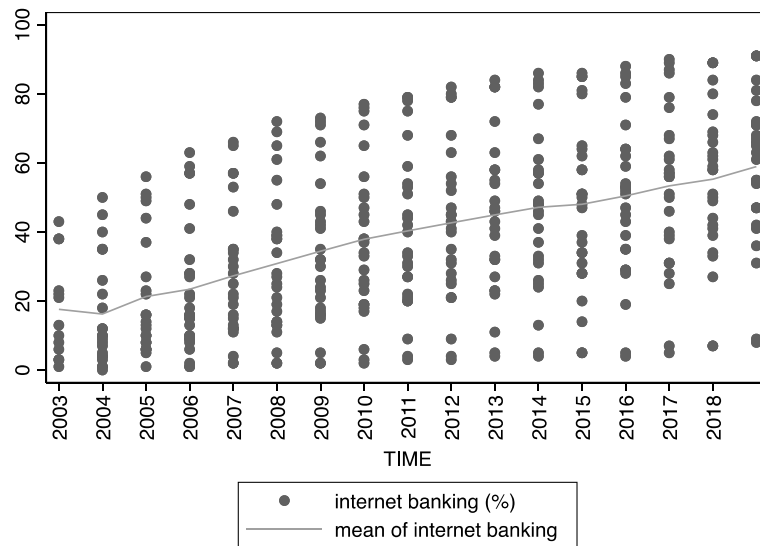
<sup>13</sup> Of course, the same  $R$  can come about from many different distributions of HQs; e.g., one region with 68% and three with 11% each would give  $R = 0.5$ .

<sup>14</sup> [https://competition-policy.ec.europa.eu/sectors/financial-services/sector-inquiry-retail-banking\\_en](https://competition-policy.ec.europa.eu/sectors/financial-services/sector-inquiry-retail-banking_en) (last checked February 24, 2023).

<sup>15</sup> The recently acceded member Croatia is not included in the sample.

<sup>16</sup> An overview of the dataset constructed by Eurostat can be found using the following link: <https://ec.europa.eu/eurostat/web/products-datasets/-/tin00099> (last checked February 24, 2023). Internet banking includes electronic transactions with a bank for payment, transfers, etc., or for looking up account information. Further details regarding the





**Figure 1.** Consumers using internet banking (by EU member state). *Data Source:* Eurostat.

our measure of the user uptake of internet banking. Similarly, data on internet usage is collected from the same survey to measure the number of individuals who have used the internet in the past 3 months.

Figure 1 summarizes the range of internet banking experiences across countries and the general trend. Each dot represents a Member State.

Two features stand out. First, the international variation is strikingly large. Second, the average increasing trend appears broadly consistent with an S-shaped diffusion.

Our assumption that the maximum possible uptake (saturation level) of internet banking is a proportion,  $\lambda$ , of current internet usage,  $IU_{it}$ , ( $M_{it} = \lambda * IU_{it}$ ) is consistent with the data on internet banking and internet usage for all country-year pairs plotted in Figure 2, which shows that  $IB$  is bounded by  $IU$ .

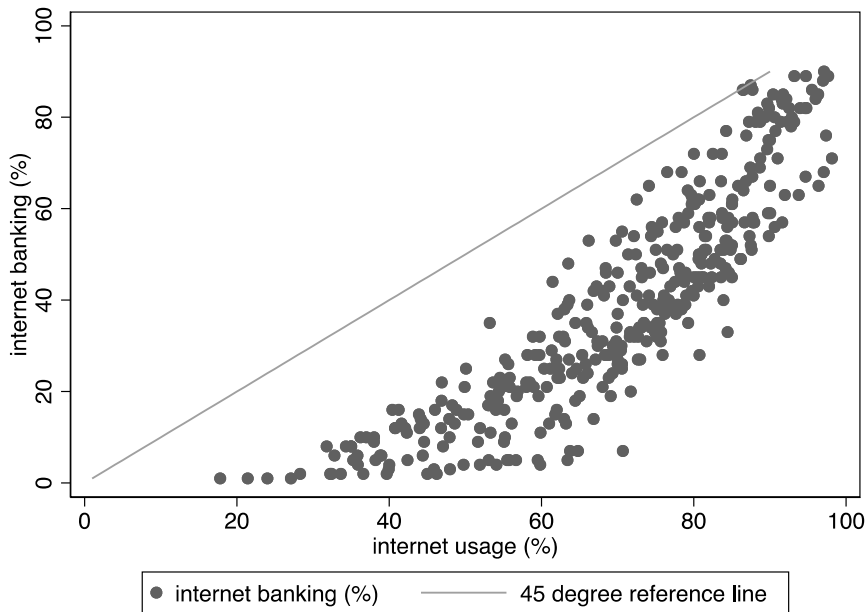
We use data on national concentration collected by the European Central Bank (ECB), which publishes systematic data on banking activities for each EU member state (whether or not it is in the Eurozone). The ECB data are for “credit institutions” defined as businesses that either (i) receive deposits or other repayable funds from the public and grant credit on their own account or (ii) issue means of payment in the form of electronic money.<sup>17</sup> We call these “banks” for short.

The ECB measures bank size by total assets. Importantly, total assets are calculated on a residence basis. Hence, this includes the activities of foreign banks in a particular Member State and excludes the foreign activities of domestic banks. The number of banks is similarly measured to include all credit institutions under the jurisdiction of each country, regardless of national or international ownership. The downside to using the ECB data is that it does not disaggregate by the type of activity (e.g., retail versus investment banking), but we still consider the ECB data to be the most meaningful available in the context of consumer choice.

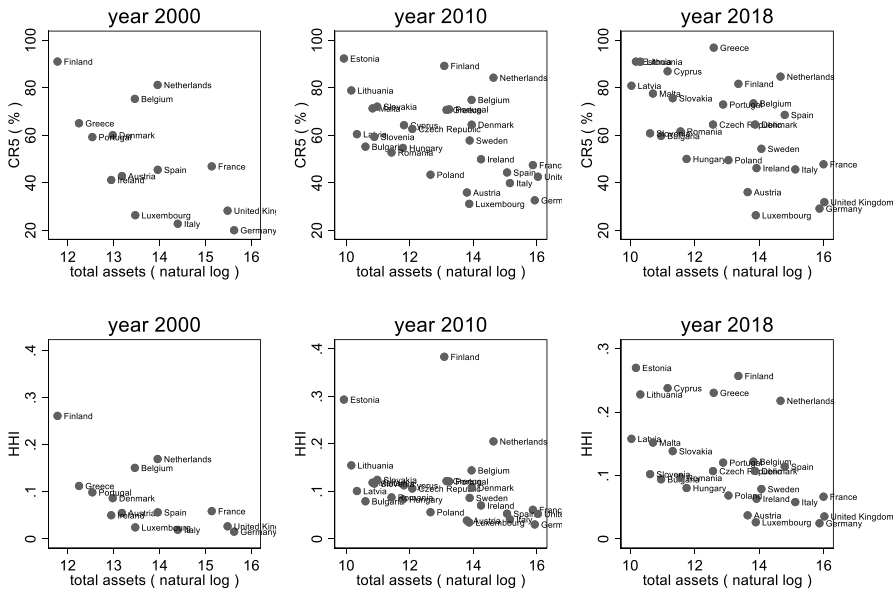
The ECB also measures market size by total assets and calculates two standard measures of national market structure: five-firm concentration ratio ( $CR5$ ) and Herfindahl index ( $HHI$ ). The relationship between the concentration ratio and market size is shown in Figure 3.

survey questionnaire can be found in the Methodological Manual of Eurostat’s Digital economy and society database: <https://ec.europa.eu/eurostat/web/digital-economy-and-society/methodology> (last checked February 24, 2023).

