Journal of Financial Services Research
https://doi.org/10.1007/s10693-024-00433-x

The Early Days of Neobanks in Europe: Identification, Performance, and Riskiness

Alberto Citterio[1] · Bernardo P. Marques[2,3] · Alessandra Tanda[4]

Received: 17 March 2022 / Revised: 17 May 2024 / Accepted: 11 June 2024

Abstract
This paper identifies banks born with a digital business model ('neobanks') and examines their performance and riskiness *vis-à-vis* traditional peers. We propose a novel approach to identify neobanks, based on non-financial hand-collected data, and identify 65 neobanks operating in Europe. We show that neobanks perform worse than their traditional peers, while recording a similar level of risk. Namely, neobanks charge higher interest income, record higher impairment charges, and face higher non-staff expenses. Further analysis suggests the presence of economies of scale and scope in digital banking. Our findings are robust to endogeneity concerns and changes to our baseline specification.

Keywords Neobanks · Digital banks · Business models · Bank performance

JEL Classification G20 · G21 · G28 · G32

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

✉ Alessandra Tanda
alessandra.tanda@unipv.it

Alberto Citterio
alberto.citterio@uninsubria.it

Bernardo P. Marques
bmarques@fep.up.pt

[1] Department of Economics, University of Insubria, Varese, Italy

[2] CEF.UP and Faculdade de Economia, Universidade do Porto, Porto, Portugal

[3] Católica Porto Business School, Portuguese Catholic University, Porto, Portugal

[4] Department of Economics and Management, University of Pavia, Pavia, Italy

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, generally identified by the previous literature as 'neobanks' (wherein 'neo' stands for 'new') or 'challenger banks' (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? How have they performed relative to traditional peers? What drives the differences in performance?

While a recent strand of the literature has discussed the emergence of these new types of banks (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.[2] Indeed, with reference to the identification of neobanks, the literature has seldomly tackled this issue, mainly relying on the digital nature of the channel employed to offer financial services (DeYoung 2005; Delgado et al. 2007). More recently, the 'Cambridge Centre for Alternative Finance' (CCAF) has identified a set of digital banks labelled as 'fully digitally native banks' (CCAF 2022) via different sources, including company websites, company statements, interviews, and news articles. Nevertheless, by having a closer look at the list of digital banks identified by CCAF, we find evidence that not all entities hold a banking licence.

With respect to the performance of neobanks, theory offers mixed predictions: on one hand, the literature suggests that neobanks may be (i) subject to the 'winners curse' (Broecker 1990; Shaffer 1998), according to which newcomers are likely to be exposed to a riskier pool of potential borrowers, previously rejected by incumbent banks, and (ii) may be more exposed to information asymmetries (Herpfer 2021) and have lower abilities to price discriminate (Degryse and Ongena 2005), due to their transactional lending business model. On the other hand, an alternative theory suggests that neobanks may benefit from a simpler hierarchical structure, coupled with superior IT capabilities, which are likely to lead to a shortened decision process and more timely response to customer requests (Tanda and Schena 2019; Boot et al. 2021; Williams 2021; Berg et al. 2022). Moreover, the empirical 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 and Yao 2021; Carmichael 2017). For instance, findings by Di Maggio and Yao (2021) indicate that fintech lenders seem to be tapping into lower quality borrowers in the personal loans market, previously rejected by the incumbents.

[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 Allen et al. (2002)

[2] Relatedly, the existing literature on banking business models (Mergaerts and Vander 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.

Effectively, given the innovative nature of neobanks, research about the identification 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, we develop an identification procedure that starts from the list of all supervised banking institutions, and explicitly apply four identification filters, related with (i) the business profile of the banks (size, asset and funding structures, and ownership type), (ii) the propensity to adopt a digital banking model (less than or equal to 5 branches, younger or equal to 20 years), (iii) the stakeholders' perception (keywords search in Factiva news), and (iv) the ability to open and account or apply for a loan online (henceforth, 'online functionalities'). In our view, such step-by-step approach has the advantages of being transparent and fully replicable, as well as allowing us to focus on a level-playing field comprised only of supervised banking entities. Nevertheless, it also requires the hand-collection of data from the banks' websites, which is quite costly, and the definition of thresholds for certain features (i.e., size, asset and funding structures, branches, age). To mitigate the latter limitation, we perform a set of robustness checks which provide evidence on the stability of the neobank sample for different thresholds of these features.

To assess the performance of neobanks, we apply 'Propensity Score Matching' (PSM) to a large set of traditional banks and find 313 suitable counterfactuals ('traditional peers') with strong similarities to neobanks with respect to size, asset and funding structures, income diversification, liquid assets, and total equity. Then, we run several analyses to compare the performance of neobanks and traditional peers, including OLS regressions on decomposed elements of Return on Assets (ROA). Furthermore, we uncover the potential presence of economies of scale, experience, and scope, by exploring the heterogeneity of our sample with respect to size, age, and diversification, as well as check whether the baseline results are driven by specific product lines. Finally, we compute a large number of robustness checks, including 2SLS regressions (wherein we develop two instruments related to the distance and quality of the closest technology-related knowledge centres), alternative identification procedures (e.g., including age as matching variable) and regression specifications (e.g., country fixed effects).

Our results indicate the existence of 65 neobanks in the EU-28. The tests for the quality of matching indicate that neobanks are quite similar to traditional peers, except for the features used to identify the neobanks (i.e., age, branches, online functionalities). With respect to performance, our results suggest that neobanks perform worse than traditional peers. In our attempt to uncover the potential drivers of these results, we find that, on one hand, neobanks record a higher level of 'interest income minus impairment charges' than traditional banks do. This finding, in addition to the similar NPL ratio (Herpfer 2021) and the presumably better 'online customer experience' (Buchak et al. 2018; Tanda and Schena 2019; Di Maggio and Yao 2021; Williams 2021; Berg et al. 2022), seems consistent with the 'transactional lending-better customer experience' narrative.

On the other hand, we find that neobanks are more inefficient than traditional peers with respect to non-staff expenses (i.e., IT, advertising, reporting). Further analyses indicate that this effect fades away as we remove very small banks (with less than EUR 600 million in total assets) and specialised banks (with less than 6 product lines) – suggesting the potential presence of economies of scale and scope in digital banking, in line with previous literature (DeYoung 2005; Delgado et al. 2007). Finally, our evidence suggests that the underperformance of neobanks seems mainly driven by banks that offer personal loans – which seems consistent with the notion that hard information can be used more effectively to assess borrower credit worthiness in collateralised loans (e.g., mortgage loans) than in non-collateralised loans (e.g., personal loans) (Stein 2002; DeYoung et al. 2007).

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 and van Ewijk 2011; Mergaerts and Vander 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, our sample of neobanks is significantly larger than that of previous studies. Namely, our paper analyses the performance of 65 neobanks, which compares with 12 and 15 digital banks for DeYoung (2005) and Delgado et al. (2007), respectively. Such increase in sample size 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 and Ongena 2005; Puri and Rocholl 2008; Drechsler et al. 2021; Thakor 2020; Boot et al. 2021). On the information frictions side, the fact that we find neobanks to record 'higher interest income-higher impairment charges', while recording a similar NPL ratio, indicates that such banks are effectively not subject to a 'winners curse' (Shaffer 1998) but are likely to be charging an interest premium as a compensation for higher information asymmetries, while also offering a better customer experience. Regarding communication frictions, we do not find any significant difference in the performance of neobanks relative to traditional peers with respect to interest expenses or non-interest income, suggesting that the 'digital spatial capture' narrative (Boot et al. 2021) cannot be confirmed in the 'early days' of 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 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 peers 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. This paper is accompanied by an Appendix and online supplementary materials.

2 Literature Review

2.1 The Definition and Identification of Neobanks

The literature offers a variety of definitions of 'neobanks'.[3] For instance:

-"Neobanks 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);

-"[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);

-"Neobanks (...) 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; Oehler et al. 2021) –, to build competitive advantage over traditional peers in specific business lines (e.g., lending, payments, trading and asset management), via the provision of a superior customer experience and/or lower costs.

More precisely, in this paper we define 'neobanks' as licensed banking institutions, born during the post internet-era (i.e., after 2000), focused on retail banking activities (as shown by their involvement in lending and deposit-taking activities), and with a strong orientation towards digital distribution channels, rather than physical branches (Ehrentraud et al. 2020; Boot et al. 2021). Our focus on banks with a strong retail orientation allows us to set a level-playing field, in terms of regulatory and supervisory treatment, between the digital banks and traditional peers in our sample, as opposed to using a broader definition of digital banks, closer to 'fintech banks' (Gelis 2016). These would include, for instance, institutions that are mostly focused on offering services to other banks, despite have a banking license, e.g., 'Banking as a Service' (BaaS), offered by Solarisbank (Germany), or Railsbank (UK).

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. Moreover, a set of advances has been made regarding the identification of banking business models, particularly regarding the methods used.

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

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. Using a larger sample of European banks, Mergaerts and Vander 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 characterised 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). With reference to the end of 2021, the database covered 2915 entities, from 108 countries, classified in 14 market segments, 63 sub-segments and 118 categories. Such classification was 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.[4] 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 management 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

[4] 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).

(Merton 1995),[5] 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 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[6] are, in fact, superior – which is corroborated by the empirical findings from Shaffer (1998). An alternative explanation is put forward by the literature on 'relationship-transactional lending', wherein neobanks are likely to adopt a transactional banking model. Under this framework, on one hand, neobanks are likely to offer arms-length contractual relationship, with less access to soft information on borrower creditworthiness, and hence greater exposure to information asymmetries (Herpfer 2021) and lower ability to price discriminate (Degryse and Ongena 2005); on the other hand, the relatively simpler hierarchical structure of neobanks, coupled with superior IT capabilities, allows neobanks to shorten the decision process, resulting in a more timely response to customer requests, for which they are willing to pay an interest rate premium (Tanda and Schena 2019; Williams 2021; Berg et al. 2022). In order for the latter narrative to hold true in our empirical setting, we would expect neobanks to record higher interest income (net of impairment charges), while showing a similar level of riskiness; on the contrary, the 'winner's curse' hypothesis would be consistent with a higher level of risk recorded by the new entrants (neobanks).

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 cross-selling (Puri and Rocholl 2008) and access to cheap and stable funding (Drechsler et al. 2021). 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

[5] An alternative framework to understand the impact of digital banking is the 'functional perspective' put forward by Merton (1995), according to which the effectiveness of financial systems can be assessed based on their ability to perform core financial functions: provide payment services, perform maturity, amount, and spatial transformation, perform risk management, provide price discovery, and reduce information asymmetries. Effectively, at the micro-level, one can pose the question of whether digital banking has specific advantages in performing certain functions, wherein a financial system populated by more digital banks would (in theory) lead to greater social welfare – e.g., by increasing the access to finance, lowering costs for consumers, creating new financial products, improving risk management, reducing systemic risk. For instance, Fuster et al. (2019) study US fintech lenders and find that they do not target borrowers with low access to finance – and hence, in the 'functional perspective' they appear not to improve social welfare via increased access to finance. We note, however, that despite the relevance of this topic, it falls outside the scope of our paper.

[6] See Tseng and Guo (2022) for a comparison of screening incentives in loans by Fintech and traditional intermediaries.

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. In our analysis, such 'digital capture' would be backed by findings consistent with the ability of neobanks to charge higher fees and pay lower interest expenses for customer funding.

Another framework that may help us explain the performance of neobanks is provided by strategic management literature – namely, strategic groups theory (SGT) (Caves and Porter 1977) and resource-based theory (RBT) (Wernerfelt 1984). According to SGT, 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 SGT is that mobility barriers may play an important role in explaining intra-industry performance heterogeneity (Porter 1979). In the same vein, according to RBT an incumbent firm's performance strongly depends on the uniqueness of its resources and capabilities (Wernerfelt 1984), i.e., whether they are rare, difficult to imitate, and difficult to transfer between firms, at least in the short-run (Barney 1991). In banking, some examples of 'unique resources and capabilities' that are difficult to transfer between banks include risk management processes and customer knowledge – in line with the 'soft information' concept, previously discussed. In general, such theories provide us with mixed predictions regarding the performance of neobanks: on one hand, the branch network and longstanding customer relationships of traditional peers may be seen as an entry barrier (or an unique resource) 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 a resource uniquely available to neobanks, or a difficult barrier for traditional peers 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 studying their characteristics and especially focused on analysing their performance relative to traditional peers (Claessens et al. 2002). 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 size only for internet banks. Respectively, such results are interpreted as consistent with the existence of 'economies of scale' 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 loans extended by 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 'better customer experience-higher rates' nexus.

