Department of Economics

University of Reading

Consumer Uptake of Internet Banking, Endogenous Market Structure and Regional Integration in Europe

by Bruce Lyons and Minyan Zhu

Discussion Paper No. 2019-14

Department of Economics
University of Reading
Whiteknights
Reading
RG6 6AA
United Kingdom

www.reading.ac.uk

Consumer Uptake of Internet Banking, Endogenous Market Structure and Regional Integration in Europe

Bruce Lyons* and Minyan Zhu**

*CCP and School of Economics, University of East Anglia, Norwich, NR4 7TJ, UK

Email: b.lyons@uea.ac.uk

**Department of Economics, University of Reading, Reading, RG6 6AA, UK;

Email: minyan.zhu@reading.ac.uk

Abstract

This paper examines how market structure influences the early introduction and consumer uptake of a digital service that is a convenient alternative to traditional service delivery. Digital provision also has "extended geographic reach" and "lower sunk costs" as compared with bricks-and-mortar service provision. We further examine how these affect market structure. Internet banking provides an important example that also allows us to separate regional integration and national concentration dimensions of market structure. We develop an econometric model of the effects of market structure on the introduction and consumer uptake of internet banking. We estimate using panel data for all EU Member States and find that both concentration and regionalisation bring these forward. Next, we examine how consumer uptake of the digital product then begins to impact on banking market structure. We find a substantial de-concentrating effect in large non-regionalised markets and indirect evidence of integration in previously regionalised markets. This is consistent with internet banking having enhanced competition in both integrated markets and, despite little change in national concentration, also in previously regionalised markets.

Keywords: Internet Banking, Digital Markets, Endogenous Market Structure, Market Integration, Consumer Diffusion

JEL Codes: L11, O33, F15, G21, L81

Version date: 25 April 2019

1. Introduction

Internet technology has changed how consumers participate in markets, and this has important implications for market competition. Some key issues have been much studied in the recent literature; for example, how the internet facilitates price comparisons for consumers and data collection by firms. Much less research has addressed the competitive effects of how goods and services are delivered, which can have very substantial impacts on both consumer welfare and supply-side competition. In particular, some services previously requiring a visit to a shop or bank can now be provided over the internet. Our focus in this paper is on the relationship between retail banking market structure and consumer take-up of a new way to deliver banking services ("internet banking"). We provide an empirical investigation of both lines of causation. First, we ask how the existing market structure, and associated networks of bricks-and-mortar branches, have affected the introduction of internet banking and the speed of its uptake by consumers. Second, we investigate how, once internet banking has reached a sufficient level of consumer penetration, it begins to change market structure.

Internet banking provides potentially very large consumer benefits. Use of a branch bank to conduct standard transactions requires time, travel and queuing costs, and severely limits convenience as branches have only limited opening hours. In addition to these direct costs, transactions may be held back until there are sufficient to make a trip to the bank 'worth it'. In contrast, internet banking is available on demand and can be used with almost unlimited frequency. The early provision of attractive digital services may therefore be as important for consumer welfare as the effects of market structure on price. Security concerns weigh heavily with some potential users and some individuals value face-to-face banking, so the uptake of new internet-based services will never be instantaneous. These trade-offs can be influenced by bank investments in technology, design, security and marketing, which in turn may be determined by competition and market structure. More generally, there has been very little econometric research on the effect of market structure on consumer adoption of new consumer products or services, with mobile phones providing a rare exception.[1]

The supply-side effects of internet banking are also significant because internet provision of services facilitates both entry and market integration. It also provides a case study of a disruptive technology that is not limited to the original innovator or restricted by licence agreements. The branch banking business model has a very expensive cost structure because it requires a heavy investment in property and people, with both dispersed widely across the potential market. Internet banking is constrained by national regulatory regimes but, unlike branch banking, it is not geographically tied to where the customers are. In principle, this facilitates both de novo entry and cross-entry by previously regional banks. This can be expected to lead to long-term changes in market structure. There is a considerable body of research on technology and market structure (e.g. Sutton, 1998) but

[1] See, for example, Gruber and Verboven (2001a, 2001b) and Li and Lyons (2012).

we are aware of no previous work on how internet service delivery affects concentration and regional integration.

This paper presents an econometric investigation of internet banking and market structure using a panel of data for EU member states. Even within the Eurozone, national banking regulations mean that the geographic market for competition in retail banking is no wider than national. Furthermore, some countries have a very strong tradition of regionally segmented banking markets. This is reflected in wide differences in national concentration. 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 much more regionalised Germany and Italy this concentration ratio (C5) was only 25% and 34%.[2] National concentration measures would therefore be very partial and potentially misleading measures of market structure if we could not control for the degree of regionalisation of markets. In the absence of relevant regional market data, our approach is to consider two dimensions of market structure: national concentration and a proxy for the degree of regionalisation.

In section 2.1, we set out our modelling approach to the consumer uptake of internet banking, separating the influences of market structure on the timing and speed of consumer uptake. Section 2.2 similarly sets out our modelling approach to concentration. Section 3 discusses the data and provides some descriptive statistics and basic empirical relationships. Section 4 details our model specification and estimation strategy in the presence of endogenous market structure, before presenting our estimation results for the consumer uptake of internet banking. Section 5 does the same for the effect of internet banking on market structure. Section 6 presents robustness checks and section 7 concludes.

2. Internet banking and market structure

2.1. Consumer uptake of internet banking

Our approach is to adapt a contagion model of the consumer 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 begin by considering the two key elements of a contagion model: consumer information via positive feedback from current users; and individual susceptibility. Consumers can only take up internet banking once it becomes accessible, and do so only when they consider that the expected advantages outweigh their initial concerns about the usability and security of the technology.[3] Individuals differ in their risk aversion,

[2] Concentration ratios are taken from the European Central Bank database. Average concentration levels have also been rising. The average five firm concentration ratio for EU15 (i.e. the 15 member states of the EU prior to the 2004 enlargement) rose from 49% in 1997 to 54% in 2006, 55% in 2009 and 60% in 2014.

[3] Takieddine and Sun (2015) review 41 individual consumer survey studies of internet banking, which mainly use correlation and cluster analysis. These suggest a range of issues that influence people, including "perceived usefulness, perceived ease of use, perceived relative advantages, compatibility, observability, trialability,

other characteristics and general susceptibility to adopting a new way of conducting personal banking.

We assume that an individual's balance of these considerations is influenced by feedback from previous adopters on the advantages and potential concerns about internet banking. The greater the proportion of people who have already tried internet banking, the more likely each non-user is to have received positive and low variance feedback.[4] The existing proportion of users is the number of internet banking users, IB_{t-1}, divided by the maximum number of potential users M_{t-1}.

Holding such feedback constant, individual susceptibility may be influenced by income, education and demographic factors, and by bank investments in marketing and in the quality of online experience and security. Previous bank investments in branch networks also affect the availability and opportunity cost of using a close substitute for internet banking, and there may be different prices associated with internet banking. We summarise these factors influencing the probability of individual consumer adoption (conditional on having had sufficient positive feedback) as the susceptibility parameter, β. β is typically referred to as the speed of diffusion. A central focus of this paper is to understand how market structure affects β.

The probability of an individual taking up internet banking in each period is individual susceptibility times the proportion of people who have already tried internet banking, $\left[\frac{IB_{t-1}}{M_{t-1}}\right]$. The expected number of new users in each period is this probability multiplied by the number of individuals who could potentially sign up but who have not yet done so, $[M_t - IB_{t-1}]$. These considerations lead us to the classic contagion model of diffusion for a new service whose users continue to use it once they have started:

$$IB_t - IB_{t-1} = \beta \left[\frac{IB_{t-1}}{M_{t-1}}\right][M_t - IB_{t-1}]$$

Integrating the continuous time version of this diffusion process generates the standard logistic function first used by Griliches (1957),

$$\frac{IB_t}{M_t} = \frac{1}{1 + exp(-[\alpha + \beta t])}$$

---

income, education, age, gender, and marital status… security, privacy, self-efficacy, computer experience, Internet experience, complexity, lack of awareness, lack of knowledge, trust, risk, legal support, governmental support, perceived credibility, availability, fees and charges" [p. 362]. More recently, Laukkanen (2016) reports the results of a postal survey of consumer attitudes to internet banking in Finland, finding that age and expectations about the value of the service and preferences for human contact affect consumer uptake.

[4] For example, each person may influenced by a given number of contacts. If these contacts are a random sample of the population of potential users, it is the share of uptake that matters. We assume that average feedback is positive because otherwise usage would fall.

This equation shows how consumer uptake evolves over time. α shifts the logistic curve horizontally and is derived from the constant of integration: $\alpha = \log\left(\frac{IB_0}{M_0}\right) - \log\left(1 - \frac{IB_0}{M_0}\right)$. It locates how far the diffusion process has evolved by time $t = 0$, and is known as the timing parameter. Its economic interpretation is that if we compare two countries at the start of our observation period, for similar β the one which started internet banking earlier will have a higher $\frac{IB_0}{M_0}$ and so a higher α.

The above discussion suggests a range of ways consumer uptake may be influenced by market structure. These 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 approach is to bypass these proximate influences, which anyway are almost impossible to measure at the market level, and estimate how market structure affects α (i.e. timing) and β (i.e. speed). In this specific respect, our approach can be seen as a reduced form approach.

Considering α, the early introduction of internet banking may be facilitated by the immediate scale possibilities of converting existing customers to the internet service in a concentrated market. A concentrated banking sector might also provide an incentive for new entrants who could not previously enter without a huge investment in a branch network, although the protection of an oligopolistic position may similarly encourage incumbent banks to pre-empt such entry. Once introduced, the speed of consumer uptake, β, will depend on the ability and incentive for banks to market their internet services. Inasmuch as high concentration is associated with less competition, this may then reduce the speed of uptake as compared with a diversity of banks with different internet banking business models and creative ideas.

