Competition and Innovation in the Financial Sector: Evidence from the Rise of FinTech Start-ups

Doina Caragea[1] · Theodor Cojoianu[2] · Mihai Dobri[1] · Andreas Hoepner[3] · Oana Peia[4] · Davide Romelli[5]

Abstract

This paper provides new evidence on the effects of entry on incumbents' incentives to innovate by examining the rise of FinTech innovations over the period 2000-2016. We employ machine learning algorithms to classify a large sample of patent applications into five types of FinTech innovations. We then show that greater competition from innovators outside the financial sector increases the probability that incumbent financial firms will innovate. Our identification strategy exploits the variation over time in the share of FinTech patent applications by non-financial start-ups relative to incumbent financial firms, as a proxy for competitive pressures from outside the financial industry. We also find that this increased competition results in a higher number of FinTech patent applications by financial incumbents relative to non-financial ones, especially when the FinTech innovations are more important, as proxied by the number of their future patent citations.

✉ Oana Peia
  oana.peia@ucd.ie

  Doina Caragea
  dcaragea@ksu.edu

  Theodor Cojoianu
  theodor.cojoianu@ed.ac.uk

  Mihai Dobri
  mihaidobri@ksu.edu

  Andreas Hoepner
  andreas.hoepner@ucd.ie

  Davide Romelli
  romellid@tcd.ie

[1] Department of Computer Science, Kansas State University, Manhattan, USA

[2] University of Edinburgh Business School, Edinburgh, Scotland

[3] College of Business, University College Dublin, Dublin, Ireland

[4] School of Economics, University College Dublin, Dublin, Ireland

[5] Department of Economics, Trinity College Dublin, Dublin, Ireland

1 Introduction

The use of technology to provide new and improved financial services, or FinTech, has long been a central characteristic of the finance industry. Since the introduction of wire transfers and ATM machines, technology has had a large impact on how the financial sector operates. Yet, in recent years, a new FinTech revolution of unprecedented speed and magnitude has brought major innovations that have the potential to disrupt the financial intermediation sector (Philippon 2016; Thakor 2020).[1] A key characteristic of this last wave of FinTech innovations is the fact that it is coming from outside the financial sector, as technology start-ups attempt to disrupt incumbent financial institutions.

In this paper, we study whether this increased competition from entering FinTech start-ups affects the incentives to innovate of incumbent financial firms. The causal link between competition and innovation is complex. On the one hand, an increase in competition can discourage innovation by reducing the rents for innovators (Romer 1990; Aghion and Howitt 1992). On the other hand, when competition is weak and pre-innovation rents are high, firms may have less incentive to innovate (Arrow 1962; Gilbert and Newbery 1982). These two competing views are rationalized in Aghion et al.'s (2005) Schumpeterian growth model in which competition and innovation display an inverted U-shaped relationship: at low levels of competition, the entry of new competitors stimulates innovation, while in markets with already high competition, a further increase in competition has a less positive or even a negative effect on innovation as firms with lower market power can extract fewer rents from new ideas and patents. The former effect, known as the "escape competition effect" means that, in industries where competition is weak, incumbents will try to innovate in order to escape from a situation in which competition constrains profits.

Given that the lack of competition has been an endemic problem in the financial sector, the recent FinTech revolution offers a unique experiment to test these theories (Philippon 2016). As such, our main testable hypothesis is that the threat of entry by FinTech innovators from outside the finance industry leads to an increase in innovation by incumbent financial firms.

One empirical challenge in addressing this question is the lack of an established taxonomy to define a FinTech innovation. As such, we first provide an objective, data-based classification of FinTech innovations by exploiting patent filing. We construct a novel dataset of FinTech patent applications based on patents published in the PATSTAT Global dataset over the period from 2000-2016. We apply several machine learning algorithms to the textual data of patent abstracts to classify patents into five categories reflecting key technological characteristics of FinTech innovations: (1) Data Analytics, (2) Fraud, (3) Insurance, (4) Investments and (5) Payments. Our final classification of patents is based on BERT (Bidirectional Encoder Representations from Transformers) models that are a family of machine learning techniques for natural language processing pre-trained on a large corpus of un-annotated text such as Wikipedia (Devlin et al. 2018).

---

[1] These recent FinTech innovations range from mobile payments, money transfers, peer-to-peer lending, and crowdfunding, to more radical innovations such as blockchain, cryptocurrencies, and robo-investing.

We identify a sample of around 19,000 patents corresponding to these categories and match them with firm level information from various sources. We find that US-based private firms are responsible for the large majority of FinTech patents and that these firms are predominantly non-financial.[2] Start-ups account for 13% of patent applications, and payment and data analytics represent the largest categories of FinTech technologies (accounting for 62% of patent applications).

We use this sample of FinTech innovators to examine whether the rise of non-financial start-ups spurs innovation in the financial sector. A key identification challenge in testing the role of competition in innovation is the fact that innovation by both incumbents and potential entrants can respond to the same unobservable technological shocks. We mitigate this concern by building a proxy for competition from outside the financial sector that is less likely to be driven by unobservable technology shocks alone. This measure exploits the time variation in the ratio of FinTech patent applications by non-financial startups relative to the patent applications of incumbent financial firms in the same country over the previous four years. Specifically, our main identifying assumption is that, if financial incumbents are more likely to innovate in periods when the competition from start-ups is stronger (a high ratio of patent applications by startups relative to incumbents), then this increase in innovation cannot be driven by technology shocks alone (which would raise innovation by both incumbents and outsiders thereby making the ratio relatively stable over time).

We find that competitive pressures from non-financial startups increase the probability that financial incumbents become FinTech innovators, that is, apply for a FinTech patent. Specifically, a one standard deviation increase in the ratio of patent applications by non-financial start-ups relative to incumbents over the previous four years, increases the probability that a financial incumbent applies for a FinTech patent by 0.2%. This effect is stronger when we weight the number of patent applications by their relative importance, measured by the forward citations count. These results are robust to controlling for time-varying firms characteristics such as past number of patenting, size, and revenue. We also control for financial incumbents' propensity to acquire FinTech innovators, as opposed to innovating themselves. Not surprisingly, we find that incumbent financial firms who have acquired FinTech innovators in the past are less likely to apply for FinTech patents.

Furthermore, as many of the patent filing in our sample belong to non-financial firms, we also investigate if competition from start-ups affects innovation relatively more among financial as compared to non-financial firms. We find that the total number of FinTech patent applications is higher among financial firms relative to non-financial firms when competitive pressures from start-ups are stronger. This result suggests that the rise of FinTech startups has disproportionally driven financial incumbents to innovate. Moreover, we show that competition from incumbent technology firms does not have the same effect and is not followed by an increase in FinTech patent applications by financial incumbents.

Our estimations are robust to a wide array of model specifications and alternative definitions of our main explanatory variable. We also include, across all specifications, time fixed effects to account for technology waves that spur FinTech innovations across all firms as well as firm fixed effects that allow us to obtain identification from within-firm variation over time. We further mitigate omitted variable bias by implementing an instrumental variable strategy where we instrument competition by start-ups by the beginning of sample ratio of venture capital investment to GDP in the country.

---

[2] We define a financial firm as belonging to the NACE Rev2 two-digit industry codes 64, 65 and 66, respectively.

Our work is related to several branches of literature. First, we contribute to a growing literature that uses taxonomies to automatically classify patents across different categories of innovations (see, among, others Fall et al. 2003; Benzineb and Guyot 2011; Gomez and Moens 2014; Grawe et al. 2017; Li et al. 2018; Shalaby et al. 2018; Hu et al. 2018; Abdelgawad et al. 2019; Lee and Hsiang 2019; Mann and Püttmann 2021). Chen et al. (2019) were the first to propose a typology to classify a sample of 2 million patents filed with the US Patent and Trademark Office (USPTO) into seven categories of FinTech innovations: cybersecurity, mobile transactions, data analytics, blockchain, peer-to-peer, robo-advising, and the internet of things. They train and employ several families of machine learning algorithms such as support vector machines and neural networks to identify a sample of FinTech innovations by US-based firms or individuals. Chen et al. (2019) then investigate the value of these FinTech innovations by studying stock market responses around the news of patent filings. They find that the FinTech patents that are most valuable to innovators are blockchain, cybersecurity, and robo-advising. Xu et al. (2020) also train machine learning classifiers such as random forest classifiers to identify FinTech patents in a sample of patent applications to the USPTO from 2014-2018.