On the other hand, Di Maggio and Yao (2021) 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 Parboteeah 2016; Claessens et al. 2018) are 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 US fintech 'Prosper' provides access to monthly loan-level data, regarding the characteristics and performance of loans. In this context, while

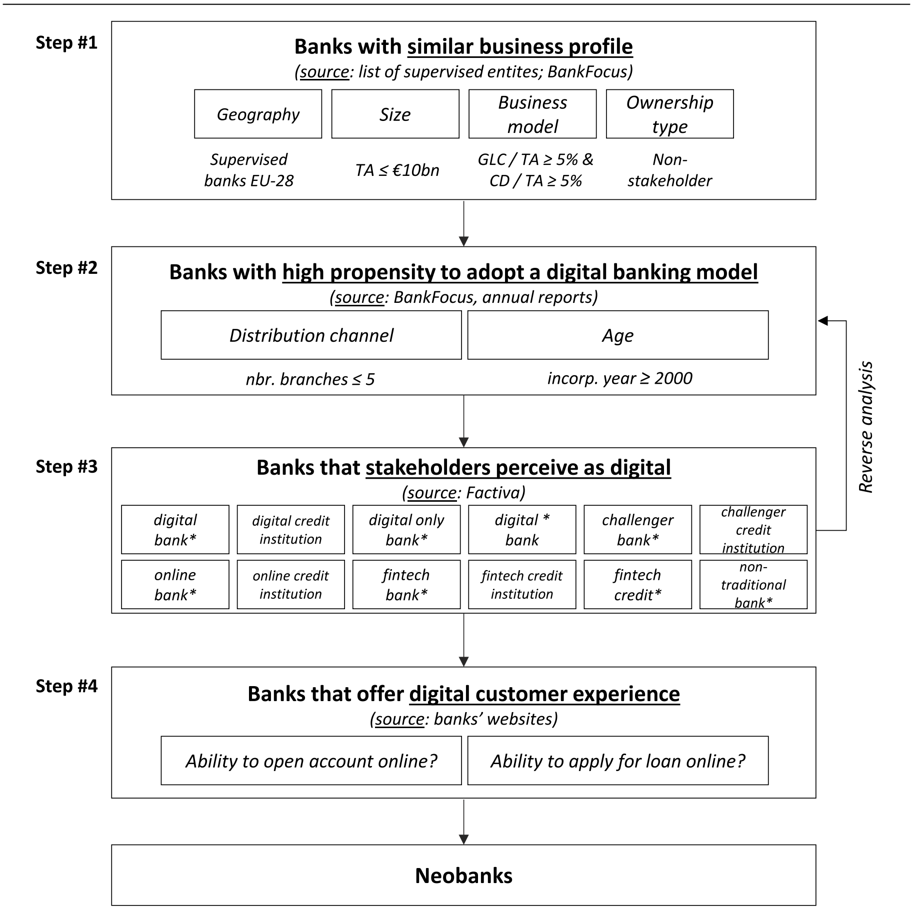

Fig. 1 Method to identify neobanks. This figure presents the step-by-step method applied to identify neobanks. Note that TA refers to total assets, GLC stands for gross loans to customers, and CD is customer deposits

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 supported by the literature. As presented in Fig. 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 neobanks to be newcomers and, as such, to exhibit a relatively small size. Hence, we only consider banks with total assets smaller than EUR 10 bn. The ecosystem of neobanks includes a variety of business lines, including retail lending, payments, trading, 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.[7]

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[8] have a low propensity to adopt a digital banking model, and as such exclude them from the neobanks 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 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.[9]

The third filter identifies banks that stakeholders perceive as being digital. This is done via the systematic analysis of news in Factiva.[10] 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" (FT 2019, 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

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

[8] 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).

[9] Please note that in Section 6 we test for the stability of the identification of neobanks using alternative thresholds of size, asset and funding structure, branches, and age (i.e., filters #1 and #2).

[10] 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).

related to digital banking (*vide* list of keywords in Fig. 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, specialisation, or number of branches).

The fourth and final filter is related to 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 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 Peers

The estimation of the impact of adopting a digital business model on the performance of a bank is bound to be affected by endogeneity issues (Clougherty et al. 2016; Shipman et al. 2017). Namely, the difficulty is to find a suitable counterfactual that allows us to estimate the effect of being born with a fully digital business model (i.e., being a neobank) *vis-à-vis* operating a traditional banking model. First, given that we cannot directly check what the performance of the neobank would be had it not been born as a digital bank. Secondly, because it is likely that bank-specific features may simultaneously determine the likelihood of a bank being born as digital and its potential performance.

To mitigate these issues, we conduct 'Propensity Score Matching' (PSM) on the full sample of banks, wherein the PSM score is computed using the logit model and the matching is performed using the Epanechnikov-Kernel matching function. The propensity score (p) is defined as "the conditional probability of assignment to a particular treatment given a vector of observed covariates" (Rosenbaum and Rubin 1983: p.41) for each bank i:

$$p\left(\mathrm{X}_i\right) = \Pr\left(\,\Omega_i = 1 \mid \mathrm{X}_\mathrm{i}\right) \tag{1}$$

wherein Ω_i is a dummy that takes the value 1 if the bank is classified as 'neobank' and 0 otherwise; and X_i is the set of matching variables which we expect to cumulatively bear some impact on the treatment (neobank dummy) and outcome variables (performance and riskiness), namely: size, gross loans to customers, customer deposits, income diversification, liquid assets, and total equity. Moreover, we force exact matching on bank specialisation and a dummy which takes on the value 1 if the bank cumulatively meets the criteria referred in filter #1 of the neobank identification process (Fig. 1), i.e., total assets <10 bn, gross loans to customers / total assets >5%, customer deposits / total assets >5%, and non-cooperative or savings bank.

Alternatively, one may consider the possibility of matching neobanks with traditional peers also on the number of branches, age, and digital orientation. However, when doing so, the number of matched observations is heavily reduced, thereby raising model overfitting issues; perhaps more importantly, the matched counterfactuals become populated by banks with very specialised business models (e.g., corporate banks, consumer credit banks, subsidiaries of foreign owned banks), hence failing to ensure the external validity of our sample. The latter issue is discussed by Shipman et al. (2017: p.216): "in settings with limited overlap, PSM systematically excludes observations that lack counterfactuals, compromising the degree to which [results] can be generalized outside of the sample". Effectively, in our empirical setting these critical features (branches, age, online functionalities) exhibit a limited overlap between the treated (neobanks) and the pool of potential controls (traditional peers), making it challenging to find similar banks on all relevant covariates without sacrificing external validity. In this context, following the suggestion made by Shipman et al. (2017), we privilege external validity while also ensuring that "alternative specifications yield similar inferences" (p.216) – namely, we iteratively include age and branches as additional matching variables in the PSM, and remove control banks with age or branches beyond specific data-driven thresholds (see Sections 6.2 and 6.3).

Finally, to assess the quality of our matching process, we test whether the mean differences between the treated and controls are significantly reduced in the post- versus pre-matching samples, hence providing evidence on the suitability of the counterfactuals (Rosenbaum and Rubin 1983).

3.3 Measuring the Performance of Neobanks

We assess the performance of neobanks *vis-à-vis* traditional peers by estimating the following model on the matched sample, using cross-section OLS with White (1980) cluster-robust standard errors:

$$Y_i = \alpha_0 + \beta Neobank_i + \gamma BC_i + \delta CL_i + \varepsilon_i \qquad (2)$$

wherein Y_i is the outcome variable, including the Return on Assets (ROA) and the sub-elements of ROA (interest income, interest expenses, net interest income, non-interest income, staff expenses, other non-staff expenses, impairment charges); α_0 is the model constant; $Neobank_i$ is a dummy which takes on the value 1 if bank i is defined as a "neobank" according to our methodology[11]; BC_i is the mean vector of individual bank control variables of bank i (size, income diversification, gross loans to customers, liquid assets, customer deposits, total equity, non-performing loans, number of products); CL_i is a vector of country-level controls (GDP growth, GDP per capita, concentration ratio); β, γ and δ are the regression coefficients' vectors; and ε_i is the disturbance term. We also perform several alternative specifications concerning the choice of outcome variables (Return on Equity - ROE) and country controls (country fixed effects).

Given the relatively small size of our matched sample ($n = 376$), concerns may be raised regarding potential model overfitting and, hence, the validity of statistical inference (Harrel

[11] Additionally, we re-run the baseline regression using the three dummies employed to classify neobanks as the main variables of interest (age equal or below 20 years, number of branches below or equal to 5, and the possibility to open an account or apply for a loan fully online). This approach provides evidence on the precise sub-criteria which drive our neobanks results.

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 (376/13 = 28.9) is above the minimum threshold set in the literature (10).

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 'neobank' is also done during 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 Buchak 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 winsorise 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.[12]

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 (van Ewijk and Arnold 2014) and (ii) our sample comprises only small/medium sized banks (total assets < EUR 10 bn).

Diversification The level of *income diversification* reflects the ability to make money beyond interest-generating services (i.e., credit granting), namely via

[12] The detailed description of variables is presented in Table A, included in the online supplementary materials.

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). In this paper, *income diversification* is computed as a Herfindahl-Hirschman Index; following Elsas et al. (2010), we take the absolute value of each component of total operating revenues, i.e., net interest income, net fees and commissions, net trading income, and other income. However, the literature has also pointed out some drawbacks to this proxy, related to measurement problems. For instance, Laeven and Levine (2007) argue that most credit products algo generate fee income. Hence, for a typical retail focused bank the income diversification measure may overestimate the level of diversification compared to the actual business model operated by the bank. As an additional proxy for diversification, we employ the *number of product lines* referred in the banks' websites. For instance, Fuertes et al. (2010) analyse the interest rate transmission to UK bank deposits and credit, and find significant non-linear patterns with respect to the retail pricing, which can be explained partly by product range – measured as the number of products offered by each bank.[13]

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

[13] Note also that the correlation between *income diversification* and the *number of products* is 0.35 for our matched sample (see Table B, supplementary materials), suggesting that the information content of both proxies does not entirely overlap.

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 level 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 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 concentration ratio, defined as the total market share held by the top3 banks in terms of total assets held. In this regard, Bikker and Haaf (2002) provide some evidence suggesting that bank concentration impairs bank competition, hence improving the overall performance for incumbents.

Economic Development Finally, as proxies for the economic development of each country, we use the *GDP* per capita. Following Grigorian and Manole (2006), we suspect that higher income countries in the EU may be expected to generate higher savings and deposits, which could influence banks' performance.

5 Results

5.1 How Many Neobanks Exist in Europe and What Are their Key Features?

Table 1 shows the step-by-step results of our identification process. We start with a total of 3342 supervised banks in EU-28 countries as of 2019, with available financial data. 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 specialisation in order built a sample of banks with similar business profile and ownership type. These criteria lead to a sample of 553 banks – corresponding to a reduction of approximately 80% of the initial sample – mainly due to the exclusion of cooperative and saving banks, which contribute to a reduction of 70% of the starting list of banks. Second, we further restrict the sample excluding banks with more than 5 branches or incorporated before the year 2000. Step 2 then allows us to define a group of potential neobanks comprised of 135 financial institutions. The group of 68 banks identified in the third step reflects both the direct and the reverse analysis performed via Factiva. Namely, in the direct analysis of the market perception we find 43 neobanks; whereas the reverse analysis yields an additional 25 neobanks. Finally, to further confirm the digital footprint of these banks, we assess the bank's online functionalities by excluding banks that do not offer the possibility to open an

Table 1 Identification of neobanks: results. This table reports the number of banks filtered in each step of the neobank identification process, as described in Fig. 1

Criterion		Nbr. of banks
Step 0	List of all supervised entities in EU-28 countries as of 2019 with available financial data	3342
Step 1	Banks with homogeneous business profile (source: BankFocus) *Size*: Total assets $\leq$€10 bn *Business model*: gross loans to customers / total assets $\geq$5% & customer deposits / total assets $\geq$5% *Ownership type*: exclude cooperative and savings banks	553
Step 2	Banks with high propensity to adopt a digital banking model (source: BankFocus, annual reports) *Distribution channel*: number of branches $\leq$5 *Age*: incorporation year $\geq$2000	135
Step 3	Banks that stakeholders perceive as adopting a digital banking model (source: Factiva)	68
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.	65
	Neobanks	**65**