We see the net balance of effects as very much an empirical issue because there are no clear-cut theoretical predictions about market structure and consumer uptake, not least because market structure is an ambiguous indicator of the intensity of competition. Nevertheless, for the purpose of clarity and future reference, we set out our expected balance of theoretical effects as summary hypotheses.

Hypothesis 1: Internet banking will be introduced earlier in more concentrated markets, but the speed of diffusion will be slower (i.e. facilitating catch-up).

Some European countries have much more regionalised 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, Italy). As shown in section 3, these countries tend to have lower national concentration than countries of similar size but with nationally integrated banks (e.g. France). Hypothesis 1 should ideally be tested at the level of the competitively relevant market, which may be the region in some countries or national in others. In the absence of consistent data at the level of

relevant geographic markets, we use a measure of pre-internet banking regionalisation as a control.

Internet banking provides new opportunities for integrating banking markets that had become ossified by a previous reliance on local branch banking. Banking brands in neighbouring regions also provide a pool of entrants with established reputations and a greater number of consumers to attract from outside their home territories. Unless regional loyalties remain strong, such banks may also more rapidly provide consumer reassurance in taking up internet banking than would de novo entrants. After controlling for national concentration, this suggests the following.

Hypothesis 2: Internet banking will be introduced earlier in more regionalised markets, and the speed of diffusion will be faster than in previously integrated markets.

Investments in a branch network will have been endogenously determined by geography and market structure, so care is needed in estimating and interpreting empirical relationships. Even taking this into account, there are no clear predictions. A dense existing branch network provides a substitute for internet banking from the consumer perspective. This may slow the speed of consumer uptake, although it is possible that consumers are more comfortable experimenting with internet banking if there is a local branch to provide advice and cash services (i.e. internet and branches could be complements, at least initially). From the perspective of the banks with dense existing branch networks, they may adopt an internet service earlier to exploit such complementarities, or later if they believe their customers are already catered for.

We are aware of no previous econometric research into the relationship between consumer uptake of internet banking and market structure over time. Nickerson and Sullivan (2003) and Sullivan and Wang (2013) investigate the timing of the initial adoption of internet banking technology, but this is adoption by US *banks* (i.e. by firms as opposed to by consumers) across US regions. In terms of our consumer uptake model, their results may be relevant to α, but not to β. 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. Takieddine and Sun (2015) consider consumer *usage* across 33 European countries but only as a cross-section in 2013 so they cannot distinguish speed from location effects, and they do not consider the supply side.[5]

[5] 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 is of relevance to our later assumption that the maximum consumer uptake is determined by internet access. In particular, they found that per capita income, internet speed and security were associated with internet access, but not directly with internet banking. Only internet access was significantly associated with internet banking, and this relationship was strong and substantial. They further find a significant difference between north and west European countries on the one hand, and those in the south and east of Europe, but this difference applies to both internet banking and internet access so it is not possible to appraise whether the latter fully explains the former.

2.2. Impact of internet banking on the evolution of market structure

Our approach to modelling endogenous market structure builds on a reduced form relationship between concentration and market size. The slope and position of this relationship depends on the toughness of price competition.[6] For any given toughness of price competition and degree of economies of scale, and in the absence of other substantial entry barriers, a sufficiently larger market facilitates long-run entry. In most markets, this results in a negative relationship between concentration and market size in free entry equilibrium. 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. Complementing this concentration relationship, for a given toughness of competition, prices are decreasing in market size as entry fragments market structure. Horizontal product differentiation moderates these effects by reducing the toughness of price competition, but does not change the basic relationships.

If consumer valuations of a product or service increase with quality-enhancing investments that benefit all consumers (e.g. denser branch networks, or better online interface or security), and if horizontal differentiation is limited, competition can be channelled into the escalation of endogenous sunk costs. Such investments increase the degree of economies of scale, often without raising marginal costs, and an increase in market size may then have less effect on concentration (and prices) than on enhanced quality.[7] 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.[8] Additionally, it is sometimes possible to develop direct measures of quality enhancing investment to bring out the quality competition mechanism.[9]

Although the relationship between concentration and market size reveals significant information about price and quality competition, it must be expected to be noisy, especially when applied to international cross-sections and time series with macroeconomic volatility and changing technologies. For example, historical and cultural features may have affected the regional fragmentation of markets, regulation may have controlled entry and the toughness of competition, and entry barriers, subsidies, mergers and other factors mean

[6] e.g. Bertrand or Cournot. See, for example, Shaked and Sutton (1987), Sutton (1991), Bresnahan & Reiss (1991) and Berry (1992).

[7] 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.

[8] See Sutton (1991, 1998). Sutton (2007) reviews the literature.

[9] For example, Dick (2007), Ellickson (2007), Berry and Waldfogel (2010).

markets are rarely in long-run competitive equilibrium. In particular, banking markets have long had a rich diversity in Europe, despite attempts economically to integrate the member states of the EU.[10]

Dick (2007) was the first to apply the concentration-market size approach to banking. She uses a cross-section of US MSAs for market definition and finds that the lower bound to concentration is insensitive to market size. She proceeds to provide evidence that banks invest more in local branch networks (service quality) when the local population is larger. The importance of the branch network, at least pre-internet banking, has been found in other work. Cohen and Mazzeo (2010) provide results that are consistent with a potential entry-deterring effect of bank branch investments. Kim and Vale (2001) provide Norwegian evidence that the relative size of a bank's branch network is associated with its market share, but has no external effect on overall market demand. Temesvary (2015) found that Hungarian banks could charge a premium on loan interest rates in the presence of relative branch network dominance.[11]

Internet banking requires a significant investment in IT, security and the consumer interface. However, four technological characteristics stand in contrast to investments in branch banking: a) it is much cheaper to set up an internet interface than a branch; b) internet service has much lower operating costs than branch service; b) there may be limits to the extent that quality can continue to be enhanced by further internet investment while opening another branch (pre-internet banking) always increases convenience for at least some consumers; and c) the internet is a non-local technology, which can reach potential consumers outside a bank's home region ('extended geographic reach'). We expect that these technological characteristics could have a substantial effect on bank market structure as consumers take up internet banking, though it is likely to take an extended period of time for market structures to adjust. The combination of lower and less escalating fixed costs network ('lower sunk costs') suggests Hypothesis 3.

Hypothesis 3: Internet banking will reduce concentration and result in a more negative relationship between concentration and market size.

Regionalised countries may appear less concentrated at the national level, but if consumers mostly use a regional bank, concentration at the competitively relevant regional market level may be much higher. Put another way, a large 'national market' may actually be an aggregation of small regional markets. Previous research using single country data has not

[10] Shaked and Sutton (1987) and Sutton (1991) develop a bounds approach to retain focus on the underlying regularities without having to take full account of the presence of numerous institutional details that make each country unique. For an application of this approach, see Lyons et al (2001) which compares European manufacturing markets with and without endogenous sunk costs by separating those with and without high advertising or R&D.

[11] See also Cesari et al (2002) who use a sample of banks across nine European countries, and Cohen and Mazzeo (2010) for an analysis of endogenous branching.

had to address different degrees of regionalisation across countries. However, there was an illuminating experiment in branch network integration in the USA in the 1990s. The 1994 Riegle-Neal Act legislated for the full removal of previously tight geographic restrictions and so allowed nationwide branching. Dick (2006) finds that banks responded actively and concentration at the multi-state region level increased. However, most importantly for competition, this deregulation left the structure of metropolitan markets almost unchanged, each with two or three dominant banks controlling over half of market deposits. She argues that consumers benefitted from improved service and lower spreads.

The 'extended geographic reach' of internet banking suggests Hypothesis 4.

Hypothesis 4: Internet banking will result in national concentration in regionalised markets converging towards levels associated with nationally integrated markets.

Finally, complementary to these hypotheses, we expect dense branch networks to be associated with high concentration, but that concentration is more likely to decline with internet banking in such markets as the previous advantages of branching are eroded.

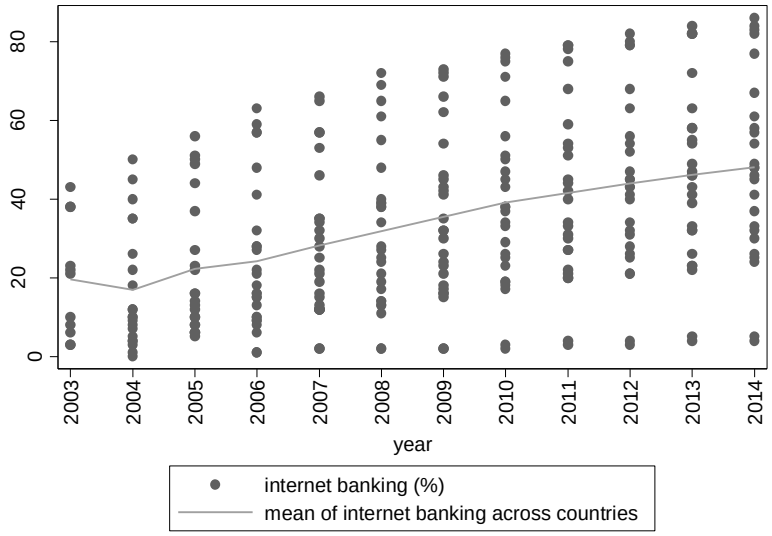
We are aware of no previous research on the effects of internet banking (or any other digital product) on either concentration or regionalisation.

3. Data and stylised facts on consumer uptake of internet banking and on market structure

An annual Eurostat survey since 2003 has reported the percentage of surveyed individuals by EU Member State who have used internet banking in the past three months.[12] We use this as our measure of the consumer uptake of internet banking. Figure 1 summarises the range of 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 curve.

[12] 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 18 Dec 2018). Internet banking includes electronic transactions with a bank for payment, transfers, etc. or for looking up account information. Further details regarding the survey questionnaire can be found in the Methodological Manual of Eurostat's Digital economy and society database : https://circabc.europa.eu/faces/jsp/extension/wai/navigation/container.jsp; (last checked 18 Dec 2018).