We extend these efforts to automatically classify FinTech innovations in several directions. For example, we provide an enhanced FinTech patent classification by training and evaluating different types of deep learning classifiers, with a focus on BERT models (Devlin et al. 2018). Specifically, while other studies that examine patent texts have used traditional machine learning approaches such as linear support vector machines or neural network models, recent research in computer science has shown that deep learning approaches outperform such methods (Abdelgawad et al. 2019). We thus aim to provide a new and improved taxonomy of FinTech innovations by using deep learning methods.

Second, we employ our classification of FinTech patents to investigate how competition drives innovation in the financial industry. The relationship between innovation and competition is of long-standing theoretical and empirical interest (Aghion et al. 2001, 2005; Aghion and Griffith 2008). A large Schumpeterian growth literature models two effects of entry on incumbent innovation. In industries with high levels of competition, entry reduces incumbents' incentives to innovate as it decreases innovation rents and reduces the probability of surviving entry (*Schumpeterian effect*). In industries with low levels of competition, where leaders compete *neck-to-neck*, an increase in competition has a positive effect and encourages incumbents to innovate in order to acquire a lead over their rival in the sector (*escape competition effect*) (Aghion et al. 2014).

Empirical works have documented these two effects across different countries, time periods, and industries. Studies on the effects of competition-enhancing reforms or foreign competition find an overall positive effect (see, for example Aghion et al. 2005; Ayyagari et al. 2011; Bloom et al. 2016), but negative effects are also found (Liu et al. 2014; Dorn et al. 2020). Moreover, the effect depends on the level of technological advancement in the industry (Aghion et al. 2009; Liu et al. 2014) or type of innovation (Tang 2006). A key empirical challenge in this literature is that entry is endogenous to the innovation of incumbents. The FinTech revolution offers an interesting setting to test this relationship, as the entry of start-ups was not likely driven by innovations of large financial incumbents. For example, Cojoianu et al. (2020) study the determinants of the emergence of FinTech start-ups across 21 OECD countries and find that creating a FinTech venture is positively related to the regional productivity and new knowledge created in the information technology (IT) sector, but not that created in the financial sector.

Our work also relates to a growing literature that looks at how the dramatic growth of FinTech start-ups has shaped the financial landscape (see Aaron et al. 2017; Alt et al. 2018;

Temelkov 2018; Stulz 2019; Allen et al. 2020). For example, Hornuf et al. (2021) study how banks interact with FinTech start-ups using detailed information on strategic alliances made by the 100 largest banks in Canada, France, Germany, and the UK. They show that banks are more likely to form alliances when they have a well-defined digital strategy or employ a chief digital officer (see also Li et al. 2017; Tseng and Guo 2022). Kowalewski and Pisany (2020) document the rise of FinTech start-ups across a large cross-section of countries and highlight the importance of technological advancements and university-industry collaborations in FinTech creation and activity.

The remainder of the paper is organized as follows: Section 2 presents the patent data and the machine learning techniques used to classify the FinTech patents, while we the identification strategy and methodology in Section 3. Section 4 presents the results, and Section 5 concludes.

2 Patent data and classification of FinTech innovations

In this section, we describe the process followed to create a FinTech taxonomy as well as to train the machine learning and deep learning models used to classify FinTech patents. We also provide some descriptive statistics of the identified dataset of FinTech patent applications.

2.1 Patent data

Patent data is obtained from the BvD Orbis database that sources information from PATSTAT Global[3]. We retrieve all patents applications filed between 2000 and 2016 that belong to Classes G and H of the International Patent Classification (IPC), and cover areas related to digital computing that underlie the FinTech technologies classified in this paper. This results in a sample of 6.8 million patents. To narrow down the search, we use a text-based filtering to identify patents that are plausibly related to financial services. We obtain from Chen et al. (2019) a list of 516 financial terms based on Campbell R. Harvey's Hypertextual Finance Glossary and the online Oxford Dictionary of Finance and Banking. We select patents that contain at least one keyword from this list in their abstract. After this filtering, we identify a potential number of 38,228 patents that are related to financial services.[4]

The distribution of the 100 most frequent financial terms in our dataset is illustrated in Fig. 1, where the size of each term is proportional to the number of patent documents in which that term appears. The most frequent terms include *payment*, *compliance*, *trading*, *banking*, *insurance*, *investment*, *money*, among others.

2.2 FinTech innovation taxonomy

There is a wide range of financial products and services that fall under the FinTech umbrella. Currently, there is no comprehensive, well-accepted taxonomy to analyze this sector. Hence, we build a FinTech taxonomy by comparing the taxonomies from several academic works, industry reports and market maps (such as Mellon 2015; Levy 2015; Young and Treasury

---

[3] PATSTAT is maintained by the European Patent Office (EPO) and comprises patent applications with the EPO as well as national patent offices of large advanced and developing countries.

[4] Appendix Table 12 shows how our dataset of FinTech patents compares to previous classifications in Chen et al. (2019) and Xu et al. (2020).

Fig. 1 WordCloud of the 100 most frequent financial terms. The figure shows the 100 most frequent financial terms in the set of 38,228 patents. The size of a term is proportional to the number of patents in which the term appears

2016; CBInsights 2017; Eckenrode and Friedman 2017; Chen et al. 2019; Haddad and Hornuf 2019).

Our taxonomy aims to capture innovations that pursue the integration of more sophisticated IT tools and data science solutions in financial products. We include five broad FinTech categories: *Data Analytics*, *Fraud*, *Insurance*, *Investments* and *Payments*. Applications corresponding to these categories, together with an example of a patent filing abstract in each category are shown in Table 1. Our FinTech taxonomy is aligned with that of Chen et al. (2019), although we use a broader range of financial services, such as including insurance. We also exclude some applications which are not necessarily specific to the financial sector such as *blockchain* or the *internet-of-things*.

2.3 Machine learning algorithms to classify FinTech patents

To train machine learning and deep learning models to identify FinTech patents, we manually labeled a subset of our patent dataset. Specifically, we manually labeled 2,350 FinTech patents (500 patents in each of the following categories, insurance, fraud, investments and payments; and 350 patents in the category of data analytics). Furthermore, we manually labeled a subset of 1,500 non-FinTech patents. We use the 3,850 manually labeled patents to train and test different types of text classifiers.

We focus on BERT models as our main text classifier, given the success of the deep learning approaches on patent classification tasks (Li et al. 2018; Shalaby et al. 2018; Sun et al. 2019; Lee and Hsiang 2019). BERT is a language model that uses a deep bidirectional transformer encoder architecture to encode sentences and their tokens into dense vector representations (Vaswani et al. 2017; Devlin et al. 2018). BERT models are pre-trained on a large corpus

Table 1 FinTech Taxonomy. The table presents examples of the abstracts of the patent filings for the five FinTech categories we identify

FinTech category	Applications	Examples of patent filing abstracts
Data Analytics	Software, data infrastructure and analytics for financial services	"Computer programs are provided for automated generic and parallel aggregation of characteristics and key figures of unsorted mass data being of specific economic interest, particularly associated with financial institutions, and with financial affairs in banking practice."
Fraud	Fraud detection, security infrastructure, identity verification & compliance	"This invention provides a system and method for reducing the fraud related to remittance transactions initiated at web portals. [..] For example, a funding agency computer that enables a remittance transaction can request that a mobile platform computer verify a customer with a mobile personal identifier. The mobile platform computer can request the mobile personal identifier from a customer via the customer's mobile handset device."
Insurance	Life, general & (re)insurance software analytics	"An automated assignment system may operate with a computer to automatically assign insurable events to one or more organizational entities associated with an insurance organization. The automated assignment system may categorize the insurable event."
Investments	Portfolio management, lending and investing platforms, portfolio analytics	"A visual interactive multi-criteria decision-making method and computer-based apparatus for portfolio management. The method/apparatus supports partitioning of a portfolio of physical or other assets into two mutually exclusive categories, such as assets recommended for sale and assets recommended for retention."
Payments	Mobile payments & transfer	"A mobile payment platform and service provides a fast, easy way to make payments by users of mobile devices. The platform also interfaces with nonmobile channels and devices such as e-mail, instant messenger, and Web. In an implementation, funds are accessed from an account holder's mobile device such as a mobile phone or a personal digital assistant to make or receive payments."
	Digital wallets	"A system and a method are provided for generating a digital receipt for purchases made utilizing a digital wallet or with other payment procedures. The digital receipt is stored in the cloud in a digital receipts repository for later retrieval. The digital receipt can be standardized to facilitate data processing of the data contained in data fields of the digital receipt."