Table 2 Matching quality. This table reports the results of the t-test for the comparison of means between neobanks and traditional banks pre- and post-PSM. We report the results for the six variables used as input in the PSM procedure: *Size* is the natural logarithm of total assets (EUR thousands), *Income diversification* is computed as a Herfindahl-Hirschman Index (HHI). Following Elsas et al. (2010), we take the absolute value of each component of total operating revenues (TOR), i.e., net interest income (NII), net fees and commissions (NFC), net trading income (NTI), other income (OTH) and apply the following formula: $[1 - [(NII/TOR)^2 + (NFC/TOR)^2 + (NTI/TOR)^2 + (OTH/TOR)^2]]$. The remaining variables, i.e., *Gross loans to customers*, *Liquid assets*, *Customer deposits* and *Total equity*, are divided by total assets and multiplied by 100

	Pre-PSM			Post-PSM		
	Neobanks	Traditional	Diff.	Neobanks	Traditional	Diff.
Number of banks	65	3277		63	313	
Size	14.1	13.8	0.3	14.0	14.0	0.0
Income diversification	32.8	47.0	−14.2***	33.7	44.9	−11.2***
Gross loans to customers	53.2	59.8	−6.6**	54.0	55.0	−1.0
Liquid assets	36.2	24.1	12.1***	35.2	35.4	−0.2
Customer deposits	72.4	65.7	6.7***	71.9	67.7	4.2
Total equity	11.8	11.2	0.6	11.6	12.1	−0.5

Statistical significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively

account or apply for a loan via their website. In the end we obtain a total of 65 neobanks operating in Europe.[14]

Regarding the identification of controls, the implementation of the PSM produces a sample of 313 traditional peers, significantly reducing the original pool of potential controls ($n = 3277$). On the other hand, the matching process leads us to drop 2 unmatched neobanks, resulting in 63 neobanks in the post-PSM sample. Moreover, the results for the quality of matching, presented in Table 2, show that the PSM improves the overall quality of controls. Namely, the comparison of mean values of the matching variables of neobanks and traditional peers, pre- and post-matching, shows that PSM reduces the mean differences in all variables, almost all of which becoming statistically insignificant, except one (income diversification). In general, we take the results of the PSM procedure as providing an important source of confidence regarding the similarity of the neobanks and traditional peers with respect to key determinants of bank performance and riskiness.

In Table 3 we provide descriptive statistics for the key features of neobanks and traditional peers. As expected, on average, neobanks exhibit almost no branches (0.7 branches), are relatively young (10.8 years), and they all provide online functionalities. Moreover, neobanks have a relatively small balance sheet (EUR 2.5 billion of total assets) and are mostly oriented towards lending (54.0% of total assets) and deposit-taking (71.9%). Coherently, they show a low product diversification (4.5 products offered). When comparing neobanks to traditional peers, as expected we do not find significant differences in most of the key features – *vide* column (1). An important exception is related to the three variables used in the neobanks' identification process: branches, age, and online functionalities. In particular, traditional peers record a significantly larger branch network (18.2 branches) and are significantly older (47.8 years) than neobanks. Additionally, only few traditional

[14] The list of neobanks is available in the Appendix (Table 11).

Table 3 Descriptive statistics of the matched sample. This table reports the descriptive statistics for key features of neobanks and traditional peers. In the column "Mean", we show the results of the t-test for the comparison of means between neobanks and traditional peers. *Branches* refers to the number of branches. *Age* is the number of years between the bank's incorporation year and 2020. *Online* is a dummy which takes the value 1 if the bank offers the possibility for customers to open an account or apply for a loan fully online, i.e., without human intervention. We define *Size* as the natural logarithm of total assets (EUR thousands). *Income diversification* is computed as a Herfindahl-Hirschman Index (HHI). Following Elsas et al. (2010), we take the absolute value of each component of total operating revenues (TOR), i.e., net interest income (NII), net fees and commissions (NFC), net trading income (NTI), other income (OTH) and apply the following formula: $[1 - [(NII/TOR)^2 + (NFC/TOR)^2 + (NTI/TOR)^2 + (OTH/TOR)^2]]$. The variables *Gross loans to customers*, *Liquid assets*, *Customer deposits* and *Total equity* are divided by total assets and multiplied by 100. *Number of products* is computed as the number of products reported in each bank's website, from the following list: accounts (current, savings, time), payments (debit cards, credit cards), loans (mortgages, personal, corporate), and services (insurance, trading/broker/custody, advisory). *Non-performing loans* is divided by gross loans to customers and multiplied by 100

	Obs.	Mean	SD	Min	Median	Max
Panel A: Neobanks						
Branches	63	0.7***	1.2	0	0	5
Age	63	10.8***	6.0	0	10	20
Online	63	1.0***	0.0	1	1	1
Size	63	14.0	1.3	10.6	14.4	16.2
Gross loans to customers	63	54.0	26.5	5.0	59.2	94.1
Customer deposits	63	71.9	19.6	7.3	78.2	94.6
Number of products	63	4.5***	2.1	1.0	4.0	11.0
Income diversification	63	33.6***	19.0	0.1	31.4	67.2
Liquid assets	63	35.4	25.5	4.1	29.5	96.5
Total equity	63	12.0	11.8	2.7	8.9	65.7
Non-performing loans	63	8.7	15.8	0.0	2.9	91.1
Panel B: Traditional peers						
Branches	313	18.2***	46.2	0	4	560
Age	313	47.8***	47.8	2	30	383
Online	313	0.1***	0.3	0	0	1
Size	313	14.0	1.2	10.2	14.1	16.1
Gross loans to customers	313	55.2	23.2	5.0	56.8	138.1
Customer deposits	313	67.6	20.7	5.0	73.5	95.2
Number of products	313	6.3***	2.8	1.0	6.0	11.0
Income diversification	313	44.9***	16.3	0.2	48.8	74.2
Liquid assets	313	35.4	20.7	1.8	33.4	91.9
Total equity	313	12.3	9.1	1.8	10.7	94.4
Non-performing loans	313	6.3	9.1	0.0	3.7	57.9

Statistical significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively

peers offer online functionalities (9.3%). Importantly, we do not find any statistically significant difference in the non-performing loans of neobanks *vis-à-vis* traditional peers – which is consistent with the findings provided by Buchak et al. (2018) and Fuster et al. (2019), according to which default rates for US borrowers are statistically similar between fintech and traditional lenders.

Table 4 Performance and riskiness of neobanks and traditional banks: mean comparisons. This table reports the results of the t-test for the comparison of means between neobanks and traditional peers for the dependent variables. All variables are divided by total assets and multiplied by 100. *ROA* is computed using pre-tax returns, i.e., the sub-components of ROA identified below: net interest income plus non-net interest income (net fees and commission, net trading income, other income) minus staff expenses, non-staff expenses (administrative expenses, other operating expenses), and impairment charges. See Table A, in the supplementary materials, for detailed description of each variable

	Neobanks	Traditional peers	Diff.
ROA	−0.125	0.396	−0.521**
Interest income	4.294	2.400	1.894***
Interest expenses	0.795	0.626	0.169**
Net interest income	3.454	1.760	1.694***
Non-net interest income	1.526	1.885	−0.359
Staff expenses	1.540	1.320	0.220
Non-staff expenses	2.476	1.502	0.974***
Impairment charges	1.095	0.375	0.720***

Statistical significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively

5.2 How Do Neobanks Perform Vis-à-Vis Traditional Peers?

Table 4 shows the mean values of neobanks and traditional peers for Return on Assets (ROA) and the sub-components of ROA. In general, the results suggest that neobanks perform worse than their traditional peers. This is shown by the negative average ROA recorded by neobanks (−0.125%), which is significantly lower than the ROA registered by traditional banks (0.396%). This result is backed by the regressions results reported in Panel A of Table 5, particularly by the negative and significant coefficient of the ‘neobank dummy’ on ROA, in column (1).

To further shed light on the reasons that determine such underperformance, we analyse the relative performance of neobanks in each of the ROA sub-components – in Table 4 and Table 5, Panel A, columns (2–8). First, we find that neobanks record significantly higher interest income than their traditional peers (shown by the positive and significant coefficient of interest income in Table 5, Panel A, as well as the higher mean value in Table 4), while also recording a significantly higher cost of risk (as per mean value and coefficient of impairment charges). Additionally, we find that such ‘high interest-high impairment charges’ nexus is driven solely by the online functionalities dummy (Table 5, Panel B). As discussed in the literature review, we envision two alternative explanations for this phenomena: the ‘winners curse’ narrative (Broecker 1990; Shaffer 1998; Di Maggio and Yao 2021), according to which neobanks may be tapping into a pool of risky borrowers previously rejected by traditional peers; or, alternatively, the ‘transactional model’ narrative, wherein neobanks may be charging higher interests and recording more loan provisions as a reflection of the information asymmetries that arise from the adoption of a transactional banking model, which impairs the exploitation of soft information on borrower creditworthiness (Herpfer 2021). In our view, the fact that we control for the riskiness of the loan portfolios in the regressions, coupled with the similar levels of non-performing loans in neobanks and traditional peers, suggests that the ‘transactional model’ narrative is a more suitable explanation for the ‘higher interests–higher impairments’ recorded by neobanks.

Table 5 Performance and riskiness of neobanks and traditional peers: baseline regressions. This table presents the regression results for the performance and riskiness variables, using two approaches. In the first approach (*Panel A*) we use the neobank dummy as the main independent variable of interest; in the second approach (*Panel B*) we use the three dummies employed to classify neobanks as the main variables of interest. Namely, age equal or below 20 years, number of branches below or equal to 5, and the possibility to open an account or apply for a loan fully online, i.e., without human intervention. For brevity reasons, we do not report the coefficients of controls in Panel B, however the specifications include the same bank and country controls as Panel A. In addition to the variables described in the previous tables, the specifications include the following country controls retrieved from the World Bank's Global Financial Development Database (2021): *GDP growth* is the annual growth rate of GDP. *GDP* per capita is the GDP divided by total population. *Concentration ratio* is the market share of the top 3 banks in terms of total assets. The country controls are computed as the mean value between 2019 and 2020. The regressions are performed using White-robust standard errors

	ROA	Interest income	Interest expenses	Net interest income	Non-net interest income	Staff expenses	Non-staff expenses	Impairment charges
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Classification								
Neobank	−0.540*	1.355***	0.017	1.306***	−0.324	0.249	0.699*	0.600***
Size	0.115	−0.116	−0.005	−0.116	−0.219	−0.397***	−0.071	−0.009
Income diversification	−0.009*	−0.034***	−0.006***	−0.028***	0.012	0.005	−0.006	−0.006
Number of products	0.002	−0.078**	−0.037***	−0.037	−0.061	−0.027	−0.089**	−0.021
Gross loans to customers	0.010*	0.029**	−0.000	0.028**	−0.061***	−0.005	−0.040**	0.015***
Liquid assets	0.005	−0.005	−0.002	−0.003	−0.054**	−0.007	−0.044**	0.006
Customer deposits	−0.001	−0.012**	−0.005***	−0.006	−0.001	0.000	0.000	−0.003
Total equity	0.044***	0.044*	−0.004	0.043**	0.064*	0.028**	0.046*	0.009
Non-performing loans	−0.028**	0.041**	0.006**	0.032**	0.018	0.019**	0.045***	0.038***
GDP growth	0.173***	0.307***	0.050***	0.263***	−0.140*	−0.038	−0.064	0.001
GDP per capita	0.059	−0.281	−0.156**	−0.129	0.246	0.005	−0.063	−0.016
Concentration ratio	0.000	−0.020**	−0.008***	−0.014*	0.016	0.005	0.010	0.001
Observations (neobanks)	376 (63)	376 (63)	376 (63)	376 (63)	376 (63)	376 (63)	376 (63)	376 (63)
Adjusted R^2	0.125	0.417	0.182	0.399	0.137	0.252	0.173	0.237
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Panel B: Criteria								
Branches ≤5	0.010	−0.092	0.161**	−0.241	−0.897**	−0.792***	−0.342	−0.219**

Table 5 (continued)

	ROA	Interest income	Interest expenses	Net interest income	Non-net interest income	Staff expenses	Non-staff expenses	Impairment charges
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Age ≤ 20 years	−0.057	0.145	0.096*	0.057	0.096	−0.100	0.196	0.138
Online	−0.461**	1.086***	−0.087	1.150***	−0.126	0.492**	0.541	0.552***
Observations (neobanks)	376 (63)	376 (63)	376 (63)	376 (63)	376 (63)	376 (63)	376 (63)	376 (63)
Bank controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.122	0.412	0.195	0.396	0.145	0.301	0.171	0.246

Statistical significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively

Relatedly, such 'high interest-high impairments' approach to lending may generate concerns regarding the sustainability of the neobanks' business model. To shed light on this aspect, we build a new dependent variable, 'net interest income minus impairment charges' and still find a positive and significant coefficient of the neobanks dummy at the 5% significance level (untabulated) – suggesting that neobanks seem to charge sufficient interests to cover the higher cost of risk, thereby alleviating the concerns on the potential lack of sustainability. Importantly, we note that these results seem to contradict the standard literature on lending technology (i.e., relationship vs transactional lending) (Boot 2000; DeYoung 2010; van Ewijk and Arnold 2014), which typically relates transactional lending to lower net interest margins, due to the lack of product differentiation. However, an emerging strand of the literature suggests that in the digital era customers place significant value on the quality of the customer experience (e.g., accessibility, usability, flexibility) and are willing to pay an interest rate premium, especially for faster loan decisions (Buchak et al. 2018; Tanda and Schena 2019; Di Maggio and Yao 2021; Williams 2021; Berg et al. 2022). For instance, Di Maggio and Yao (2021) study a large sample of US personal loans and show that personal loans from fintech firms are approximately 3 p.p. more expensive than traditional peers. Also, Buchak et al. (2018) find that fintech lenders tend to charge higher spreads than traditional banks for comparable borrowers and interpret this finding as evidence that borrowers are willing to pay for the superior customer experience offered by fintech lenders. In general, our results seem to align with the predictions put forward by this emerging strand of the literature.

