The early days of neobanks in Europe: identification, performance, and riskiness

Alberto Citterio[1], Bernardo P. Marques[2], Alessandra Tanda[3]

Abstract

This paper identifies banks that are born with a digital business model ("neobanks") and examines their performance and riskiness *vis-à-vis* traditional incumbents. We propose a novel approach to identify neobanks, based on non-financial data, hand-collected from different sources, and document the existence of 55 neobanks in 17 European countries. Our findings show that, on average, neobanks perform worse and are riskier than their incumbent peers. On one hand, neobanks seem to adequately price the risk of lending to high-risk borrowers and record staff efficiencies. On the other hand, they charge significantly lower fees and commissions and record higher non-staff expenses. Further analysis suggests that such non-staff inefficiencies disappear when we consider banks that are older than 8 years or that offer at least 3 types of products. Our findings are robust to endogeneity concerns and changes to our baseline specification.

JEL classification: G20; G21; G28; G32.

Keywords: neobanks; digital banks; business models; bank performance.

---

[1] (corresponding author) Department of Economics, University of Insubria, Varese (Italy), email: a.citterio2@uninsubria.it

[2] Católica Porto Business School, Portuguese Catholic University (Portugal), email: bmarques@ucp.pt.

[3] Department of Economics and Management, University of Pavia, Pavia (Italy), email: alessandra.tanda@unipv.it.

This research is financed by the European Union and Portuguese public funds through the FCT (Fundação para a Ciência e a Tecnologia, I.P.) and the European Social Funds (Operational Program Norte 2020) under projects number UIDB/00731/2020 (CEGE).

1. Introduction

In the aftermath of the 2007-08 global financial crisis, the financial services industry was hit by a wave of disruptive digitalisation, driven by demand and supply side factors (Arner *et al.*, 2017; OECD, 2020). On the demand side, customer preferences shifted (more rapidly) from 'brick-and-mortar' to digital banking (FSB, 2019); on the supply side, the availability of big data (Boot *et al.*, 2021), advances in key technologies (e.g., application programming interfaces, distributed ledger technology, cloud computing), and the entry of BigTech in the financial services industry (Frost *et al.*, 2019), helped pave the way for a new era of digitalisation in banking[1].

An interesting outcome of this shift in paradigm has been the arrival of a new player in the banking market, born in a fully digital environment, i.e., the so-called 'neobanks' (BCBS, 2018; Tanda and Schena, 2019; Boot *et al.*, 2021; Carbó-Valverde *et al.*, 2021). In a nutshell, these banks may be seen as having "a business model in which the production and delivery of banking products and services are based on technology-enabled innovation" (ECB, 2018: p.3). The fact that we are still in the 'early days' of neobanking naturally poses a number of questions, such as: how many 'neobanks' currently exist? Where are they located? What distinguishes their business model from that of traditional banks? How have they performed? What market factors help (or hinder) their performance?

While a recent strand of literature has discussed the emergence of neobanks (BCBS, 2018; Tanda and Schena, 2019; Boot *et al.*, 2021; Carbó-Valverde *et al.*, 2021), little is known regarding their nature, business model and identification. Relatedly, the existing literature on banking business models (e.g., Mergaerts and Vennet, 2016; Marques and Alves, 2020, 2021) has not yet made progress in mapping innovative banking business models, such as those of 'neobanks', presumably due to data availability issues, as most studies use exclusively financial data as proxies. Also, with respect to the performance of neobanks, empirical literature is scant, and theory offers mixed predictions: on one hand, Boot *et al.* (2021) argue that neobanks are in a prime position to "spatially" capture banking customers by setting up efficient distribution channels, targeting customers via direct marketing tools, and drawing on flexible IT infrastructures (e.g., cloud services); on the other hand, the results for the performance of US fintech lenders are ambiguous, and seem to depend crucially on the type of business line – while the evidence points towards a positive performance of fintech lenders in the online mortgage market (Buchak *et al.*, 2018; Fuster *et al.*, 2019), an opposite effect is found for the personal loans and small business loans (Di Maggio

[1] An historical perspective on how technological progress have influenced key elements of the bank intermediation activity can be found in Boot *et al.* (2021).

and Yao, 2020; Carmichael, 2017). For instance, findings by Di Maggio and Yao (2020) indicate that fintech lenders seem to be tapping into lower quality borrowers in the personal loans market, previously rejected by the incumbents, which could indicate that, as newcomers to an already saturated banking market (ESRB, 2014), neobanks may be subject to the ‘winners curse’ (Shaffer, 1998).

Effectively, given the innovative nature of neobanks, research about the business model and performance of these new players seems to crucially depend on gaining access to non-financial data regarding the type of products and innovative customer experience offered by banks. In this paper we address this issue by using a unique set of hand-collected data from the banks’ websites and Factiva news, regarding the business lines, online functionalities, and stakeholder perception of neobanks. Particularly, to ensure the viability of our approach, we identify neobanks by imposing a set of *a priori* filters to our initial sample (comprised of all supervised banks in EU-28 countries) related to size, business model, ownership structure, age, and branch network – yielding a total of 172 banks for which we hand-collect data. Additionally, to assess the performance of neobanks, we establish as counterfactuals a set of traditional banks with a similar size, business model and ownership structure, but with a significantly lower digital orientation. To compare the performance of neobanks and traditional banks, we run several analyses, including univariate comparison of means, OLS regressions on decomposed elements of ROA, and sub-sample regressions of banks that operate within specific business lines (e.g., credit cards, personal loans). Furthermore, to check whether market factors matter for the performance of neobanks, we compare the regression results of banks operating in countries with larger banking sectors (potentially overbanked), with those operating in countries with smaller banking sectors, where presumably there may exist more space in the market for new entrants. Finally, we compute a set of robustness checks, including 2SLS regressions, where we use as instruments the proximity of the bank to relevant knowledge centres, as well as the quality of such knowledge centres (as measured by their patents’ output).

Our results indicate the existence of 55 neobanks operating in 17 European countries, wherein the UK (12), Italy (8) and Sweden (7) are the countries with the highest number of neobanks. Besides the differences between neobanks and traditional banks (n=158) that occur by “design” – i.e., age, branches, online functionalities –, we find that neobanks tend to be smaller and more specialized; they also tend to dedicate a smaller fraction of assets to gross loans to customers and hold more liquid assets in their balance sheet; on the funding side, no significant differences emerge between neobanks and traditional banks, as both types rely mostly on customer deposits; regarding solvency, we find that neobanks operate, on average, with a higher total capital ratio than incumbents; importantly, we document a significantly higher NPL ratio for neobanks than for traditional banks.

With respect to performance, our results suggest that neobanks perform worse than traditional incumbents. We find that neobanks record a higher level of interest income and higher cost of risk than traditional banks do. This finding, in addition to the higher NPL ratio, suggests that neobanks may have entered the market by lending to higher-risk borrowers, which would be line with extant literature (Di Maggio and Yao, 2020). Additionally, the results show that neobanks tend to generate significantly less non-interest income than their traditional peers – which seems to corroborate anecdotal evidence regarding the lower commissions charged by neobanks (FT, 2019). Also, while our results suggest that neobanks are more efficient than traditional banks with respect to staff expenses, an opposite result is found regarding non-staff operating costs (i.e., IT, advertising, reporting). Further analysis indicates that this effect fades away as we consider older banks or banks with a larger number of product lines, which suggests the presence of economies of experience and scope in digital banking (while no such effect is found for increases in size).

This paper contributes to the literature in several ways. First, we contribute to the literature on the performance of banking business models in Europe (Delgado *et al.*, 2007; Arnold & Ewijk, 2011; Mergaerts and Vennet, 2016; Marques and Alves, 2020, 2021). Namely, we update the literature on the performance of digital banks by covering the 2019-2020 period, which compares to the period between 1997 and 2002 covered by Delgado *et al.* (2007). Moreover, we expand the type of analyses performed by assessing the cost of risk and the potential for economies of scope of neobanks, which had not been addressed so far in the literature. Finally, the fact that our sample of neobanks is significantly larger than that of previous studies (55 vs 15), allows us to perform a greater variety of, and more robust, empirical analyses.

Second, the paper speaks to the literature related to the effects of digitalisation on information and communications frictions in banking intermediation (Diamond, 1984; Merton, 1995; Broecker, 1990; Degryse & Ongena, 2005; Puri & Rocholl, 2008; Drechsler *et al.*, 2018; Thakor, 2020; Boot *et al.*, 2021). On the information frictions side, the fact that we find neobanks to reflect the higher risk into their loan pricing, indicates that such banks are effectively able to escape the 'winners curse' (Shaffer, 1998). Regarding communication frictions, the underperformance recorded by neobanks relative to traditional banks indicates that, for now, the 'digital spatial capture' narrative (Boot *et al.*, 2021) cannot be confirmed for neobanks.

Our third, and final, contribution is related to the development of valid instruments for performance related research. As argued by Clougherty *et al.* (2016: p.308), such studies are often "characterized by the difficulty of finding strong IVs". In this regard, we identify two IVs: proximity to knowledge centres and the quality of knowledge centres. The former consists in the road distance (in hours) between the bank's headquarters and the nearest top50 university in the 'Scimago Institutions Ranking'; and the latter measures the total number of ICT patents recorded in the region of the nearest top50 university. In a nutshell, both aim to reflect the banks' access to the knowledge necessary to purse certain digital strategies. Importantly, we discuss why, in our view, such knowledge spillovers are expected to impact the performance of banks mainly via the business model channel.

The findings in this paper bear relevant policy implications. First, our findings related to the riskiness of the neobanks' credit portfolio are an indication of potential underlying issues related to their credit provisioning policies of these banks. Second, the non-staff inefficiencies of neobanks may be linked, among other factors, to the reporting burden. In this sense, our results emphasize the need for more advances on risk-based regulatory and supervisory practices (EBA 2013). Third, and final, by finding that higher risk neobanks tend to operate in potentially 'overbanked' markets, we see this paper as contributing to the current debate on the consolidation and integration of the European banking market (Enria, 2021).

The remainder of the paper is organised as follows. In **Section 2** we present the literature review on the definition and identification of neobanks, the theoretical framework, and recent empirical literature. **Section 3** describes the methodology used to identify neobanks, traditional banks and assess their performance. **Section 4** provides an overview of the data. In **Section 5** we present and discuss the results. Robustness checks are performed in **Section 6**, while **Section 7** concludes.