Figure 1. Consumers using internet banking (by EU member state)



Source: Eurostat

Meaningful measures of market structure depend on the appropriate definition of both product and geographic markets. Ideally, we would want product measures based on retail banking and with a geographic dimension determined by the location of consumers and the choice of banks they see as reasonable alternatives. Financial regulation essentially restricts that choice of banks to those operating in the same country (though many will also have international operations). We start with measures of national market structure before considering regionalisation of banking within countries.

The most widely used data on banking for the period we study was collected by BankScope, which included the global activities of domestically owned banks disaggregated by type of activity (i.e. by product dimension). However, this dataset could not be disaggregated by location of activities and so is problematic for geographic market definition. For example, a measure for France based on these data would include BNP Paribas's international activity and so make it appear to have a disproportionately large share of French retail banking compared to a French bank without an international presence, while a foreign bank operating in France would have no measured market share.[13]

Instead, we use data on market size and 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 which either (i) receive deposits or other repayable funds from the public and

[13] BankScope no longer exists. It has been replaced by Orbis BankFocus which does have information on unconsolidated subsidiaries but the time series only starts in 2013.

grant credit on their own account, or (ii) issue means of payment in the form of electronic money.[14] We call these 'banks' for short.

The ECB measures bank size by total assets. Importantly, total assets are calculated on a residence basis so 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 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.

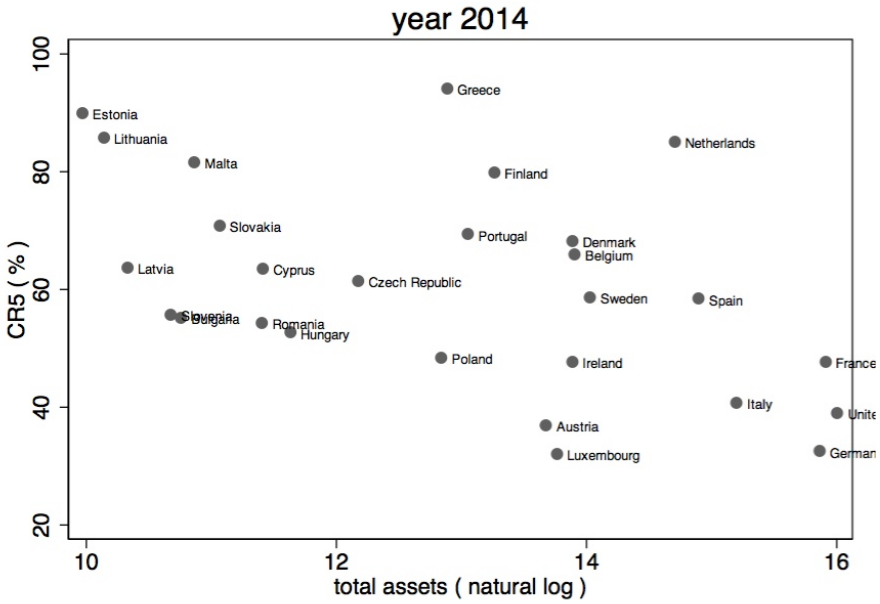
The ECB also measures market size by total assets and calculates two standard measures of national market structure: five-firm concentration ratio (C5) and Herfindahl index (HHI). We are able to construct a panel of concentration measures for the years 1997-2014 for the fifteen EU member states at the start of the period, and our panel increases to 27 countries from 2001 (i.e. including the new members who acceded in 2004).[15]

The relationship between the concentration ratio (C5) and market size in 2014 is shown in Figure 2. 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 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 excluded).

[14] 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 18 Dec 2018).

[15] The most recently acceded member Croatia is not included in the sample.

Figure 2 Bank Concentration and market size in 2014



Source: ECB

A distinctive feature of some European banking markets is that they are highly regionalised, with local banks historically serving their 'home' regions (e.g. German 'Landesbank's). Sector knowledge and inspection of lists of leading banks is strongly suggestive of international differences in the degree of regionalisation (e.g. retail banking in Italy is clearly more regionalised than in France) but there is no previously published objective measure available for our econometric analysis, so we had to create a new measure.

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 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. We proceeded by collecting information on each bank's headquarters location (including city and postcode) from the Banker database.[16] We include

[16] 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 19 Dec 2018) . 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.

all banks at the group level (bank holding companies) and use the postcode to identify the NUTS level 2 region in which each bank is headquartered.[17]

Consider a country with *K* regions. We require an index of regionalisation, *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 similar to a Herfindahl index of regional concentration of bank HQs, and the "one minus" converts this to an index of regionalisation. *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: a) there are two equal sized regions and the remaining $K - 2$ regions have no HQs, in which case $R = 0.5$, and b) there are four regions containing 40%, 30%, 20% and 10% shares of HQs, in which case $R = 0.7$.[18] It is straightforward that *R* satisfies the first desirable property. The second and third follow from a standard property of the Herfindahl that $\sum_{k=1}^{K} \left(\frac{S_k}{S}\right)^2 = \frac{1+v^2}{K}$, where *v* is the coefficient of variation.

Table 1 reports our index of regionalisation based on bank assets for each country in our dataset.[19] 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,

[17] 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

[18] Of course, the same *R* can come about from different distributions of HQs; e.g. one region with 68% and three with 11% each would give *R*=0.5.

[19] 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.

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 regionalised 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 regionalisation around a dominant financial capital.

Table 1: Regionalisation within countries

Country	*R index*
Austria	0.48
Belgium	0
Bulgaria	0
Cyprus	0
Czech Republic	0
Denmark	0.20
Estonia	0
Finland	0.07
France	0
Germany	0.69
Greece	0.00
Hungary	0
Ireland	0
Italy	0.70
Latvia	0
Lithuania	0
Luxembourg	0
Malta	0
Netherlands	0.50
Poland	0.04
Portugal	0.53
Romania	0.26
Slovakia	0
Slovenia	0.38
Spain	0.68
Sweden	0.00
United Kingdom	0.40

Source: authors' calculations. Full descriptions of all variables, their measurement and sources are given in Appendix 1 (Table A1 and A2).

Descriptive statistics and definitions for all variables used in this paper can be found in Appendix 1.

4. Econometric estimation of the consumer uptake of internet banking

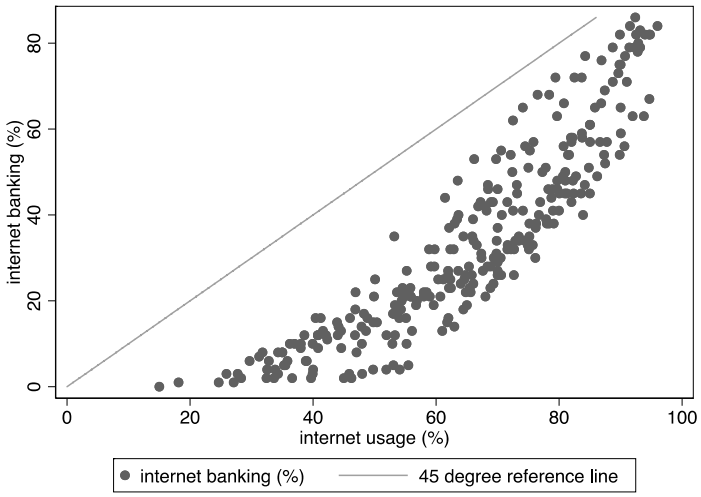
4.1. Model specification

The logistic diffusion function developed in section 2 can be written as

$$IB_{it} = \frac{M_{it}}{1 + exp\big(-(a_{it} + b_{it}t)\big)}$$

where the *i* subscripts identify different countries. We assume that the maximum possible uptake (saturation level) is a proportion, λ, of current internet usage, IU_{it}, so $M_{it} = \lambda * IU_{it}$. This is consistent with the data on internet banking and internet usage for all country-year pairs plotted in Figure 3, which shows that *IB* is bounded by *IU*.[20]

Figure 3 Internet Banking vs. Internet Usage



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

$$\log\left(\frac{IB_{it}}{\lambda * IU_{it} - IB_{it}}\right) = a_{it} + b_{it} * t + u_{it} \qquad (1)$$

For the reasons set out in section 2, both the location and speed parameters are allowed to be affected by market structure, including *concentration* (C_{it}) and *regionalisation* (R_i). Note that there is insufficient intertemporal variation in our regionalisation measure to include a time dimension. Both diffusion parameters are also allowed to vary by *branch density* (B_{it}) and a vector of controls (X_{it}). The latter include *education* (*E*) measured as percentage of population aged 25-74 who have obtained tertiary education (Eurostat) and *GDP per capita* (*G*) from Eurostat. *t* is the time trend and u_{it} is the error term.

[20] This is also consistent with Takieddine and Sun (2015), as discussed earlier.

4.2. Estimation strategy

To allow for unobserved heterogeneity in equation (1), we follow the correlated random effects approach proposed in Wooldridge (forthcoming) which is an extension of the Chamberlain-Mundlak approach for balanced panel data to unbalanced cases. Under this approach, unobserved heterogeneity is allowed be correlated to the history of selection and the selected covariates.

Let $\{s_{it}: t = 1, \dots, T\}$ be a sequence of "selection indicators" where $s_{it} = 1$ if and only if the full set of data for unit *i* in period *t* is observed; otherwise $s_{it} = 0$. This characterises a typical unbalanced panel data set. The number of periods available for unit i is: $T_i = \sum_{r=1}^{T} s_{ir}$. It is assumed that observed data in any time period cannot be systematically related to the idiosyncratic errors, but it can be correlated to observed covariates and time invariant unobserved heterogeneity.

Define $\overline{X_\iota} = T_i^{-1} \sum_{r=1}^{T} s_{ir} X_{ir}$, which are the averages of covariates over time where we observe a full set of data on the dependent and independent variables. Note that these time averages differ across countries as different countries have a different number of observed time periods. The average of time trend $\overline{t_\iota}$ is also included for each country and varies depending on the number of time periods available for each. To control for unobserved heterogeneity equation (1) becomes:

$$\log\left(\frac{IB_{it}}{\lambda * IU_{it} - IB_{it}}\right) = a_0 + a_1 C_{it} + a_2 B_{it} + a_3 R_i + a_4 X_{it} + b_0 \mathrm{t} + b_1 C_{it} * t + b_2 B_{it} * t + b_3 R_i * t + b_4 X_{it} * t + c_i + v_{it} \quad (2)$$

where the unobserved heterogeneity is assumed to be $c_i = \varphi + a_{10} * \overline{X_\iota} + a_{11} * \overline{t_\iota} + a_i$.