With respect to the mitigation of 'communication frictions', the findings reported in Tables 4 and 5, indicate that, overall, neobanks exhibit a similar funding cost and ability to generate fees to that of traditional peers – hence, falling to provide support for the 'digital capture' hypothesis (Boot et al. 2021). Furthermore, the results show that neobanks record significantly higher non-staff expenses (e.g., IT costs, advertising costs, reporting costs) than traditional banks (as shown by the negative and significant coefficient of the neobank dummy in Table 5, column (7), as well as the statistically significant mean difference in Table 4). This result seems in line with the empirical literature which relates the relative underperformance of internet banks *vis-à-vis* traditional peers to cost inefficiencies (DeYoung 2005; Delgado et al. 2007). Note, however, that this strand also argues that digital banks seem to show a greater 'depth' of economies of scale than traditional peers – as will be discussed in the next sub-section.

Finally, in Panel B of Table 5 we report the results of the regressions with the criteria used to identify neobanks (i.e., number of branches below or equal to 5, age below or equal to 20 years and online functionalities) as the main independent variables, instead of the neobank dummy. In general, the findings suggest that the baseline results (higher interest income, non-staff expenses and impairment charges) are driven by the online functionalities, and not the other criteria. Moreover, this analysis allows us to uncover several results that are coherent with our prior expectations. For instance, banks with smaller branch networks and younger banks seem to record relatively higher funding costs than other banks. The fact that no such effect is found for the neobank dummy (Panel A) suggests the potential presence of a 'digital discount' on funding costs, possibly linked to the superior online customer experience, as previously discussed. Another interesting result is related to staff expenses, as we uncover two apparently contradicting effects for neobanks: while operating with smaller branch networks seems to reduce staff expenses, the implementation of online functionalities seems to require a greater investment in such expenses, which may be linked to the need to invest relatively more in specialised human resources (Tanda and Schena 2019; Williams 2021).

Table 6 Economies of scale, scope, and experience: interaction terms. This table reports the results for regressions using interaction terms between the neobank dummy and three mediating variables: size (scale, Panel A), age (experience, Panel B), and income diversification (scope, Panel C). Namely, for each mediating variable we construct a dummy representing banks with low values (below the 10th percentile) and high values (above the 90th percentile). For brevity reasons we report the regression coefficients only for the main variable of interest, although the specification includes bank and country controls. In Panel C, ID stands for income diversification. The regressions are performed using White-robust standard errors

	ROA	Interest income	Interest expenses	Net interest income	Non-net interest income	Staff expenses	Non-staff expenses	Impairment charges
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Scale								
Neobank	−0.738**	0.982**	−0.035	0.990***	−0.240	0.084	0.687*	0.571***
Neobank * Size (<10p)	1.346	3.982	0.474	3.440	−0.967	1.811*	0.690	0.417
Neobank * Size (>90p)	1.236***	−0.020	0.133	−0.135	0.123	−0.087	−0.927*	−0.176
Observations (neobanks)	376 (63)	376 (63)	376 (63)	376 (63)	376 (63)	376 (63)	376 (63)	376 (63)
Bank controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.135	0.445	0.186	0.427	0.133	0.271	0.172	0.235
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Panel B: Experience								
Neobank	−0.747**	1.301***	0.074	1.214***	−0.773*	0.315	0.801*	0.536**
Neobank * Age (<10p)	1.215	−0.597	−0.174	−0.586	2.190	−0.438	−1.343	−0.649
Neobank * Age (>90p)	0.450	0.863	−0.257*	1.121	1.337***	−0.104	0.383	0.978
Observations (neobanks)	376 (63)	376 (63)	376 (63)	376 (63)	376 (63)	376 (63)	376 (63)	376 (63)
Bank controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.130	0.418	0.182	0.403	0.143	0.250	0.176	0.257
	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)
Panel C: Scope								
Neobank	−0.659*	0.938**	−0.001	0.915**	0.150	0.264	0.809**	0.573**

Table 6 (continued)

	ROA	Interest income	Interest expenses	Net interest income	Non-net interest income	Staff expenses	Non-staff expenses	Impairment charges
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Neobank * ID (<10p)	0.341	2.001*	0.116	1.858*	−1.987***	−0.066	−0.157	0.171
Neobank * ID (>90p)	0.817*	−0.356	−0.140	−0.259	−0.742	−0.000	−1.379***	−0.196
Observations (neobanks)	376 (63)	376 (63)	376 (63)	376 (63)	376 (63)	376 (63)	376 (63)	376 (63)
Bank controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.156	0.454	0.210	0.439	0.178	0.276	0.203	0.262

Statistical significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively

Table 7 Economies of scale, experience, and scope: sub-sample regressions. This table reports the regression results for sub-samples, using the same specification as column (7) of Table 5, i.e., the dependent variable is non-staff expenses: in Panel A, the regressions are performed on different sub-samples of size, allowing us to test for the presence of economies of scale; in Panel B, the sub-samples are created based on age thresholds, designed to test the occurrence of economies of experience; Panel C includes six sub-samples comprising banks with different levels of product diversification, wherein we test the prevalence of economies of scope in digital banking. NP stands for number of products. For brevity reasons we report the regression coefficients only for the main variable of interest, although the specification also includes bank and country controls. The regressions are performed using White-robust standard errors

	Total assets > € 200 M	Total assets > € 400 M	Total assets > € 600 M	Total assets > € 800 M	Total assets > € 1000 M	Total assets > € 2000 M
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Scale						
Neobank	0.712*	0.777**	0.550	0.620	0.688	0.600
Observations (neobanks)	337 (57)	305 (50)	269 (43)	235 (38)	215 (36)	143 (29)
Bank controls	Yes	Yes	Yes	Yes	Yes	Yes
Country controls	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.165	0.186	0.198	0.221	0.239	0.301
	Age > 1y	Age > 2y	Age > 3 y	Age > 5y	Age > 10y	Age > 15y
	(7)	(8)	(9)	(10)	(11)	(12)
Panel B. Experience						
Neobank	0.711*	0.852**	0.895**	0.916**	1.026*	1.402*
Observations (neobanks)	374 (61)	369 (57)	365 (56)	354 (47)	319 (31)	286 (16)
Bank controls	Yes	Yes	Yes	Yes	Yes	Yes
Country controls	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.173	0.200	0.204	0.203	0.203	0.226
	NP > 1	NP > 3	NP > 5	NP > 6	NP > 7	NP > 8
	(13)	(14)	(15)	(16)	(17)	(18)
Panel C. Scope						
Neobank	0.678*	0.762*	0.489*	0.134	−0.005	0.163
Observations (neobanks)	370 (62)	287 (42)	200 (15)	162 (10)	128 (6)	91 (3)
Bank controls	Yes	Yes	Yes	Yes	Yes	Yes
Country controls	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.167	0.136	0.176	0.182	0.215	0.355

Statistical significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively

5.3 Are There Economies of Scale, Experience, and Scope in Neobanking?

To examine this question, we re-run the baseline regressions including two interaction terms with the neobank dummy, representing banks with low values (below the 10th percentile) and high values (above the 90th percentile) of size (Table 6, Panel A), age[15] (Panel

[15] Given that above the 90th percentile of age for the full sample (101 years) there are no neobanks, in the case of 'age' we build the interaction dummies using the 10th and 90th percentiles of the neobank sample (3 and 19 years, respectively).

B), and income diversification (Panel C); and we split the original sample into buckets of size (Table 7, Panel A), age (Panel B) and number of products (Panel C), and run sub-sample regressions using 'non-staff expenses' as the dependent variable (DeYoung 2005).

Regarding the potential presence of *economies of scale*, the results in Table 6 show that neobanks in the top decile of size significantly outperform other neobanks, and this can be traced to their ability to record lower non-staff expenses, as reported in column (7). Moreover, the net effect on non-staff expenses (i.e., the sum of coefficients for the neobank dummy and neobank dummy*top decile of size) becomes negative, suggesting that larger neobanks may be more efficient than traditional peers. With respect to the sub-sample regressions, the results reported in Table 7 (Panel A) suggest that the underperformance of neobanks in terms of non-staff expenses fades away as we remove very small banks (less than EUR 600 million) from the sample. In general, the results suggest that in our sample of relatively small banks (total assets below EUR 10 bn), neobanks enjoy some 'depth' of economies of scale, in line with earlier evidence collected by DeYoung (2005) and Delgado et al. (2007).

With respect to *economies of experience*, the results in Panel B of Tables 6 and 7 suggest the performance gap towards traditional peers is not smaller for older neobanks. As a case in point, the regression results reported in columns (7–12) of Table 7 indicate that the underperformance of neobanks persists as we exclude younger banks, at different age thresholds. Notwithstanding, we notice that older neobanks seem to charge higher fees and commissions – as observed in column (13) of Table 6. In our view this finding is consistent with the 'digital spatial capture' narrative (Boot et al. 2021), according to which as customers become more acquainted with a bank's technology, their switching costs become higher, hence allowing digital banks to charge fees for their services.

As for *economies of scope*, in general, we find similar results to the economies of scale. Namely, the coefficients reported in Table 6 (Panel C) show that (i) neobanks in the top decile of income diversification record a better performance in terms of ROA than less diversified neobanks, (ii) this effect seems to be exclusively driven by efficiencies in terms of non-staff expenses, and (iii) the overall effect (coefficient of neobanks plus the coefficient of neobanks*top decile of diversification) is negative and sizable – suggesting that more diversified neobanks may be more efficient than traditional peers. Coherently, the findings in Table 7 (Panel C) reveal that neobanks no longer exhibit non-staff inefficiencies after broadening the scope of products offered (namely 6 or more). In our view, such results provide support regarding the presence of economies of scope in neobanking.

5.4 Are the Baseline Results Driven by Specific Product Lines?

We provide answers to this question by re-running the baseline regressions in three sub-samples of banks based on whether they offer payments (credit and debit cards), loans (personal and mortgage loans) and services (broker & advisory services and insurance). Moreover, to improve the comparability of coefficients on the interaction terms, we remove from each sub-sample the banks that offer all products of that specific business line. For instance, in column (1) we remove 147 banks from the baseline sample (of which, 17 neobanks) that offer both credit and debit cards. The results are presented in Table 8.