<sup>17</sup> Further details relating to the data description can be found from the database of Structural Financial Indicators constructed by ECB Statistical Data Warehouse: <https://sdw.ecb.europa.eu/browse.do?node=9691551> (last checked March 6, 2023).



**Figure 2.** Internet banking versus internet usage. *Data Source:* Eurostat.



**Figure 3.** Bank concentration and market size. *Data Source:* ECB.

Three observations stand out. First, there is a very wide range of market sizes. Much of this is consistent with differences in population and the very different histories of the accession countries, but Luxembourg stands out as disproportionately large. This is likely to be due to the wide definition of banking used by the ECB, so we test our later results for sensitivity to excluding

**Table 1.** Regionalization within countries

Country	R index
Belgium, Bulgaria, Cyprus, Czech Republic, Estonia, France, Greece, Hungary, Ireland, Latvia, Lithuania, Luxembourg, Malta, Slovakia Sweden	0
Poland	0.04
Finland	0.07
Denmark	0.2
Romania	0.26
Slovenia	0.38
United Kingdom	0.4
Austria	0.48
Netherlands	0.5
Portugal	0.53
Spain	0.68
Germany	0.69
Italy	0.7

*Source:* Authors' calculations.

Luxembourg. Second, even markets of a similar size demonstrate a considerable range in concentration. Third, there appears to be a broadly negative relationship between concentration and market size, with a fairly well-defined lower bound (especially if Luxembourg is excluded).

To measure regionalization, as defined in Section 3.3, we collect information on each bank's HQs' location (including city and postcode) from the Banker database.<sup>18</sup> We include all banks at the group level (bank-holding companies)<sup>19</sup> and use the postcode to identify the Nomenclature of Territorial Units for Statistics or NUTS level 2 region in which each bank is headquartered.<sup>20</sup>

Table 1 reports our index of regionalization based on bank assets for each country in our dataset.<sup>21</sup> In Cyprus, Estonia, Latvia, Lithuania, Luxembourg, and Malta, there is only one NUTS2 region, so the index is zero. In Belgium, Bulgaria, Czech Republic, France, Hungary, Ireland, and Slovakia, all banks covered in the sample are headquartered in one region, so our index is also zero for these countries. The countries with the most regionalized banking are Germany, Italy, and Spain, each with  $R \approx 0.7$ . Austria, the Netherlands, Portugal, and, to a lesser extent, the UK and Slovenia, have two strong HQ locations ( $R \approx 0.5$ ) and the remaining countries have very asymmetric regionalization around a dominant financial capital.

It is worth noting that the reported variation in the degree of regionalization is consistent with the data presented in the European Commission's Sector Inquiry on retail banking.<sup>22</sup>

Full descriptions of all variables, their measurement and sources are given in Appendix 1 (Tables A1 and A2).

<sup>18</sup> The ECB does not publish information on individual banks. The Banker Database was created as part of The Banker magazine's regular rankings of the world's largest banks. Owned by the Financial Times, it provides coverage of the leading banks in more than 190 countries. <https://www.thebankerdatabase.com/index.cfm?fuseaction=lite.overview> (last checked March 6, 2023). While not fully comprehensive, banks covered by the database represent more than 90% of the banking assets in each European country. Note that the coverage of banks in a given country may vary slightly over time due to merger, entry, and exit. Our reported country index, as used in our econometrics, is the average over time.

<sup>19</sup> For instance, if a savings bank in Italy (such as Cassa di Risparmio di Carpi, Cassa di Risparmio di Trento e Rovereto [Caritro], and Cassa di Risparmio di Trieste) joined Unicredit group, it would not be counted as a separate bank even though it might still have a local headquarter.

<sup>20</sup> Where postcode information was lacking, we matched the bank city with NUTS regions directly. The EU defines level 2 regions to mirror the territorial administrative divisions of Member States, each with populations generally in a band of 800,000 to 3,000,000. See <https://ec.europa.eu/eurostat/web/nuts/principles-and-characteristics> (last checked March 6, 2023).

<sup>21</sup> An alternative index based on the number of banks (rather than bank assets) was also constructed. It made no material difference to our descriptive or econometric results.

<sup>22</sup> [https://ec.europa.eu/competition/sectors/financial\\_services/inquiries/retail.html](https://ec.europa.eu/competition/sectors/financial_services/inquiries/retail.html) (last checked March 6, 2023).

## 5. Estimation and identification

### 5.1 Estimating the user uptake of internet banking

To allow for unobserved heterogeneity in [equation \(1\)](#), we follow the correlated random effects approach proposed in [Wooldridge \(2019\)](#), which is an extension of the well-established Chamberlain–Mundlak approach for balanced panel data to unbalanced cases. Under this approach, unobserved heterogeneity is allowed to be correlated to the history of selection and the selected covariates. The averages of covariates over time, where we observe a full set of data on the dependent and independent variables,  $\bar{X}_i$  are therefore constructed.

[Equation \(1\)](#) becomes:

$$\log\left(\frac{IB_{it}}{\lambda * IU_{it} - IB_{it}}\right) = a_0 + a_1 C_{it} + a_2 B_{it} + a_3 R_i + a_4 X_{it} + b_0 t + b_1 C_{it} * t + b_2 B_{it} * t + b_3 R_i * t + b_4 X_{it} * t + c_i + v_{it} \quad (4)$$

where the unobserved heterogeneity is assumed to be  $c_i = \varphi + a_{10} * \bar{X}_i + a_{11} * \bar{t}_i + a_i$  and  $v_{it}$  is the time-varying unobserved heterogeneity.

If the panel data are balanced, [Mundlak \(1978\)](#) shows that the fixed effect (FE) estimator can be computed using pooled ordinary least squares (OLS) regressions from the original data with the time averages of the covariates added as additional explanatory variables. [Wooldridge \(2019\)](#) extends this result to the unbalanced cases. Estimating [equation \(4\)](#) with additional explanatory variables in  $c_i$  (i.e., the time averages of time-varying covariates), using pooled OLS with all observations that have a full set of data on the dependent and independent variables, the coefficient vector on the time varying covariates is the same as that obtained from the FE estimator.

$C_{it}$  and  $B_{it}$  must be viewed as potentially endogenous. We follow the control function (CF) approach in [Wooldridge \(2015\)](#) to eliminate the potential bias. First, we assume a linear reduced form for the endogenous variable  $Y_{it}$ , which could be  $C$  or  $B$ :

$$Y_{it} = \gamma_0 + \gamma_1 Z_{it} + \gamma_2 \bar{Z}_i + r_{it} \quad (5)$$

and obtain the OLS residual  $\hat{r}_{it}$ .  $Z_{it} = (Z_{it1}, Z_{it2})$  where  $Z_{it1}$  are the exogenous variables in [equation \(4\)](#) including  $t$ ,  $R$ , and  $X$ , and  $Z_{it2}$  are instruments for  $C$  or  $B$  which are excluded from [equation \(4\)](#).

As already argued, the theoretical and empirical relationship between  $C$  and market size is long established and robust. Also, the essential nature of the retail banking product means that market size depends on the size of the population, so we use (the natural log of) population as our instrument for  $C$ . Furthermore, a bank's decision to open a branch in a particular location depends on the number of potential customers who live nearby, so we use (the natural log of) population density as our instrument for  $B$ . Note that there is no obvious reason why individual decisions to take up internet banking should be *directly* determined by either population size or population density, so the exclusion restrictions are likely to be satisfied.