If the panel data are balanced, Mundlak (1978) shows that the FE estimator can be computed using pooled OLS from the original data with the time averages of the covariates added as additional explanatory variables. Wooldridge (forthcoming) extends this result to the unbalanced cases. Estimating equation (2) with additional explanatory variables in c_i (i.e. the time averages of time varying covariates), using pooled OLS with all $s_{it} = 1$ observations, the coefficient vector on the time varying covariates is the same as the that obtained from the FE estimator.

As discussed in section 2 of this paper, C_{it} and B_{it} must be viewed as potentially endogenous. We follow the Control Function 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*:

$$Y_{it} = \gamma_0 + \gamma_1 Z_{it} + \gamma_2 \overline{Z_\iota} + r_{it} \quad (3)$$

and obtain the OLS residual $\hat{r}_{it}$. $Z_{it} = (Z_{it1}, Z_{it2})$ where Z_{it1} are the exogenous variables in (2) including *t, R* and *X*, and Z_{it2} are instruments for *C* or *B* which are excluded from (2).

We have already argued that concentration is negatively related to market size and that this is a very strong empirical relationship in the literature. It is also confirmed for our data in the following sub-section. Furthermore, the local and essential nature of the retail banking product means that market size depends on the size of population, so we use the natural log of population as our instrument for *C*. Also, 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 consumer 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. The first stage reduced form estimates reported in Appendix 2 confirm the strong empirical significance of these instruments.

We thus estimate the following equation:

$$\log\left(\frac{IB_{it}}{\lambda * IU_{it} - IB_{it}}\right) = a_0 + a_1 C_{it} + a_2 B_{it} + a_3 R_i + a_4 X_{it} + b_0 \mathrm{t} + b_1 C_{it} * t + b_2 B_{it} * t + b_3 R_i * t + b_4 X_{it} * t + d_1 \hat{r}_{it}^{C} + d_2 \hat{r}_{it}^{B} + c_i + v_{it} \quad (4)$$

where $\hat{r}_{it}^{C}$ is control function (OLS residual) obtained from the reduced form equation for *C* and $\hat{r}_{it}^{B}$ is the control function obtained from the reduced form equation for *B*. Since the estimation of equation (4) 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 below are based on bootstrapped standard errors.

4.3. Results

First, consider econometric identification. Our first stage regression results are reported in the first two columns of Table A3 in Appendix 2. These 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. As shown towards the bottom of Table 2, the control function for bank concentration $\hat{r}_{it}^{C}$ is significant, which confirms the value of investigating the endogeneity of concentration more deeply in the next section. Although the endogeneity of branch density is not confirmed, as $\hat{r}_{it}^{B}$ is insignificant, we report our results including this control function. Excluding the latter makes no substantive difference to the results in Table 2.

Turning to our main results, we start with the complete specification "Spec 1" as shown in equation (4) 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 "Spec 2". Both are reported in table 2, and for both concentration measures. In discussing our results, we focus on Spec 2. There is very little difference in our *CR5* and *HHI* specifications, so we focus on the former. We proceed by discussing significance before returning to quantitative effects.

Table 2 Estimation Results for the Diffusion of IB

	C measured by *CR5*		*C* measured by *HHI*	
	Spec 1	**Spec 2**	Spec 1	**Spec 2**
C	0.023 (0.0108)**	0.033 (0.0050)***	7.629 (2.4122)***	8.034 (1.3786)***
R	1.712 (0.6135)***	1.829 (0.6352)***	1.232 (0.6039)**	1.213 (0.6017)**
B	0.071 (0.1314)	0.051 (0.1159)	0.208 (0.1356)	0.191 (0.1112)
E	0.012 (0.0252)	-0.020 (0.0154)	-0.007 (0.0256)	-0.021 (0.0173)
G	0.925 (1.1393)	1.331 (0.9806)	1.458 (1.1269)	1.614 (0.9550)*
t	-0.119 (0.4093)	0.105 (0.0469)**	0.099 (0.4080)	0.101 (0.0461)**
*C*t*	0.001 (0.0007)		0.033 (0.1583)	
*R*t*	-0.094 (0.0471)**	-0.101 (0.0464)**	-0.085 (0.0464)*	-0.082 (0.0463)*
*B*t*	-0.016 (0.0091)*	-0.014 (0.0080)*	-0.016 (0.0093)*	-0.015 (0.0075)*
*E*t*	-0.002 (0.0015)		-0.001 (0.0017)	
*G*t*	0.022 (0.0397)		0.002 (0.0417)	
constant	-14.710 (5.3935)***	-17.054 (1.8887)***	-11.381 (5.1049)**	-11.281 (1.5239)***
$\hat{r}_{it}^{C}$	-0.011 (0.0063)*	-0.011 (0.0053)**	-2.947 (1.6121)*	-2.902 (1.5512)*
$\hat{r}_{it}^{B}$	0.086 (0.1098)	0.079 (0.1207)	-0.025 (0.1129)	-0.020 (0.1099)
F test to compare spec 1 and spec 2	0.77		0.14	
Adjusted R^2	0.82	0.82	0.82	0.82
No. of Obs.	286	286	286	286

***, **, and *indicate statistical significance at 1%, 5% and 10% respectively; standard errors are in brackets.

Market structure has a highly significant and nuanced effect on consumer uptake of internet banking. National concentration has a significant positive effect on the 'location' parameter, which can be interpreted as bringing forward the start of the diffusion process. This supports the first part of Hypothesis 1, but concentration has no significant effect on the subsequent speed of adoption, so the second part finds no support. Our second dimension of market structure, regionalisation, also brings forward the start of the process, but it then has the effect of reducing the speed of adoption. Thus, we find support for the first part of Hypothesis 2, but quite the opposite for the second part – unlike national concentration, regionalisation has a negative impact on the speed of adoption. A possible explanation is that regional banks have an early incentive to take customers from other regions with this new digitial service, but they find it difficult to break down consumer loyalty to their regional banks.

For a given level of national concentration, more regionalised markets imply increased local concentration, so the location parameter effects consistently suggest bank investment in internet products was brought forward in concentrated markets. This is consistent with either the importance of established incumbency advantages such as a large customer base, or consumer confidence in major brands being important when considering the first adoption of an unknown, possibly risky, service.

Table 3 combines both location and speed effects to show the accumulated impact over time. It reveals that the early positive impact of regionalisation was eliminated by the end of our period.

Table 3 Coefficient of R varying with t

t	CE	SE
7 (2003)	1.120***	0.324
8 (2004)	1.019***	0.282
9 (2005)	0.918***	0.241
10 (2006)	0.816***	0.203
11 (2007)	0.715***	0.179
12 (2008)	0.614***	0.143
13 (2009)	0.513***	0.129
14 (2010)	0.412***	0.130
15 (2011)	0.310**	0.147
16 (2012)	0.209	0.175
17 (2013)	0.108	0.209
18 (2014)	0.007	0.248

Having proved its statistical significance, the following figures illustrate the substantial quantitative effect of market structure, and the combined effects of concentration and regionalisation. Figure 3 uses 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 regionalisation (R = 0 and R = 0.7). First consider Figure 3a. Panel g shows how uptake in a non-regional, low concentration market proceeds over the period, increasing from 4% to 39% over 11 years. Panels h and i show how uptake increases with concentration (low to medium), and increases again from medium to high. By 2015, high concentration markets could expect twice as much internet banking as low concentration markets. A similar effect of concentration in regionalised markets is seen by comparing across panels a, b and c.

The effects of regionalisation are most clearly seen in Figure 3b. Regional markets are shown in the first row, and the second row compares them with non-regionalised markets. For all levels of concentration, it is clear that regionalised markets had an early start on internet banking, but by 2015 this advantage had completely disappeared.

Figure 3a: Predicted internet banking uptake: the impact of concentration given each level of regionalisation