Overall, we only find statistically significant results in the 'Loans' sub-sample. Namely, the results in column (2) indicate that neobanks offering personal loans perform

Table 8 Performance of neobanks and traditional peers: business lines. This table reports the regression results for banks that operate in the following business lines: payments (1), loans (2) and services (3). Namely, we visit the website of each bank and check which products are offered by banks. To reduce the noise introduced by banks with a wider scope of products and facilitate the interpretation of the results, we remove from each sample the banks that offer all products of that specific business line. For instance, in column (1) we remove 147 banks from the baseline sample (of which, 17 neobanks) that offer both credit and debit cards. The dependent variable in all regressions is ROA. For brevity reasons, we report the regression coefficients only for the main variables of interest, although the specification also includes bank and country controls. The regressions are performed using White-robust standard errors

	Payments	Loans	Services
	(1)	(2)	(3)
Neobank	0.128	0.577	−0.574
Neobank*Credit cards	−0.939		
Neobank*Debit cards	−1.136		
Neobank*Personal loans		−2.297***	
Neobank*Mortgage loans		−1.204	
Neobank*Broker & advisory services			0.141
Neobank*Insurance sale			−0.143
Observations (neobanks)	215(46)	249(50)	267(59)
Product dummies	Yes	Yes	Yes
Bank controls	Yes	Yes	Yes
Country controls	Yes	Yes	Yes
Adjusted R^2	0.194	0.245	0.179

Statistical significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively

significantly worse than other neobanks, whereas the same does not occur for those offering mortgage loans. This result is backed by untabulated results, which show that top performing neobanks (i.e., with above median ROA) are significantly less likely to offer personal loans than worse performing neobanks (41.9% vs 67.7%), whereas such difference is much smaller for mortgage loans (29.0% vs 38.7%). Moreover, the overall effect (i.e., the sum of coefficients for the neobank dummy and neobank dummy*personal loans) becomes negative, suggesting that neobanks that offer personal loans perform significantly worse than traditional peers. In our view this finding seems in line with the literature suggesting that hard information can be used more effectively to assess borrower creditworthiness in collateralised loans (e.g., mortgage loans) than in non-collateralised loans (e.g., personal loans) (Stein 2002; DeYoung et al. 2007).

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 instrumental variables (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 centres*, calculated as the natural log of the road distance between the bank's headquarters and the nearest top50 university,[16] 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 centres are more likely to access the specialised resources (human and technological) necessary to adopt certain digital strategies (Tanda and Schena 2019; Williams 2021). An opposing argument, however, could be made that the location of certain knowledge centres 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 centres, 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 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 centres*. This IV is computed using the total number of patents in ICT, as published by the OECD, for the NUTS2 region where the nearest top50 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 centres 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.

The results of the 2SLS estimations are presented in Table 9. The post-estimation tests show that the Cragg-Donald F-test is 11.30 which is above the rule of thumb of 10, as suggested by Staiger and Stock (1997) and just below the 15% TSLS bias critical value of 11.59 as defined by Stock and Yogo (2005). Additionally, we report the Anderson LM statistic of 22.094, which suggests the rejection of underidentification at the 1% level. Regarding the first-stage regression, the proximity to knowledge centres 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 centre. As for the second stage result, the neobanks dummy is found to negatively affect ROA, lending support to our baseline regressions.

6.2 Are Age and Number of Branches Driving the Underperformance of Neobanks?

Given that we do not include age or branches as matching criteria between neobanks and traditional peers, one legitimate concern is whether the performance differences are chiefly related to the fact that neobanks are younger and have relatively less branches than traditional peers, rather than to the fact that they adopt a digitally native business model. Indeed, as shown by the distribution of banks per age and branch buckets (Fig. 2), one can observe that, as expected: (i) with reference to age, neobanks are all in the '0–20 years' bucket, whereas the interquartile (Q3-Q1) of traditional peers is located in the '20–50 years' bucket, with

[16] For countries without any university in the top50, we include the highest ranked university in the dataset. This occurs for Cyprus (#380), Estonia (#172), Hungary (#194), Latvia (#407), Lithuania (#341), Luxembourg (#202), Malta (#502), Poland (#118), Romania (#430), and Slovakia (#299).

Table 9 Endogeneity in the choice of business model: IV regressions. This table reports the results of the regressions using two instrumental variables (*distance to knowledge centre* and *quality of knowledge centre*) to curb endogeneity concerns regarding the choice of the neobank model. *Distance to knowledge centre* is the number of road hours between the city where the bank is headquartered and the nearest top50 universities in the EU-28, according to the '2021 Scimago Institutions Ranking' (natural log). *Quality of knowledge centre* is measured as the total number of ICT-related patents issued by the nearest top50 universities in the EU-28, as reported by the OCED in 2013 (natural log). All controls included in the first stage regression are also included in the second stage regression. For brevity reasons we report the regression coefficients only for the main variables of interest. The regressions are performed using White-robust standard errors. The regressions are performed using White-robust standard errors

	First stage regression	Second stage regression
	Neobanks dummy	ROA
	(1)	(2)
Instrumental variables		
Distance to knowledge centre	−0.058**	
Quality of knowledge centre	0.039***	
Instrumented variables		
Neobank (dummy)		−1.732*
Observations (neobanks)	376 (63)	376 (63)
Anderson LM statistic	22.094	
Cragg-Donald Wald F-test	11.300	
Stock-Yogo weak id critical values	20% TSLS bias: 8.75 15% TSLS bias: 11.59	
Sargan overidentification statistic	0.275	
Adjusted R^2		0.118
F-statistic		4.89

Statistical significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively

particular concentration in three age buckets: '15–20', '20–25', and '25–30' years; and (ii) regarding branches, all neobanks are located in the '0–5' branches bucket, whereas 54.0% of traditional peers in the matched sample ($n = 169$) are also positioned in this bucket.

To address this potential concern, we adopt several strategies that allow us to compare the performance of neobanks with similarly young and relatively branchless traditional peers. Regarding age, the results of the tests are presented in columns (1–4) of Table 10. First, we include age as an additional matching variable in the PSM. As expected, this step significantly reduces the age gap between neobanks and controls, even though we still find a statistically significant difference. Namely, the mean age gap moves from 37.0 years (neobanks: 10.8 years vs traditional: 47.8 years) to 9.3 years (10.7 years vs 20.0 years, respectively). Also, as expected the number of matched observations significantly reduces (from $n = 376$ to $n = 226$). Second, we retain in the sample only banks born in 2000 or after and re-run the PSM. Again, this narrows the age gap significantly to just 2.2 years, at the cost of a significantly smaller matched sample ($n = 132$). Third, based on the notion that the first neobanks were born in 1995 (namely, Nest bank, Boursorama and Skandiebanken) we remove from the original matched sample ($n = 376$) banks older than 25 years. In line with the previous strategies, this reduces the age gap in an important way (to 4.9 years) at the cost of a significantly smaller sample ($n = 157$). Fourth, we re-run the baseline regression

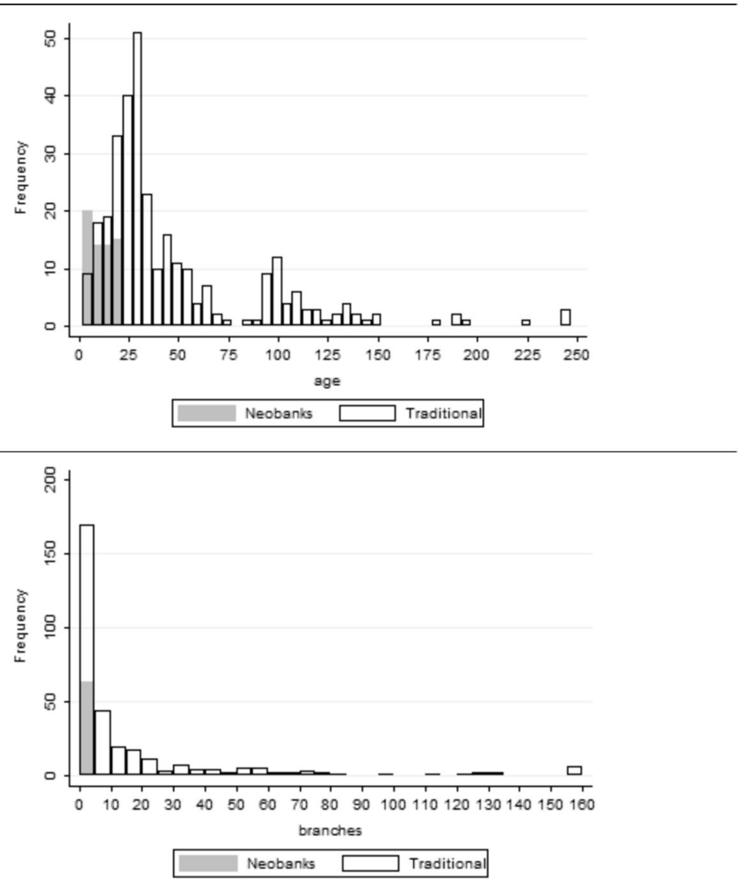

Fig. 2 Distribution of neobanks and traditional peers per age and number of branches (matched sample)

removing banks that are less than 5 years old, and the baseline results are confirmed. Overall, the baseline results hold for all specifications.

With respect to branches, the results are shown in columns (5–7) of Table 10. In line with the age-related tests, we use branches as a matching variable in the PSM (column 5), and we remove banks with more than 5 branches before and after running the PSM (column 6 and 7). These procedures significantly reduce the gap between neobanks and traditional peers with respect to the number of branches, from 17.5 branches (neobanks: 0.7 vs traditional: 18.2) to 0.5 branches (1.0 vs 1.5), respectively. Our main findings are confirmed.

6.3 Are Results Driven by the Thresholds Used to Identify Neobanks?

An additional possible concern is related to the sensitiveness of the neobanks identification to the thresholds of size, asset and funding structures, age, and branches. To address this potential issue, we adjust the thresholds of size (from 0-10b to 0-30b), the share of loans to assets and deposits to assets, (from a minimum of 5% to 10%), the age (from maximum of 20 years to 30 years) and the

Table 10 Robustness checks. In this table we report the regression results for the robustness tests related with age, branches, identification of neobanks, sample period, country fixed effects, and winsorisation. Regarding *age* we perform the following tests: (1) we include age as an additional matching variable in the PSM, (2) we keep in the sample only banks born in 2000 or after and re-run the PSM, (3) we remove from the original matched sample banks older than 25 years, and (4) we remove banks younger than 5 years. Concerning *branches*: (5) we use branches as a matching variable in the PSM, and we remove banks with more than 5 branches (6) before and (7) after running the PSM. With respect to *neobank identification*, we adjust the thresholds of (8) size (from 0-10b to 0-30b), (9) the share of loans to assets and deposits to assets (from a minimum of 5% to 10%), (10) the age (from maximum of 20 years to 30 years) and (11) the number of branches (from maximum of 5 to 20); (12) we combine the new thresholds defined in columns to identify neobanks (8–11); and (13) we combine the new thresholds defined in columns (8–9) to identify traditional peers. Regarding the *sample period*, (14) we use financial data exclusively from 2019 (and not the mean value of 2019–2020). We re-run the baseline model using *country fixed effects* (15) and *unwinsorised* data (16). For brevity reasons we report the regression coefficients only for the main variables of interest. The regressions are performed using White-robust standard errors

	Age: additional PSM variable	Age: max 20 years before PSM	Age: max 25 years after PSM	Age: min 5 years after PSM	Branches: additional PSM variable	Branches: max 5 branches before PSM	Branches: Max 5 branches after PSM	Neobanks: Size <30 bn
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Neobank	−0.543*	−0.632*	−0.564*	−0.537*	−0.680**	−0.715**	−0.637**	−0.540*
Observations (neobanks)	226 (60)	132 (52)	157 (63)	360 (53)	273 (48)	276 (62)	232 (63)	376 (63)
Adjusted R^2	0.121	0.143	0.142	0.140	0.122	0.126	0.097	0.125
	Neobanks: CD & GLC > 10%	Neobanks: Age < 30	Neobanks: Branches <20	Neobanks: combination new criteria (8–11)	Traditional peers: combination new criteria (8–9)	Sample period: Only 2019	Country FE	Non-winsorised Data
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Neobank	−0.518*	−0.548*	−0.540*	−0.599**	−0.619*	−0.800**	−0.793**	−0.854*
Observations (neobanks)	372 (59)	376 (66)	376 (63)	372 (59)	617 (65)	375 (62)	376 (63)	376 (63)
Adjusted R^2	0.121	0.126	0.125	0.126	0.111	0.184	0.183	0.171

Statistical significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively

number of branches (from maximum of 5 to 20). We check for changes in the pool of neobanks by applying the new thresholds to each variable individually, and to all variables cumulatively.

Overall, we find that the identification of neobanks remains notably stable after changing the referred thresholds. Namely, the criteria which affects the identification the most is related to customer lending and deposits (4 neobanks are removed); increasing the size threshold allows the inclusion of 2 additional neobanks; and increasing the age to 30 years facilitates the inclusion of 3 neobanks. Interestingly, changing the number of branches does not affect the total number of neobanks identified. Alternatively, applying the four changes to criteria cumulatively leads to off-setting effects: we identify 7 additional neobanks, but also remove 4 neobanks.[17]

Moreover, we re-run the baseline regressions for each of the new samples of neobanks (Table 10, columns 9–12), and find that the baseline results remain unchanged. Relatedly, we apply the new thresholds on size and asset and funding structure to the sample of traditional banks and find that the main baseline results hold – as reported in column (13). In our view, both results (stability of neobank identification and robustness of 'neobank-performance' linkage) provide a strong case for the robustness of our approach.