Table 2 Performance results of traditional and deep learning models. Table shows the out-of-sample performance of each machine-learning method. Accuracy is one minus the ratio of incorrect category predictions to total observations. Precision is the ratio of true positives to the sum of true positives and true negatives. Recall is the ratio of true positives to the sum of true and false positives. F1 score is the harmonic mean of Precision and Recall. Columns (1)-(7) correspond to traditional machine learning models that have been commonly employed to classify patent data. These are: Linear Support Vector Machines (Linear-SVM), Gaussian Support Vector Machines (Gaussian-SVM), Naïve Bayes (NB), k-nearest neighbour (kNN), random forest (RF), gradient boosting (GB) and Multilayer Perceptron (MLP). Column (8) is a voting classifier that aggregates predictions of the Linear-SVM, Gaussian-SVM, and neural network models and uses the prediction made by the MLP, in case of ties. The deep learning methods employed are: Convolutional Neural Networks (CNN), recurrent neural networks (RNN) and Bidirectional Encoder Representations from Transformers (BERT). The best result for each criteria and category is marked in bold

Category	Metric	Traditional machine learning models								Deep learning models		
		Linear SVM	Gaussian SVM	Naïve Bayes	kNN	Random Forest	GB	MLP	Voting classifier	CNN	RNN	BERT
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Insurance	Precision	0.779	0.857	**0.878**	0.689	0.824	0.864	0.848	0.854	0.904	**0.922***	0.891
	Recall	0.844	0.812	0.750	0.875	0.635	0.792	**0.927**	0.792	0.979	**0.990***	**0.990***
	F1-score	0.810	0.834	0.809	0.771	0.718	0.826	**0.886**	0.822	0.940	**0.955***	0.938
Payments	Precision	0.827	**0.852**	0.645	0.636	0.802	0.800	0.835	0.851	**0.990***	0.884	**0.906***
	Recall	0.910	0.920	0.800	0.840	0.650	0.840	**0.960**	0.860	0.934	**0.990***	0.970
	F1-score	0.867	0.885	0.714	0.724	0.718	0.820	**0.893**	0.856	0.934	0.934	**0.937***
Investment	Precision	0.837	0.850	0.765	0.725	0.778	0.796	**0.870**	0.861	**0.943***	0.934	0.924
	Recall	0.870	**0.960**	0.910	0.870	0.770	0.860	0.940	0.930	**0.990***	**0.990**	0.980
	F1-score	0.853	0.901	0.831	0.791	0.774	0.827	**0.904**	0.894	**0.966***	0.961	0.918
Fraud	Precision	0.788	0.790	0.676	0.617	0.813	0.815	0.849	**0.884**	0.951	**0.961***	0.960
	Recall	0.780	**0.830**	0.750	0.740	0.610	0.750	0.790	0.760	**0.980***	**0.980***	**0.980***
	F1-score	0.784	0.810	0.711	0.673	0.697	0.781	**0.819**	0.817	0.966	**0.970***	**0.970***
Data Analytics	Precision	0.530	0.581	0.351	0.420	**0.700***	0.565	0.581	0.513	0.635	0.635	**0.652**
	Recall	0.700	0.720	0.660	0.420	0.280	0.700	0.720	**0.800**	0.800	0.800	**0.900***
	F1-score	0.603	0.643	0.458	0.420	0.400	0.625	**0.643**	0.625	0.708	0.708	**0.753***

Table 2 continued

Category	Metric	Traditional machine learning models								Deep learning models		
		Linear SVM (1)	Gaussian SVM (2)	Naïve Bayes (3)	kNN (4)	Random Forest (5)	GB (6)	MLP (7)	Voting classifier (8)	CNN (9)	RNN (10)	BERT (11)
Non-FinTech	Precision	0.939	0.955	0.968	**0.980**	0.700	0.880	0.962	0.923	0.992	0.996	**1.000***
	Recall	0.823	0.850	0.697	0.660	**0.927***	0.853	0.843	0.873	0.857	**0.863**	0.850
	F1-score	0.877	**0.899**	0.810	0.789	0.798	0.866	**0.899**	0.897	0.919	**0.925***	0.918
Average	Precision	0.783	0.814	0.714	0.678	0.770	0.787	**0.824**	0.814	0.885	**0.889***	**0.889***
	Recall	0.821	0.849	0.761	0.734	0.645	0.799	**0.863**	0.836	0.933	0.935	**0.945***
	F1-score	0.799	0.829	0.722	0.695	0.684	0.791	**0.840**	0.819	0.905	0.909	**0.912***
	Accuracy	0.830	0.858	0.751	0.735	0.745	0.820	**0.867**	0.849	0.921	**0.924***	0.923

* marks the best result in a row

Table 3 Filtering steps in the construction of the dataset of patent applications. The table summarizes the number of patent filing selected in each step of the construction of the dataset of patent applications

(A) Source of patents: PATSTAT	
(B) Years covered: 2000-2016	
(C) Legal jurisdiction of patents: US + Europe	
(D) IPC classes used: G and H	
(1) Initial number of patents based on criteria (A)-(D) above	6.8M
(2) Financial terms for filtering financial patents	516
(3) Number of patent applications after filtering with financial terms (2)	38,228
- Number of FinTech categories considered: 5	
- Number of manually annotated patents used for training: 3,850	
(4) Total number of FinTech patents identified out of (3)	19,055

of un-annotated text (such as Wikipedia) using two self-supervised learning tasks: masked word prediction and next sentence prediction. A generic BERT model can then be further pre-trained and fine-tuned for specific natural language processing tasks.

As it is common in the literature, we do not simply rely on a single approach, but rather use several different families of algorithms, which have been successfully applied to classification problems in other domains. Specifically, we compare the results obtained from BERT models with text classifiers such as CNN (Convolutional Neural Networks) and RNN (Recurrent Neural Networks) that use BERT embedding as input. Finally, to allow comparison with previous FinTech patent classifications in Chen et al. (2019) and Xu et al. (2020), we also contrast the results of the deep learning models with traditional machine learning approaches such as support vector machines (SVM), neural networks, Naïve Bayes, k-nearest neighbour (kNN), random forest, and gradient boosting. The Online Appendix presents details on the machine learning methods we employ.

To evaluate the performance of the various models that we train, we use four standard metrics: accuracy, precision, recall and F1-score.[5] We also report these metrics for each of our five FinTech categories to determine what categories are the easier or harder to identify. The performance results on the test data are presented in Table 2. The first eight columns correspond to traditional machine learning models, while the last three depict deep learning models. The best result for each method and evaluation criteria is highlighted in bold, while the best overall result in a row is marked with a star. Based on the average score for all performance criteria, we find that the best machine learning model is the MLP model, while the best deep learning model is the BERT model.

We thus use the BERT model to classify our sample of patents. Table 3 summarizes the number of patent filings removed at each step of our methodology. The classifiers we employ allow us to identify a sample of 19,055 unique FinTech patent applications.

2.4 Descriptive statistics

Table 4 shows the frequencies of FinTech patent applications filed by various groups of innovators. From the sample of 19,055 patents identified by the machine learning algorithms

---

[5] Accuracy is one minus the ratio of the number of incorrect category predictions to the total number of observations. Precision is the ratio of true positives to the sum of true positives and true negatives. Recall is the ratio of true positives to the sum of true and false positives. F1 is the harmonic mean of precision and recall.

Table 4 FinTech innovation by innovator type and region. This table shows the frequencies of FinTech patent filings by various types of innovators and across regions. Financial firms represent those firms belonging to the NACE Rev2 two-digit industry codes 64, 65 and 66, respectively. Start-ups are defined as firms with a founding date of no more than eight years prior to the patent filing date

	Number of FinTech patent filing
Innovator type	
Individual	3,166
Public company	5,028
Private company	10,861
Company type	
Financial services firms	3,236
Monetary intermediation & financial services (NACE code 64)	1,893
Insurance companies (NACE code 65)	457
Auxiliary fin. services & fund management (NACE code 66)	886
Non-financial firms	12,653
Start-ups	2,012
Geographical location of company	
US	12,322
Europe	1,636
Rest of the world	1,931

described in the previous section, 3,166 (20%) were filed by individuals. This leaves a sample of 15,889 patent applications by firms that will be the focus of our analysis. Among these, private firms are the most important group of FinTech innovators by accounting for 68% of the applications. Non-financial firms are also responsible for a large share of innovations in FinTech, while startups account for 13% of the applications.