We thus estimate the following equation:

$$\log\left(\frac{IB_{it}}{\lambda * IU_{it} - IB_{it}}\right) = a_0 + a_1 C_{it} + a_2 B_{it} + a_3 R_i + a_4 X_{it} + b_0 t + b_1 C_{it} * t + b_2 B_{it} * t + b_3 R_i * t + b_4 X_{it} * t + d_1 \hat{r}_{it}^C + d_2 \hat{r}_{it}^B + d_3 \hat{r}_{it}^{C*t} + d_4 \hat{r}_{it}^{B*t} + c_i + v_{it} \quad (6)$$

where  $\hat{r}_{it}^C$  and  $\hat{r}_{it}^{C*t}$  are CFs (OLS residuals) obtained from the reduced form equations for  $C_{it}$  and  $C_{it} * t$ , respectively. Similarly,  $\hat{r}_{it}^B$  and  $\hat{r}_{it}^{B*t}$  are the CFs obtained from the reduced form equation for  $B_{it}$  and  $B_{it} * t$ .<sup>23</sup>

<sup>23</sup> The instrument for the interactive term of the endogenous variable with time is the selected instrument variable interacted with time.

The control variable vector  $X_{it}$  includes *education* (measured as percentage of population aged 25–74 years who have obtained tertiary education), *Gross Domestic Product (GDP) per capita*, *median income*, *income inequality (Gini coefficient)*, *median age*, and *unemployment*. A full list of variables and data sources is shown in [Appendix 1](#). We cannot estimate  $\lambda$  but expect it to be close to one and conduct a sensitivity analysis for lower values.

Since the estimation of [equation \(6\)](#) uses the estimated  $\hat{r}_{it}$  instead of the true  $r_{it}$ , this extra source of variation has to be taken into account. To do so we implement the bootstrap as suggested by [Wooldridge \(2015\)](#). The significance levels reported later are based on bootstrapped standard errors.

It should be noted that the above estimation strategy employing the CF approach produces results identical to two-stage least-squares (2SLS) estimation. As indicated by [Wooldridge \(2015\)](#), the advantage of using the CF approach is that it allows for a simple and robust test of endogeneity of the relevant (potentially endogenous) variables.

## 5.2 Estimating endogenous national concentration

To estimate [equation \(3\)](#), we construct the dummy variable  $D = 1$  if  $IB > \widetilde{IB}$  and 0 otherwise.  $\widetilde{IB}$  is a threshold value we use to indicate whether an observation is  $IB$  intensive or not. We cannot directly estimate  $\widetilde{IB}$  but use alternative candidate thresholds to test sensitivity. In our data sample,  $IB$  is only available from 2003. In year 2003, if the observation has  $IB$  below the threshold, then  $D = 0$  for all previous years. There are, however, a few countries whose level of  $IB$  is already above the threshold in 2003. In these cases, we use the first-stage probit model estimation (on the sample where we can determine all the values of  $D$ ) to obtain the predicted probability of  $D = 1$  for all observations. If the predicted probability is greater than 50%, we set  $D = 1$  for these observations before 2003 and 0 otherwise. We then apply the first-stage regression again, but with all observations, to obtain the CF to be used for the second-stage estimation.

As in [Section 5.1](#), we apply the correlated random effects approach to control for unobserved heterogeneity and the CF approach to control for potential endogeneity related to  $S$  and  $B$ . As explained earlier, we use the natural log of population size and population density as identifying instruments. All are highly significant in their respective first-stage regressions. The CF enters [equation \(3\)](#) as  $\delta_0 \hat{e}_{it}^S$ ,  $\delta_1 \hat{e}_{it}^{S^*R}$ , and  $\delta_2 \hat{e}_{it}^B$ , to control for endogeneity of  $S$ ,  $S$  interacted with  $R$ , and  $B$ , respectively, where  $\hat{e}_{it}^S$ ,  $\hat{e}_{it}^{S^*R}$ , and  $\hat{e}_{it}^B$  are estimated residuals from reduced forms of  $S$ ,  $S$  interacted with  $R$ , and  $B$ , respectively.

We also need to control for a potential bias arising from the endogeneity of  $D$ , since  $D$  is constructed using  $IB$ . The following CF is adopted following [Wooldridge \(2015\)](#) using a first-stage probit regression to obtain a “generalized error” term defined as:

$\hat{e}_{it}^D = D\lambda(Z_{it}\delta) - (1-D)\lambda(-Z_{it}\delta)$  where  $\lambda(\cdot) = \phi(\cdot)/\Phi(\cdot)$  is the inverse Mills ratio and internet usage,  $IU$ , is used as an identifying instrument for  $D$ .<sup>24</sup>

It should be noted that  $D$  interacts with other variables in [equation \(3\)](#). In models with multiple, nonlinear functions of endogenous variables (like our models, where endogenous variables are interacted with exogenous variables), [Wooldridge \(2015\)](#) points out the advantage of the CF approach in terms of its parsimonious way to account for endogeneity by including only one CF, compared with the standard IV approach using additional instruments (the interaction between instruments and the exogenous variables). He finds that the IV estimator is generally consistent, but the CF approach is more efficient.

For the above reason, we first adopt the CF approach including one CF  $\hat{e}_{it}^D$  to exploit the efficiency of using one CF only. We then adopt the standard IV approach of 2SLS (equivalent to the CF approach including additional CFs associated with the endogenous variables interacted

<sup>24</sup> All exogenous variables and instruments (including the time averages of covariates) used for reduced form of the  $IB$  equation are also included in the reduced form of the  $B$ ,  $S$  and the interaction between  $R$  and  $S$  equations here in the first-stage regression. Four regressions were run to obtain four CFs in the first stage following [Wooldridge \(2015\)](#). All four regressions use the same set of exogenous regressors/instruments.

with other variables) as a consistency check. The results are both reported in Table 4, which suggest high consistency across the two approaches.<sup>25</sup>

## 6. Results

### 6.1 The user uptake of internet banking

First, consider econometric identification. First-stage regression results (not reported) confirm that our identifying instruments for  $C$  and  $B$ , population size and population density, respectively, are both highly significant and contribute substantially to explaining the variance in these endogenous variables. Our estimates of equation (6) are reported in Table 2. As shown toward the bottom of the table, the CF for bank concentration  $\hat{\gamma}_{it}^C$  is significant. This confirms the value of investigating the endogeneity of concentration in Section 5.2. The endogeneity of branch density is not confirmed as  $\hat{\gamma}_{it}^B$  and its interaction with time are insignificant. Although we report that our results with both CFs, excluding those which are not statistically significant, make no substantive difference to the results in Table 2.

Turning to our main results, we start with the complete specification “Spec 1” and then eliminate each insignificant interactive term one by one using F-tests to compare how well each reduced specification fits the data. We end up with the more parsimonious “Spec 2” on which we focus. There is very little difference whether we measure concentration by  $CR5$  or  $HHI$ , so we report both but focus our discussion on the former. We proceed by discussing significance before returning to quantitative effects.<sup>26</sup>

Market structure has a highly significant and nuanced effect on user uptake of internet banking. National concentration has a significant positive effect on the “location” parameter (early adoption by banks). This early adoption of internet banking may be facilitated by incumbents’ response to competitive threats and their immediate scale possibilities of converting existing customers to the internet service in a concentrated market.