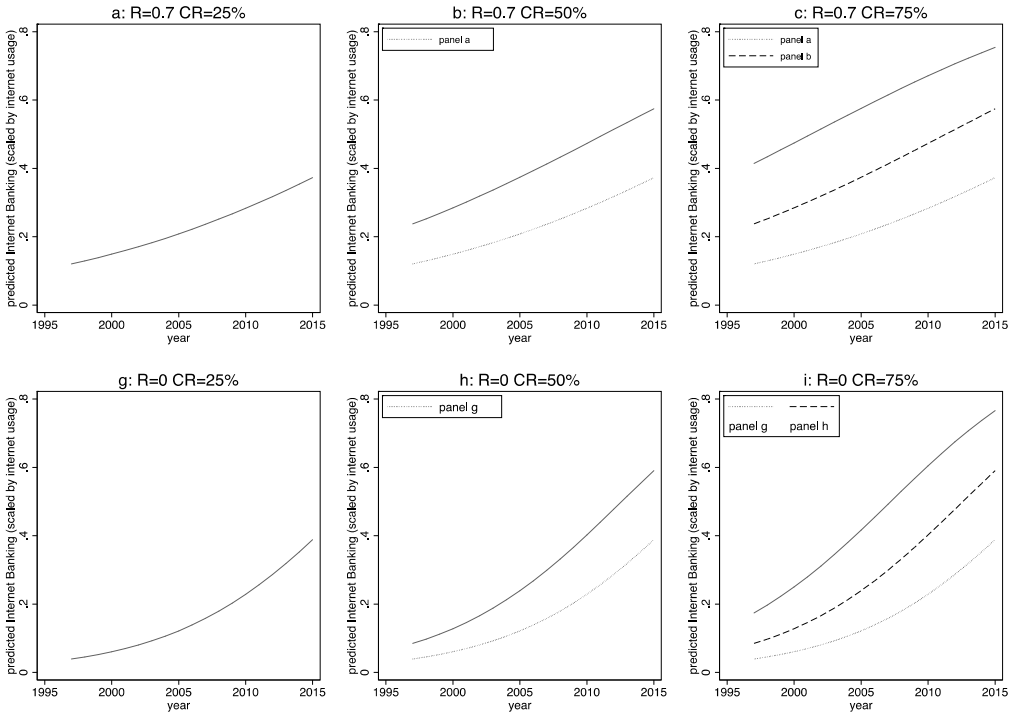
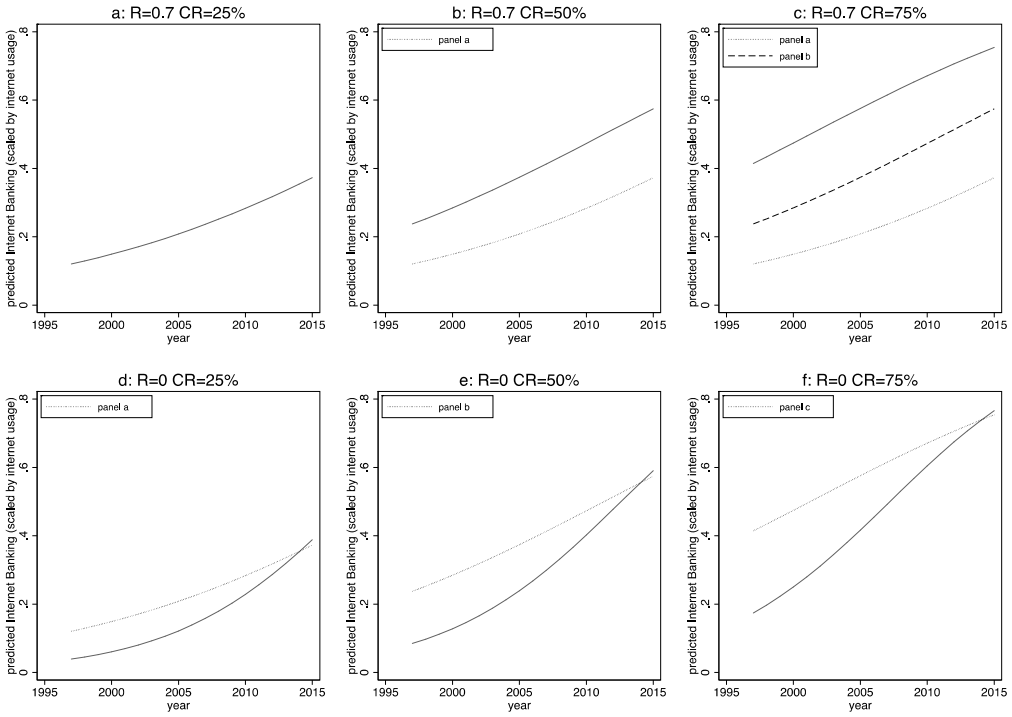


Figure 3b: Predicted internet banking uptake: the impact of regionalisation given each level of concentration



Looking only at Table 2, there is some weakly significant evidence that high branch density, *B*, slows down the speed of consumer uptake, though the sign of the insignificant location effect hints that it may bring forward the introduction of the service. In Table 3, we combine the two effects in each year and consider the cumulative significance of the effects. This clarifies that, once the diffusion process is under way, a dense branch network has a significantly negative effect on consumer uptake, which is consistent with a substitution effect – consumers are ultimately slower to take up internet banking if there is dense branch network nearby.

Table 4 Coefficient of B varying with t

t	CE	SE
7 (2003)	-0.048	0.067
8 (2004)	-0.062	0.061
9 (2005)	-0.076	0.055
10 (2006)	-0.091*	0.051
11 (2007)	-0.105**	0.047
12 (2008)	-0.119***	0.044
13 (2009)	-0.133***	0.042
14 (2010)	-0.147***	0.042
15 (2011)	-0.161***	0.043
16 (2012)	-0.175***	0.046
17 (2013)	-0.190***	0.050
18 (2014)	-0.204***	0.055

Of our remaining variables, the independent time trend is positive, as expected in any diffusion model. Our other control variables are only marginally significant if at all. GDP per capita has a positive sign and education has a negative sign on the location parameters, and neither has any effect on the speed of adoption.

5. Endogenous national concentration

5.1. Specification

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

national level while some markets are regionalised, we allow both coefficients to vary with regionalisation, *R*.

As argued in section 2, high branch density makes it more difficult for an entrant to attract customers by finding a niche market, especially pre-internet banking. Branch proliferation may therefore be used strategically by incumbent banks and the incentive to do so could be influenced by market structure. As in section 4, we treat branch density, *B*, as an endogenous variable with population density, *P*, as the identifying instrument on the grounds that banks have a greater incentive to invest in branches where population is dense.

Adding a time trend, *t*, gives

$$\log(\frac{C_{it}}{100-C_{it}}) = \theta_0 + \theta_1 t + \theta_2 \frac{1}{\ln S_{it}} + \theta_3 R_i + \theta_4 R_i * \frac{1}{\ln S_{it}} + \theta_5 B_{it} + u_{it} \qquad (5)$$

where C_{it} is the five-firm concentration ratio, R_i is our regionalisation index, B_{it} is branch density and S_{it} is total national banking assets.

Our principal aim in this section is to explore whether increasing consumer uptake of internet banking has had a de-concentrating effect, as proposed in Hypotheses 3 and 4. 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 cannot simply add *IB* to the right hand side of equation (5).

Instead, we test a model that allows the relationship in equation (5) 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 (5), 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 (5) as $\boldsymbol{\gamma X}_{it} + u_{it}$, we estimate equation (6):

$$\log\left(\frac{C_{it}}{100-C_{it}}\right) = \boldsymbol{\gamma X}_{it} + \boldsymbol{\delta} D_{it} \boldsymbol{X}_{it} + u_{it} \qquad (6)$$

5.2. Estimation strategy

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. In our data sample, *IB* is only available from year 2003. In year 2003, if the observation has *IB* below the threshold in 2003, 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 control function to be used for the second stage estimation.

As in the previous section, we apply the correlated random effects approach to control for unobserved heterogeneity and the control function 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. Both are highly significant in their respective first stage regressions (see Appendix 2). The control function enters equation (5) as $\delta_0 \hat{e}_{it}^S$ and $\delta_1 \hat{e}_{it}^B$, to control for endogeneity of *S* and *B* respectively, where $\hat{e}_{it}^S$ and $\hat{e}_{it}^B$ are estimated residuals from reduced forms of *S* and *B* respectively.[21]

We also need to control for a potential bias arising from the endogeneity of *D*, since *D* is constructed using *IB*. The following control function is adopted following Wooldridge (2015) using a first-stage probit regression to obtain a "generalised error" term defined as:

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

5.3. Results

Table 5 reports the results for a threshold of $\widetilde{IB} = 25\%$. Results for $\widetilde{IB} = 30\%$ are reported in section 6 alongside other robustness checks. The alternative thresholds do not materially change our estimates.

First, consider the significance of the control function errors reported at the bottom of Table 5. $\hat{e}_{it}^B$ 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 control function error term for the internet banking threshold is not significant.

Next, consider our "pre-internet banking" results (i.e. for *D*=0). All variables are highly significant. As expected, concentration falls with both market size and regionalisation, and regionalisation 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 'lower sunk cost' effect in Hypothesis 3 is supported by the statistical and quantitative significance of *D* and its interaction with market size – internet banking has reduced national concentration and made it more sensitive to market size. The effect of regionalisation is more complex – internet banking has reduced concentration and reversed its sensitivity to national size. We return to Hypothesis 4 after introducing the following figures.

[21] 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* and *S* equations here. The first stage regression results are reported in Appendix 2.

Table 5: Estimation results for national concentration in banking

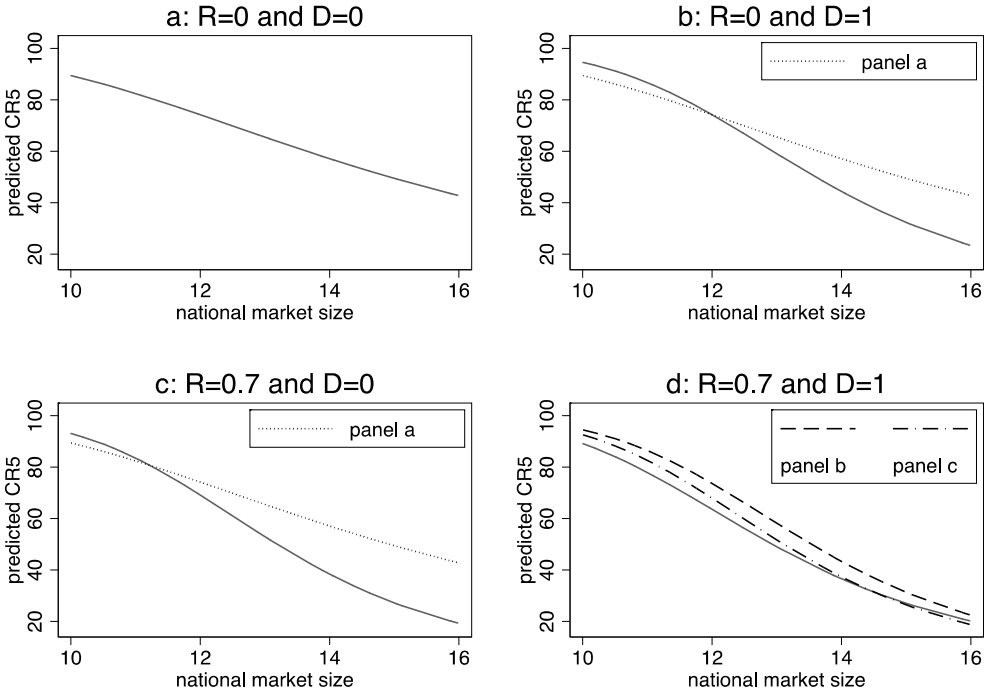
Dependent variable: $\log\left(\frac{C_{it}}{100-C_{it}}\right)$	*Estimated coefficients*
t	0.102 (0.0113)***
$\frac{1}{\ln S_{it}}$	64.677 (4.957)***
R_i	-5.462 (1.3282)***
B_{it}	0.331 (0.0547)***
$R_i * \frac{1}{\ln S_{it}}$	61.311 (18.5850)***
D	-3.252 (0.4631)***
$D*t$	-0.080 (0.0154)***
$D * \frac{1}{\ln S_{it}}$	43.069 (5.7795)***
$D * R_i$	6.859 (3.0005)**
$D * B_{it}$	-0.167 (0.0565)***
$D * R_i * \frac{1}{\ln S_{it}}$	-86.829 (43.7195)**
$\hat{e}_{it}^{S}$	-0.143 (0.0748)**
$\hat{e}_{it}^{B}$	-0.459 (0.0876)***
$\hat{e}_{it}^{D}$	0.141 (0.1238)
Constant	9.390 (1.6747)***
Adjusted R^2	0.73
No of Obs.	364

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

Standard errors are in the brackets.