2. Literature review

2.1. The definition and identification of neobanks

The literature offers a variety of definitions of 'neobanks'[2]. For instance:

- "Neo-banks make extensive use of technology in order to offer retail banking services predominantly through a smartphone app and internet-based platform." (...) "They leverage scalable infrastructure through cloud providers or API-based systems to better interact through online, mobile and social media-based platforms." (...) "[They] may adopt big data technologies and advanced data analytics" (BCBS, 2018; p. 16);

[2] The literature uses several *alias* for the term 'neobanks' (BCBS, 2018; Boot *et al.*, 2021; Carbo-Valverde *et al.*, 2021), including 'fintech banks' (ECB, 2018), 'challenger banks' (Gontarek, 2021) and 'digital-only banks' (ECB, 2020; Nel and Boshoff, 2021).

- “[Digital-only banks] are recently established (...) [and] innovative banks that use primarily digital channels (e.g. online, mobile apps, etc.) to serve their existing and new clients. These banks do not have branches nor maintain a network of private bankers, while relying on new technologies for managing interactions with their customers.” (p.42) (...) “They tend to specialise in certain primary business lines, such as payment systems, trading and asset management”. (ECB, 2020: p.45);
- “Digital-only banks are branchless banks, meaning that their customers can only transact with them using digital banking channels such as online banking and mobile banking” (Nel and Boshoff, 2021: pp. 429-430);
- “Neo banks (...) are offering customer-friendly interfaces and employing more efficient IT processes.” (Boot *et al.*, 2021: p.14).

In general, such citations suggest that, while there is no unique definition in the literature, ‘neobanks’ may be seen as a ‘new’ type of banking business model, that uses innovative digital technologies – including blockchain, smart contracts, robo-advisors, advanced analytics and big data (Arner *et al.*, 2017; Frost, 2020; Carbó-Valverde *et al.*, 2021) –, to build competitive advantage over incumbent banks in specific markets (e.g., lending, payments, trading and asset management), via the provision of a superior customer experience and/or lower costs.

Interestingly, the management literature has often placed the onset of academic interest for the term ‘business model’ in the context of the dot-com boom of the late 1990’s (Zott *et al.*, 2011). Thus, in a sense, describing ‘neobanks’ as a ‘business model’ may be seen as a full-circle for business model literature. On the other hand, a set of advances has been made regarding the identification of banking business models, particularly regarding the methods used. For instance, Ayadi *et al.* (2011) apply hard clustering techniques to the financial data (e.g., customer deposits, trading assets, loans to banks) of 26 banks and group them into three banking business models: retail, investment, and wholesale. On the other hand, using a larger sample of European banks, Mergaerts and Vennet (2016) apply factor analysis to financial data and find two main factors: retail and diversification. More recently, Marques and Alves (2020, 2021) combine both approaches by using the retained principal components as inputs to an ensemble of three alternative clustering techniques (fuzzy c-means, self-organizing maps, partitioning around medoids); yielding a total of four business models: retail focused, retail diversified funding, retail diversified assets, and large diversified. While such developments may be noteworthy, it is also striking that none of the cited works use non-financial data to identify business models (e.g., business lines or customer experience offered). We argue that, while such shortcomings may be less severe when mapping a sample of traditional banks, it becomes more striking when the goal is to identify the 'neobanks' business model, which, as described above, is focused on specific segments and on offering an innovative digital customer experience. However, gaining access to non-financial data can be very costly, especially for large samples.

As such, it is not surprising that the identification of 'neobanks' has seldomly been attempted by the literature. For instance, Delgado *et al.* (2007) study the performance of 15 'primarily internet banks' in Europe between 1994 and 2002, which are characterized as being "heavily reliant on the Internet as their most important delivery channel" (p.650). Similarly, DeYoung (2005) identifies 12 internet banks operating in the US, applying a set of criteria related to age, size and range of banking products offered through the Internet. More recently, the 'Cambridge Centre for Alternative Finance' (CCAF) set up the 'Cambridge FinTech Ecosystem Atlas' with the goal to "systematically identify, classify and visualize FinTech entities" (CCAF, 2022). The database currently covers 2,915 entities, from 108 countries, which are classified in 14 market segments, 63 sub-segments and 118 categories. Such classification is done in a collaborative way by academics and industry participants, using data from a variety of sources such as company websites, company statements, interviews, and news articles. One of the classifications presented by CCAF is the 'fully digitally native banks', which seems to fit our definition of neobanks. However, when checking the entities operating in EU-28 countries, only 23 entities are currently classified as 'fully digitally native bank', in 7 countries. Perhaps more importantly, we find evidence that not all these entities hold a banking license[3]. In our view, while the collaborative and evolving nature of the project is bound to eliminate such imprecisions, this contributes to the perception that a gap effectively exists in the banking literature regarding the identification of neobanks.

2.2. Theoretical framework

The performance of 'neobanks' may be understood using two conceptual frameworks from the management and banking literature: bank intermediation theory, and strategic groups theory.

A key theoretical framework for our study is bank intermediation theory. In general, when performing its role in the efficient allocation of savings in investment opportunities (Merton, 1995), banks address two types of frictions: (i) information frictions, which are related to moral hazard and adverse selection, and are mitigated via the screening and monitoring of risky investments on behalf of savers (Diamond, 1984); and (ii) communication frictions, which are linked to search, switching

[3] For instance, the entity 'Saffe', which is identified by CCAF as a 'fully digitally native bank', describes itself as providing "world-class facial recognition technology" (Saffe, 2022).

and transportation costs, and historically have been overcome by setting up physical branches which enabled customer relationships to arise (Boot *et al.*, 2021). Regarding the information frictions, a theoretical result that seems particularly timely for our empirical context is the notion of 'winner's curse'. Namely, this phenomenon relates to effects, for *de novo* banks, that emerge from the possibility that loan applicants may indefinitely apply for loans, after being previously rejected. The model by Broecker (1990) suggests that this feature of the loan market makes new entrants prone to have a riskier loan portfolio than those of incumbents unless their screening abilities are, in fact, superior – which is corroborated by the empirical findings from Shaffer (1998). This creates an interesting set-up for our empirical work: all other relevant factors constant, unless the screening ability of neobanks is superior to that of incumbents we may expect to find their loan portfolio to be riskier.

As for communication frictions, theoretical and empirical works have historically used the term "spatial capture" to depict the relationship banks' ability to profit from their physical proximity to customers via price discrimination (Degryse & Ongena, 2005), cross-selling (Puri & Rocholl, 2008), and access to cheap and stable funding (Drechsler *et al.*, 2018). However, according to Boot *et al.* (2021) the current wave of digitalisation may have shifted the "spatial capture" of customers from the physical to the digital domain. Particularly, according to Boot *et al.* (2021) such 'digital spatial capture' could result from the neobanks' ability to "set up efficient communication channels via web portals and mobile apps at very low cost (…), reach targeted audiences via direct marketing tools, including social media (…), source flexible and cost-effective IT infrastructure through cloud services (…), [and] facilitate payments for online purchases" (p.6). As such, whether neobanks have been able to 'digitally capture' bank customers, remains an open empirical question.

Another framework that may help us explain the performance of neobanks, is provided by strategic groups theory (Caves and Porter, 1977). According to this theory, the firms of a given market are likely to make decisions regarding a set of strategic dimensions, such as the distribution channel, the type of products offered, or the level of value chain integration, resulting in the creation of groups of firms that exhibit similarity of strategic choices within each group, and dissimilarity from other groups. Such strategic choices are often related to costly investments which (i) could impede firms from easily changing their group membership in the short run, i.e., the so-called 'barriers to mobility'; and (ii) may protect incumbent firms from new entrants (McGee, 2006). As such, one of the key propositions of strategic groups theory is that mobility barriers may play an important role in explaining intra-industry performance heterogeneity (Porter, 1979). This theory provides us with mixed predictions regarding the performance of neobanks: on one hand, the branch network and longstanding customer relationships of traditional banks may be seen as an entry barrier impeding neobanks from tapping into certain potentially lucrative markets (e.g., SME lending); on the other hand, the ability to take full advantage of innovative technologies may be seen as a difficult barrier for certain traditional banks to overcome, particularly given that many of these banks face IT legacy issues (Stulz, 2019).

2.3. Empirical literature: internet-only banks and fintech lenders

The emergence of internet banks in the late 1990s and early 2000s ignited a strand of literature focused on analysing their performance relative to traditional peers. The seminal work in this subject was developed by DeYoung (2005), wherein the performance of 12 internet-only US banks, incorporated between 1997 and 2001, is compared to that of 644 branching banks set up in the same period. According to the author, the focus on banks born as internet-only banks allows "a clean test of the internet-only business model (…) unaffected by any production structure or client relationships left over from a preexisting business model" (DeYoung, 2005: p.894). The results show that, on average, internet-only banks underperform relative to the traditional peers, due to overhead inefficiencies that more than offset the better pricing abilities. According to the author, such efficiencies may be linked to the fact that "internet-only banks have access to deeper scale economies than branching banks" (DeYoung, 2005: p.895) – an indication that internet-only banks could become more profitable than their traditional peers, after reaching a certain size.

Bearing this aspect in mind, Delgado *et al.* (2007) extend DeYoung's (2005) framework to a sample of 15 'primarily internet banks' operating in 6 European countries, between 1994 and 2002. The authors find that "internet banks performance lies below both the newly chartered traditional banks and the small established traditional banks for all size categories, but this performance gap diminishes for larger Internet banks" (Delgado *et al.*, 2007: p.654). Moreover, the results regarding the drivers for the underperformance of US internet banks are confirmed for European banks. Namely, it is found that such underperformance is mainly related to the overhead costs. Importantly, however, the authors find evidence that the rate at which overhead costs reduce with size is greater for internet banks than for traditional peers, and that the net margin increases with age only for internet banks. Respectively, such results are interpreted as consistent with the existence of 'economies of scale' and 'economies of experience' in internet banking.

Strikingly, however, the analysis of the cost of risk of internet banks has fallen outside the scope of existing studies on internet bank performance. As such, next we draw on the literature regarding the performance of fintech lenders, where emphasis has been placed on their screening abilities.

Overall, the literature on the performance of fintech lenders provides mixed results regarding the merits and shortcomings of fintech lenders, depending on the type of business line. On one hand, using a large sample of mortgage loans in the US between 2007 and 2015, Buchak *et al.* (2018) find that default rates are statistically similar between fintech lenders and traditional banks. Moreover, fintech lenders tend to charge higher spreads than traditional banks for comparable borrowers, which is interpreted as a premium that borrowers are willing to pay for the superior customer experience offered by fintech lenders. In the same vein, using a similar database to Buchak *et al*. (2018), Fuster *et al.* (2019) also document similar default rates between fintech and non-fintech lenders. Also, the evidence collected points towards a significantly faster pace at which mortgage applications are processed in fintech lenders, which seems to provide additional reasoning for the 'higher convenience-higher rates' nexus.