However, concentration has a negative effect on the subsequent speed of uptake by consumers. This is consistent with concentrated banks most invested in an expensive brick-and-mortar network to be less aggressive in encouraging their customers to take up internet banking.

Regionalized countries appear less concentrated at the (measured) national level, but if consumers mostly use a regional bank, concentration at the competitively relevant regional market level will be much higher. After controlling for national concentration, we therefore expect the impact of regionalization on the user adoption of internet banking to be qualitatively similar to that of market concentration. This is exactly what we find. An additional incentive for early adoption by regional banks is that internet banking gives access to a much larger pool of customers outside their home region. But this pool of customers may have less familiarity and trust in banks outside their region, which adds to delayed uptake (lower speed). Table 3 combines both location and speed effects to show the accumulated impact over time. It reveals that the early positive impact of regionalization was eliminated by the end of our period.

Figure 4a and b illustrate the substantial quantitative effect of market structure, and the combined effects of concentration and regionalization. The figures use the estimated coefficients from Table 2 to predict internet banking uptake for three illustrative levels of the concentration ratio (25%, 50%, and 75%) and two of regionalization ( $R=0$  and  $R=0.7$ ). The higher value is a natural choice since it applies to Germany, Italy, and Spain.

First consider Figure 4a. We start with the second row. Panel *d* shows how uptake in a non-regional, low-concentration market proceeds over the sample period, increasing from 8% in 2003 to 60% in 2018. Panels *e* and *f* show how uptake increases with concentration (from low to

<sup>25</sup> The size of the banking sector was much more limited under the pre-1990 communist regimes of Eastern and Central Europe, and this might have had an effect on more recent levels of concentration. Although markets were opened up a decade before our sample period begins, we created a dummy variable for these countries and included it as an additional instrument for bank market size alongside population and our other exogenous variables. The effect of this dummy variable is significant in the first-stage regression, justifying its inclusion.

<sup>26</sup> We also ran 2SLS regressions for “Spec 2” using  $CR5$  and  $HHI$  as measures of market concentration. The estimated coefficients and significance levels are identical to the CF approach, as expected, except that the standard errors differ slightly. These estimates are not reported to avoid repetition.

**Table 2.** Estimation results for the diffusion of *IB*

Dependent variable: user uptake of <i>IB</i>	Concentration measured by <i>CR5</i>		Concentration measured by <i>HHI</i>	
	Spec 1	Spec 2	Spec 1	Spec 2
<i>C</i> (concentration)	0.047*** (0.0151)	0.048*** (0.0079)	7.199* (3.8037)	8.375*** (2.7339)
<i>R</i> (regionalization)	1.805*** (0.5700)	1.723*** (0.3625)	0.792 (0.6020)	0.887* (0.3646)
<i>B</i> (branch density)	0.126 (0.1319)	0.073 (0.1001)	0.288* (0.1272)	0.217*** (0.1023)
<i>E</i> (education, %)	-0.013 (0.0216)	-0.038*** (0.0113)	-0.001 (0.0233)	-0.036*** (0.0117)
<i>G</i> ( <i>GDP per capita</i> , log)	-0.170 (0.7579)	-	-0.142 (0.6947)	-
<i>I</i> (median income, log)	0.325 (0.6031)	0.528* (0.1936)	0.355 (0.5880)	0.571* (0.2099)
<i>Gi</i> ( <i>Gini income</i> , 0–100)	0.071** (0.0283)	0.082*** (0.0255)	0.053 (0.0278)	0.078*** (0.0246)
<i>PA</i> (median age, log)	-2.584 (3.4694)	-	-5.739* (3.4353)	-
<i>U</i> (unemployment, %)	-0.009 (0.0313)	0.025*** (0.0062)	-0.020 (0.0313)	0.025*** (0.0061)
<i>T</i>	0.194 (0.5827)	0.407*** (0.0530)	-0.084 (0.6437)	0.297*** (0.0468)
<i>C</i> * <i>t</i>	-0.001** (0.0007)	-0.001*** (0.0004)	-0.117 (0.2435)	-0.209 (0.1690)
<i>R</i> * <i>t</i>	-0.089*** (0.0359)	-0.083*** (0.0211)	-0.038 (0.0391)	-0.045* (0.0220)
<i>B</i> * <i>t</i>	-0.031*** (0.0108)	-0.027*** (0.0087)	-0.036** (0.0099)	-0.031*** (0.0082)
<i>E</i> * <i>t</i>	-0.002 (0.0014)	-	-0.002 (0.0015)	-
<i>G</i> * <i>t</i>	0.006 (0.0408)	-	0.017 (0.0378)	-
<i>I</i> * <i>t</i>	0.018 (0.0432)	-	0.015 (0.0418)	-
<i>Gi</i> * <i>t</i>	-0.008*** (0.0018)	-0.009*** (0.0017)	-0.007* (0.0018)	-0.008*** (0.0016)
<i>PA</i> * <i>t</i>	0.027 (0.1297)	-	0.017 (0.1501)	-
<i>U</i> * <i>t</i>	0.002 (0.0020)	-	0.003 (0.0020)	-
Constant	-67.915** (34.6492)	-67.814** (34.3305)	-13.007 (32.3030)	-12.594 (31.7151)
$\hat{\tau}_t^C$	-0.066*** (0.0169)	-0.068*** (0.0141)	-12.829** (4.9540)	-14.005*** (4.1566)
$\hat{\tau}_t^B$	0.003*** (0.0011)	0.004*** (0.0009)	0.690* (0.3094)	0.782*** (0.2567)
$\hat{\tau}_t^E$	-0.029 (0.2938)	-0.024 (0.2844)	-0.187 (0.3082)	-0.117 (0.3028)
$\hat{\tau}_t^I$	0.027 (0.0221)	0.024 (0.0212)	0.033 (0.0230)	0.028 (0.0224)
Adjusted <i>R</i> <sup>2</sup>	0.92	0.92	0.92	0.92
No. of Obs.	381	381	381	381
<i>F</i> test to compare spec 1 & spec 2 (5% critical value: 1.84)	0.67	0.67	1.51	1.51

1. \*\*\*, \*\*, and \* indicate statistical significance at 1%, 5%, and 10%, respectively; standard errors are in brackets.  
 2. CFEs for the two endogenous variables *B* and *C* and their interactive terms with time are obtained in the first stage following Wooldridge (2015).<sup>27</sup>  
 3. All regressions include the time averages of all time varying explanatory variables in spec 1 to account for any FEs, except for the two variables of main interest which are *C* and *B*. *R* is of main interest too, but it is fixed over time. By using the Mundlak approach, we are able to control for FEs but can still investigate the effect of time fixed variables such as *R*.  
 4. Recall that we assume the maximum possible uptake (saturation level) is a proportion,  $\lambda$ , of current internet usage,  $IU_{it}$ , so  $M_{it} = \lambda * IU_{it}$ . The results reported here are for  $\lambda = 1$ . We repeated our estimations using  $\lambda = 0.95$  and  $\lambda = 0.90$ . The latter necessitated a slightly reduced sample, since our sample's maximum ratio of internet banking to internet usage is 0.93. There was no substantial difference from our reported results.

<sup>27</sup> In the first stage, the regressions were run as follows. Our endogenous variables include *C* and *B*, and the time interactions with *C* and *B*. All the exogenous variables include instruments for *C* and *B* (which are population and population density), the time interactions with exogenous variables/instruments, and the time average of all these variables. Four regressions were run to obtain four CFEs in the first stage. All four regressions use the same set of exogenous regressors/instruments.