Furthermore, while the 15,889 patents are filed at the European and US Patent offices by firms across 58 countries, the large majority of filing (77.5%) belong to US-based firms. The geographical split is similar if we organize the data based on the priority country of the patent (first country in which the patent application was filed), as opposed to the headquarters of the inventor, with 82.1% (13,050) of applications having the US as priority country, 8.5% (1,351) Europe, and 9.4% (1,488) in the rest of the world.

Table 5 shows the distribution of innovators across the different types of technologies classified as FinTech innovations. Data analytics and Payments represent the largest share of FinTech innovations (62%). Table 5 also shows that private and non-financial firms dominate all types of FinTech technologies. Similarly, startups innovate evenly across the 5 types of technologies considered.

Next, we show some trends in the evolution of FinTech innovation over time. Figures 2 and 3 display the 6-months moving averages of patent filing from 2000 to 2016. Figure 2 shows the evolution for financial versus non-financial firms. Non-financial firms dominate the patent filing numbers in all periods, but the pace at which applications are filed by these two groups of firms shows considerable variation over time.

Figure 3 plots the evolution of FinTech patents by company age and sector. It shows that while non-financial start-ups and financial incumbents have similar patent filing rates prior

Table 5 FinTech innovation activity by technology category and innovator type. The table shows the frequencies of FinTech patent filing by various types of innovators for each technology category. Financial firms represent those firms belonging to the NACE Rev2 two-digit industry codes 64, 65 and 66, respectively. Start-ups are defined as firms with a founding date of no more than eight years prior to the patent filing date

FinTech Classification	Number of applications	Innovator type		Company type		
		Private firms (%)	Public firms (%)	Financial firms (%)	Non financial firms (%)	Start-ups (%)
Data analytics	4,790	64	36	18	82	11
Fraud	1,075	65	35	9	91	13
Insurance	1,343	81	19	35	65	11
Investment	3,657	76	24	30	70	15
Payments	5,024	64	36	14	86	13
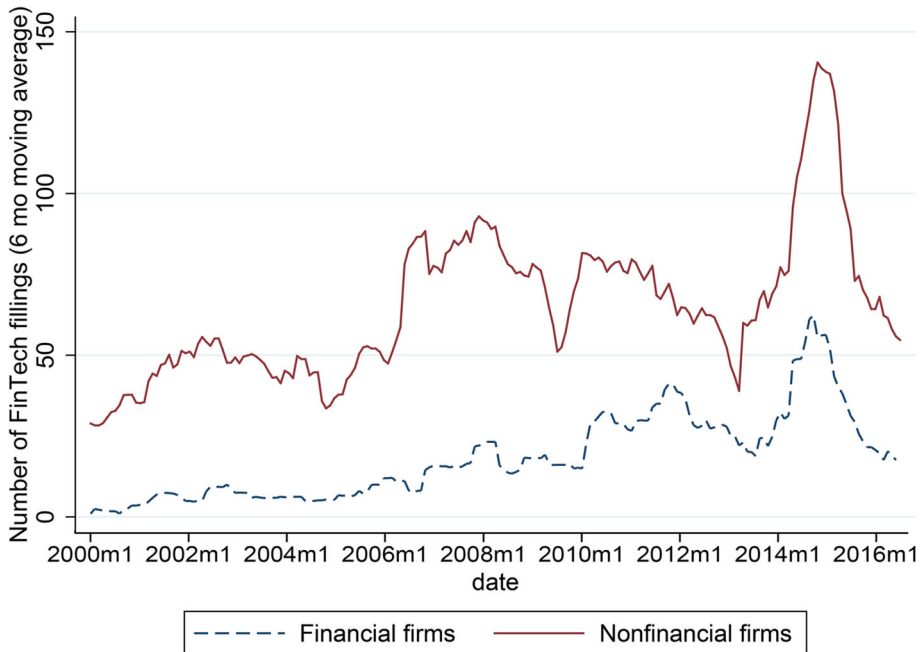
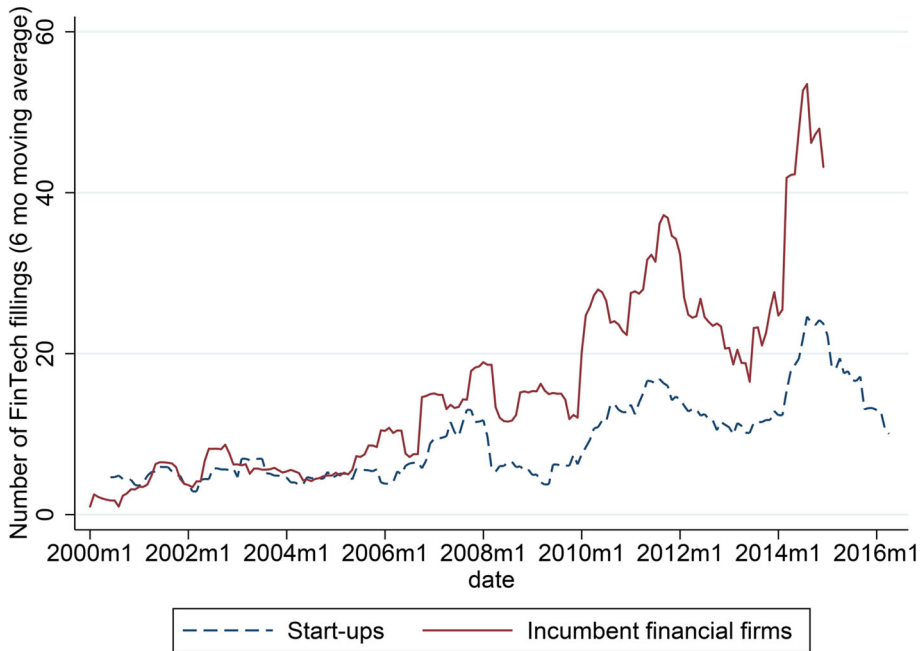**Fig. 2** FinTech innovation over time by financial versus non-financial firms. The figure shows the 6-month moving average of patent applications between 2000 and 2016 by financial firms (corresponding to NACE 2 digits codes 64, 65, and 66) versus non-financial firms

to 2007, incumbents dominate filing in the second half of the sample. Our empirical strategy exploits this time variation in patent filings across these two sectors.

Finally, Fig. 4 shows the evolution of patent applications by type of FinTech technology. It is apparent that the increase in FinTech innovations occurred across all types of technologies.

3 Identification strategy

Our empirical strategy aims to identify how competition from FinTech startups affects innovation efforts by incumbent financial firms. This strategy faces two main identification challenges. The first is the fact that entry can be endogenous to innovation. In other words, the innovations and potential productivity gains in the finance industry can drive the innovations in financial technologies by entrants. Several arguments suggests that this is not a very likely concern. First, there is ample anecdotal evidence that the most recent FinTech innovations, such as digital advisory and trading systems, artificial intelligence, machine learning, peer-to-peer lending, equity crowdfunding and mobile payment systems, have originated outside the incumbent financial industry (Philippon 2016; Vives 2019).

Second, the financial industry is rather inefficient. Philippon (2015) estimates that the unit cost of financial intermediation in the US has remained around 2% for the past 130 years. As such, innovations by financial incumbents have done little to increase productivity and translate into lower costs for the end users. Therefore, the argument that the most recent

Fig. 3 FinTech innovation over time by non-financial start-ups versus incumbent financial firms. The figure shows the 6 months moving average of patent applications between 2000 and 2016 by incumbent financial firms (corresponding to NACE 2 digits codes 64, 65 and 66) versus start-ups outside the finance industry. Start-ups are defined as firms less than eight years old at the time of the patent application

FinTech revolution is not driven by innovations already happening in the financial sector is plausible.

The second threat to identification is that innovations by both start-ups and financial incumbents are driven by the same unobservable technology shocks. This is a more plausible concern and ideally one would use an exogenous shock to identify FinTech innovation among start-ups. However, in the absence of such a shock, we address this concern by controlling for the amount of patent applications by incumbent competitors in the financial industry. The argument is that, if unobservable technology shocks are driving innovation in both incumbents and outsiders, then the rate of innovation between the two groups would be relatively stable over time. This, however, does not seem to be the case. Figure 5 shows the evolution of the share of patent applications by (non-financial) start-ups as a fraction of applications by financial incumbents. We observe a higher share of FinTech patents by start-up firms in the beginning of the sample (2000-2008), which reverses in the second part of the sample. Moreover, there is significant variation in this ratio across time, which is key to our identification strategy. We exploit this time variation as a source of competitive pressure from outside the financial sector. As such, if financial incumbents are more likely to innovate when the innovation by outsiders is higher than the one from insiders (a high ratio of patent applications by startups vs incumbents) then it is less plausible that the higher patent output by financial incumbents is driven by technology shocks alone. We can thus argue that incumbents respond to competition by innovating themselves (escape competition effect).