On the other hand, Di Maggio and Yao (2020) focus on a large sample of US personal loans and show that fintech lenders enter the market by lending to higher-risk borrowers. Also, the evidence suggests that borrowers with similar characteristics are more likely to default when borrowing from fintech lenders than from traditional banks. Despite this, the results suggest a high correlation between interest rates and default rates, which is interpreted as evidence of fintech lenders' pricing ability. Similarly, Balyuk *et al.* (2020) analyse a large sample of US small business loans and find evidence suggesting that fintech lending tends to replace the riskier loans granted by large and out-of-market banks. Finally, Carmichael (2017) find that fintech borrowers in the US personal loans market who were previously rejected by another fintech competitor are twice as likely to default as borrowers who were not rejected, which is seen as evidence of the 'winners curse' (Shaffer, 1998) in the online lending market.

One feature of this strand of literature is the lack of research studies regarding European fintech lenders. The few exceptions that provide a picture of European fintech lenders (e.g., Milne and Parbooteeah, 2016; Claessens *et al.*, 2018) is focused mostly on country-level data and do not offer a comparative performance of fintech and non-fintech lenders. We argue that this phenomenon is presumably linked to the relative lack of microdata in Europe *vis-à-vis* the US, where some fintech lenders provide open access to their data. For instance, the fintech 'Prosper' provides access to monthly loan-level data, regarding the characteristics and performance of loans. In this context, while the use of bank-level data is not ideal, we argue that it remains a relevant contribution to the scant literature on emerging digital business models in Europe.

3. Methodology

3.1. Identification of neobanks

Our method to identify neobanks consists of four sequential filters applied to the data, wherein each step uses different sources of data and is sustained in the literature. As presented in **Figure 1**, in the first filter, our goal is to find banks that share a similar business profile in terms of geography, size, business model and ownership type. Regarding *geography*, we focus on supervised banks operating in EU-28 countries. This is done by retrieving the list of supervised banks from each national supervisory authority, including the ECB for the Euro Area, with reference to 2019. With respect to *size*, we expect digital challenger banks to be newcomers and, as such, to exhibit a relatively small size. Hence, we only consider banks with total assets smaller than €10bn. The ecosystem of neobanks includes a variety of business lines, including retail lending, payments, trading and asset management, and B2B (ECB, 2020). To improve the comparability of our sample, we narrow our focus on retail-oriented neobanks by requiring banks in our sample to have at least 5% of total assets dedicated to customer lending and customer deposits. Finally, we exclude stakeholder banks (i.e., cooperative and savings) from our sample, given that these are *ex-ante* more likely to operate a 'brick-and-mortar' model due to their strong presence in rural areas, where customers are typically less digitally mature[4].

In the second filter, our aim is to identify banks with a high propensity to adopt a digital banking model, typical of neobanks. This is achieved by analysing the distribution channel and the age of banks. Regarding the *distribution channel*, as reported by Ehrentraud *et al.* (2020: p.8), digital banks may be seen as "delivering banking services primarily through electronic channels instead of physical branches". Hence, we consider that banks with more than 5 branches[5] have a low propensity to adopt a digital banking model, and as such exclude them from the challenger sample. With respect to the *age* of the bank, we expect that customer preferences towards digital banking services as well as the technology necessary to offer such services, became more intense after the

---

[4] Besides cooperatives and savings banks, other types of specialization are also excluded from our analysis, such as promotional/development banks and clearing/custody banks, due to their unique business models. The retained specialization codes are commercial banks, investment banks, real estate and mortgage banks, and finance companies.

[5] Note that the BankFocus database has well-known data coverage issues, notably regarding less standard data points such as the number of branches. In our initial dataset, branch data was missing ("n.a.") for 62.4% of the banks for 2019 (reference year). To mitigate this issue, whenever possible we use the data for the years immediately before and after the reference year (2017, 2018, 2020); also, we decide to keep all banks with "n.a." in the sample (in other words the criterion at this stage is "< 5 branches" or "n.a."). For banks with "n.a." that meet all the subsequent criteria (i.e., Factiva news and website functionalities), we manually retrieve the number of branches from the annual reports. The same principle is applied to other criteria with low data coverage in BankFocus (e.g., year of incorporation).

internet diffusion reached a critical mass. According to Boot *et al.* (2021) such tipping point occurred during the early 2000s. As such, we consider that banks that were incorporated before the year 2000 are less likely to adopt a digital banking model, and hence exclude them from this sample.

[Figure 1 near here]

The third filter identifies banks that stakeholders perceive as being digital. This is done via the systematic analysis of news in Factiva[6]. Importantly, this procedure allows us to complement the analysis of hard data (filter #2) with information on the perception (of peers, industry associations, media) regarding which banks may be considered as "neobanks" (e.g., FT, 2020). Theoretically, such type of analysis bears support from the cognitive perspective of strategic groups (Reger and Huff, 1993: p. 103), according to which "industry participants share perceptions about strategic commonalities among firms" which are likely to influence decision making. To undergo the systematic analysis of news in Factiva, we run individual searches for each bank with a high propensity to operate a digital banking model, combining the name of the bank with twelve alternative keywords related to digital banking (*vide* list of keywords in **Figure 1**). Also, when analysing the news reports, in some cases we identify new candidate digital banks (i.e., banks not included in the initial list of banks with high propensity to operate a digital banking model). For such banks, we run a reverse analysis, wherein we go back to filter #2 and re-check the reason for exclusion. Such analysis allows us to correct several issues in the BankFocus database (e.g., incorrect year of incorporation, specialization, or number of branches).

The fourth and final filter is related with the availability of certain online functionalities that may be seen as typical of a neobank. We view this step as complementing the previous two (hard data and stakeholder perception). In other words, even if the number of branches of a bank is very low, and/or stakeholders identify the bank as a neobank, one could argue that neither of these indicators constitute direct evidence of the ability of the bank's IT systems to offer "customer-friendly interfaces and employ efficient IT processes" that enable it to reduce communication frictions – an important feature of digital banking, as suggested by Boot *et al.* (2021: p. 6). In line with this, Buchak *et al.* (2018: pp.458-459) classify as fintech lenders those that show "a strong online presence and if nearly all of the mortgage application process takes place online with no human involvement from the lender". Our approach in this regard is to check the website of each bank for two specific features: (i) the ability to open an account fully online and (ii) the ability to apply for a

[6] Factiva is a global news database owned by Dow Jones & Company, which allows for advanced search of keywords in nearly 33,000 news sources and is often cited in banking and finance literature (e.g., Bertay *et al.*, 2015).

loan online. Given that some banks in our sample have only a small orientation towards credit granting, we impose as sufficient condition for a bank to be considered as digital that it offers at least one of these online functionalities.

3.2. Identification of traditional banks

A key aspect of our empirical approach is the identification of traditional banks that may be considered credible counterfactuals, and hence allow us to estimate the performance effect of adopting a digital banking model.

Bearing this in mind, our general approach is to find banks that have a similar business profile as neobanks, while also revealing significant differences with respect to the main channel of distribution. We do so in the following way. First, we take the pool of banks with a homogeneous business profile identified in filter #1 of **Figure 1** and consider only those operating in the sub-set of countries where we find at least one neobank. Second, mindful of the fact that our empirical approach requires the manual collection of data from multiple sources (e.g., banks' websites, annual reports), we restrict our sample to 1/3 of the total number of banks per country. By applying this criterion at the country-level, we ensure that the distribution of banks per country is preserved[7]. However, we observe that for countries with smaller banking sectors and a significant presence of neobanks (e.g., Lithuania, Malta), the implementation of the previous criterion would impair the inclusion of all identified neobanks and/or any traditional banks in the sample. To mitigate this issue, we impose that the final sample must include at least 5 banks per country[8] and the number of traditional banks must be at least equal to the number of neobanks per country.

Finally, we select the traditional banks to include in our sample by identifying the TB_j banks with the highest number of branches per country– wherein TB_j refers to the number of traditional banks of country j to include in the sample, as identified in the previous steps. In case of ties (i.e., banks with the same number of branches), we select the banks with the highest ratio of customer deposits and gross loans to customers to total assets.

[7] The focus attributed to the geographical dimension of bank competition may be explained by the fact that, in our view, most neobanks still compete chiefly at the local/regional level (possibly also due to the fact the European Banking Union is still incomplete), despite their ability to provide banking services outside national boundaries. As such, our assessment is that a balanced representation of digital vs traditional banks at the country level is paramount to ensure the credibility of the counterfactuals in our sample.

[8] This condition is met for all countries in the sample except for Greece, for which only 4 banks meet the "business profile" criteria.

3.3. Measuring the performance of neobanks

To assess the performance of neobanks *vis-à-vis* traditional banks we estimate the following model, using cross-section OLS with White (1984) cluster-robust standard errors:

$$Y_i = \alpha_0 + \beta DB_i + \gamma \boldsymbol{BC_i} + \delta \boldsymbol{CL_i} + \varepsilon_i$$

wherein Y_i is the outcome variable, including the ROA and the sub-elements of ROA (interest income, interest expenses, net interest income, non-interest income, staff expenses, other non-staff expenses, cost of risk); α_0 is the model constant; NB_i is a dummy which takes on the value 1 if bank *i* is defined as a neobank according to our methodology; $\boldsymbol{BC_i}$ is the mean vector of individual bank control variables of bank *i* (size, income diversification, gross loans customers, liquid assets, customer deposits, total equity, non-performing loans, number of products line); $\boldsymbol{CL_i}$ is a vector of country-level controls (GDP growth, total banking sector assets on GDP, GDP per capita, sovereign yield,); β, γ and δ are the regression coefficients' vectors; and $\varepsilon_{i,t}$ is the disturbance term.

Given the relatively small size of our sample (n=213), concerns may be raised regarding potential model overfitting and, hence, the validity of statistical inference (Harrel, 2001). In this regard, the standard rule of thumb is that a "fitted regression model is likely to be reliable when the number of predictors (…) is less than n/10 or n/20" (Harrel, 2001: p.61). This means that the n/p ratio in our baseline regressions (213/12=17.8) is above the minimum threshold defined in the literature (10). As for our sub-sample regressions, we adjust our baseline specification, by removing two country-level controls, which ensures that the minimum threshold is not violated in any specification.