**Table 3.** Cumulative effect of regionalization over time

$t$ (year)	Coefficient	Standard Error
7 (2003)	1.143 <sup>***</sup>	0.2307
8 (2004)	1.060 <sup>***</sup>	0.2136
9 (2005)	0.977 <sup>***</sup>	0.1973
10 (2006)	0.895 <sup>***</sup>	0.1820
11 (2007)	0.812 <sup>***</sup>	0.1680
12 (2008)	0.729 <sup>***</sup>	0.1556
13 (2009)	0.646 <sup>***</sup>	0.1452
14 (2010)	0.563 <sup>***</sup>	0.1374
15 (2011)	0.480 <sup>***</sup>	0.1324
16 (2012)	0.398 <sup>**</sup>	0.1308
17 (2013)	0.315 <sup>*</sup>	0.1326
18 (2014)	0.232 <sup>*</sup>	0.1376
19 (2015)	0.149	0.1455
20 (2016)	0.066	0.1560
21 (2017)	-0.016	0.1684
22 (2018)	-0.099	0.1825

1. \*\*\*, \*\*, and \* indicate statistical significance at 1%, 5%, and 10%, respectively.

2. The above are coefficients of  $R$  (regionalization) varying with time, estimated from Specification 2 with CR5 in Table 2.

medium and to high). For example, for three hypothetical countries with the same characteristics other than concentration, panel  $f$  shows that the high-concentration country would achieve 50% consumer uptake in 2008, the medium-concentration country in 2013, and the low-concentration country in 2016. The earlier introduction in higher concentration countries is not overtaken by the faster speed of uptake in lower concentration countries before they converge near market saturation. Meanwhile, a 25% point higher concentration ratio could bring forward consumer benefits from internet banking by 3–5 years.

A similar effect of concentration in regionalized markets is seen by comparing across panels  $a$ ,  $b$ , and  $c$ . Note that in all cases the accumulated years of delay due to market structural factors, and consequently foregone consumer surplus, are substantial.

The effects of regionalization are most clearly seen in Figure 4b. Regionalized markets are shown in the first row, and the second row compares them with non-regionalized markets. For all levels of concentration, regionalized markets had an earlier adoption of internet banking, but this advantage had disappeared by the end of our sample period.

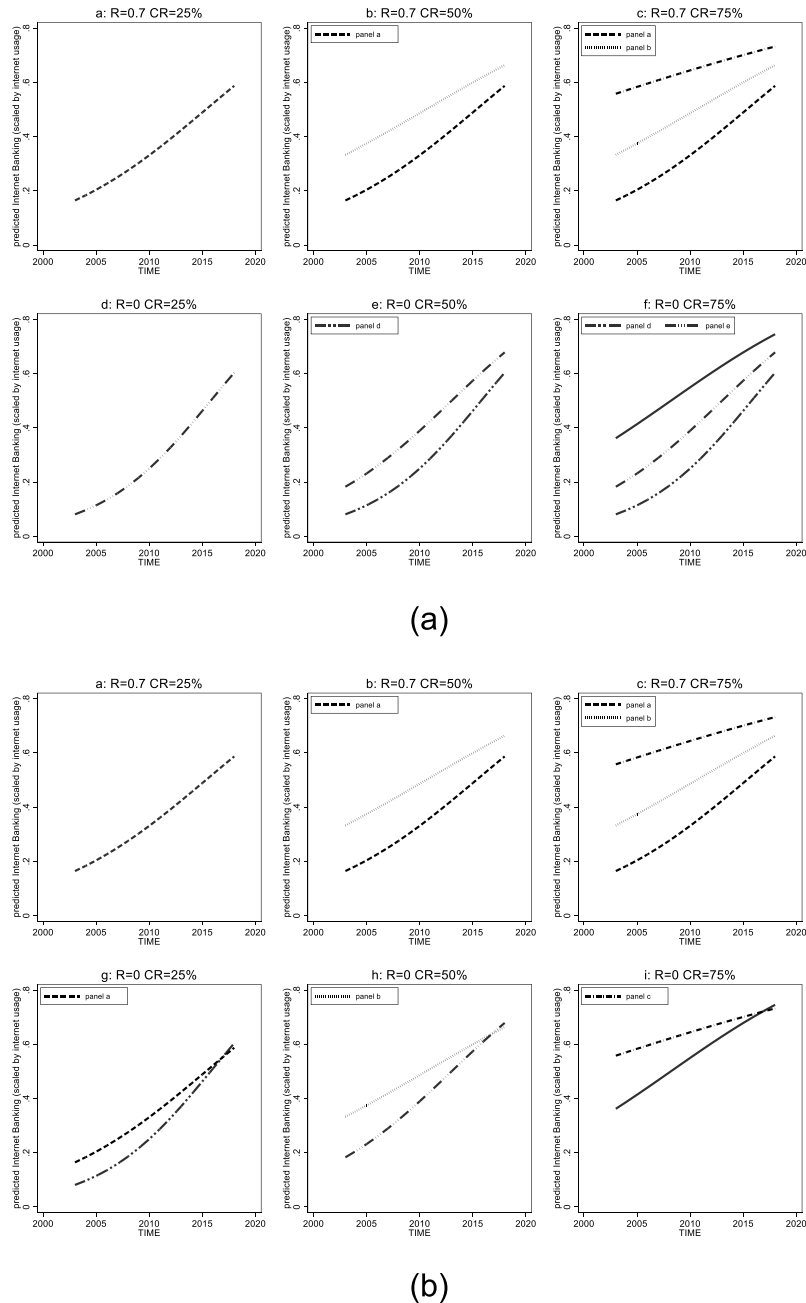
Returning to Table 2, there is significant evidence that high branch density,  $B$ , slows down the speed of user uptake. This is consistent with consumers, who have easy access to a local branch, also seeing less advantage in internet banking. The effect of  $B$  on the initial introduction is inconclusive. It is insignificant when CR5 is used as the concentration measure, but the  $HHI$  results suggest a positive effect on the earlier introduction.

Of our remaining variables, the independent time trend is positive, as expected in any diffusion model. Median income brings forward the introduction of internet banking but has no significant effect on the subsequent speed of uptake. Income inequality similarly incentivizes introduction but thereafter reduces the speed of uptake. In combination, these consumer income effects suggest that richer consumers are earlier adopters. Having controlled for income effects, education and employment both have a negative impact on the initial introduction with insignificant effects on the speed of adoption.

## 6.2 Endogenous market concentration

Table 4 reports the results from our estimation of equation (3). We report results for a candidate threshold effect for internet banking to have a substantial effect on concentration of  $\widetilde{IB} = 25\%$ . Results for  $\widetilde{IB} = 30\%$  are reported in Supplementary Appendix 3 alongside other robustness checks. The alternative thresholds do not materially change our estimates.





**Figure 4.** (a) Predicted internet banking uptake: the impact of concentration given each level of regionalization. (b) Predicted internet banking uptake: the impact of regionalization given each level of concentration. *Notes:* (i)  $R$  refers to regionalization;  $CR$  refers to the five-firm concentration ratio; (ii) the predicted adoption of internet banking is based on the model estimates from Spec 2 in Table 2.