We further deal with omitted variable bias by considering not only the number of patent applications, but also their importance captured by the citation count (Akcigit and Kerr 2018).
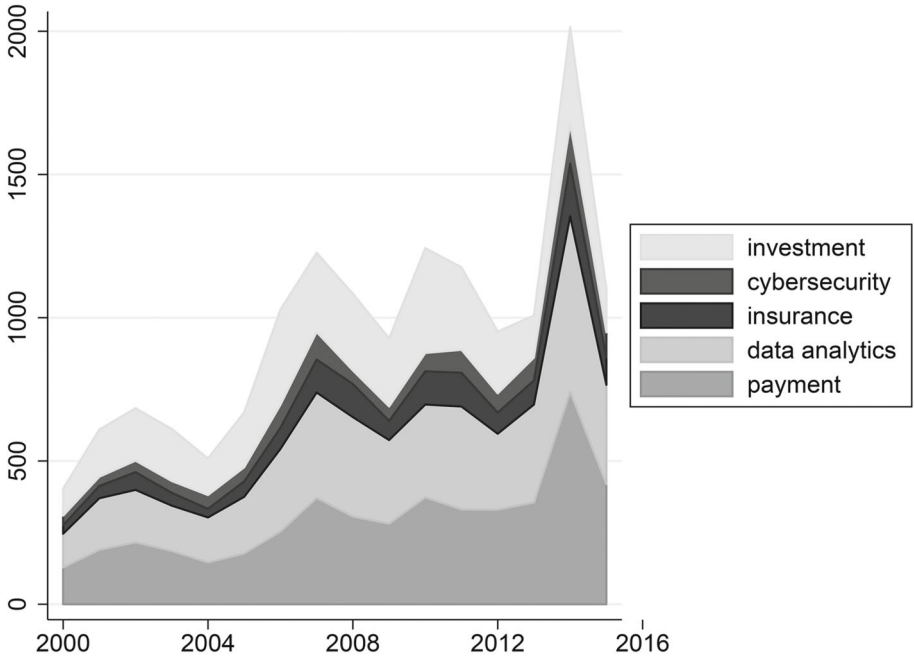
Fig. 4 FinTech patents by technology type. The figure shows the total number of applications by year and type of FinTech innovation
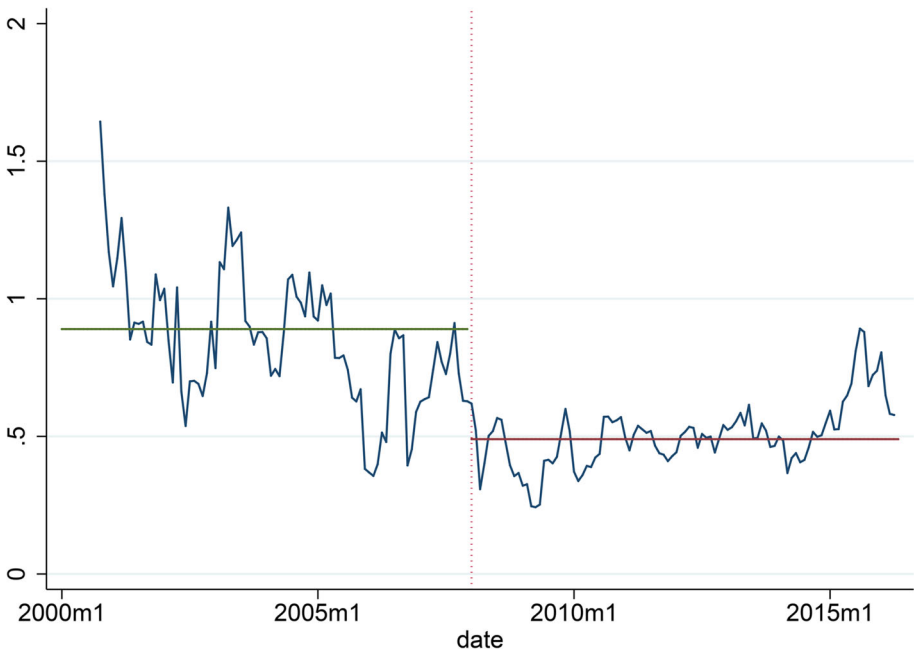

Fig. 5 Ratio of patent applications by start-ups relative to incumbent financial firms (6 months moving average)

This is done by scaling our measure of competition by the future citation count of start-ups'
patents at the end of the sample. This measure of competition weighted by forward citations
is meant to distinguish small incremental FinTech patents from those that bring significant
innovations to the sector. As such, we expect competition to be stronger when the innovation
by outsiders is more important.

Our baseline strategy is as follows:

$$\text{FinTech Innovation}_{i,t} = \alpha_i + \gamma_t + \beta_1 \text{Competition Startups}_{c,t-4 \rightarrow t-1} + \theta' X_{i,t} + \epsilon_{i,t}, \quad (1)$$

where FinTech Innovation$_{i,t}$ is an indicator variable equal to one if firm i is an incumbent
financial institution and has applied for at least one FinTech patent in year t. We define an
incumbent financial firm as one having the 2-digit NACE Rev 2 codes of 64, 65 and 66, and
is not a start-up, that is, founded more than eight years prior to the patent application date.

Our main independent variable is the ratio of FinTech patent applications by start-ups rel-
ative to financial incumbents (see also Etro 2004; Cockburn and MacGarvie 2011; Czarnitzki
et al. 2014), that is computed as:

$$\text{Competition Start-ups}_{c,t-4 \rightarrow t-1} = \frac{\sum_{\tau=t-4}^{t-1} \text{Startup Patents}_c}{\sum_{\tau=t-4}^{t-1} \text{Financial incumbents}_{-i,c}}. \quad (2)$$

In the baseline model, we consider the sum of patents in the previous four years (excluding
the current year) by firms in the same country as financial incumbent i. The sum in the
denominator in (2) excludes the patent applications by incumbent i, which we add as a
separate explanatory variable.[6] One concern with the specification in (2) is that the number
of innovations by incumbent i is positively correlated with the sum of innovations of other
financial incumbents. However, this correlation would put a downward bias on our estimations
that would lessen the effect of competition by startups. Another concern is that the shock
affects incumbents versus outsiders to a different extent over time. In particular, periods
in which the unobservable shock drives more innovation among incumbents leads to lower
competition ratios resulting in a downward bias. On the other hand, periods where the shock
leads to more innovations among start-ups create an upward bias. It is not a priori clear which
direction the bias would be across the entire time period, and, as a result, it is difficult to
argue that the bias is consistently upwards or downwards over the entire period.

The model in Eq. (1) also controls for time varying and invariant covariates at the firm
and country level. The vector $X_{i,t}$ includes measures of the past innovation and financial
performance of the firm. We control for the lagged sum of patent applications by firm i over
$[t-4, t-1]$ to account for the propensity of an incumbent to produce FinTech innovations. We
also include measures of firms size, such as log of Total Assets and Operating Revenue. We
also control for the propensity of a firm to acquire FinTech innovators as opposed to innovating
itself. This is captured by the sum of acquisitions, mergers, follow-on funding, management
buyout (MBO), or joint ventures by financial incumbent i with a FinTech innovator over the
period $[t-4, t-1]$. Finally, we control for time-invariant firm characteristics through firm
fixed effects (α_i), as well as year fixed effects that capture waves of innovation across all
countries in a given year.

---

[6] Note also that, if the sum of patents by other incumbents, $\sum_{\tau=t-4}^{t-1} \text{Financial incumbents}_{-i,c}$, is zero, then
we assume that the measure of competition is just the sum of start-up patents, $\sum_{\tau=t-4}^{t-1} \text{Startup Patents}_c$.