4. Data

4.1. Data description

Our financial data corresponds to yearly financial statements information, at the unconsolidated level, obtained via Orbis BankFocus. We retrieve data for 2019 and 2020. Regarding non-financial data, we collect the information from business lines and online functionalities by checking each bank's website in October 2021. The Factiva news search for media references equating each bank as a 'digital bank' is also done in this period. To narrow the period mismatch and mitigate the difference in data frequency between the financial and non-financial data, we perform two data treatments. First, following the approach by Bachuk *et al.* (2018) we assess the historical accuracy of the website analysis using the 'Wayback Machine', which provides access to archived webpages from 2019 and 2020; this allows us to confirm our initial results. Second, we compute the average values of financial data across 2019 and 2020, resulting in a 2019-2020 cross-section database. The country-level data is obtained from the World Bank database for 2019. We winsorize bank-level data at the 1st and 99th percentile.

4.2. Bank-level controls

We follow the literature on bank performance and include bank-level controls related to the size, diversification, asset and liability structures, leverage, and risk. Below we provide the reasoning for our choices of proxies.

Size. According to the ‘efficiency hypothesis’, profitability may increase with bank size, via economies of scale and scope (Scholes *et al.*, 1976). However, consistent with agency theory, increases above a certain optimal size are often evidence of managerial empire building (Jensen and Meckling, 1976). Moreover, the ‘soft information’ argument, according to which the flow of soft information within larger banks may be impaired due to the presence of complex hierarchical structures (Liberti and Mian, 2008), may be seen as less relevant for our empirical context, given that (i) neobanks are expected to perform mostly transaction-based retail banking (Ewijk and Arnold, 2012) and (ii) our sample comprises only small/medium sized banks (total assets < €10bn).

Diversification. The level of *income diversification* as well as the *number of product lines* referred in the banks’ websites, reflects the ability to make money beyond interest-generating services (i.e., credit granting), namely via fee-based services (e.g., insurance products, investment advisory, credit cards) (Elsas *et al.*, 2010), which may improve the screening and monitoring of customers, as well as diversify risks (Diamond, 1984).

Asset structure. The ratio of *gross loans to customers to total assets* provides a measure of engagement in traditional, ‘originate to hold’ lending (Diamond, 1984). Moreover, the ratio of *liquid assets to total assets* allows us to proxy for the exposure to liquidity risk – wherein, in principle, a higher ratio reflects the existence of a buffer for the bank to pursue investment opportunities whenever they arise; but, on the other hand, it could also reflect the general lack of opportunities to generate revenues for the bank.

Liability structure. The ratio of *customer deposits to total assets* reflects whether the bank’s funding is obtained via customer deposits or other debt instruments. Whether the presence of customer deposits may result in cheap funding seems to depend crucially on the type of bank’s customer base: retail deposits (i.e., household and SMEs) are typically seen as a stable and cheap source of funding due to the presence of deposit guarantee schemes (Diamond and Dybvig, 1983); whereas the reliance on deposits from large corporates and institutions may result in concentration risk and, as such, be costlier for the bank.

Leverage. Besides the more obvious role of capital as safety net against negative earnings, the ratio of *equity to total assets* may also be expected to reflect the ability of banks to pursue business opportunities and to constitute a means for "banks that expect to have better performance to credibly transmit this information through higher capital" (Athanasoglou *et al.*, 2008: p. 127).

Risk culture. We employ the ratio of *non-performing loans to gross loans to customers* (NPL ratio) to reflect the credit risk culture of banks, wherein the expected effect on profitability is negative, as a high NPL ratio is associated to higher impairments and loss provisions, and the consumption of costly capital.

4.3. Country-level controls

According to the literature, country levels factors may influence the performance of banks in a variety of ways, and as such we control for their effects.

Business cycle. We use the *GDP growth* in order to control for the business cycle. For instance, Albertazzi and Gambacorta (2009) show that the pro-cyclicality of banks' profits affects is mostly derived from changes in the lending activity (which affect net interest income) and credit portfolio quality (impacting loan loss provisions).

Market structure. As proxy for the structure of the banking market we employ the total domestic assets of banks divided by GDP. In this regard, Demirguc and Huizinga (2001) show a bank that operates in a mature environment may have to deal with more competition. Moreover, the report by the ESRB (2014) finds evidence suggesting that countries with larger banking sectors relative to GDP are linked to greater risk-taking by banks.

Institutional environment. Finally, as proxies for the quality of the financial system environment we use the *sovereign yield* and *GDP per capita*. Regarding the former, Altavilla *et al.* (2018) find that reducing the sovereign yield spread is associated with higher bank profitability, presumably due to the positive impact on the overall credit quality of the portfolio, that is anchored to the country's overall credit rating, but also due to the appreciation of assets recorded by banks at fair value.

5. Results

5.1. How many neobanks exist in Europe and where are they located?

Table 1 shows the step-by-step results of our identification process. We start with a total of 3,582 supervised banks in EU-28 countries as of 2019. The countries with most supervised banks are located in Germany, France and Italy, which together represent 57% of total banks. In the first step we apply filters on bank's size (total assets), level of retail-orientation (gross loans to customer on total assets, customer deposits on total assets) and specialization in order built a homogenous sample of banks with similar business profile and ownership type. These criteria lead to a sample of 696 banks – corresponding to a reduction of 80% of the initial sample – mainly due to the exclusion of cooperative and saving banks, which alone contribute to a reduction of 70% of the starting list of banks. Second, we further restrict the sample excluding at first banks with a well-developed branch network and, secondly, the older banks, i.e., incorporated before the 2000. Step 2 then allows us to define a group of potential neobanks comprised of 172 financial institutions. The group of 58 banks identified in the third step reflects both the direct and the indirect analysis performed via Factiva. Directly, the analysis of the market perception results in 43 matches. Indirectly, the analysis of news reports allows us to identify further 42 candidates, which are subsequently reduced to 12 after the application of the quantitative criteria (via BankFocus and/or bank's website/financial reports). To further confirm the strong digital footprint of these banks, in step 4 we assess the bank's online functionalities, by excluding banks that do not allow to open an account or apply for a loan via their website. In the end we obtain a total of 55 neobanks operating in Europe.

[Table 1 near here]

Table 2 shows the composition of neobanks per country. The results reveal that the countries with the highest number of banks in the final sample are Germany (18.3%), France (16.9%) and UK (15.0%). The fact that none of these countries represents more than 20% of the final sample is seen as evidence that no country is overrepresented. On the other hand, untabulated results for the ratio between the number of banks in the final sample and the total number of banks identified as having a similar business profile (step 2) per country, show that the minimum ratio is 22% (Luxembourg), which we interpret as an indication that country representativeness has been ensured.

Our final sample consists of 213 banks, amounting to €612.4 bn in total assets in 2019-2020 (neobanks: € 142.6 bn, traditional banks: € 469.8 bn), which represents approximately 6.7% of the total banking assets in the countries they operate. In general, this shows that neobanks still represent only a small fraction of the European sector – at least for now.

[Table 2 near here]

5.2. What are the key distinctive differences of neobanks vis-à-vis traditional banks?

The descriptive statistics per type of bank and the results of the mean comparison tests are presented in **Table 3**. The findings confirm that neobanks are on average younger, with a higher propensity to offer services via online platform and with an almost inexistent branch network. These results are not surprising, since they are strongly influenced by the criteria used for the sample identification. However, despite the use of the same threshold in terms of total assets (less than 10 bn €), neobanks are on average smaller in size than traditional banks. We relate these results to the fact that neobanks are significantly younger and often still in a start-up phase (25% of the neobanks have less than 5 years).

Interestingly, **Table 3** shows that the two groups have a similar funding structure (oriented towards customer deposits) and only a weak difference in terms of gross loans to customers. This condition suggests that the identification strategy was successful in identifying banks with a similar orientation and business model. Regarding the asset structure, the lower exposure to gross loans to loans and higher proportion of liquid assets indicate that neobanks may be facing difficulties in transforming liquidity into productive lending. Coherently, despite the similarity of 'total equity to total assets' between neobanks and traditional banks, untabulated results show that the level of total capital ratio[9] is significant higher for neobanks.

[Table 3 near here]

On the income side, neobanks show a lower level of income diversification, which is consistent with the fewer number of products offered to their customers when compared to traditional banks. According to Diamond (1984), a lower diversification may reduce the ability to screen and monitor customers, as well as to diversify risks – which could help to explain the higher level of NPL ratio recorded by neobanks. Alternatively, the excess NPL ratio suggests that the late entry of neobanks to the market makes them subject to a pool of lower quality potential borrowers, given that the higher quality borrowers should already be locked-in with traditional banks.

5.3. How do neobanks perform when compared to traditional banks?

Table 4 shows the mean values of neobanks and traditional banks for ROA and the sub-components of ROA. In general, we find that neobanks perform worse than their traditional peers.

[9] We were able to obtain the total capital ratio for a subset of 179 (out of the 213) banks.

This is shown by the negative average ROA recorded by neobanks (-0.225%), which is significantly lower than the ROA registered by traditional banks (0.371%). This result is backed by regressions results reported in **Table 5**, particularly by the negative and significant coefficient of the 'neobank dummy' on ROA.

[Table 4 near here]

To shed light on the reasons that underline such underperformance, we analyse how the neobank dummy relates to the sub-components of ROA. First, we find that neobanks record significantly higher interest income than their traditional peers (shown by the positive and significant coefficient of interest income, as well as the higher mean value), while also recording a significantly higher cost of risk (as per mean value and coefficient of impairment costs) – which is consistent with the higher NPL ratio observed in **Table 3**. In our view, such 'high interest–high impairments' linkage seems consistent with the empirical findings by Di Maggio and Yao (2020), according to which US fintech lenders entered the market by lending to higher-risk borrowers. Additionally, to understand whether such effect may fall under the 'winners curse' (Shaffer, 1998), we assess whether the credit screening abilities of *de novo* banks have allowed them to circumvent the effects of tapping into a riskier pool of borrowers (Broecker, 1990). To test this, we build a new dependent variable, 'interest income minus impairment charges divided by total assets'. If the neobank dummy is positive and significant it indicates that, on average, the neobanks cover the higher impairment costs with additional interests charged. Indeed, untabulated results show a positive and significant coefficient of the neobanks dummy at the 1% significance level – which suggests that, despite lending to high-risk borrowers, this is offset by the ability of neobanks to correctly price the risk, which ultimately suggests that neobanks are able to escape the 'winners' curse' trap.

Second, we find that neobanks generate lower non-interest income than their traditional peers, which is observable in the lower mean values and the negative and significant coefficient of the neobanks dummy. Furthermore, untabulated results indicate that such findings are mostly driven by net fees and commission and other income (rather than net trading income). In our view this suggests that neobanks may be attracting customers by lowering their level of fees and commissions – which seems to be in line with anecdotal evidence regarding the lower fees and commissions charged by neobanks (FT, 2019). An alternative explanation may be linked to the fact that several fee generating services are chiefly provided via physical branches, such as safekeeping or financial advisory (Chiorazzo *et al.*, 2018).