We include a control variable ( $A$ ) motivated by the financial crisis that arose in the middle of our sample period. This could potentially have been a confounding factor with an impact on concentration at a time of rising internet banking uptake. There was considerable variation in

**Table 4.** Estimation results for national concentration in banking

Dependent variable: $\log\left(\frac{C_{it}}{100-C_{it}}\right)$	Estimated coefficients CF	Estimated coefficients 2SLS
$T$	0.090 <sup>***</sup> (0.0082)	0.081 <sup>***</sup> (0.0113)
$\frac{1}{\ln S_{it}}$	49.706 <sup>***</sup> (7.4198)	58.581 <sup>***</sup> (7.6082)
$R_i$	-11.674 <sup>***</sup> (2.1856)	-8.038 <sup>***</sup> (1.7908)
$B_{it}$	0.286 <sup>***</sup> (0.0670)	0.259 <sup>***</sup> (0.0688)
$R_i * \frac{1}{\ln S_{it}}$	149.774 <sup>***</sup> (30.1833)	98.528 <sup>***</sup> (24.1566)
$D$	-2.058 <sup>***</sup> (0.5838)	-1.368 <sup>*</sup> (0.8381)
$D * t$	-0.090 <sup>***</sup> (0.0101)	-0.077 <sup>***</sup> (0.0144)
$D * \frac{1}{\ln S_{it}}$	31.362 <sup>***</sup> (6.4704)	20.863 <sup>***</sup> (9.8031)
$D * R_i$	13.119 <sup>***</sup> (2.5263)	7.686 <sup>***</sup> (2.4002)
$D * B_{it}$	-0.159 <sup>***</sup> (0.0615)	-0.154 <sup>**</sup> (0.0700)
$D * R_i * \frac{1}{\ln S_{it}}$	-179.961 <sup>***</sup> (36.7303)	-101.104 <sup>***</sup> (33.7451)
A (crisis)	9.572 <sup>***</sup> (2.0891)	8.355 <sup>***</sup> (2.7722)
Constant	3.706 (3.1020)	3.884 (4.0957)
$\hat{e}_{it}^S$	-8.587 (16.3207)	
$\hat{e}_{it}^B$	-0.981 <sup>***</sup> (0.0883)	
$\hat{e}_{it}^D$	-0.128 (0.0937)	
$\hat{e}_{it}^{R*S}$	-199.214 <sup>***</sup> (35.9062)	
Adjusted R <sup>2</sup>	0.77	0.65
No of Obs.	430	430

Standard errors in brackets.

\*\*\*, \*\*, and \* indicate statistical significance at 1%, 5%, and 10%, respectively.

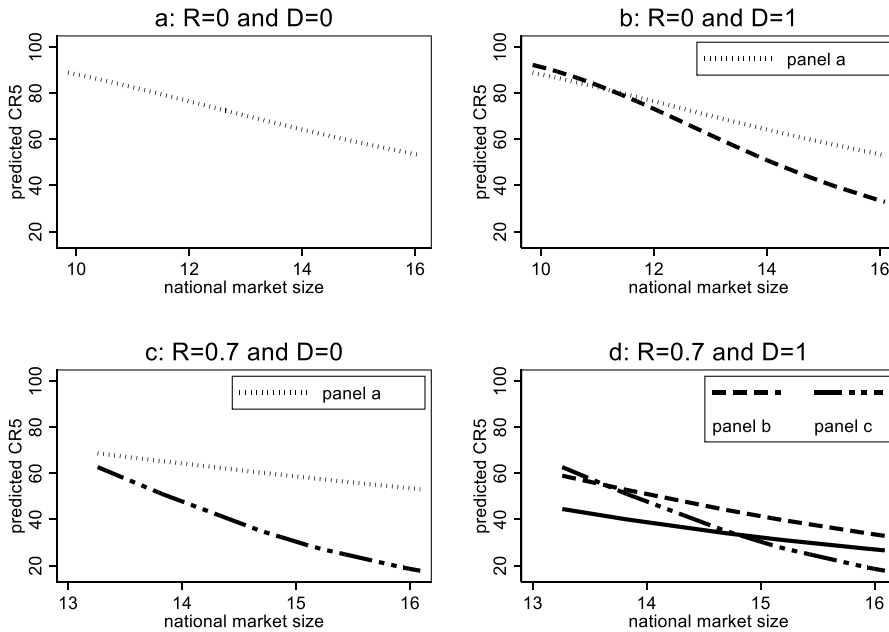
the extent to which the financial crisis hit different European banks and at what time. Identification is aided by the fact that this pattern was not closely correlated with internet banking uptake. We measure the extent of the crisis by the total amount of state aid used by EU member states, as published by the European Commission. More precisely, our measure is the cumulative total amount of aid in the form of recapitalization and impaired asset relief relative to market size (measured by total assets). As expected, this control variable does significantly increase the concentration in our sample.

Next, consider the significance of the CF errors reported at the bottom of the first set of results in Table 4.  $\hat{e}_{it}^B$  is highly significant and with a negative coefficient, which confirms the endogeneity of branch density. There is similar evidence that the size of the banking sector (measured by assets) is also endogenous. The CF error term for the internet banking threshold is not significant. The second column reports 2SLS estimates. The results are similar to our CF estimates so we focus on the first column.

Consider our “pre-internet banking” estimates (i.e., for  $D=0$ ). All variables are highly significant. As expected, concentration falls with both market size and regionalization, and regionalization has a stronger effect in larger countries. We also find that a dense branch network appears to have been an entry barrier and is associated with high concentration. Bearing in mind that we use an inverse measure of market size in our estimation, the “de-concentrating” effect of internet banking (discussed later) is supported by the statistical significance of  $D$  and its interaction with market size.

In order to understand the quantitative effects of internet banking, it is helpful to consider Figure 5. We use Table 4 estimates to compute and plot the predicted relationship between concentration and national market size, with  $t$  and  $B$  set at their mean values. Each panel combines two values of  $D$  and  $R$ . We compare combinations of pre internet banking ( $D=0$ ) and post internet banking ( $D=1$ ) and non-regionalized markets ( $R=0$ ) and regionalized markets ( $R=0.7$ ).

Consider the non-regionalized markets (comparing panels *a* and *b*). Internet banking has substantially reduced concentration but only in larger markets. This “de-concentrating effect” is consistent with internet banking eroding incumbency advantages (e.g., branch network) and encouraging entrants and/or expansion of challengers. More generally, following Sutton (1991),



**Figure 5.** Predicted market concentration varying with different levels of regionalization and maturity of internet banking. *Note:* national market size is measured by the natural log of total assets (instrumented by population);  $R$  refers to regionalization and  $D$  refers to the dummy constructed to indicate pre or post internet banking era;  $D=1$  indicates post-internet banking and 0 otherwise; predicted  $CR5$  (concentration ratio) is based on model estimates from Table 5. The size of regionalized markets in our sample is relatively large. To enable the comparison without too much extrapolation, we have adjusted the scale of national market size for panels  $c$  and  $d$ .

the steeper slope of the concentration–market size curve implies that quality competition after internet banking involves less costly escalation of sunk costs as compared with a competitive process that focuses on a brick-and-mortar network.

Next, consider regionalized markets. The comparison between panels  $c$  and  $d$  shows that internet banking has reduced concentration in small and medium markets but not in the larger markets (such as Germany, Italy, and Spain). This can be interpreted in terms of two distinct effects working against each other: a within-region de-concentrating technology effect with reduced importance of the branch network (as for non-regionalized countries) and an “extended geographic reach” effect out of the home region raising competition, which results in bank consolidation and exit. Comparing panels  $c$  and  $d$ , we find that the net effect is that internet banking reduces national concentration in smaller countries with regionalized markets but raises national concentration in larger regionalized markets.