6.4 Additional Robustness Checks

In this sub-section we report four additional robustness checks, which provide further evidence on the reliability of our results. First, our sample includes 2019 and 2020 data. This period can raise concerns given that the Covid-2019 pandemic (which developed in 2020) significantly accelerated the development and usage of digital banking channels (Abramson et al. 2020; Dadoukis et al. 2021). We tackle this potential issue by running a robustness check using financial data exclusively from 2019 (and not the mean value of 2019–2020). By doing so, we drop one neobank from the sample, which is born in 2020. We re-run the baseline regressions and find similar results (Table 10, column 14).

Second, our baseline regressions include three country-level controls. One possible issue is the presence of omitted factors that are not captured in our country-level controls, hence affecting our results. To test whether this is the case, we re-run the baseline model including country fixed effects, and no changes occur to the baseline results (see Table 10, column 15). Nonetheless, our preferred specification remains the inclusion of country controls, rather than fixed effects, due to model overfitting issues for some of the sub-sample tests we perform throughout the paper.

Third, in our baseline approach we winsorise the dataset to tackle data quality issues. However, in general, winsorisation can create bias in the results, by affecting relatively more one of the groups than the other. To address such concerns, we analyse the distribution of neobanks and traditional peers below (above) the p1(p99) thresholds, for each of the bank controls. Overall, in the 1.4% of the full sample that is changed across the control variables, only 0.3% of values are related to neobanks – clearly indicating that most changes are related to traditional peers. As an additional test, we re-run the baseline regression using non-winsorised data and, as expected, the baseline results hold (column 16).

Fourth and final, while our preferred dependent variables, ROA and its decomposed elements, have been widely studied in the banking literature (e.g., Mergaerts and Vander Vennet 2016; Marques and Alves 2021), other proxies have also been used to provide insights into the performance of banks. As such, next we re-run our baseline regressions using Return on Equity (ROE) as an alternative dependent variable. Untabulated results show a negative coefficient of the neobank dummy (−0.080), significant at the 5% level, which confirms the previous findings.

[17] The changes to the baseline list of neobanks (as provided in Table I in the Appendix) is available upon request.

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 seems 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 65 neobanks operating in EU-28 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 deepen our understanding of 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 information asymmetries that emerge from performing transactional lending (Boot 2000; DeYoung 2010; van Ewijk and Arnold 2014) and providing better customer experience (Buchak et al. 2018; Tanda and Schena 2019; Di Maggio and Yao 2021; Williams 2021; Berg et al. 2022). Second, we show that neobanks do not pay lower funding costs or generate higher fees and commissions, providing a lack of support for the 'digital spatial capture' hypothesis (Boot et al. 2021). Finally, neobanks record significantly higher non-staff expenses than traditional peers. By running additional analyses, we observe the non-staff inefficiencies fade away as we remove very small banks (with less than EUR 600 million in total assets) and specialised banks (with less than 6 product lines).

This paper bears relevant contributions to the literature and to policy. On the literature side, we update the literature on the identification and 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. We also test various bank intermediation theories and uncover two suitable IVs to study bank performance.

On the policy side, our paper contributes to the supervisory assessment of non-significant banks – under the EU framework, the so-called 'Less Significant Institutions' (LSI). Namely, we believe that our proposed process to identify neobanks can make a significant contribution to supervisor's efforts to monitor digital players (ECB 2020), and more broadly to the task of business model identification undergone by bank regulators and supervisors (Cernov and Urbano 2018). Moreover, while our findings regarding the 'high interest-high impairments', similar NPL ratio, and potential for economies of scale and scope, seem to alleviate concerns on the potential lack of sustainability of neobanks, we note that our results are based on a very stable, low interest-rate environment. In this context, an interesting policy-oriented addition to our paper would be to track the performance of neobanks in a more stressed environment (e.g., higher interest rates period). Finally, we also note that our paper paves the way for future research to shed light on digital banks' specific exposures to operational risk, that arise from operating fully in a digital setting, i.e., IT outsourcing and IT security/cyber risk (ECB 2022).

Appendix

Table 11 List of neobanks. This table shows the list of 65 neobanks identified using our proposed method. In the table we focus exclusively on the distinguishing features of each bank retrieved mainly from the banks' websites, annual reports and Orbis/Bankfocus. Indeed, as described in the paper, the method ensures that all the identified neobanks have retail business model (i.e., lending and deposit-taking activities), digital orientation (i.e., they offer the possibility to apply for a loan or open an account fully online) and were incorporated in the post-Internet era. The fact that we use an objective set of criteria to identify neobanks in the sample allows us to propose a procedure that can be replicated without subjectivity over time by other researchers. As shown in the table, the method is successful in identifying well-known digital banks, such as N26 (DE), Illimity Bank (IT), Bunq (NL), BNI (PT) and Monzo Bank (UK). Interestingly, this approach also allows us to understand that certain entities that are commonly perceived as digital banks, actually operated with a non-retail business model at the time of the analysis (e.g., in 2019 the lending activity of Revolut (LT) represented only 1.6% of its total assets) or had a relatively large branch network (e.g., in 2019 Activo Bank (PT) had 16 branches)

#	Neobank	Description
1	Advanzia Bank (LU)	Luxembourg-based bank specialised in credit cards and payment solutions for individuals and corporates. The bank actively uses online marketing channels and has an omni-channel UX programme in place that offers a fully digital customer and payment experience via mobile apps and mobile payment solutions.
2	Aigis Banca (IT)	Specialist lender that provides online credit, accounts and savings solutions to retail customers (individuals and SMEs). The bank delivers its products and services through digital processes that rely on the use of the latest generation technologies (e.g., AI) and cooperates with fintech companies.
3	Aion Bank (BE)	Full-scope bank focused on retail banking services. The bank operates with a subscription only model, and offers several high-end technology services such as an AI-powered service that allows customers to manage their savings, households bills, and online purchases (MoneyMax), it provides a fully digital customer experience, and online concierge and ETF counselling services.
4	Air Bank (CZ)	Czech bank focused on the full range of retail products, operating mainly through online and mobile channels. The bank was a pioneer in integrating instant payment solutions in its banking platform (e.g., Apple Pay, Google Pay, Garmin Pay, Fitbit Pay), as well as a fully digital personal identification and investment functionalities (e.g., Zonky, Portu service).
5	Ancoria Bank (CY)	Retail bank from Cyprus that operates mainly online. The bank allows customers to open an account via its mobile app (myAncoria), as well as digital loans and overdrafts. It has received awards for its digital capabilities in several categories such as fintech solutions and Banking as a Service (BaaS).
6	Arkea Direct Bank (FR)	Operating under the brand 'Fortuneo Banque', the bank is focused on providing the full range of digital retail banking services. The bank offers a 100% online experience (no branches), including "fast and user friendly" account opening and loan application processes; the bank has won several awards in the "online banking" category, and has an open banking (API) policy.

Table 11 (continued)

#	Neobank	Description
7	Atom Bank (GB)	UK's first app-based bank and the first digital-only bank to be granted a regulatory license. It offers the full range of retail products, including savings accounts, mortgages, and loans tailored for both individuals and businesses. Atom Bank has earned multiple mentions in KPMG's Fintech100 as a frontrunner in global fintech innovation.
8	Avanza Bank (SE)	Swedish digital bank that operates exclusively via online channels, focused on savings and investment solutions. The bank is considered the largest stockbroker in Sweden. Avanza's trading and investment platform is considered one of the top platforms in the industry, allowing investors to tailor it to their needs; the bank has won several awards related to its digital capabilities (e.g., best internet broker, banking technology award for best use of IT).
9	Banca 5 (IT)	The bank describes itself as the first Italian online bank. Banca 5 focuses on digital retail banking services (e.g., payment accounts, prepaid cards, personal loans, insurance services) to individuals and companies. The bank actively partners with fintech companies as part of its open banking policy, for instance 'Oval Money' (personal savings management app) and 'Yolo' (on-demand insurance).
10	Banca Progetto (IT)	Italian digital bank specialised in retail banking services to SMEs and individuals (e.g., deposit accounts, medium- to long-term loans, factoring). The bank is considered a pioneer in cloud-based banking (Amazon Web Services, AWS), allowing it to apply advanced data analytics and machine learning, to automate financial operations and to enhance customer-facing web/mobile applications.
11	Banca Sistema (IT)	The bank provides the full range of retail services, with a primary focus on consumer credit (namely, salary/pension-backed loans) and factoring. Banca Sistema allows customers have a fully digital experience when opening an account and applying for consumer loans and factoring (the latter is provided via the brand 'QuintoPuoi'); as part of its business model, the bank actively participates in online deposit platforms (Raisin).
12	Banca Widiba (IT)	One of the most recognised banks in Italy with a fully digital orientation, Widiba covers the entire range of retail banking products. Some of its innovative and user-friendly technologies include an app that allows customers to have face-to-face interaction with the bank, an online AI-based robot advisor, a paperless mortgage loan process (the first in Italy). Widiba is part of one of the main Open Banking platforms (CBI Globe).

Table 11 (continued)

#	Neobank	Description
13	Banco Atlantico Europa (PT)	The bank has a universal product offer and customer base, and is focused on providing banking services via its digital channels. Besides offering the possibility to open an account fully online, the bank has been reported as the first cloud-based bank in Portugal, has an open banking (API) policy, and an AI-based chatbot (ALICE).
14	Banco Best (PT)	Primarily focused on saving and financial investment solutions, Banco Best is a pioneer online bank in Portugal (founded in 2001). The bank allows customers to open accounts online and provides a fully digital experience via its website, mobile app, and AI-powered chatbot (BEA, Best Electronic Assistant).
15	Bank11 (DE)	German bank focused on providing the full scope of retail banking services to individuals and SMEs. The bank has implemented blockchain technology, which allows it to offer enhanced transactions privacy (via PoS, Proof-of-Stake), and innovative financing products such as Crypto Asset-Secured loans, P2P lending, P2P exchange, and stablecoins-based payments.
16	Basisbank (DK)	Engaged in the provision the typical range of retail banking services (e.g., accounts, loans, investment solutions, insurance), Basisbank has in place a state-of-the-art digital credit risk management process (FICO Blaze Advisor), powered by AI/ML analytics.
17	Bforbank (FR)	French retail bank that provides the full range of banking products via its online platforms, including brokerage services. Bforbank has a cloud-based core system (Temenos) and a Google based solution for API management (Apigee API Management).
18	Bigbank (EE)	Retail bank based in Estonia, with a wide scope of retail services, as well as specialised corporate lending solutions such as leasing. The bank has partnerships with fintech companies in several areas of its activities, including the automation of the loan granting process (Provenir Cloud), advanced payments solutions such as virtual credit cards and tokenisation (Nets/Nexi).
19	Binckbank (NL)	One of the first banks in Netherlands with a fully digital business model, Binckbank is focused on online brokerage services and investment solutions. The bank offers robo-advisory (after acquiring the fintech startup Pritle), as well as automated investment management services (e.g., Binck Forward, Binck Fund-Coach, Binck Pensioen). Bincbank has won several prizes as best online broker.
20	BNI Europa (PT)	Portuguese digital bank focused on offering its retail customers a modern, low-cost digital platform with innovative products. The bank has set up partnerships with international fintech firms in different areas, e.g., card payment system (PT Parcela Já), factoring (Belgian's EDEBEX), SME loans (German's Creditshelf) and P2P lending (UK's MarketInvoice).