In order to understand the quantitative effects of internet banking, it is helpful to consider Figure 4. We use Table 5 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-regionalised markets (*R*=0) and regionalised markets (*R*=0.7). The latter value is a natural choice since it applies to Germany, Italy and Spain.

Figure 4: Predicted market concentration varying with different levels of regionalisation and maturity of internet banking



Note: national market size is measured by the natural log of total assets (instrumented by population)

Comparing panels a and c in Figure 4, it is clear that pre-internet banking, regionalisation reduced national concentration particularly substantially in large national markets. In the largest European countries, the predicted 5-firm concentration ratio was around 40% in a non-regionalised market, but only around 20% under strongly regionalised banking. Note that this says little about the level of competition because, in the latter case, the competitively relevant regional market would be much more concentrated than this predicted national level.

Comparing panels a and b (i.e. non-regionalised markets), internet banking has had little effect on predicted concentration in smaller countries, but has substantially reduced predicted concentration in large markets. The quantitative effect is not much smaller than the effect of regionalisation just discussed. In fact, panel d shows that the 'extended geographic reach' effect in previously regionalised markets has more or less cancelled the effect of regionalisation in the sense that there is convergence with previously integrated markets post internet banking. This supports Hypothesis 4. It is consistent with the view that once a strong uptake of internet banking has been established, this allows the erosion of regional loyalties, with some regional (or national) banks entering across different

regional markets, gaining market shares in other regions, and others being forced to consolidate or exit.

6. Robustness checks

We considered a number of robustness checks relating to an additional control variable, maximum internet banking take-up, the cut-off for internet banking in the concentration estimation, a possible outlier country, an additional instrument, and the impact of the financial crisis. In each case, we re-ran our estimations using alternative assumptions and compared the results with those presented in the main text. Detailed results are available in Appendix 3 and 4.

It might be expected that the uptake of internet banking could be affected by demographic factors in addition to our reported controls. On closer reflection, this is not so obvious because our measure of the uptake of internet banking is scaled by internet usage which may be similarly affected. Nevertheless, our robustness checks included the use median age of the population in each country over time as an additional demographic variable in our internet banking model. This variable varies quite significantly across countries and over time (as shown in Figure A1 in Appendix 3). However, median age is never significant and its inclusion as an additional control variable does not change the main pattern described above (see Table A4 in Appendix 3).

Recall that we assume the maximum possible uptake (saturation level) is a proportion, λ, of current internet usage, IU_{it}, so $M_{it} = \lambda * IU_{it}$. The results reported in section 3c 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 maximum ratio of internet banking to internet usage is 0.93. There was no substantial difference from our reported results.

We similarly explored different threshold values for construction of the dummy variable *IB* used in the concentration estimations. Table A5 and Figure A2 in Appendix 4 report the results using a different cut-off point: *D* = 1 if *IB > 30%* and 0 otherwise. It is clear that the results are essentially the same as those reported in section 5.

We have already noted that Luxembourg might be an outlier in the concentration equation given its small population size and large total bank assets as measured by the ECB. If we exclude Luxembourg from the sample for estimating concentration, the only change of significance is that the control function for *D* becomes significant (see Table A6). Figure A3 confirms that the quantitative effects are similar to Figure 4 except that small and medium sized countries continue to have lower concentration than already integrated countries even when *D* = 1..

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. Table A7 and Figure A4 in Appendix 4 confirm that this also has no effect on our results in the second stage concentration regression.

Our final robustness check is 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. However, there was considerable variation in the extent to which the financial crisis hit different European banks and at what time, and this pattern was not closely correlated with internet banking uptake. We measure the extent of 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 recapitalisation and impaired asset relief relative to market size (measured by total assets). This variable is never significant in our concentration estimations. See detailed results in table A8 in Appendix 4.

7. Conclusion

Internet banking in Europe provides an important example of a digital service that is a convenient alternative to bricks-and-mortar service delivery. This convenience means that consumer welfare is enhanced if the digital product is introduced earlier and if the supporting bank investments in the digital interface, security and marketing are conducive to more rapid consumer uptake. Internet banking also permits us to investigate how digital service delivery in turn affects two dimensions of market structure: national concentration and regional integration.

On the demand side, our econometric model of consumer uptake shows that national concentration brings forward the introduction of the digital product. Furthermore, there is no evidence that apparently more competitively structured markets catch up, so this advantage persists. Our second dimension of market structure, regionalisation, also brings forward the start of the process, but it then significantly reduces the speed of consumer uptake so integrated markets catch up. A possible explanation is that regional banks have an early incentive to take customers from other regions by providing a digital service, but they find it slow to break down consumer loyalty to local-region banks.

On the supply side, digital provision has a very different cost structure and geographic reach to bricks-and-mortar service provision. This led us to examine the endogeneity of both concentration and regional integration. Once internet banking is adopted by consumers, we find a substantial de-concentrating effect in large, non-regionalised markets. This suggests that the 'lower sunk cost' technology of internet banking is introducing competition into banking markets by fragmenting national concentration.

In previously regionalised markets, we find a second, market-integration mechanism also at work. The 'extended geographic reach' of the digital product would normally be expected to increase national concentration as cross-regional competition drives out weaker banks or forces consolidation. Empirically, we find that the 'lower sunk cost' effect and the 'extended geographic reach' effect more or less cancel out in terms of national concentration. When compared with the 'lower sunk cost' effect in already integrated markets, this provides indirect evidence that the 'extended geographic reach' effect has already been an integrating force in previously regional markets.

Overall, there has been a substantial post internet banking convergence in the relationship between national concentration and market size in already-integrated and previously-regionalised markets. Our findings are consistent with internet banking having enhanced competition in both, despite little measured change in national concentration in previously regionalised markets. These supply-side competition benefits of digital service delivery should add long-term advantages to the more immediate demand-side benefits of a more convenient consumer service.

References:

Berry, S.T., 1992. Estimation of a Model of Entry in the Airline Industry. Econometrica, 60(4): 889-917

Berry, S., and Waldfogel, J., 2010. Product Quality and Market Size. The Journal of Industrial Economics LVIII.1: 1–31.

Bresnahan, T.F. and Reiss, P.C., 1991. Entry and Competition in Concentrated Markets, Journal of Political Economy 99: 977–1009.

Cerasi, V., Chizzolini, B. and Ivaldi,M., 2002. Branching and competition in the European banking industry. Applied Economics 34: 2213–2225.

Cetorelli, N., 2002. Entry and competition in highly concentrated banking markets. *Economic Perspectives* Q IV, 26, (4): 18-27

Cohen, A.M. and Mazzeo, M.J., 2007. Market Structure and Competition among Retail Depository Institutions, Review of Economics and Statistics 89(1):60-74.

Cohen, A.M. and Mazzeo, M.J., 2010. Investment strategies and market structure: an empirical analysis of bank branching decisions. Journal of Financial Services Research 38, 1–21.

Dick, A., 2006. Nationwide branching and its impact on market structure, quality, and bank performance. Journal of Business 79: 567–592.

Dick, A., 2007. Market size, service quality, and competition in banking. Journal of Money, Credit, and Banking 39: 49–81.

Ellickson, P., 2007. Does Sutton Apply to Supermarkets? RAND Journal of Economics 38: 43-59.

Genakos, C., Valletti,T. and Verboven. F., 2018. Evaluating market consolidation in mobile communications. Economic Policy, 33 (93): 45–100

Griliches, Z., 1957. Hybrid Corn: An Exploration in the Economics of Technological Change. Econometrica 25: 501–522.

Gruber, H., Verboven, F., 2001a. The Diffusion of Mobile Telecommunications Services in the European Union. European Economic Review 45, 577–588.

Gruber, H., Verboven, F., 2001b. The evolution of markets under entry and standards regulation – the case of global mobile telecommunications. International Journal of Industrial Organization 19, 1189–1212.

Hausman, Jerry and Leibtag, Ephraim. 2007. Consumer Benefits from Increased Competition in Shopping Outlets: Measuring the Effect of Wal-Mart. Journal of Applied Economics. 22: 1157–1177

Laukkanen, T., 2016. Consumer adoption versus rejection decisions in seemingly similar service innovations: The case of the Internet and mobile banking. Journal of Business Research 69, 2432–2439

Lyons, B., C. Matraves and P. Moffatt. 2001,'Industrial Concentration and Market Integration in the European Union' *Economica,* February, vol 68, #269, 1-26

Li, Y. and Lyons, B. 2012. Market structure, regulation and the speed of mobile network penetration. International Journal of Industrial Organization 30, 697-707

Mundlak, Y., 1978., On the Pooling of Time Series and Cross Section Data. *Econometrica* 46 (1), 69-85

Nickerson, D. and Sullivan, R., 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 WP 03-01.