Finally, comparing panels  $a$  and  $c$ , we can see the very large difference between concentration in integrated markets and regionalized markets pre internet banking, particularly in large markets. Further comparison with panels  $b$  and  $d$  shows how internet banking has substantially closed that gap.

## 7. Robustness checks

We considered a number of robustness checks relating to the cut-off for internet banking in the concentration estimation, a possible outlier country (Luxembourg), and 2SLS estimation. Considering that branch density (as measured in our study) can vary significantly within a country dependent on the size of the urban versus rural areas, we included an additional instrument for

both models (the natural log of metropolitan areas in each Member State). In each case, we re-ran our estimations using alternative assumptions and compared the results with those presented in the main text. Detailed results and explanations are available in [Supplementary Appendices 2, 3, and 4](#) (available online).

## 8. Conclusions

In this paper, we have tried to understand the market effects of digital product disruption by examining the co-evolution of market structure and consumer uptake of a new product technology. Our econometric models investigate the consequences of market concentration and bank regionalization for the adoption of internet banking and vice versa. This required us to use a range of identification strategies as discussed in [Section 5](#). We are not aware of any previous literature that has tried to examine this co-evolution, but our results suggest that it is important to do so.

Our first set of results relate to how market structure affects both digital product introduction and consumer acceptability of the new product. We find that banks introduce internet banking earlier in more concentrated markets, but then the speed of consumer uptake is slower than in less concentrated markets. We interpret this as firms with loyal customer bases in concentrated markets seeking to pre-empt entry or expansion of smaller incumbents but then not investing so creatively to convert their existing customers to the internet. Our simulations show that the former effect (early introduction) outweighs the latter (slower uptake), with catch-up only as the market becomes saturated. Consequently, within the range of observed concentration levels, there may be as much as an 8-year delay in achieving a 50% consumer uptake in low-concentration markets. Although this estimate assumes that all other relevant variables are held constant and compares only hypothetical countries, it does suggest a very substantial loss of potential consumer surplus. In reality, this effect is partially mitigated in that low-concentration banking markets (measured at the national level) also tend to be very regionalized, and the latter is shown to have a similar effect to higher concentration.

Our second set of results addresses the longer-term impact of internet banking on the evolution of market structure. This is an ongoing process as market structure evolves slowly, and the versatility of and trust in internet banking develop over time. Our sample period is also short of observations for when consumer uptake is saturated, so our results are best interpreted as directions of travel. We find a “deconcentrating effect” of internet banking as it facilitates the entry and expansion of smaller banks that can compete without needing to invest so heavily in expensive brick-and-mortar networks.

We also find significant differences in the market structure effects of internet banking in countries with initially regionalized banks as compared with those that started from a more unified national market. We identify an “extended geographic reach effect” of internet banking because different regional banks are enabled to compete in each other’s regions. This drives a process of consolidation and exit that tends to increase concentration at the national level, even as it strengthens competition within each region. We find that the deconcentrating effect outweighs the extended geographic reach effect in small countries, but the reverse holds for large countries.

Taken together, our results show how disruptive a new product technology can be, even when that technology is equally available to incumbents and entrants. In the case of internet banking, concentration facilitated a more rapid rollout of this fairly generic new product, but this new product ultimately resulted in the erosion of that concentration (or dominance).

## Supplementary Data

[Supplementary materials](#) are available at *Industrial and Corporate Change* online.

## References

Abernathy, P. M. 2020, ‘News Deserts and Ghost Newspapers: Will Local News Survive?’ The University of North Carolina Press, <https://tinyurl.com/AbernathyNewsDeserts>.

- Berry, S. T. (1992), 'Estimation of a model of entry in the airline industry,' *Econometrica*, 60(4), 889–917.
- Berry, S. and J. Waldfogel (2010), 'Product quality and market size,' *The Journal of Industrial Economics*, LVIII(1), 1–31.
- Bresnahan, T. F. and P. C. Reiss (1991), 'Entry and competition in concentrated markets,' *Journal of Political Economy*, 99(5), 977–1009.
- Christensen, C. M. (1997). *The Innovator's Dilemma: When New Technologies Cause Great Firms To Fail*, Harvard Business School Press: Boston, MA.
- Christensen, C. M., M. E. Raynor and R. McDonald (2015), 'What is Disruptive Innovation?' *Harvard Business Review*, <https://hbr.org/2015/12/what-is-disruptive-innovation>.
- Christensen, C. M. and R. S. Rosenbloom (1995), 'Explaining the attacker's advantage: technological paradigms, organizational dynamics, and the value network,' *Research Policy*, 24(2), 233–257.
- Cohen, A. M. and M. J. Mazzeo (2010), 'Investment strategies and market structure: an empirical analysis of bank branching decisions,' *Journal of Financial Services Research*, 38(1), 1–21.
- Dick, A. (2007), 'Market size, service quality, and competition in banking,' *Journal of Money, Credit, and Banking*, 39(1), 49–81.
- Gans, J. (2016), 'The other disruption,' *Harvard Business Review*.
- Griliches, Z. (1957), 'Hybrid corn: an exploration in the economics of technological change,' *Econometrica*, 25(4), 501–522.
- Gruber, H. and F. Verboven (2001a), 'The diffusion of mobile telecommunications services in the European Union,' *European Economic Review*, 45(3), 577–588.
- Gruber, H. and F. Verboven (2001b), 'The evolution of markets under entry and standards regulation—the case of global mobile telecommunications,' *International Journal of Industrial Organization*, 19(7), 1189–1212.
- Ho, J. C. and H. Chen (2018), 'Managing the disruptive and sustaining the disrupted: the case of Kodak and Fujifilm in the face of digital disruption,' *Review of Research Policy*, 35(3), 352–371.
- Hoenig, J. M. and D. M. Heisey (2001), 'The abuse of power: the pervasive fallacy of power calculations in data analysis,' *The American Statistician*, 55(1), 19–24.
- Igami, M. (2017), 'Estimating the innovator's dilemma: structural analysis of creative destruction in the hard disk drive industry, 1981–1998,' *Journal of Political Economy*, 125(3), 798–847.
- Jenkins, J. and R. K. Nielsen (2018), 'The digital transition of local news,' *Reuters Institute Digital News Report*.
- Kim, M. and B. Vale (2000), 'Non-price Strategic Behavior: The Case of Bank Branches,' *International Journal of Industrial Organization*, 19(10), 1583–1602.
- Li, Y. and B. Lyons (2012), 'Market structure, regulation and the speed of mobile network penetration,' *International Journal of Industrial Organization*, 30(6), 697–707.
- Mundlak, Y. (1978), 'On the pooling of time series and cross section data,' *Econometrica*, 46(1), 69–85.
- Nickerson, D. and R. Sullivan (2003), 'Financial innovation, strategic real options and endogenous competition: theory and an application to internet banking, payments system research,' *Federal Reserve Bank of Kansas City Working paper 03-01*.
- O'Reilly, C. A. and M. L. Tushman (2004), 'The ambidextrous organization,' *Harvard Business Review*, 82(4), 74–81.
- Pelletier, A., S. Khavul and S. Estrin (2020), 'Innovations in emerging markets: the case of mobile money,' *Industrial and Corporate Change*, 29(2), 395–421.
- Reinganum, J. F. (1989), 'The timing of innovation: research, development, and diffusion,' in R. Schmalensee and R. Willig (eds), *Handbook of Industrial Organization*, Vol. 1, 1st edn, chapter 14. : Elsevier, pp. 849–908.
- Schulz, K. F. and D. A. Grimes (2005), 'Sample size calculations in randomized trials: mandatory and mystical,' *The Lancet*, 365(9467), 1348–1353.
- Shaked, A. and J. Sutton (1987), 'Product differentiation and industrial structure,' *Journal of Industrial Economics*, 36(2), 131–146.
- Sullivan, R. and Z. Wang (2013), 'Internet banking: an exploration in technology diffusion and impact,' *The Federal Reserve Bank of Richmond Working Paper No. 13-10*.
- Sutton, J. (1991), *Sunk Costs and Market Structure: Price Competition, Advertising, and the Evolution of Market Structure*. MIT Press: Cambridge, MA.
- Sutton, J. (1998), *Technology and Market Structure*. MIT Press: Cambridge, MA.
- Sutton, J. (2007), 'Market structure: theory and evidence,' in M. Armstrong and R. Porter (eds), *The Handbook of Industrial Organization*, Vol. 3. Elsevier, pp. 2301–2368.
- Takieddine, S. and J. Sun (2015), 'Internet banking diffusion: a country-level analysis,' *Electronic Commerce Research and Applications*, 14(5), 361–371.
- Teece, D. J. (2007), 'Explicating dynamic capabilities: the nature and micro foundations of (sustainable) enterprise performance,' *Strategic Management Journal*, 28(13), 1319–1350.
- Temesvary, J. (2015), 'Dynamic branching and interest rate competition of commercial banks: evidence from Hungary,' *International Journal of Industrial Organization*, 43, 98–110.