Table 11 (continued)

#	Neobank	Description
21	Bunq (NL)	Dutch bank offering the full range of retail services. Bunq is considered as the first European digital bank to offer mortgage loans online (via the platform Tulp) and has an Open Banking (API) policy, which allows the bank's customers to perform real-time payments, among other capabilities. The bank has been named the Best Digital Bank by the FinTech Breakthrough awards.
22	Charter Court Financial Services (GB)	Specialised in mortgages, bridging loans, and savings products, CCFS uses a fully automated decision-making platform to manage mortgage applications (Precise Mortgages). The bank has won several industry awards for its online capabilities, including the 'Moneyfacts Consumer Award for Online Savings Provider of the Year' (2016–2018).
23	Chetwood Financial (GB)	UK-based bank focused on providing online savings solutions (SmartSave) and loans (BetterBorrow and LiveLend) to retail customers. The bank operates a fully cloud-based core system (AWS/Yobota), uses advanced AI/ML data analytics, and has won several prizes for its online capabilities (e.g., Best use of Data, Fintech Awards, and top 20 UK FinTech Disrupters).
24	Collector Bank (SE)	Primarily focused on SME financing, Collector provides innovative B2B payment solutions, and has set up several partnerships with fintech firms (Trustly, NFT Ventures) to develop and incorporate innovative payment solutions.
25	Ekspres Bank (DK)	Ekspress is a Danish retail bank focused on providing online personal loans. The bank operates a cloud-based core system, has an Open Banking (API) Policy, as well as a cloud-based application performance management system (Ymonitor/Appdynamics/Sentia) that allows the bank to monitor the performance of third-party web services.
26	EVO Banco (ES)	Based in Spain, EVO is a full-scope retail bank offering, for instance, online mortgage loans. The bank has an AI-powered chatbot (Maria) based on Google Kubernetes Engine, as well as advanced mobile payment solutions (EVO Bizum). EVO has been awarded several prizes for its online presence (e.g., Most Innovative Bank in Europe, Best Consumer Digital Bank).
27	Ferratum Bank (MT)	Focused on providing online consumer loans in several European countries, Ferratum has a cloud-based core banking system (Fusion Essence/Fusion Digital Channels), has an Open Banking (API) Policy (e.g., the bank has a partnership with Nordigen for data analytics) and is an active participant in online deposit platforms (Raisin).

Table 11 (continued)

#	Neobank	Description
28	Fidor Bank (DE)	Besides the traditional retail banking products, Fidor offers innovative services such as digital currencies, P2P lending and crowdfunding. The bank operates under an Open Banking platform, and benefits from partnerships with fintech firms in several areas, such as sales and marketing (Salesforce), advanced online payment solutions (PayDo), and online authentication (Zwipe); the bank has won several digital banking awards (e.g., Best Neobank EFM Accenture Award).
29	Flatexdegiro Bank (DE)	Specialised in brokerage and investment solutions, Flatexdegiro provides online robo-advisory services (via its partnership with Whitebox), as well as a crowd-based investment advisory app (Flatex Next 3.0); the bank has won several prizes for its online functionalities (e.g., Best Online Broker, Champion of Digital Transformation).
30	GF Bankas (LT)	GF Bankas is based in Lithuania and offers the typical range of retail banking services via its online channels. The bank has a cloud-based operating system (Mambu), a fully integrated IT system (Inventi), and is part of Lithuania's fintech start-ups list.
31	Guber Banca (IT)	Italian bank focused on distressed SME credit management, as well as other retail services (e.g., the bank offers online deposit accounts). Guber Banca has a digital platform specialised in reverse factoring (Anticipo102), as well as a platform focused on online auctions of movable/ immovable assets (Reperform).
32	ICA Banken (SE)	ICA offers the complete range of retail banking services via a wide range of digital channels (website, apps, chatbot, social media). The bank has a cloud-based operating infrastructure (Google Cloud) and an Open banking (API) policy, which allows it to partner with fintech firms, such as Oberthur (contactless payments).
33	IDEA Bank (PL)	With a primary orientation towards SME financing, IDEA Bank offers customers the access to a cloud-based transactional platform (Idea Cloud) which integrates a typical business management system (firm's accounts, communications, bills, budgets) with bank accounts, enabling SMEs to manage their bank finances and business operations in a single online platform.
34	Igea Digital Bank (IT)	Igea Digital Bank is an Italian retail bank that offers the full scope of retail banking services exclusively through its digital channels. Also, the bank offers partners an innovative 'Banking as a Service' (BaaS) solution, enabling the integration of financial services into customers' platforms and applications.
35	Ikano Bank (SE)	Specialised on consumer loans, credit cards and retail savings solutions, Ikano's operating infrastructure platform is cloud-based and event-driven, and offers a digital credit solution that allows customers to receive loan application responses in under 60 seconds.

Table 11 (continued)

#	Neobank	Description
36	Illimity Bank (IT)	Illimity is an Italian retail bank that offers the full scope of retail products via its cloud-based digital platform (Amazon Web Services), with AI/Machine Learning capabilities, and integrated with third-party services via Open Banking (API) policy. The bank has a set of partnerships with fintech firms related to payments (HYPE), real-estate (Quimmo), and SME finance (b-ilty).
37	Inbank (EE)	Based in Estonia, Inbank is a retail bank with a focus on payment services. The bank has partnerships with two payment-related fintech firms: Maksekeskus (payment solutions) and Paywerk (Buy-Now-Pay Later platform). The bank is also an active participant in online deposit platforms (Raisin).
38	Keytrade Bank (LU)	Focused on online brokerage services to individual and corporate investors, Keytrade Bank allows customers to open accounts digitally in under 10 minutes, and offers a range of online trading and investing services, which can be accessed through its digital platforms. The bank has a cloud-based core system and an Open Banking (API) policy.
39	Klarna Bank (SE)	Klarna is a Swedish bank focused on providing credit, payments and savings solutions to retail and corporates exclusively via its online channels, in more than 40 countries. Klarna's app integrates traditional banking products as well as the ability directly access the shopping platform; the bank has an Open Banking (API) policy.
40	LHV Pank (EE)	Estonian retail bank offering the full range of products to individuals and SMEs (including leasing), LHV Pank has partnerships with fintech firms in several areas, such as acquiring and issuing services (Nets), bitcoin-based blockchain wallet application (ChromaWay), and cloud-based computing (Tuum). The bank has implemented advanced AI/ML processes in card fraud prevention, and has won several awards for digital innovation (e.g., 'Payments Tech of the Year' in Estonia).
41	Mano Bankas (LT)	Among the first retail banks in Lithuania to operate a digital business model, Mano Bankas is specialised in financially underserved SMEs and individuals. The bank is part of the Currencycloud B2B platform, a well-known international payment platform for digital banks and fintech companies.
42	Marginalen Bank (SE)	Marginalen is a Swedish retail bank endowed with a cloud-based 'Software-as-a-Service' (SaaS) lending platform, which allows an automated risk scoring for loan approvals. The bank has an Open Banking (API) policy, and actively partners with fintech firms focused on online direct-targeted marketing (Match2One) and advanced customer experience (Optimizely).

Table 11 (continued)

#	Neobank	Description
43	Masthaven Bank (GB)	Originally focused on bridging loans and secured lending, Masthaven Bank is a full-scope retail bank offering a cloud-based broker portal (Microsoft ASP.NET Core), B2C and intermediary websites (Kentico), and an Open Banking (API) Policy. Its partnership with Knowledgebank (a fintech firm) allows brokers to have instant access to the bank's full list of lending criteria.
44	Medirect Bank (MT)	Maltese bank offering the full scope of banking products to individuals and SMEs, Medirect has a cloud-based operating system (Snowflake) and provides customers with an innovative digital wealth management platform (WealthTech). The bank has won several awards for its digital capabilities (e.g., 'Best use of Technology in Business Transformation').
45	Monabanq (FR)	Specialised in providing banking services to individuals and self-entrepreneurs, Monabanq offers an AI-powered budget management assistance app (&Vous), an Open Banking (API) Policy, and has won several awards for its digital capabilities (e.g., in 2017–2019 the bank won the 'Customer Service of the Year' prize in the 'Online Banking' category).
46	Monzo Bank (GB)	Monzo is a UK bank providing the full array of retail banking products exclusively through its online channels. The bank operates a fully cloud-based banking system (AWS, including Amazon Elastic Compute Cloud, Block Store and Simple Storage Service), and partners with fintech firms to employ AI/ML analytics tools, including a customer demand forecasting system (with Google's BigQuery).
47	N26 Bank (DE)	One of the most well-known German retail banks operating under a digital business model. N26's core system is cloud-based (Amazon Web Services), has an Open Banking (API) policy, and offers flexible and innovative web-based services such as Perks (cash back rewards for debit spending) and Spaces (personalised subaccounts that allow customers to focus on savings goals).
48	Net-M Privatbank 1891 (DE)	Net-m is German retail bank with a special focus on corporates (e.g., the bank offers online factoring solutions). The bank has developed an Open Banking (API) platform in cooperation with Fiducia, and has set-up partnerships with fintech firms to offer innovative banking products, including 'Banking-as-a-Service' (BaaS).
49	Nordax Bank (SE)	The Swedish bank exclusively uses online platforms to offer the typical range of retail products, primarily focused on individuals (consumer loans, savings solutions, mortgages). Nordax operates with a cloud-based core banking system (Banqsoft) and is an active participant in online deposit platforms (Raisin).
50	Nordnet Bank (SE)	Specialised in investment, trading and pension savings, Nordnet is a Swedish bank that provides digital investment counselling services via a robo-advisor (Robosave), and operates a social investment network (Shareville), where customers can follow the trades made by other investors.

Table 11 (continued)

#	Neobank	Description
51	Northmill Bank (SE)	Northmill is a retail bank based in Sweden, exclusively operating via its online channels. The bank uses state of the art cloud-based solutions, namely Amazon Web Services (AWS) Beanstalk to support financial applications, and customer-facing website and uses Amazon Simple Storage Service (Amazon S3) to store customer data.
52	Oaknorth Bank (GB)	UK-based retail bank targeting primarily SMEs, Oaknorth Bank runs a fully cloud-based operating system (Salesforce/MuleSoft), and an Open Banking (API) policy that has potentiated partnerships with several fintech firms, such as TrueLayer (payments) and Mambu (SME lending).
53	Orange Bank (FR)	Orange Bank is a full-scope retail bank based in France. The bank has a cloud-based infrastructure (Amazon Web Services) and offers customers an AI-powered virtual advisor (Djingo) developed by IBM Watson and Salesforce.
54	Paypal Bank Europe (LU)	Specialised in online payments, Paypal Europe is a retail bank based in Luxembourg that offers credit cards, merchant accounts, and trade credit. The bank operates a cloud-based operating system (Google Cloud), has an Open Banking (API) policy, and applies advanced AI/ML tools in several areas of its business, including fraud detection.
55	Payray Bank (LT)	Lithuanian bank offering online banking services primarily to SMEs (e.g., short-term loans, factoring, leasing). Payray Bank has a fully automated decision-making process for SME lending (BankingLab) and actively participates in online deposit platforms (Raisin). The bank is a member of Lithuania's fintech enterprise cluster (fintech Lithuania).
56	Redwood Bank (GB)	Primarily focused on SMEs, Redwood Bank is a UK-based retail bank with cloud-based operating systems developed by Microsoft (Azure/DPR). The bank has won several awards for its digital capabilities (e.g., Best Challenger Bank and Best Use of Technology).
57	Resurs Bank (SE)	Swedish bank specialised in payment solutions and consumer loans, Resurs Bank has a cloud-based banking platform (Intellect Design Arena) and an innovative app that allows customers to make one-click invoice payments (Resurs). The bank's Open Banking policy (Nordic API Gateway), that has enabled partnerships with fintech companies in serveal areas, such as payments (Aiia) and credit risk scoring (Provenir).
58	Shawbrook Bank (GB)	Shawbrook Bank is a retail bank based in UK, primarily targeting SMEs. The bank has a cloud-based core operating system (Microsoft Azure), as well as advanced AI/Machine Learning capabilities applied to risk and product development (SAS Viya). The bank has an Open Banking (API) policy, and has established partnerships with fintech firms in several areas, including credit risk scoring (ClearScore).

Table 11 (continued)

#	Neobank	Description
59	Starling Bank (GB)	One of the most well-known digitally-oriented banks in the UK, Starling Bank offers the full range of typical banking services to individuals and SMEs. The bank's operating system is cloud-based (Amazon Web Services), uses advanced AI/ML data analytics (Google BigQuery), and has an Open Banking (API) policy fostering partnerships with fintech firms in areas such as trading (moneybox), merchant payments (sumup), international payments (Currencycloud) and treasury (vitesse).
60	Tandem Bank (GB)	Tandem offers retail banking services exclusively via its online channels, with emphasis on sustainability. The bank has a cloud-based operating system (Amazon Web Services), offers an AI-powered personal finance adviser (Ada), employes AI/ML data analytics (intive/AWS) and has an Open Banking (API) policy (via provider TrueLayer).
61	Triodos Bank (GB)	Specialised in ethical and sustainable retail banking services, Triodos Bank is based in the UK and has a cloud-based infrastructure (Microsoft Azure), an automated investment process (via a partnership with fintech Flowable), as well as an Open Banking (API) policy.
62	Vivabank (GR)	Greek retail bank providing services exclusively via online channels and with a primary focus on lending and payment solutions. Vivabank operates entirely cloud-based systems (Microsoft Azure), which enables the use of AI/Machine Learning capabilities in several areas of the bank, including security (Face API). The bank has an Open Banking (API) policy.
63	Vivibanca (IT)	Focused on SME financing and consumer lending products, including salary-secured loans ('Cessione del Quinto'), Vivibanca's onboarding process is fully digital (based on 'Robotic Process Automation' and 'Optical Character Recognition'). The bank has an AI-powered chatbot (ViV-IA) and has implemented an Open Banking (API) policy.
64	Wizink Bank (ES)	Spanish retail bank specialised on online payments, credit cards and savings solutions. Wizink Bank has a cloud-based operating system (Microsoft Azure) and customer relationship management app (Adobe Experience Cloud). The bank has partnerships with several fintech firms to boost its digital capabilities, including in authentication/validation systems (indigitall) and collections (CGI).
65	Zenith Bank (GB)	UK retail bank with the full range of banking services, exclusively provided via its online channels. The bank has an AI-powered chatbot (ZiVA), has a cloud-based operating system (Microsoft Azure), an Open Banking (API) policy, and has won several awards related to its digital capabilities (e.g., Best Digital Bank).