[Table 5 near here]

Third, we find that neobanks record higher operating expenses than traditional banks. Particularly, while the results show that, on average, neobanks operate with lower staff expenses (as shown by the negative and significant coefficient of the neobank dummy in **Table 5**), the evidence also indicates that non-staff operating expenses (e.g., IT costs, advertising costs, reporting costs) are significantly higher for neobanks. This seems to be in line with the empirical literature on the performance of internet banks (DeYoung, 2005; Delgado *et al.*, 2007). However, this strand also argues that digital banks seem to show a greater 'depth' of economies of scale and experience than traditional peers.

To test this hypothesis, we split the original sample into buckets of size and age and run sub-sample regressions regarding the main variable of interest on the costs side ('non-staff expenses to total assets'). Furthermore, we expand the literature by also testing the existence of a greater 'depth' for economies of scope among neobanks, by splitting the sample also according to the number of products offered by each bank. Inspired by the framework laid out by DeYoung (2007), we test whether the positive and significant 'neobanks dummy' becomes less significant as size, age and/or the number of products increase. The results of the analysis are presented in **Table 6.** Our findings reveal that, as banks increase in size, the 'neobanks' coefficient remains positive and significant, and as such our data does not suggest the presence of economies of scale specific to the digital model. On the other hand, we find that, with experience, neobanks seem the reduce their relative inefficiencies (untabulated results show that this occurs when banks operate for at least 8 years). Similarly, we find that similar effect after banks broaden the scope of products offered (namely 3 or more). Under our framework, such results suggest that neobanks enjoy some 'depth' of economies of experience and scope.

[Table 6 near here]

Next, we check whether the baseline results change when considering specific product lines. To do so, we focus on three products: credit cards, personal loans, and broker/advisory services. Particularly, we run sub-sample regressions, each comprising exclusively the banks that offer each product. For the sake of simplicity, **Table 7** shows only the coefficients of the 'neobanks dummy'. With respect to ROA, we find a negative and significant coefficient only for the 'personal loans' sub-sample. Taking a closer look at the decomposed elements of ROA, we observe that the coefficient of the interest income (0.978) is significantly higher than that of impairment charges (0.604). As such, we reject the hypothesis that such underperformance of neobanks may be attributed to risk mispricing. Alternatively, the evidence indicates that neobanks experience material cost inefficiencies in the provision of personal loans. To understand whether neobanks may experience economies of scale, experience, and scope in this segment, we re-run the results shown in **Table 6** only for banks that offer personal loans. The untabulated evidence suggests that, over time and by broadening their offer, neobanks reduce the efficiency gap towards traditional banks. With respect to fee-based activities, we find that, compared to traditional peers, neobanks record significantly less non-interest income associated to the provision of credit card and advisory services – which is consistent with anecdotal evidence that neobanks attract new customers in these market segments by charging lower fees and commissions. Interestingly, the lower level of staff expenses identified in **Table 5** seems to be broadly explained by the broker/advisory product line, which is likely related to the relatively low need for specialized financial advisors when providing online brokerage services, as well as the advent of new advisory related technologies, such as the 'robo-advisor' (Oehler *et al.*, 2021).

[Table 7 near here]

Another striking result is related to the dispersion of performance of neobanks. Namely, the coefficient of variation of ROA for digital neobanks is 10.5 (-0.23/2.41) which compares with 3.2 for traditional banks. Such finding seems to indicate the presence of significant heterogeneity within our sample of neobanks, which leads us to question whether there may be some distinctive features that separate top and bottom performing neobanks. **Table 8** shows that top neobanks (the first quartile in terms of ROA) perform better due to their higher level of interest income and lower level of non-staff operating expenses. Coherently with the results shown in **Table 7**, untabulated results indicate that, on average, bottom performing neobanks are more involved in offering personal loans (67.8% vs 37% of high performing neobanks), while top neobanks are primarily focused on broker and advisory activities (48.1% vs 32.1%).

[Table 8 near here]

5.4. Are market factors important for the performance of neobanks?

In this section we investigate if market factors are relevant for the performance of neobanks. Particularly, considering our baseline results, we are interested in understanding whether the riskiness of neobanks may be traced to country-specific factors, such as the size of the banking sector (ESRB, 2014) and/or the growth of credit (Castro, 2013).

To address this question, we split countries according to two proxies for the size of the banking sector: total banking assets to GDP and the ratio of bank credit to deposits. In detail, we considered as "Highly intermediated" countries with above median values of total banking assets to GDP

(Panel A) or bank credit to deposits (Panel B). We restrict the analysis to ROA (as measure of performance) and the level of impairment charges (as a measure of risk). For brevity, **Table 9** shows only the coefficients associated to the neobanking dummy. Coherently with our expectations, the results show that neobanks record a lower level of ROA and a higher cost of risk in countries with larger banking sectors and/or with higher credit intermediation. With respect to smaller banking sectors, the coefficient of the neobank dummy is not statistically significant. Such findings indicate that the 'higher interest-higher risk' profile of neobanks is driven by the banks operating in 'overbanked' countries.

[Table 9 near here]

6. Robustness checks

6.1. Endogeneity issues in business model performance

A tangible concern regarding our empirical setting is related to the possibility that some unobserved features of banks may simultaneously affect their propensity to follow a business model and their performance (Clougherty *et al*., 2016). To address this issue, we apply 2SLS estimation using two IVs that reveal the banks' access to the knowledge necessary to purse certain digital strategies; and such knowledge spillovers are expected to impact the performance of banks mainly via the digital business model channel.

Our first IV is the *proximity to knowledge centers*, and is calculated as the natural log of the road distance between the bank's headquarters and the nearest top-50 university[10], according to the '2021 Scimago Institutions Ranking'. The distance is measured in number of road hours and refers to the distance between the NUTS2 regions of the bank and the university (Persyn *et al.*, 2020). Our rationale for the choice of this IV is that banks with headquarters positioned closer to knowledge centers are more likely to access the specialized resources (human and technological) necessary to adopt certain digital strategies. An opposing argument, however, could be made that the location of certain knowledge centers may, in turn, be a function of the proximity to employers, such as banks, which would entail a problem of reverse causality. While to the best of our knowledge there is no work that studies the relationship between the location of bank headquarters and knowledge centers, anecdotal evidence would suggest that most top universities are centenary-old institutions, which face very significant relocation costs (and would hardly make such costly decision based on the

[10] For countries without any university in the top50, we include the highest ranked university in the dataset. This occurs for Czech Republic (#65), Estonia (#150), Greece (#91), Latvia (#235), Luxembourg (#166), Malta (#295) and Poland (#107).

proximity, or lack thereof, to specific employers). As such, we are confident that this may be considered a suitable instrument.

The second instrument is related to the *quality of knowledge centers*. This IV is computed using the total number of patents in ICT, as published by the OECD, for the NUTS2 region where the nearest top-50 university is located. While the time series for the patents' data has not been updated by the OECD since 2013, we argue that the quality of top research centers is bound to be stable overtime. In our view this IV complements the previous one in a relevant way: considering two banks located at similar distances from two top50 universities, the bank located to the most research productive university is likely to enjoy the greatest technology spillovers – which may, in turn, facilitate its adoption of a digital banking model.

[Table 10 near here]

The results of the 2SLS estimations are presented in **Table 10**. The F-test of (weak) instruments is rejected at the 1% level and the overidentifying restrictions are not rejected. Regarding the first-stage regression, the proximity to knowledge centers increases the likelihood to follow a digital banking model. We also find that the inclination to adopt a digital model is positively affected by the level of excellence of the nearest knowledge center. As for the second stage result, the neobanks dummy is found to negatively affect ROA, lending support to our baseline regressions.

6.2. Alternative specifications

Another source of estimation bias may lie in our dependent variable. Namely, while ROA, and the decomposed elements of ROA, have been widely studied in the banking literature (e.g., Mergaerts and Vennet, 2016; Marques and Alves, 2021), other proxies have also been used to provide insights into the performance and riskiness of banks. As such, next we re-run our baseline regressions using two alternative dependent variables: Return on Equity (ROE) and Total Capital Ratio (TCR)[11].

Regarding ROE, untabulated results show a negative coefficient of the neobank dummy (-7.322), significant at the 10% level, which confirms the previous findings. With respect to TCR, we find a positive and significant coefficient of the neobank dummy (8.485, significant at the 5% level). Apparently, this result may seem to contrast with our main findings. A range of literature claims that higher capital is followed by higher earnings. More in detail, better capital adequacy helps to (i) achieve a reduction in interest rates (Berger, 1995), (ii) have a higher ability to attract funds (Holmstrom and Tirole 1997), (iii) build longer-term customer relationships (Allen *et al.*, 2011),

[11] We limit our analysis to banks for which was possible to collect at least one value of TCR during the period 2019-2020. The sub-sample is thus composed by 179 banks, of which 51 neobanks.

and (iv) carry risks essential to lending (Calem and Rob 1999; Perotti *et al.*, 2011). However, Dagher *et al.* (2020) find that capital ratios in the range of 15% to 23% are likely to absorb losses in most past banking crises, while further capital may entail higher costs which could negatively affect profits (Dietrich and Wanzenried, 2011; Tabak *et al.*, 2017). In support to this view, untabulated results show that, on average, capital adequacy is equal to 19.5% for traditional banks and 31.9% for neobanks in comparison to an average value equal to 17.2% in the EU-27[12] (EBA Risk Dashboard, 2021). In this context, our findings may be seen coherent with our previous results, i.e., consistent with our view of lower borrow opportunity/ability for neobanks.

7. Conclusions

The high degree of disruption of technology-driven innovation in the finance industry has required a complete rethink of the way banks approach customers. Furthermore, the birth of neobanks, together with the entry of new fintech and Bigtech companies in the banking market will probably have a material impact in the industry – although the size of such effects remains unclear (BCBS, 2018). As such, the present topic has increasingly attracted the attention of authorities, managers, and academics. However, tackling this topic is empirically challenging as there is no common database of neobanks, and describing the business model of these banks seem to crucially depend on gaining access to non-financial data.

In this paper we address these issues by developing a methodology to identify neobanks that draws on hand-collected data from banks' websites and Factiva, and performing several analyses on the performance and riskiness of neobanks *vis-à-vis* traditional incumbents. Our results indicate the existence of 55 neobanks operating in 7 European countries. In general, we find that neobanks perform worse than their traditional peers, and this finding is robust to endogeneity concerns and changes to the baseline specification. To deeper understand the source of such under performance, we analyse the contribution of each sub-components of ROA on the overall result. First, we find that neobanks charge sufficiently high interest income to cover the excess cost of risk of lending to higher risk-borrowers (Di Maggio and Yao, 2020), allowing them to escape the "winner curse". Second, we show that neobanks generate lower fees and commissions, which is seen as consistent with their entry strategy. Finally, neobanks record lower staff expenses but significantly higher non-staff expenses than traditional banks. By running additional analyses, we observe the non-staff inefficiencies fade away as we consider older banks (age > 8 years) or banks with more product lines (>3), which suggests the presence of economies of experience and scope in digital banking.