Table 6 Descriptive statistics. The table presents the descriptive statistics for the two main independent variables, i.e., the *Competition Start-ups* ratio and the *Competition Start-ups (citations)* ratio, as well as the firm control variables

	Full Sample			Financial incumbents			Non-financial start-ups			Non-financial firms		
	Mean	St dev	N	Mean	St dev	N	Mean	St dev	N	Mean	St dev	N
ln(Total Assets)	7.35	4.26	20,790	10.69	4.86	2,607	4.74	4.04	4,129	7.60	3.56	13,430
ln(Operating Revenue)	6.92	3.90	20,258	8.81	4.13	2,528	4.55	3.90	3,866	7.40	3.45	13,268
FinTech Acquisitions	0.00	0.04	20,790	0.01	0.04	4,883	0.00	0.02	11,011	0.00	0.04	20,790
Past Patents	1.36	9.20	20,790	3.22	17.34	4,628	0.60	2.21	9,848	1.34	8.91	20,790
Competition Start-ups	0.234	0.699	20,790									
Competition Start-ups (Citations)	0.207	0.628	20,790									

3.1 Financial data

We match our new database of FinTech patent applications by public and private firms with several datasets containing financial data. The primary source of company data is Orbis (Bureau van Dijk) that we complement with data from Capital IQ to obtain a longer time series for a subset of listed firms. We keep in our sample countries that have at least one patent filing from a financial incumbent that yields a final dataset of 15,889 patent applications from 3,446 unique firms over the period from 2000-2016.

To control for the propensity to acquire FinTech innovators, we also collect data on the universe of acquisitions, mergers, follow-on funding, MBO, or joint ventures, where at least one of the members of the deal is a FinTech innovator identified by our machine learning algorithm. We obtained this data from Zephyr.

A summary of the financial information by types of FinTech innovators is presented in Table 6. Financial incumbents tend to be larger and have higher operating revenue than the average firm in the overall sample. They also have, on average, more patents and acquire more FinTech innovators. At the end of Table 6 we also present the descriptive statistics for our two main independent variables the *Competition Start-ups* ratio and the *Competition Start-ups (citations)* ratio, with the latter being the forward citations weighted variant of the former.

4 Results

The results from our baseline model in Eq. (1) are presented in columns (1)-(3) of Table 7, in which the dependent variable is an indicator equal to one if a financial incumbent applied for a FinTech patent in year t.

Our main explanatory variable is competition by start-ups (Column 1) measured by the ratio of patent applications by start-ups relative to those by other financial incumbents over the period $[t-4, t-1]$. We find evidence that following periods with relatively higher Fin-Tech innovations by non-financial start-ups, financial incumbents are more likely to innovate themselves. To ease interpretation of the point estimates, we have standardised throughout all estimations the measure of Competition Start-ups. As such, the point estimate in Columns (1) suggests that a one standard deviation increase in the Competition Start-ups ratio increases the probability of being an innovating financial incumbent by 0.2%.[7]

We consider next the importance of FinTech innovations by start-ups by weighting the number of patent application by their forward citations count (Column 2). Again, we find that financial incumbents are more likely to apply for FinTech patents particularly when the quality of patent applications by outsiders is higher. The point estimate in Column (2)

---

[7] Our baseline estimations employ a linear probability model (LPM). We have opted for the LPM for several reasons, which include the ease of interpretation of the results and its flexibility in including two-way fixed effects. One important shortcoming of LPM vis-à-vis non-linear models for binary response such as probit or logit is that the LPM estimates are not constrained to the unit interval, which can lead to biased and inconsistent estimates. Moreover, Horrace and Oaxaca (2006) show that the potential bias of the LPM increases with the proportion of LPM predicted probabilities that fall outside the unit interval. However, for the model estimated in Table 7 column (1) less than 0.1% of predicted probabilities fall outside the unit interval. At the same time, as Angrist and Pischke (2009) argue, when the goal of the econometric model is to estimate the partial effect of a single independent variable on the response probability, the LPM does a good job at estimating the linear conditional expectation function (CEF) for a (saturated) model. If the CEF is non-linear, then LPM provides the best linear approximation to it. There is no guarantee that a logit/probit estimator provides a better fit to a non-linear CEF than the linear approximation (Angrist and Pischke 2009).

Table 7 FinTech patent applications and competition from start-ups. The dependent variable in columns (1)–(3) is an indicator variable equal to one if a financial firm applied for a FinTech patent in year t, while in columns (4)–(6) it is the total number of FinTech patent applications by a firm in year t. Financial firms correspond to the 2-digit NACE Rev 2 codes: 64, 65 and 66. Competition Start-ups is the ratio of the sum of FinTech patent applications by non-financial start-ups relative to that of financial incumbents over the period $[t-4, t-1]$. Competition Start-ups (Citations) is the Competition Start-ups ratio scaled by the number of forward citations of non-financial start-ups. Number Start-ups is the ratio of the total number of start-ups applying for a FinTech patent over the total number of incumbents during the previous four years. Finance is a dummy variable for financial firms. Past Patents is the sum of FinTech patent applications by incumbent firm i over $[t-4, t-1]$. FinTech Acquisitions is the sum of acquisitions, mergers, follow-on funding, MBO, or joint ventures by firm i of a FinTech innovator over the period $[t-4, t-1]$. Standard errors clustered at the country level are reported in parentheses

Dependent variable:	FinTech innovation			Number FinTech Patents		
	(1)	(2)	(3)	(4)	(5)	(6)
Competition Start-ups	0.002***			-0.003**		
	(0.001)			(0.001)		
Competition Start-ups (Citations)		0.011***			0.000	
		(0.002)			(0.008)	
Number Start-ups			0.002**			-0.001
			(0.001)			(0.002)
Competition Start-ups × Finance				0.032**		
				(0.013)		
Competition Start-ups (Citations) × Finance					0.065***	
					(0.003)	
Number Start-ups × Finance						0.027**
						(0.011)

Table 7 continued

Dependent variable:	FinTech innovation			Number FinTech Patents		
	(1)	(2)	(3)	(4)	(5)	(6)
Past Patents	0.000***	0.000***	0.000***	0.006***	0.006***	0.006***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
FinTech Acquisitions	-0.041***	-0.041***	-0.041***	-0.147***	-0.126***	-0.147***
	(0.013)	(0.013)	(0.013)	(0.016)	(0.012)	(0.016)
Total Assets	0.001	0.001	0.001	0.003	0.004*	0.003
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Operating Revenue	0.002	0.002	0.002	0.013***	0.013***	0.013***
	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)	(0.003)
Observations	18,613	18,613	18,613	18,613	18,613	18,613
R-squared	0.468	0.468	0.468	0.532	0.533	0.532
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

*/**/*** represent significance at 10, 5 and 1% level, respectively

suggests a 1.1% higher likelihood of financial incumbent innovation for a one standard deviation increase in the *Competition Start-ups (Citations)* variable.

Our main proxy for outside competitive pressures counts the number of patent applications. This approach, however, does not account for the fact that these applications might come from a small number of very successful innovators. Consequently, in column (3), we replace the main independent variable with an alternative one that considers the number of startup firms applying for a FinTech patent as opposed to the total number of applications in a period. Specifically, we recompute the ratio in Eq. (2) by considering the number of startup firms as opposed to the number of patent applications in the numerator. The results are very similar, suggesting that the applications counted in the *Competition Start-ups* measure come from different firms, making the two alternative proxies of competition almost identical in size.

In columns (4)-(6) of Table 7, we extend our baseline model by using the the total number of FinTech patent applications by a firm i in year t as the dependent variable. Since our sample is comprised of both financial and non-financial firms that are FinTech innovators, we modify the model in Eq. (1) by considering the effect of competition from start-ups on financial firms relative to non-financial firms. Specifically, the model tested is as follows:

$$
\text{FinTech patents}_{i,t} = \alpha_i + \gamma_t + \beta_1 \text{Competition Start-ups}_{c,t-4\to t-1} \\
\beta_2 \text{Competition Start-ups}_{c,t-4\to t-1} \times \text{Finance} + \theta' X_{i,t} + \epsilon_{i,t},
\tag{3}
$$

where FinTech patents$_{i,t}$ is now the total number of patents applications by firm i in year t, measured as $log(1+\text{Total Patents}_{i,t})$.[8] The coefficient of interest, β_2, is that of the interaction between the Competition Start-ups$_{c,t-4\to t-1}$ ratio and a dummy equal to one for financial firms (Finance). This interaction term captures the increase in the number of patents applications by financial incumbents relative to that of non-financial firms. We control separately for Competition Start-ups$_{c,t-4\to t-1}$, while the Finance dummy is absorbed by firm fixed effects.

The results in column (4) of Table 7 show that this interaction term is positive and statistically significant, suggesting that competition from start-ups results in a relatively higher increase in the number of FinTech patent applications by financial firms as compared to non-financial ones. We next repeat the same exercise as in columns (2)-(3) and replace the number of patents by start-ups as the main explanatory variable with (i) its value weighted by citations (column (5)) and (ii) the number of firms as opposed to the number of patent applications (column (6)). The results are robust across all specifications and indicate that competitive pressures from start-ups push financial incumbents to innovate relatively more than non-financial firms.