Petrin, A. and Train, K., 2009. A control function approach to endogeneity in consumer choice models, Journal of Marketing Research XLVI.

Shaked, A. and Sutton, J., 1987. 'Product Differentiation and Industrial Structure,' Journal of Industrial Economics, 36.

Sullivan, R. and Wang, Z., 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.

Sutton, J., 2007. Market Structure: Theory and Evidence, in *The Handbook of Industrial Organization*, Volume 3, Armstrong, M. and Porter, R., eds. Elsevier, 2301-2368.

Takieddine,S. and Sun, J., 2015. Internet banking diffusion: A country-level analysis, Electronic Commerce Research and Applications, 14(5), 361-371

Temesvary, J., 2015. 'Dynamic branching and interest rate competition of commercial banks: Evidence from Hungary' International Journal of Industrial Organization, 43, 98–110

Wooldridge, J.M., 2015. Control Function Methods in Applied Econometrics. The Journal of Human Resources 50 (2).

Wooldridge, J. M., forthcoming. Correlated random effects models with unbalanced panels. Journal of Econometrics: https://doi.org/10.1016/j.jeconom.2018.12.010

Appendices

Appendix 1

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

	mean	min	max	sd	cv	N	definition	source
C5 (%)	58.91	21.63	98.55	17.28	0.29	286	Concentration ratio of the largest 5 banks	Structural financial indicators database: ECB Statistical Data Warehouse: https://sdw.ecb.europa.eu/browse.do?node=9691551
HHI	0.11	0.02	0.40	0.07	0.67	286	Herfindahl-Hirschman Index	Structural financial indicators database: ECB Statistical Data Warehouse
R	0.19	0.00	0.70	0.25	1.34	286	Bank Regionalisation	Calculated using bank headquarter information from Banker Database
B (natural log)	-3.33	-5.86	-0.99	1.11	-0.33	286	Branch density: number of branches per km^2	Structural financial indicators database: ECB Statistical Data Warehouse

IB (%)	34.18	0.01	86.00	22.37	0.65	286	Internet usage: the proportion of individuals who used the Internet in the last three months.	Eurostat's digital economy and society database
IU (%)	-3.33	15.00	95.99	17.94	0.28	286	Internet usage: the proportion of individuals who used the Internet in the last three months.	World Telecommunication Union's statistics: https://www.itu.int/en/ITU-D/Statistics/Pages/datacollection/default.aspx#questionnaires
PD (natural log)	11.54	9.64	14.11	0.92	0.08	286	Population density: population per km^2	Eurostat: General and regional statistics
Population (natural log)	15.84	12.91	18.23	1.44	0.09	286	population	Eurostat: General and regional statistics
G (natural log)	10.05	8.92	11.20	0.40	0.04	286	GDP per capita	Eurostat: Economy and finance
E (%)	26.20	10.60	45.90	8.33	0.32	286	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
C5 (%)	57.40	18.95	94.75	18.05	0.31	364	Concentration ratio of the largest 5 banks	Structural financial indicators database: ECB Statistical Data Warehouse
HHI	13.05	9.05	16.12	1.76	0.13	364	Herfindahl-Hirschman Index	Structural financial indicators database: ECB Statistical Data Warehouse
R	0.21	0.00	0.70	0.26	1.25	364	Bank Regionalisation	Calculated using bank headquarter information from Banker Database
B (natural log)	-3.22	-5.86	-0.99	1.07	-0.33	364	Branch density: number of branches per $Km^{2.}$	Structural financial indicators database: ECB Statistical Data Warehouse
D	-	-	-	-	-	364	Dummy variable to indicate internet banking penetration	Constructed using data on Internet banking (IB)
IU (%)	56.84	3.22	95.99	23.66	0.42	364	Internet usage: the proportion of individuals who used the Internet from any location	World Telecommunication Union: https://www.itu.int/en/ITU-D/Statistics/Pages/datacollection/default.aspx#questionnaires

							in the last three months	
PD (natural log)	11.60	9.63	14.11	0.89	0.08	364	Population density: population per km^2	Eurostat: General and regional statistics
Population (natural log)	16.08	12.91	18.23	1.40	0.09	364	population	Eurostat: general and regional statistics
G (natural log)	10.06	8.92	11.20	0.37	0.04	364	GDP per capita	Eurostat: Economy and Finance
E (%)	25.12	8.20	45.90	8.27	0.33	364	The percentage of population aged 25-74 who have obtained Tertiary education	Eurostat: General and regional statistics
A (deflated by total assets)	0.005	0	0.10	0.01	2.77	363	Accumulative total amount of state aid in the form of recapitalisation and impaired asset relief	European Commission: http://ec.europa.eu/competition/state_aid/scoreboard/financial_economic_crisis_aid_en.html

							deflated by total assets	

Appendix 2:

Table A3: First –stage regression results

	For the Internet Banking diffusion model		For the endogenous bank concentration model			
					IB cut off at 25%	*IB* cut off at 30%
	C	*B*	*Log(TA)*	*B*	*D* (probit)	*D* (probit)
P	-5.883***	0.008	0.879***	-0.017	-0.116	-0.232
PD	1.337	0.923***	0.197***	0.992***	0.428**	0.423**
R	7.770*	0.573***	0.111	0.742***	0.060	-0.621
IU	-0.040	0.011*	0.005*	0.001	0.152***	0.170***
E	0.016	0.002	-0.019**	0.000	0.070	0.211
G	-3.944	0.547	1.241***	0.523	3.906	2.199
t	0.269	-0.065***	0.025*	-0.030*	0.210	-0.092
constant	401.082***	-25.446***	-38.977***	-21.122***	-12.733	-8.231
R^2 or Pseudo R^2	0.52	0.87	0.96	0.87	0.84	0.80
No. of obs.	286	286	364	364	364	364

Appendix 3: Controlling for the population age when estimating the diffusion of internet banking

Figure A1 Median age of the population across country and over time

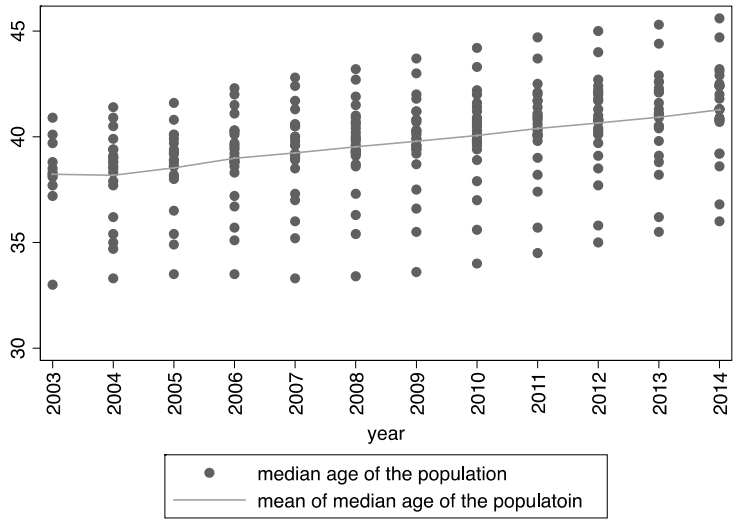


Table A4: Estimating the uptake of internet banking including the age control variable (denoted as $\boldsymbol{PA}$)

	C measured by $CR5$		C measured by HHI	
	Spec 1a	**Spec 2**	Spec 1a	**Spec 2**
C	0.022*	0.033***	7.265***	8.034***
B	0.111	0.051	0.213	0.191*
R	1.606**	1.830***	1.362**	1.212*
E	-0.004	-0.020*	0.029	-0.021
G	0.429	1.331	1.014	1.614*
$\boldsymbol{PA}$	**9.880 (9.4413)**		**8.470 (9.6432)**	
t	0.944	0.105**	0.985	0.101**
$C*t$	0.001		0.064	
$B*t$	-0.018*	-0.014*	-0.017*	-0.015**
$R*t$	-0.101*	-0.101**	-0.093*	-0.082*
$E*t$	-0.002		-0.001	

$G*t$	0.043		0.023	
$PA*t$	**-0.353 (0.2389)**		**-0.303 (0.2496)**	
constant	-34.150***	-17.054***	-22.955**	-11.281***
$\hat{r}_{it}^{C}$	-0.011**	-0.011**	-3.201**	-2.902**
$\hat{r}_{it}^{B}$	0.010	0.079	-0.062	-0.020
F test to compare spec 1 and spec 2	1.93		0.91	
Adjusted R^2	0.82	0.82	0.82	0.82
No. of Obs.	286	286	286	286

Appendix 4: Robustness check for the estimation of endogenous market concentration

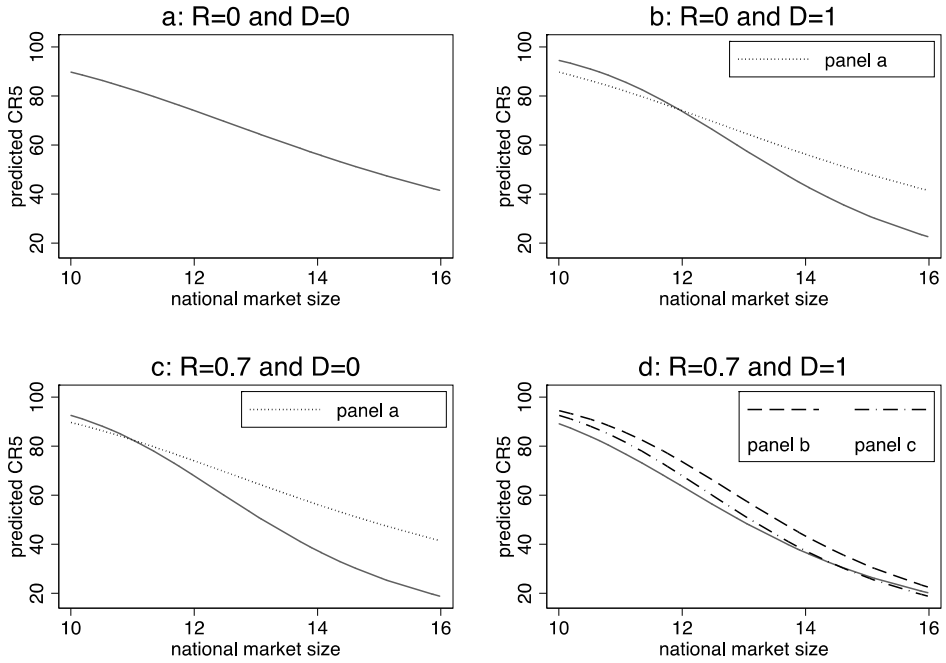
1) Sensitivity analysis relating to a different cut-off point for the internet banking dummy variable D