[12] Value at December 2020.

This paper bears relevant contributions to the literature and to policy. On the literature side, we update the literature on the performance of digital banks by covering the 2019-2020 period (vs 1997- 2002) (Delgado *et al.*, 2007); we assess the cost of risk and potential for economies of scope of neobanks, which had not been addressed so far; and we study a significantly larger universe of neobanks relative to previous studies (55 vs 15). We also test various bank intermediation theories and uncover two suitable IVs to study bank performance.

On the policy side, our findings regarding the riskiness of the neobanks' credit portfolio suggest the need for careful supervision (e.g., regarding credit provisioning); the higher non-staff expenses supported by neobanks emphasize the need to monitor the reporting costs faced by these banks, under the current risk-based regulatory and supervisory framework (EBA 2013); and third, the fact that neobanks operate mainly in more saturated (potentially overbanked) markets, may contribute to the current debate on the consolidation (and integration) of the banking market (Enria, 2021).

References

Albertazzi, U., & Gambacorta, L. (2009). Bank profitability and the business cycle. *Journal of Financial Stability, 5*(4), 393-409.

Allen, F., Carletti, E., & Marquez, R. (2011). Credit market competition and capital regulation. *The Review of Financial Studies*, *24*(4), 983–1018.

Altavilla, C., Boucinha, M., & Peydró, J. L. (2018). Monetary policy and bank profitability in a low interest rate environment. *Economic Policy, 33*(96), 531-586.

Arner, D. W., Barberis, J., & Buckley, R. P. (2017). The evolution of Fintech: A new post-crisis paradigm. *Georgetown Journal of International Law*, *47*, 1271.

Arnold, I. J., & van Ewijk, S. E. (2011). Can pure play internet banking survive the credit crisis?. *Journal of Banking & Finance*, *35*(4), 783-793.

Athanasoglou, P. P., Brissimis, S. N., & Delis, M. D. (2008). Bank-specific, industry-specific and macroeconomic determinants of bank profitability. *Journal of International Financial Markets Institutions and Money*, *18*(2), 121–136.

Ayadi, R., Arbak, E., & Pieter De Groen, W. (2011). *Business Models in European Banking: A pre- and post-crisis screening*. Center for European Policy Studies, Brussels.

BCBS. (2018). S*ound practices: Implications of fintech developments for banks and bank supervisors*. Basel Committee on Banking Supervision, Bank for International Settlements. 19 February 2018, Basel.

Berger, A. N. (1995). The Relationship between Capital and Earnings in Banking. *Journal of Money, Credit and Banking*, *27*(2), 432–456.

Berndt, A., & Gupta, A. (2009). Moral hazard and adverse selection in the originate-to-distribute model of bank credit. *Journal of Monetary Economics*, *56*(5), 725-743.

Bertay, A. C., Demirgüç-Kunt, A., & Huizinga, H. (2015). Bank ownership and credit over the business cycle: Is lending by state banks less procyclical? *Journal of Banking & Finance*, *50*, 326–339.

Bikker, J., & Bos, J. W. (2008). *Bank Performance: A theoretical and empirical framework for the analysis of profitability, competition and efficiency*. Routledge.

Boot, A. W. A. (2000). Relationship banking: What do we know? *Journal of Financial Intermediation*, *9*(1), 7–25.

Boot, A. W., & Thakor, A. V. (1994). Moral hazard and secured lending in an infinitely repeated credit market game. *International Economic Review*, 899-920.

Boot, A., Hoffmann, P., Laeven, L., & Ratnovski, L. (2021). Fintech: what's old, what's new?. *Journal of Financial Stability*, *53*, 100836.

Broecker, T. (1990). Credit-worthiness tests and interbank competition. *Econometrica 58*, 429–452.

Buchak, G., Matvos, G., Piskorski, T., & Seru, A. (2018). Fintech, regulatory arbitrage, and the rise of shadow banks. *Journal of Financial Economics, 130*(3), 453-483

Calem, P., & Rob, R. (1999). The impact of capital-based regulation on bank risk-taking. *Journal of Financial Intermediation*, *8*(4), 317–352.

Carbó-Valverde, S., Cuadros-Solas P.J., & Rodríguez-Fernández F., (2021). FinTech and Banking: An Evolving Relationship. In King, T., Lopes, F. S. S., Srivastav, A., & Williams, J. (eds). (2021). *Disruptive Technology in Banking and Finance* (pp. 161-194). Palgrave Macmillan, Cham.

Castro, V. (2013). Macroeconomic determinants of the credit risk in the banking system: The case of the GIPSI. *Economic Modelling*, *31*, 672–683.

CCAF. (2022). *Cambridge FinTech Ecosystem Atlas*. Cambridge Centre for Alternative Finance (retrieved from: https://ccaf.io/).

Chiorazzo, V., D'Apice, V., DeYoung, R., & Morelli, P. (2018). Is the traditional banking model a survivor?. *Journal of Banking & Finance*, *97*, 238-256.

Claessens, S., Frost, J., Turner, G., & Zhu, F. (2018). *Fintech credit markets around the world: size, drivers and policy issues*. BIS Quarterly Review, September 2018.

Clougherty, J. A., Duso, T., & Muck, J. (2016). Correcting for self-selection based endogeneity in management research: Review, recommendations and simulations. *Organizational Research Methods,* 19(2), 286-347.

Curado, C. (2008). Perceptions of knowledge management and intellectual capital in the banking industry. *Journal of Knowledge Management*, *12*(3), 141–155.

Dagher, J. C., Dell'Ariccia, G., Laeven, L., Ratnovski, L., & Hui, T. (2020). Bank capital: a seawall approach. *International Journal of Central Banking*, *16*(2), 249-291.

Degryse, H., & Ongena, S. (2005). Distance, lending relationships, and competition. *The Journal of Finance*, *60*(1), 231-266.

Demirgüc-Kunt, A., & Huizinga, H. P. (2000). *Financial structure and bank profitability*. World Bank Policy Research Working Paper, No. 2430. August 2000.

Dewatripont, M., & Tirole, J. (1994). *The prudential regulation of banks*. Cambridge, MA: MIT press.

Di Maggio, M., & Yao, V. (2021). Fintech borrowers: Lax screening or cream-skimming? *The Review of Financial Studies*, *34*(10), 4565–4618.

Diamond, D. W. (1984). Financial intermediation and delegated monitoring. *The Review of Economic Studies*, *51*(3), 393-414.

Diamond, D. W., & Dybvig, P. H. (1983). Bank runs, deposit insurance, and liquidity. *Journal of Political Economy*, *91*(3), 401-419.

Dietrich, A., & Wanzenried, G. (2011). Determinants of bank profitability before and during the crisis: Evidence from Switzerland. *Journal of International Financial Markets Institutions and Money*, *21*(3), 307–327.

Drechsler, I., Savov, A., & Schnabl, P. (2021). Banking on deposits: Maturity transformation without interest rate risk. *The Journal of Finance*, *76*(3), 1091-1143.

EBA (2021). *Risk dashboard data as of Q4 2020.* European Banking Authority, Paris.

ECB (2018). *Guide to assessments of fintech credit institution licence applications.* European Central Bank, March 2019, Frankfurt.

ECB. (2020). *Risk report on less significant institutions*. European Central Bank, January 2020, Frankfurt.

Ehrentraud, J., Ocampo, D., Garzoni, L., & Piccolo, M. (2020). *Policy responses to fintech a cross-country overview*. Bank for International Settlements. FSI Insights on Policy Implementation, No 23. Basel. January 2020.

Elsas, R., Hackethal, A., & Holzhäuser, M. (2010). The anatomy of bank diversification. *Journal of Banking & Finance*, *34*(6), 1274–1287.

Enria, A. (2021). *How can we make the most of an incomplete banking union?*. Speech by Andrea Enria, Chair of the Supervisory Board of the ECB, at the Eurofi Financial Forum. Ljubljana, 9 September 2021

ESRB (2014). (2014). *Is Europe overbanked?*. European Systemic Risk Board, Reports of the Advisory Scientific Committee, No. 4, June 2014.

Ewijk, S., & Arnold, I. (2014). How bank business models drive interest margins: Evidence from US bank-level data. *The European Journal of Finance*, *20*(10), 850-873.

Frost, J., Gambacorta, L., Huang, Y., Shin, H., & Zbinden, P. (2019). *BigTech and the changing structure of financial intermediation.* BIS Working Papers no. 779.

FSB (2019). *BigTech in finance: Market developments and potential financial stability implications*. Financial Stability Board. 9 December 2019.

FT. (2019). *Digital banks seek to turn popularity into profits*. Financial Times, London, UK. July 6, 2019.

FT. (2020). *Digital banks win UK-EU expats' accounts*. Financial Times, London, UK. December 18, 2020.

Fuster, A., Plosser, M., Schnabl, P., & Vickery, J. (2019). The role of technology in mortgage lending. *The Review of Financial Studies, 32*(5), 1854-1899.

Gontarek, W. (2021). Digital Disruption: How the Financial Services Landscape Is Being Transformed. In *Disruptive Technology in Banking and Finance* (pp. 221-240). Palgrave Macmillan, Cham.

Gorton, G. B., & Pennacchi, G. (1988). *Transaction Contracts*, 34-88. Wharton School Rodney L. White Center for Financial Research.

Gurley, J.G. and Shaw, E.S. (1960) *Money in Theory of Finance*. Brookings, Washington DC.

Harrel, F. E. (2001). *Regression Modeling Strategie*s. New York: Springer Series in Statistics.

Holmstrom, B., & Tirole, J. (1997). Financial intermediation, loanable funds, and the real sector. *The Quarterly Journal of Economics*, *112*(3), 663–691.

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics, 3*(4), 305–360.

King, T., Previati, D.A. (2021). FinTech Cultures and Organizational Changes in Financial Services Providers. In King, T., Lopes, F. S. S., Srivastav, A., Williams, J. (eds). (2021). *Disruptive Technology in Banking and Finance* (pp. 195-219). Palgrave Macmillan, Cham.

Leask, G., & Parnell, J. A. (2005) Integrating Strategic Groups and the Resource Based Perspective: Understanding the Competitive Process. *European Management Journal, 23*(4), 458-470.