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10693-024-00433-x.

Acknowledgements The authors gratefully acknowledge the support of the COST Action "Fintech and Artificial Intelligence in Finance - Towards a transparent financial industry" (FinAI) CA19130 funded by the Horizon 2020 Framework Programme of the EU.

Funding Open access funding provided by Università degli Studi di Pavia within the CRUI-CARE Agreement. 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) and UIDB/04105/2020 (CEF.UP).

Data availability The data that support the findings of this study are available from the corresponding author, AT, upon reasonable request.

Declarations

Competing Interests The authors have no conflicts of interest to declare. The authors do not have financial or non-financial competing interests that are directly or indirectly related to the work submitted for publication.

References

Abramson R, Kosin V, Howard-Jones D, de Lanzos CL (2020) The post-COVID opportunity for banks. Oliver Wyman 6:2020

Albertazzi U, Gambacorta L (2009) Bank profitability and the business cycle. J Financ Stab 5(4):393–409

Allen F, McAndrews J, Strahan P (2002) E-finance: an introduction. J Financ Serv Res 22(1):5–27

Arner DW, Barberis J, Buckley RP (2017) The evolution of Fintech: a new post-crisis paradigm. Georget J Int Law 47:1271

Arnold IJ, van Ewijk SE (2011) Can pure play internet banking survive the credit crisis? J Bank Financ 35(4):783–793

Athanasoglou PP, Brissimis SN, Delis MD (2008) Bank-specific, industry-specific and macroeconomic determinants of bank profitability. J Int Financ Mark Inst 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. Centre for European Policy Studies, Brussels

Balyuk T, Berger AN, Hackney J (2020) What is fueling fintech lending? The role of banking market structure. The Role of Banking Market Structure

Barney J (1991) Firm resources and sustained competitive advantage. J Manag 17(1):99–120

BCBS (2018) S*ound practices: implications of fintech developments for banks and bank supervisors*. Basel committee on banking supervision, Bank for International Settlements, Basel

Berg T, Fuster A, Puri M (2022) Fintech lending. Ann Rev Financ Econ 14:187–207

Bertay AC, Demirgüç-Kunt A, Huizinga H (2015) Bank ownership and credit over the business cycle: is lending by state banks less procyclical? J Bank Financ 50:326–339

Bikker JA, Haaf K (2002) Competition, concentration and their relationship: an empirical analysis of the banking industry. J Bank Financ 26(11):2191–2214

Boot A, Hoffmann P, Laeven L, Ratnovski L (2021) Fintech: what's old, what's new? J Financ Stab 53:100836

Boot AWA (2000) Relationship banking: what do we know? *J*. Financ Intermed 9(1):7–25

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. J Financ Econ 130(3):453–483
Carbó-Valverde S, Cuadros-Solas PJ, Rodríguez-Fernández F (2021) FinTech and Banking: An Evolving Relationship. In: King T, Lopes FSS, Srivastav A, Williams J (eds) *Disruptive Technology in Banking and Finance*. Palgrave Macmillan, Cham, pp 161–194
Carmichael D (2017) The Winner's curse in an online lending market. Available at SSRN 2931022:1–31. https://doi.org/10.2139/ssrn.2931022
Caves RE, Porter ME (1977) From entry barriers to mobility barriers: conjectural decisions and contrived deterrence to new competition. Q J Econ 91(2):241–261
CCAF. (2022). *Cambridge FinTech ecosystem atlas*. Cambridge Centre for alternative finance (retrieved from: https://ccaf.io/)
Cernov M, Urbano T (2018) Identification of EU bank business models. European Banking Authority Staff Research Paper Series 2:1–52
Claessens S, Frost J, Turner G, Zhu F (2018) Fintech credit markets around the world: size. In: *drivers and policy issues*. BIS Quarterly Review
Claessens S, Glaessner T, Klingebiel D (2002) Electronic finance: reshaping the financial landscape around the world. J Financ Serv Res 22(1):29–61
Clougherty JA, Duso T, Muck J (2016) Correcting for self-selection based endogeneity in management research: review, recommendations and simulations. Organ Res Methods 19(2):286–347
Dadoukis A, Fiaschetti M, Fusi G (2021) IT adoption and bank performance during the Covid-19 pandemic. Econ Lett 204:109904
Degryse H, Ongena S (2005) Distance, lending relationships, and competition. J Finance 60(1):231–266
Delgado J, Hernando I, Nieto MJ (2007) Do European primarily internet banks show scale and experience efficiencies? Eur Financ Manag 13(4):643–671
DeYoung R (2005) The performance of internet-based business models: evidence from the banking industry. J Bus 78(3):893–948
DeYoung R (2010) Banking in the United States. In: Berger AN, Molyneux P, Wilson J (eds) The Oxford handbook of banking. Oxford University Press, Oxford, pp 777–806
DeYoung R, Lang WW, Nolle DL (2007) How the internet affects output and performance at community banks. J Bank Financ 31(4):1033–1060
Di Maggio M, Yao V (2021) Fintech borrowers: lax screening or cream-skimming? Rev Financ Stud 34(10):4565–4618
Diamond DW (1984) Financial intermediation and delegated monitoring. Rev Econ Stud 51(3):393–414
Diamond DW, Dybvig PH (1983) Bank runs, deposit insurance, and liquidity. J Polit Econ 91(3):401–419
Drechsler I, Savov A, Schnabl P (2021) Banking on deposits: maturity transformation without interest rate risk. J Finance 76(3):1091–1143
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
ECB (2022) *SSM supervisory priorities for* 2023-2025. European Central Bank, December 2022, 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
Elsas R, Hackethal A, Holzhäuser M (2010) The anatomy of bank diversification. J Bank Financ 34(6):1274–1287
Frost J (2020) *The economic forces driving Fintech adoption across countries*. Bank for International Settlements, BIS working paper no. 838
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. https://www.fsb.org/wp-content/uploads/P091219-1.pdf
FT (2019) Digital banks seek to turn popularity into profits. Financial Times, London, UK
FT (2020) Digital banks win UK-EU expats' accounts. Financial Times, London, UK
Fuertes AM, Heffernan S, Kalotychou E (2010) How do UK banks react to changing central bank rates? J Financ Serv Res 37:99–130
Fuster A, Plosser M, Schnabl P, Vickery J (2019) The role of technology in mortgage lending. Rev Financ Stud 32(5):1854–1899
Gelis P (2016) Why FinTech banks will rule the world. In: Chishti S, Barberis J (eds) The FinTech book: the financial technology handbook for investors, entrepreneurs and visionaries. Wiley, West Sussex, pp 235–237

Gontarek W (2021) Digital disruption: how the financial services landscape is being transformed. In: *disruptive Technology in Banking and Finance*. Palgrave Macmillan, Cham, pp 221–240
Grigorian DA, Manole V (2006) Determinants of commercial bank performance in transition: an application of data envelopment analysis. Comp Econ Stud 48:497–522
Harrel FE (2001) *Regression Modeling Strategie*s. Springer Series in Statistics, New York
Herpfer C (2021) The role of bankers in the US syndicated loan market. J Account Econ 71(2-3):101383
Jensen MC, Meckling WH (1976) Theory of the firm: managerial behavior, agency costs and ownership structure. J Financ Econ 3(4):305–360
Laeven L, Levine R (2007) Is there a diversification discount in financial conglomerates? J Financ Econ 85(2):331–367
Liberti JM, Mian AR (2008) Estimating the effect of hierarchies on information use. Rev Financ Stud 22(10):4057–4090
Marques BP, Alves CF (2020) Using clustering ensemble to identify banking business models. Intell Syst Account Finance Manag 27(2):66–94
Marques BP, Alves CF (2021) The profitability and distance to distress of European banks: do business choices matter? Eur J Finance 27(15):1553–1580
McGee J (2006) Strategic groups: theory and practice. In: Campbell A, Faulkner D (eds) The Oxford handbook of strategy, a strategy overview and competitive strategy. Oxford University Press, Oxford, pp 267–301
Mergaerts F, Vander Vennet R (2016) Business models and bank performance: a long-term perspective. J Financ Stab 22:57–75
Merton RC (1995) A functional perspective of financial intermediation. Financ Manage 24:23–41
Milne A, Parboteeah P (2016) The business models and economics of peer-to-peer lending. European Credit Res Instit Res Rep 17:1–31. https://cdn.ceps.eu/wp-content/uploads/2016/05/ECRI%20RR17%20P2P%20Lending.pdf
Nel J, Boshoff C (2021) Traditional-bank customers' digital-only bank resistance: evidence from South Africa. Int J Bank Mark 39(3):429–454
OECD (2020) Digital disruption in banking and its impact on competition. Organization for Economic Cooperation and Development, Paris 5 June 2019. https://www.oecd.org/competition/digital-disruption-in-banking-and-its-impact-on-competition-2020.pdf
Oehler A, Horn M, Wendt S (2021) Investor characteristics and their impact on the decision to use a Robo-advisor. J Financ Serv Res 62:91–125. https://doi.org/10.1007/s10693-021-00367-8
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
Porter ME (1979) The structure within industries and companies' performance. Rev Econ Stat 61(2):214–227
Puri M, Rocholl J (2008) On the importance of retail banking relationships. J Financ Econ 89(2):253–267
Reger RK, Huff AS (1993) Strategic groups: a cognitive perspective. Strateg Manag J 14(2):103–123
Rosenbaum PR, Rubin DB (1983) The central role of the propensity score in observational studies for causal effects. Biometrika 70(1):41–55
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. J Financ 31(2):215–231
Shaffer S (1998) The winner's curse in banking. J Financ Intermed 7(4):359–392
Shipman JE, Swanquist QT, Whited RL (2017) Propensity score matching in accounting research. Account Rev 92(1):213–244
Staiger D, Stock J (1997) Instrumental variables with weak instruments. Econometrica 65(3):557–586
Stein JC (2002) Information production and capital allocation: decentralized versus hierarchical firms. J Financ 57(5):1891–1921
Stock J, Yogo M (2005) Asymptotic distributions of instrumental variables statistics with many instruments. In: Andrews DWK, Stock JH (eds) *identification and inference for econometric models: essays in honor of Thomas Rothenberg*. Cambridge University press, pp 109–120
Stulz RM (2019) Fintech, bigtech, and the future of banks. J Appl Corp Finance 31(4):86–97
Tanda A, Schena CM (2019) FinTech, BigTech and banks: digitalisation and its impact on banking business models. Springer: Palgrave Macmillan Studies in Banking and Financial Institutions
Thakor AV (2020) Fintech and banking: what do we know? J Financ Intermed 41:100833
Tseng PL, Guo WC (2022) Fintech, credit market competition, and Bank asset quality. J Financ Serv Res 61(3):285–318. https://doi.org/10.1007/s10693-021-00363-y
van Ewijk S, Arnold I (2014) How bank business models drive interest margins: Evidence from US bank-level data. Eur J Fin 20(10):850–873

Wernerfelt B (1984) A resource-based view of the firm. Strateg Manag J 5(2):171–180
White H (1980) A Heteroskedasticity-consistent covariance matrix estimator and a direct test for Heteroskedasticity. Econometrica 48(4):817–838
Williams J (2021) Fintech—A Perfect Day or Walk on the Wild Side? In: King T, Lopes FSS, Srivastav A, Williams J (eds) *Disruptive Technology in Banking and Finance*. Palgrave Macmillan, Cham, pp 161–194
Zott C, Amit R, Massa L (2011) The business model: recent developments and future research. J Manage 37(4):1019–1042