   Table A5 Estimating the endogenous market concentration model using a different threshold value of IB (30%) for D

Dependent variable: $\log\left(\frac{C_{it}}{100-C_{it}}\right)$	Spec 1	Spec 2
t	0.096***	0.093***
$\frac{1}{\ln S_{it}}$	67.012***	70.296***
R_i	-5.126***	-3.446**
B_{it}	0.344***	0.339***
$R_i * \frac{1}{\ln S_{it}}$	56.420***	33.742*
D	-3.343***	-2.710***
$D*t$	-0.071***	-0.068***
$D * \frac{1}{\ln S_{it}}$	41.820***	33.998***
$D * R_i$	6.350**	0.808***
$D * B_{it}$	-0.191***	-0.180***

$D * R_i * \frac{1}{\ln S_{it}}$	-79.210*	
$\hat{e}_{it}^{S}$	-0.144**	-0.111
$\hat{e}_{it}^{B}$	-0.474***	-0.517***
$\hat{e}_{it}^{D}$	0.240**	0.237**
Constant	9.337***	9.321***
Adjusted R^2	0.73	0.72
No of Obs.	364	364
F test comparing spec 1 and spec 2	8.915***	

Figure A2: Predicted market concentration varying with regionalisation and the uptake of internet banking (using estimates from spec 1 in table A5)

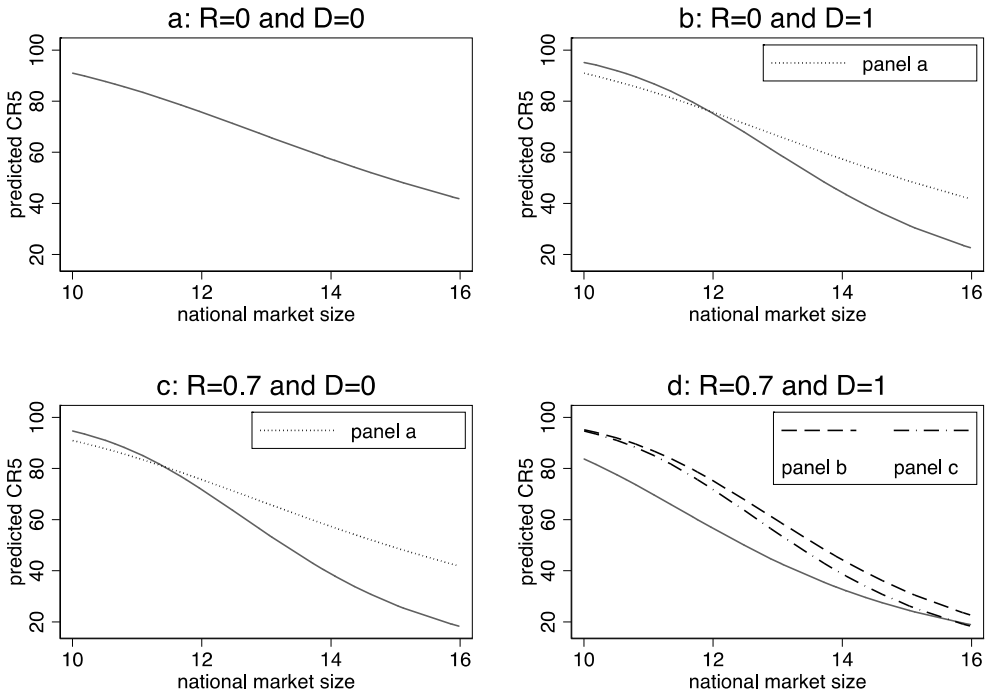


2) Sensitivity analysis by excluding Luxembourg from our sample

Table A6 Estimating the endogenous market concentration model by excluding Luxembourg

Dependent variable: $\log\left(\frac{C_{it}}{100-C_{it}}\right)$	*Estimated coefficients*
t	0.108***
$\frac{1}{\ln S_{it}}$	70.500***
R_i	-5.823***
B_{it}	0.284***
$R_i * \frac{1}{\ln S_{it}}$	66.473***
D	-3.002***
$D*t$	-0.078***
$D * \frac{1}{\ln S_{it}}$	42.013***
$D * R_i$	8.1606***
$D * B_{it}$	-0.102**
$D * R_i * \frac{1}{\ln S_{it}}$	-109.012***
$\hat{e}_{it}^{S}$	-0.142**
$\hat{e}_{it}^{B}$	-0.355***
$\hat{e}_{it}^{D}$	0.180**
Constant	0.122
Adjusted R^2	0.74
No of Obs.	350

Figure A3 Predicted market concentration varying with regionalisation and the uptake of internet banking (excluding Luxembourg)



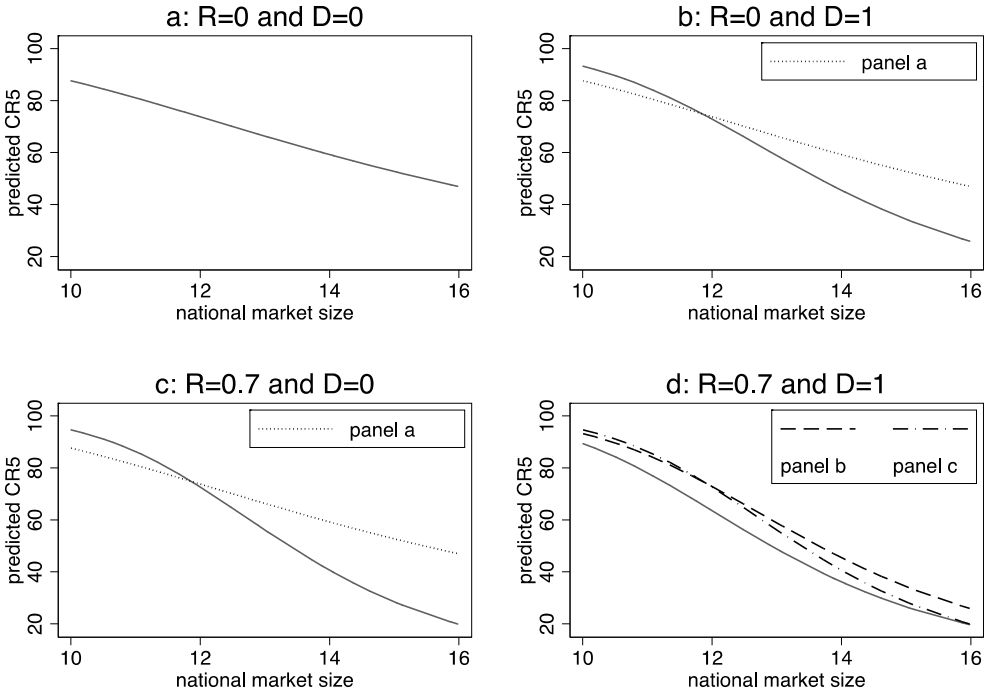
3) Including an additional instrument for market size when estimating model 2 (the relationship between concentration and market size)

Table A7 Estimating the endogenous market concentration model by instrumenting market size with an additional instrument (Central and East European Countries)

Dependent variable: $log\left(\frac{C_{it}}{100-C_{it}}\right)$	*Estimated coefficients*
t	*0.104****
$\frac{1}{ln\, S_{it}}$	*55.513****
R_i	*-7.037****
B_{it}	*0.340****
$R_i * \frac{1}{ln\, S_{it}}$	*83.469****

D	*-3.218****
*D*t*	*-0.077****
$D * \frac{1}{ln\, S_{it}}$	*42.603****
$D * R_i$	*6.859****
$D * B_{it}$	*-0.147****
$D * R_i * \frac{1}{ln\, S_{it}}$	*-86.726****
$\hat{e}_{it}^{S}$	*-0.492****
$\hat{e}_{it}^{B}$	*-0.540****
$\hat{e}_{it}^{D}$	*0.166**
Constant	*10.851****
Adjusted R2	*0.76*
No of Obs.	*364*

Figure A4 Predicted market concentration varying with regionalisation and the uptake of internet banking (using estimates from table A7)



4) Controlling for impact of the financial crisis when estimating the endogenous market concentration model

Table A8 Estimating the relationship between concentration and market size controlling for the impact of financial crisis

Dependent variable: $\log\left(\frac{C_{it}}{100-C_{it}}\right)$	IB cut off point at 25%	IB cut off point at 30%
t	0.096***	0.089***
$\frac{1}{\ln S_{it}}$	62.051***	64.444***
R_i	-5.789***	-5.511***
B_{it}	0.330***	0.345***
$R_i * \frac{1}{\ln S_{it}}$	65.409***	61.220***
D	-3.310***	-3.459***
$D*t$	-0.077***	-0.072***
$D * \frac{1}{\ln S_{it}}$	43.431***	43.544***
$D * R_i$	7.275**	7.129**
$D * B_{it}$	-0.162***	-0.193***
$D * R_i * \frac{1}{\ln S_{it}}$	-92.812*	-90.174*
A	**3.698 (4.1780)**	**4.963 (4.7601)**
$\hat{e}_{it}^{S}$	-0.144**	-0.149**
$\hat{e}_{it}^{B}$	-0.473***	-0.486***
$\hat{e}_{it}^{D}$	0.155	0.231**

Constant	10.272***	10.150***
Adjusted R^2	0.73	0.73
No of Obs.	364	364

Note: Standard errors in parenthesis for financial crisis variable A