All estimations in Table 7 control for an array of firm characteristics and fixed effects. We include a measure of Past Patents as the sum of patent applications by incumbent i over $[t-4, t-1]$. As expected, firms that produce FinTech innovations in the past are more likely to continue innovating. We also control for whether a financial incumbent has acquired or merged with a FinTech innovator over the past four years. Again, as expected, firms that buy FinTech innovators are less likely to apply for new patents themselves. Firm size and operating revenue are positively correlated with the probability of applying for a FinTech innovation, although the effect is not statistically significant across all specifications. Finally, all estimations include firm fixed effects, which means that we obtain identification from

---

[8] In robustness tests, we show that our results hold when we consider models that account for the count nature of the patent data.

changes within a firm over time. We also include year fixed effects throughout to capture innovation waves that can occur in all countries in a given year.[9]

Overall, the results in Table 7, point to an *escape competition* effect, whereby financial incumbents innovate when the competitive pressure from non-financial start-ups relative to other financial incumbents is higher. In Appendix Fig. 7, we show that this result is even stronger when we consider a longer time horizon for measuring the relative number of total patent applications by start-ups to financial incumbents. Figure 7 presents the coefficient estimates for our measures of competitive pressures in (2) over different time frames: $[t-2, t-1]$, $[t-3, t-1]$ and $[t-5, t-1]$, respectively. We find the strongest effect for the last interval, suggesting a relatively large time lag between the patent applications by competitors and financial incumbents' FinTech innovation.

The results in Table 7 are also robust to alternative model specifications. Specifically, the model in Eq. (1) was estimated using a fixed effects linear probability model that includes both financial and non-financial firms applying for FinTech patents. As such, the average marginal effect of a covariate is a linear combination of (i) the estimated coefficient of the group of financial incumbents (which innovate at different points in time), and (ii) zero, which is the corresponding coefficient of the group of non-financial companies (Beck 2020). An alternative specification is to estimate a fixed effects logistic model, which would only employ the subset of data that has variation in the dependent variable, i.e., the group of financial incumbents that apply for FinTech patents in some (but not all) periods in the sample. We present the results employing a fixed effects logistic model in Table 8, columns (1)-(3) for all definitions of FinTech competition. The results are consistent with the baseline model and the estimates in column (1) suggest that a one standard deviation increase in Competition Start-ups increases the probability of innovating of incumbent financial firms by 2.8%.

Furthermore, the dependent variable in model (1) is the log-transformation of patent count. However, while widely used, this linear transformation can produce biased results (see Cohn et al. 2021; Wooldridge 2010, page 726). We thus re-estimate Eq. (1) using a negative binomial model, which is suitable for modeling count variables, especially for over-dispersed outcome variables.[10] The results using this alternative specification are presented in Table 8, columns (4)-(6), and yield similar results as our baseline model.

Overall, the results in Tables 7 and 8 point to a strong effect of entry by non-financial start-ups on the propensity to innovate of financial firms. We investigate next whether this effect corresponds to a particular type of FinTech technology.

4.1 Type of FinTech innovation

To investigate whether the effects of start-up competition on innovation are driven by a particular FinTech technology, we re-estimate the models in (1) and (3) for each of the five FinTech technologies classified by the machine learning algorithms in Section 2. More specifically, we look at the probability that financial incumbents apply for a specific type of

---

[9] In unreported regressions, we also control for country-decade fixed effects to capture innovation waves in the same country in a given decade. The results are qualitatively similar.

[10] In the working paper version of the paper we also present Poisson estimates with separable fixed effects, which are less efficient if the model exhibits overdispersion, however they remain unbiased and consistent as long as the standard conditional mean independence assumption holds (Cohn et al. 2021).

Table 8 FinTech patents and competition: alternative model specifications. Columns (1)-(3) present the coefficient estimates of a logit model where the dependent variable is an indicator variable equal to one if a financial firm applied for a FinTech patent in year t. Columns (4)-(6) show coefficient estimates of a negative binomial model where the dependent variable is the total number of FinTech patent applications of a firm in year t. Financial firms correspond to the 2-digit NACE Rev 2 codes: 64, 65 and 66. Competition Start-ups is the ratio of the sum of FinTech patent applications by non-financial start-ups relative to that of financial incumbents during the previous four years. Competition Start-ups (Citations) is the Competition Start-ups ratio scaled by the number of forward citations of non-financial start-ups. Number Start-ups is the ratio of the total number of start-ups applying for a FinTech patent over the total number of incumbents in the previous four years. Finance is a dummy variable for financial firms. Firm controls include: the sum of patent applications and the sum of acquisitions of FinTech innovators over the last four years, as well as the log of Total Assets and Operating Revenue. Standard errors clustered at the country level are reported in parentheses

	Logit model			Negative binomial model		
	(1)	(2)	(3)	(4)	(5)	(6)
Competition Start-ups	0.239***			-0.128***		
	(0.091)			(0.023)		
Competition Start-ups (Citations)		0.608***			0.200***	
		(0.057)			(0.035)	
Number Start-ups			0.194*			-0.116***
			(0.107)			(0.022)
Competition Start-ups × Finance				0.211***		
				(0.071)		
Competition Start-ups (citations) × Finance					0.182***	
					(0.058)	
Number Start-ups × Finance						0.167**
						(0.079)
Observations	1,733	1,909	1,733	10,067	10,067	10,067
Firm controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	No	No	No

*/**/*** represent significance at 10, 5 and 1% level, respectively

FinTech patents when competition from non-financial start-ups in the same type of financial service intensifies. The model estimated is thus:

$$\text{FinTech Innovation}_{i,f,t} = \alpha_i + \gamma_t + \beta \text{Competition Start-ups}_{c,f,t-4\to t-1} + \theta' X_{i,t} + \epsilon_{i,t}, \quad (4)$$

where FinTech Innovation$_{i,f,t}$ is an indicator variable equal to one if company i is an incumbent financial institution and has applied for at least one FinTech patent in category $f \in \{$Data Analytics, Fraud, Insurance, Investments and Payments$\}$ in year t. Competition Start-ups$_{c,f,t-4\to t-1}$ is now the sum of patent applications by non-financial start-ups in a FinTech category f divided by the sum of patent applications by other financial incumbents in the same category over the past 4 years. We estimate Eq. (4) for each category f separately.

Figure 6(a) presents the point estimates of β in Eq. (4). The results are not as precisely estimated given the lower number of observations within each category. We only find a statistically significant effect for the Investment technology, which corresponds to innovations related to portfolio management, lending and investing platforms. This is confirmed when considering the importance of patents applications measured by their forwards citations (lower part of figure (a)). Applications related to insurance are also driven by the competition pressure from start-ups in this case.

We also consider the effect of innovation across different types of FinTech patents on the number of patent applications by financial firms as compared to non-financial ones within each patent category. In figure (b) we present the coefficient β_2 in Eq. (3), which is re-estimated for each type of FinTech category j. Here we find stronger evidence across other categories of FinTech applications, such as Fraud, Insurance and Investments, with all the coefficients being statistically significant when we weight the interaction term by the citation count.

The bulk of this evidence confirms that competitive threats from FinTech start-ups push financial incumbents to innovate to escape competition. This evidence is stronger across some type of FinTech innovations such as Investment technologies, in particular when we consider the magnitude of the innovation proxied by patent citations. Furthermore, the effect

(a) FinTech Applications by patent type   (b) Number of FinTech Patents by patent type
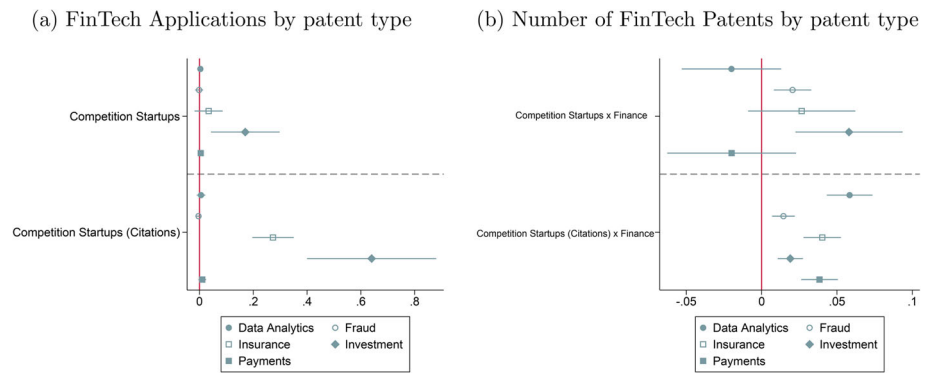
95% confidence intervals are shown.