Liberti, J. M., & Mian, A. R. (2009). Estimating the effect of hierarchies on information use. *The Review of Financial Studies*, *22*(10), 4057–4090.

Locatelli, R., Schena, C., & Tanda, A. (2021). A historical perspective on disruptive technologies. In King, T., Lopes, F. S. S., Srivastav, A., & Williams, J. (2021). *Disruptive Technology in Banking and Finance* (pp. 9-46). Palgrave Macmillan, Cham.

Mangold, R. (2013). Determinants of bank advisor revenues. *International Journal of Productivity and Performance Management*, *62*(7), 674–693.

Marques, B. P., & Alves, C. F. (2020). Using clustering ensemble to identify banking business models. *Intelligent Systems in Accounting, Finance and Management*, *27*(2), 66-94.

Marques, B. P., & Alves, C. F. (2021). The profitability and distance to distress of European banks: do business choices matter?. *The European Journal of Finance, 27*(15), 1553-1580.

Mergaerts, F., & Vander Vennet, R. (2016). Business models and bank performance: A long-term perspective. *Journal of Financial Stability*, *22*, 57-75.

Merton, R. C. (1995). A functional perspective of financial intermediation. *Financial management, 24*, 23-41.

Milne, A., & Parboteeah. P. (2016). *The Business Models and Economics of Peer-to-Peer Lending*. European Credit Research Institute Research Report, No. 17. May 5, 2016.

Nel, J., & Boshoff, C. (2021). Traditional-bank customers' digital-only bank resistance: evidence from South Africa. *International Journal of Bank Marketing*, *39*(3), 429–454.

Nouthoofd, N. & Heene, A. (1997). Strategic groups as subsets of strategic scope groups in the Belgian brewing industry. *Strategic Management Journal*, *18*(8), 653–666.

OECD. (2020). *Digital disruption in banking and its impact on competition*. Organization for Economic Cooperation and Development, Paris. 5 June 2019.

Oehler, A., Horn, M., & Wendt, S. (2021). Investor Characteristics and their Impact on the Decision to use a Robo-advisor. *Journal of Financial Services Research*, 1-35.

Perotti, E., Ratnovski, L., & Vlahu, R. (2011). Capital regulation and tail risk. *International Journal of Central Banking, 7*(4), 123-163.

Persyn, D., Diaz-Lanchas, J., Barbero, J., Conte, A., & Salotti, S. (2020). *A new dataset of distance and time related transport costs for EU regions*. Joint Research Centre, JRC Working Papers No. JRC119412. Seville.

Petralia, K., Philippon, T., Rice, T. N., & Véron, N. (2019). *Banking Disrupted?: Financial Intermediation in an Era of Transformational Technology*. ICMB International Center for Monetary and Banking Studies.

Pringle, J. J. (1975). Bank capital and the performance of banks as financial intermediaries: comment. *Journal of Money, Credit and Banking*, *7*(4), 545-550.

Puri, M., Rocholl, J., & Steffen, S. (2017). What do a million observations have to say about loan defaults? Opening the black box of relationships. *Journal of Financial Intermediation*, *31*, 1–15.

Rajan, R. G. (1992). Insiders and outsiders: The choice between informed and arm's-length debt. *The Journal of finance*, *47*(4), 1367-1400.

Reger, R. K., & Huff, A. S. (1993). Strategic groups: A cognitive perspective. *Strategic Management Journal*, *14*(2), 103-123.

Saffe. (2022). *Who we are*. (Retrieved from: https://www.saffe.ai/who-we-are/).

Scholes, M., Benston, G., & Smith, C. (1976). A transactions cost approach to the theory of financial intermediation. *The Journal of Finance*, *31*(2), 215-231.

Shaffer, S. (1998). The winner's curse in banking. *Journal of Financial Intermediation*, *7*(4), 359–392.

Sharpe, S. A. (1990). Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships. *The journal of finance*, *45*(4), 1069-1087.

Stulz, R. M. (2019). Fintech, bigtech, and the future of banks. *Journal of Applied Corporate Finance*, *31*(4), 86-97.

Tabak, B. M., Fazio, D. M., Ely, R. A., Amaral, J. M., & Cajueiro, D. (2017). The effects of capital buffers on profitability: An empirical study. *Economics Bulletin*, *37*(3), 1468-1473.

Tanda, A., Schena, C.M. (2019). *Fintech, Bigtech and Banks. Digitalisation and its Impact on Banking Business Models.* Palgrave Macmillan.

Thakor, A. V. (2020). Fintech and banking: What do we know?. *Journal of Financial Intermediation*, *41*, 100833.

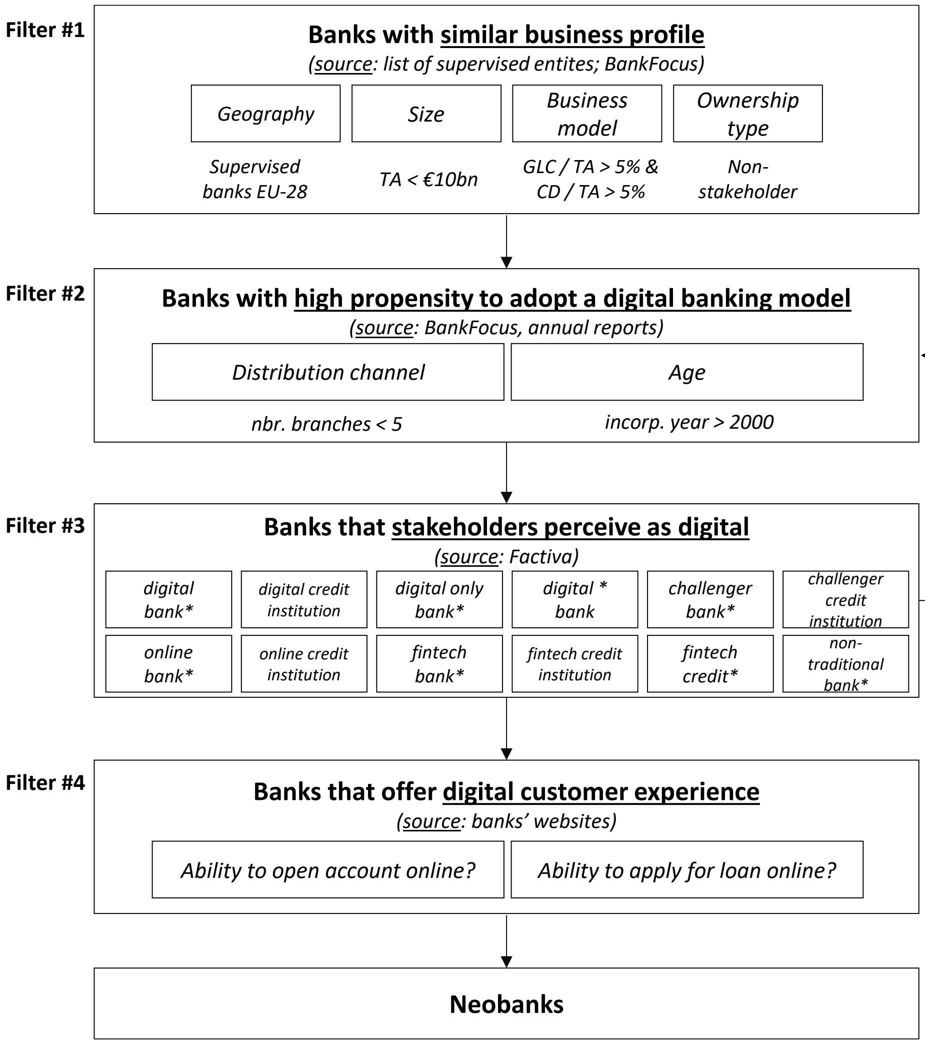
Venturelli, V., Landi, A., Ferretti, R., Cosma, S., & Gualandri, E. (2021). How does the financial market evaluate business models? Evidence from European banks. *Economic Notes*, e12184.

White, H., & Domowitz, I. (1984). Nonlinear regression with dependent observations. *Econometrica: Journal of the Econometric Society*, *52*(1), 143.

Williams, J. (2021). Fintech—A Perfect Day or Walk on the Wild Side?. In King, T., Lopes, F. S. S., Srivastav, A., & Williams, J. (eds). (2021). *Disruptive Technology in Banking and Finance* (pp. 161-194). Palgrave Macmillan, Cham.

Figures

Figure 1. Method to identify neobanks



Tables

Table 1. Identification of neobanks: results

Criterion		Nbr. of banks
Step 0	List of all supervised entities in EU-28 countries as of 2019	3 582
Step 1	Banks with homogeneous business profile (source: BankFocus) • *Size*: Total assets < 10 bn • *Business model*: gross loans to customers / total assets > 5% & customer deposits / total assets > 5% • *Ownership type*: exclude cooperatives and savings banks	696
Step 2	Banks with high propensity to adopt a digital banking model (source: BankFocus, annual reports) • *Distribution channel*: Number of branches < 5 • *Age*: incorporation year > 2000	172
Step 3	Banks that stakeholders perceive as adopting a digital banking model (source: Factiva)	58
Step 4	Banks with online functionalities typical of a digital banking model (source: banks' websites) • Ability to open account online; or • Ability to apply for loan online.	55
	Neobanks	**55**

Table 2. Distribution of banks per country

Countries	Nbr. of banks		
	Neobanks	Traditional banks	Total
Belgium	1	4	5
Czech Republic	1	4	5
Denmark	1	7	8
Estonia	2	3	5
France	4	32	36
Germany	3	36	39
Greece	1	3	4
Italy	8	15	23
Lithuania	3	3	6
Luxembourg	3	6	9
Malta	2	3	5
Netherlands	2	3	5
Poland	1	4	5
Portugal	3	3	6
Spain	1	5	6
Sweden	7	7	14
United Kingdom	12	20	32
Total	**55**	**158**	**213**

Table 3. Main features of neobanks and traditional banks: descriptive statistics and mean comparisons

	Obs.	Mean	Std. deviation	Min	Median	Max
Panel A. Digital banks						
Number of branches	55	0.6***	1.2	0.0	0.0	5.0
Age	55	11.2***	6.1	0.0	11.0	21.0
Online (account/loan)	55	1.0***	0.0	1.0	1.0	1.0
Total assets	55	14.1**	1.3	11.2	14.4	16.1
Income diversification	55	33.7**	19.6	0.1	34.5	67.2
Number of products	55	4.7***	2.2	1.0	4.0	11.0
Gross loans to customers	55	48.7*	26.5	5.1	49.3	93.8
Liquid assets	55	38.7**	25.8	4.2	31.8	85.9
Customer deposits	55	72.2	19.9	13.3	77.1	94.3
Total equity	55	10.3	6.6	2.7	8.9	30.0
Non-performing loans	55	6.7*	10.0	0.0	2.9	47.9
Panel B. Traditional banks						
Number of branches	158	23.8***	39.8	0.0	10.0	319.0
Age	158	71.1***	64.7	3.0	45.0	348.0
Online (account/loan)	158	0.2***	0.4	0.0	0.0	1.0
Total assets	158	14.5**	1.0	11.8	14.6	16.1
Income diversification	158	44.0**	15.4	0.2	46.6	70.1
Number of products	158	6.3***	2.6	1.0	10.0	11.0
Gross loans to customers	158	56.2*	23.6	5.1	59.1	93.8
Liquid assets	158	30.3**	19.6	2.2	27.3	85.0
Customer deposits	158	71.7	17.3	13.3	75.5	94.3
Total equity	158	9.5	4.7	2.7	8.2	30.0
Non-performing loans	158	4.5*	7.4	0.0	2.3	47.9

Note: In the column "Mean", we show the results of Tuckey HSD test for the comparison of means between neobanks and traditional banks. ∗∗∗, ∗∗ and ∗ indicate statistical significance at the 1%. 5% and 10% level respectively. All financial variables are divided by total assets except for income diversification and total assets (natural logarithm).