- The Furman Review. (2019), 'Unlocking Digital Competition,' *Report of the Digital Competition Expert Panel*, HM treasury, <https://www.gov.uk/government/publications/unlocking-digital-competition-report-of-the-digital-competition-expert-panel>.
- Wooldridge, J. M. (2015), 'Control function methods in applied econometrics,' *The Journal of Human Resources*, 50(2), 420–445.
- Wooldridge, J. M. (2019), 'Correlated random effects models with unbalanced panels,' *Journal of Econometrics*, 211(1), 137–150.

## Appendix 1

### Data description

**Table A1.** Variables and sample used for the diffusion of internet banking model

	Mean	Min	Max	sd	cv	N	Definition	Source
<i>CR5</i> (%)	60.70	21.99	98.55	17.86	0.29	381	Concentration ratio of the largest five banks	ECB Statistical Data Warehouse (Publications/Reports/Financial Corporations/Structural Financial Indicators)
<i>HHI</i>	0.11	0.02	0.40	0.07	0.65	381	Herfindahl–Hirschman Index	ECB Statistical Data Warehouse (Publications/Reports/Financial Corporations/Structural Financial Indicators)
<i>R</i>	0.18	0.00	0.70	0.25	1.41	381	Bank Regionalization	Calculated using bank headquarter information from Banker Database
<i>B</i> (natural log)	-3.39	-6.12	-0.99	1.12	-0.33	381	Branch density: number of branches per km <sup>2</sup>	ECB Statistical Data Warehouse (Publications/Reports/Financial Corporations/Structural Financial Indicators)
<i>IB</i> (%)	39.08	1.00	90.00	23.71	0.61	381	Internet banking usage: the proportion of individuals who used the Internet banking in the last 3 months	Eurostat: Digital Economy and Society
<i>IU</i> (%)	69.69	17.80	98.14	16.99	0.24	381	Internet usage: the proportion of individuals who used the Internet in the last 3 months.	Eurostat: Digital Economy and Society
<i>PD</i> (natural log)	4.64	2.74	7.32	0.91	0.20	381	Population density: population per km <sup>2</sup>	Eurostat: Population and Social Conditions
Population (natural log)	8.92	6.00	11.32	1.42	0.16	381	Population	Eurostat: Population and Social Conditions
GDP per capita (natural log)	10.13	9.14	11.30	0.38	0.04	381	GDP per capita	Eurostat: Economy and Finance
Median age (natural log)	3.70	3.50	3.84	0.06	0.02	381	Median age of the population	Eurostat: Population and Social Conditions
Median income (natural log)	9.43	7.93	10.37	0.49	0.05	381	Median income of the population	Eurostat: Population and Social Conditions
Gini coefficient of income (0–100)	29.86	20.90	40.20	4.00	0.13	381	Gini coefficient of equalized disposable income before social transfers	Eurostat: Population and Social Conditions
Unemployment rate	8.64	4.29	2.20	27.50	3.18	381	Unemployment rate as a percentage of active population	Eurostat: Population and Social Conditions
Education (%)	24.48	9.90	40.50	7.43	0.30	381	The percentage of population aged 25–74 years who have obtained Tertiary education	Eurostat: General and regional statistics

**Table A2.** Variables and sample used for the endogenous market concentration model

	Mean	Min	Max	sd	cv	N	Definition	Source
CR5 (%)	59.19	18.95	97.28	18.57	0.31	430	Concentration ratio of the largest five banks	ECB Statistical Data Warehouse (Publications/Reports/Financial Corporations/Structural Financial Indicators)
TA <sup>a</sup> (natural log)	13.00	9.49	16.12	1.76	0.14	430	Total assets from the aggregated balance sheet of MFIs (excluding national central banks)	ECB Statistical Data Warehouse (Publications/Reports/Money, Credit and Banking/MFI balance sheet)
R	0.20	0.00	0.70	0.26	1.32	430	Bank Regionalization	Calculated using bank headquarter information from Banker Database
B (natural log)	-3.30	-6.12	-0.99	1.12	-0.34	430	Branch density: number of branches per km <sup>2</sup> .	ECB Statistical Data Warehouse (Publications/Reports/Financial Corporations/Structural Financial Indicators)
D	62%	of observations	with D = 1	when IB threshold is at 25%		430	Dummy variable to indicate internet banking penetration	Constructed using data on Internet banking (IB)
IU (%)	69.26	16.00	98.00	18.11	0.26	377	Internet usage: the proportion of individuals who used the Internet from any location in the last 3 months	World development indicator database
PD (natural log)	4.68	2.72	7.32	0.91	0.19	430	Population density: population (in thousands) per km <sup>2</sup>	Eurostat: General and regional statistics
Population (natural log)	9.07	6.00	11.32	1.41	0.16	430	Population (in thousands)	Eurostat: general and regional statistics
A (crisis, deflated by total assets)	0.01	0.00	0.16	0.02	2.30	430	Accumulative total amount of state aid in the form of recapitalization and impaired asset relief deflated by total assets	European Commission: <a href="https://ec.europa.eu/competition/state_aid/scoreboard/index_en.html">https://ec.europa.eu/competition/state_aid/scoreboard/index_en.html</a>

<sup>a</sup>Total assets data is obtained from the aggregated balance sheet of Monetary Financial Institutions (MFIs) (excluding national central banks) by country. MFIs are defined as central banks and other MFIs, which comprise credit institutions, deposit-taking corporations other than credit institutions, and money market funds. Ideally, we would like to have total assets measured by the aggregated balance sheet of all credit institutions only in a given country to be consistent with our banking market structural measures. But due to the lack of consistent data across all years and all countries, we use the data from the aggregated balance sheet of all MFIs. The share of total assets of all credit institutions in the whole euro area accounts for about 97% of total assets of all MFIs in 2018. Therefore, we think using total assets from MFIs would be a representative and good proxy of total assets of credit institutions.