Table 4. Performance of neobanks and traditional banks: mean comparisons

	Neobanks	Traditional	Diff.
ROA	-0.225	0.371	-0.596**
Interest income	3.939	2.084	1.855***
Interest expenses	0.755	0.530	0.225***
Net interest income	3.055	1.555	1.500***
Non-net interest income	1.259	1.790	-0.531*
Staff expenses	1.527	1.475	0.052
Non-staff operating expenses	2.682	1.386	1.296***
Impairment charges	0.972	0.281	0.691***

Note: Mean values per type of bank. We perform Tuckey HSD test for the comparison of means. ∗∗∗, ∗∗ and ∗ indicate statistical significance at the 1%. 5% and 10% level respectively.

Table 5. Performance of neobanks and traditional banks: baseline regressions

	ROA (1)	II (2)	IE (3)	NII (4)	NNII (5)	SE (6)	NSE (7)	IC (8)
Neobank dummy	**-0.683****	**1.445*****	**0.108**	**1.235*****	**-0.871****	**-0.322***	**0.770***	**0.585****
Total assets	-0.012	-0.086	-0.004	-0.057	-0.378***	-0.423***	-0.226*	0.014
Income diversification	-0.015**	-0.036***	-0.005**	-0.029***	0.010	-0.001	-0.004	-0.006
Number of products	0.023	-0.066	-0.044***	-0.023	-0.087*	-0.063**	-0.130***	-0.006
Gross loans to customers	0.015**	0.041***	0.002	0.036***	-0.033***	-0.012**	-0.011	0.010
Liquid assets	0.014	-0.001	-0.004	0.002	0.000	-0.000	0.000	0.002
Customer deposits	-0.000	0.002	-0.002	0.005	-0.011	-0.005	-0.002	0.001
Total equity	0.065***	0.109***	0.003	0.096***	0.061**	0.062***	0.073***	0.018
Non-performing loans	-0.068***	0.039	0.009***	0.030	-0.034**	0.014	0.044**	0.046**
Observations (neobanks)	213 (55)	213 (55)	213 (55)	213 (55)	213 (55)	213 (55)	213 (55)	213 (55)
R-square	0.252	0.552	0.303	0.532	0.350	0.394	0.319	0.244

Note: Values presented are the coefficient estimates of cross-section OLS regressions with bank and country controls, using White-robust standard errors. All variables are divided by total assets except for income diversification and cost to income (both divided by operating revenues) and total assets (natural logarithm). ∗∗∗, ∗∗ and ∗ indicate statistical significance at the 1%. 5% and 10% level respectively. Dependent variables: II - interest income, IE - interest expenses, NII - net interest income, NNII - other income, SE - staff expenses, NSE – non-staff operating expenses, IC - total impairment charges.

Table 6. Performance of neobanks and traditional banks: economies of scale, experience, and scope

	TA > €200M (1)	TA > €400M (2)	TA > €600M (3)	TA > €800M (4)	TA > €1,000M (5)	TA > €2,000M (6)
Panel A. Scale						
Neobanking dummy	**0.838****	**0.963****	**1.227****	**1.286****	**1.480****	**1.647****
Observations (neobanks)	202 (50)	182 (40)	159 (35)	149 (33)	110 (27)	84 (21)
R-square	0.268	0.187	0.212	0.218	0.243	0.315
	Age > 1y (7)	**Age > 2y** (8)	**Age > 3 y** (9)	**Age > 4y** (10)	**Age > 5y** (11)	**Age > 10y** (12)
Panel B. Experience						
Neobanking dummy	**-0.584***	**-0.650***	**-0.702****	**-0.590***	**-0.486**	**-0.497**
Observations (neobanks)	211 (53)	208 (50)	206 (49)	202 (46)	197 (41)	182 (30)
R-square	0.229	0.247	0.256	0.264	0.254	0.315
	NP > 1 (13)	**NP > 2** (14)	**NP > 3** (15)	**NP > 4** (16)	**NP > 5** (17)	**NP > 6** (18)
Panel C. Scope						
Neobanking dummy	**0.764***	**0.754***	**0.715**	**-0.185**	**0.383**	**0.132**
Observations (neobanks)	209 (54)	190 (46)	165 (38)	133 (27)	111 (14)	91 (10)
R-square	0.311	0.249	0.220	0.301	0.282	0.259

Note: Values presented are the coefficient estimates of cross-section OLS regressions with bank and country controls, using White-robust standard errors. **The dependent variable is 'non-staff expenses to total assets'**. For brevity reasons we only show the coefficient for the main coefficient of interest. ∗∗∗, ∗∗ and ∗ indicate statistical significance at the 1%. 5% and 10% level respectively. TA – Total assets, NP – Number of Products

Table 7. Performance of neobanks and traditional banks: sub-sample regressions per business line

	ROA (1)	**II** (2)	**IE** (3)	**NII** (4)	**NNII** (5)	**SE** (6)	**NSE** (7)	**IC** (8)
Panel A. Credit cards								
Neobank dummy	**-0.800**	**1.059****	**0.079**	**0.911****	**-0.829***	**-0.360**	**0.721**	**0.656***
Observations (neobanks)	103 (23)	103 (23)	103 (23)	103 (23)	103 (23)	103 (23)	103 (23)	103 (23)
R-square	0.257	0.699	0.424	0.690	0.332	0.325	0.375	0.334
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Panel B. Personal loans								
Neobank dummy	**-1.293*****	**0.978****	**0.046**	**0.876***	**-0.556**	**-0.060**	**1.264****	**0.604***
Observations (neobanks)	118 (29)	118 (29)	118 (29)	118 (29)	118 (29)	118 (29)	118 (29)	118 (29)
R-square	0.341	0.673	0.418	0.621	0.289	0.427	0.385	0.165
	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)
Panel C. Broker and advisory								
Neobank dummy	**-0.326**	**0.439***	**0.175***	**0.263**	**-1.447*****	**-0.797*****	**-0.569****	**0.216**
Observations (neobanks)	143 (22)	143 (22)	143 (22)	143 (22)	143 (22)	143 (22)	143 (22)	143 (22)
R-square	0.245	0.470	0.135	0.482	0.435	0.426	0.403	0.082

Note: Values presented are the coefficient estimates of cross-section OLS regressions with bank and country controls, using White-robust standard errors. **The dependent variable is ROA.** For brevity reasons we only show the coefficient for the main coefficient of interest. ∗∗∗, ∗∗ and ∗ indicate statistical significance at the 1%. 5% and 10% level respectively. Dependent variables: II - interest income, IE - interest expenses, NII - net interest income, NNII - other income, SE - staff expenses, NSE – non-staff operating expenses, IC - total impairment charges.

Table 8. Performance of top neobanks and bottom neobanks: mean comparisons

	Top Neobanks	**Bottom Neobanks**	**Diff.**
ROA	1.532	-1.919	3.451***
Interest income	5.017	2.899	2.271*
Interest expenses	0.889	0.625	0.264
Net interest income	3.841	2.297	1.544*
Non-net interest income	1.209	1.308	-0.099
Staff expenses	1.365	1.683	-0.318
Non-staff operating expenses	1.889	3.446	-1.557**
Impairment charges	0.942	1.001	-0.059

Note: Mean values per type of bank. We perform Tuckey HSD test for the comparison of means. ∗∗∗, ∗∗ and ∗ indicate statistical significance at the 1%. 5% and 10% level respectively.

Table 9. Performance of neobanks and traditional banks: market factors

	ROA			**IC**		
	High	**Low**	**Diff.**	**High**	**Low**	**Diff.**
Panel A. Total bank assets to GDP						
Neobanking dummy	**-0.872***	**-0.617**	**-0.255**	**0.906****	**0.052**	**0.854**
Observations	115 (32)	98 (23)		115 (32)	98 (23)	
R-square	0.178	0.456		0.368	0.245	
Panel B: Bank credit to deposits						
Neobanking dummy	**-0.805***	**-0.556**	**-0.249**	**0.649****	**0.533**	**0.116**
Observations	110 (32)	103 (23)		110 (32)	103 (23)	
R-square	0.309	0.243		0.410	0.218	

Note: Values presented are the coefficient estimates of cross-section OLS regressions with bank and country controls, using White-robust standard errors. For brevity reasons we only show the coefficient for the main coefficient of interest. ∗∗∗, ∗∗ and ∗ indicate statistical significance at the 1%. 5% and 10% level respectively. IC – Impairment charges.

Table 10. Endogeneity in the choice of business model: IV regressions

	First stage regression	**Second stage regression**
	Neobanks dummy	ROA
	(1)	(3)
Instrumental variables		
Proximity to knowledge center	-0.0346**	
Quality of knowledge center	0.0348*	
Instrumented variables		
Neobanks dummy		-1.979*
Number of observations	213	213
R-squared	0.250	0.155
F-test of instruments (p-value)	0.000	
Stock-Yogo's min. eigenvalue	5.080	
Wald Chi-square test (p-value)		0.000
Sargan test overid. (p-value)		0.141

Notes: The values presented are the coefficient estimates of cross-section IV regressions with bank and country controls. The 'proximity to knowledge center' is computed as the natural log of road distance (in hours) between the bank's headquarters and the nearest top50 university in the Scimago Institutions Ranking; the 'quality of knowledge center' is computed as the total number of ICT patents recorded by the nearest top50 university. Results reported using robust standard errors. ***, ** and * indicate statistical significance at the 1%. 5% and 10% level, respectively.