Fig. 6 Competition and type of FinTech technologies. Figure **a** shows the point estimates of β in Eq. (4) for each category f separately, where $f \in \{$Data Analytics, Fraud, Insurance, Investments and Payments$\}$. Figure **b** shows the point estimates of β_2 in Eq. (3) for each category f separately

is stronger when we investigate the effect of competition on financial incumbents as compared to non-financial firms.

4.2 Alternative measures of competition

We check the robustness of our main results when employing alternative definitions of our measures of competitive pressures. First, we consider competition not just from non-financial start-ups headquartered in the same country as the financial incumbent, but also those that list as priority country the domicile of the financial incumbent. The priority country is the country where the patent was first filed worldwide before being extended to other countries. If FinTech start-ups register patents in the country where the financial incumbent is headquartered then they are considered to compete with the incumbent. As such, the ratio in Eq. (2) now includes the sum of patent applications that list as priority country the domicile of financial incumbent i. In many cases this corresponds to the applicant's domestic patent office, and as such, the number of patents counted would be the same as in our baseline measure. However, this extended definition can also capture competition from Start-ups domiciled abroad, but whose FinTech innovation was patented in the financial incumbent's domicile country.

The results using this alternative measure are presented in Table 9 Panel A and are qualitatively similar to our baseline results across all specifications. The magnitude of the effect is larger as compared to the estimates in Table 7, which is to be expected as the number of FinTech start-up innovators using this alternative definition is likely to be larger.

We next consider an alternative proxy for the importance of FinTech innovations. In our baseline specification we count the number of FinTech patent applications as a measure of competition. We then weight these patents by importance using a measure of forward citations count. An alternative measure of importance is to consider only the number of patent applications that are eventually granted. As such, in ratio (2) we only include the applications whose final status is "patented". The results are presented in Panel B of Table 9 and support our main hypothesis although are statistically significant only when we consider the measure weighted by citations count. This suggests that the importance of the applications, as measured by the citation count, matters more than whether the applications are merely granted.

Our results are also robust to alternative definitions of Start-ups. Our baseline model follows Chen et al. (2019) and uses an 8-year threshold to define a non-financial startup. In Appendix Fig. 9 we show that our baseline results in Table 7 are robust to start-ups being defined according to 6-year age cutoff or a 10-year age cutoff.

Finally, our main measure of competition by start-ups is scaled by other financial incumbents' patent applications in order to mitigate the role of unobservable technology shocks on both financial firms', as well as non-financial start-ups' innovations. However, to further mitigate any endogeneity concern, we also implement an instrumental variable model. Specifically, we instrument the innovative efforts by start-ups by the size of the venture capital (VC) investment to GDP in each country at the beginning of the sample. We obtain venture capital data for the sample of OECD countries, so our estimations are based on a subset of the data. Venture capital is an important source of financing for start-ups, in particular in high-tech sectors (Bijlsma and Zwart 2013; Janeway et al. 2021). As such, we expect this measure to be highly correlated with the future innovative effort of start-ups by reflecting the overall ease of obtaining financing in a country. The exclusion restriction condition for our IV strategy is that the VC investment at the beginning of the sample is not correlated with the future unobserved technology shocks in the reduced form model in Eq. (5) below (i.e., Cov

Table 9 Alternative measures of competition (based on Priority Country). The dependent variable in columns (1)-(3) is an indicator variable equal to one if a financial firm applied for a FinTech patent in year t, while in columns (4)-(6) it is the total number of FinTech patent applications a firm in year t. Financial firms correspond to the 2-digit NACE Rev 2 codes: 64, 65 and 66. Competition Start-ups is the ratio of the sum of FinTech patent applications by non-financial start-ups relative to that of financial incumbents during the previous four years. Competition Start-ups (Citations) is the Competition Start-ups ratio scaled by the number of forward citations of non-financial start-ups. Number Start-ups is the ratio of the total number of start-ups applying for a FinTech patent over the total number of incumbents in the previous four years. Finance is a dummy variable for financial firms. Firm controls include: the sum of patents acquisitions of FinTech innovators over the last 5 years, as well as the log of Total Assets and Operating Revenue. Standard errors clustered at the country level are reported in parentheses

Dependent variable:	FinTech patent application			Total number of patents		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Competition measures constructed based on priority country						
Competition Start-ups	0.035**			0.146***		
	(0.017)			(0.025)		
Competition Start-ups (Citations)		0.050***			0.150***	
		(0.004)			(0.006)	
Number Start-ups			0.037**			0.147***
			(0.016)			(0.026)
Competition Start-ups × Finance				0.185***		
				(0.038)		
Competition Start-ups (Citations) × Finance					0.146***	
					(0.006)	
Number Start-ups × Finance						0.175***
						(0.039)
Observations	18,613	18,613	18,613	18,613	18,613	18,613
R-squared	0.507	0.527	0.509	0.674	0.649	0.672

Table 9 continued

Dependent variable:	FinTech patent application			Total number of patents		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel B: Competition measures constructed based on patents granted						
Competition Start-ups	0.002			-0.003		
	(0.002)			(0.006)		
Competition Start-ups (Citations)		0.010***			-0.012	
		(0.002)			(0.009)	
Competition Start-ups × Finance				0.023		
				(0.043)		
Competition Start-ups (Citations) × Finance					0.038***	
					(0.003)	
Observations	18,507	18,507		15,027	15,027	
R-squared	0.469	0.469		0.516	0.516	
Firm controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

*/**/*** represent significance at 10, 5 and 1% level, respectively

Table 10 Instrumental variable model. The dependent variable in columns (1)-(2) is an indicator variable equal to one if a financial firm applied for a FinTech patent in year t, while in columns (3)-(4) it is the total number of FinTech patent applications a firm in year t. Financial firms correspond to the 2-digit NACE Rev 2 codes: 64, 65 and 66. Startup Patents is the sum of FinTech patent applications by non-financial start-ups during the previous four years. Finance is a dummy variable for financial firms. Firm controls include: the sum of patents acquisitions of FinTech innovators over the last 5 years, as well as the log of Total Assets and Operating Revenue. Standard errors clustered at the country level are reported in parentheses

Dependent variable	FinTech patent application		Total number of patents	
	(1) OLS	(2) IV	(3) OLS	(4) IV
Startup Patents$_{c,t-4\to t-1}$	0.032***	0.044***	0.105***	-0.044
	(0.003)	(0.007)	(0.005)	(0.037)
Startup Patents$_{c,t-4\to t-1}\times$Finance			0.085***	0.201***
			(0.011)	(0.038)
Finance			0.000	-0.074**
			(0.011)	(0.032)
Firm controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
First-stage instrument		2.518***		2.480***
		(0.226)		(0.568)
F-statistic for IV in first stage		123.81		92.80
Observations	18,651	11,978	18,651	11,978

*/**/*** represent significance at 10, 5 and 1% level, respectively

(VC/GDP$_i$, $\epsilon_{i,t}$)=0). This is a sensible assumption, as the important technological waves in FinTech such as machine learning, blockchain and cloud computing were arguably difficult to anticipate in the early 2000s.

$$\text{FinTech Innovation}_{i,t} = \alpha + \delta \text{Startup Patents}_{c,t-4\to t-1} + \theta' X_{i,t} + \gamma_t + \epsilon_{i,t} \quad (5)$$
$$\text{Startup Patents}_{c,t-4\to t-1} = \gamma + \beta \text{VC/GDP}_i + \varsigma_{i,t}. \quad (6)$$

Table 10 presents both the OLS estimations of Eq. (5) (column (1)), as well as the first and second stage IV estimations (column (2)). As seen in the bottom of column (2), the measure of VC/GDP is highly correlated with Startup Patents$_{c,t-4\to t-1}$ and this correlation is significant at the 1% level. Moreover, the high F-statistic suggests the instrument is strong enough for identification. Note also that both OLS and IV estimation in Table 10 do not include firm fixed effects, as our instrument is not time-varying. Furthermore, in columns (3) and (4) we repeat the IV strategy for the second model that looks at the total number of patents granted. Again we find that the effect of competition by start-ups leads to a higher number of patent applications by financial incumbents as compared to non-financial firms in both the OLS and IV estimations.

4.3 Competition from technology incumbents

The descriptive statistics in Table 4 showed that a large fraction of FinTech patent applications are owned by non-financial firms. Technology firms corresponding to the 2-digit NACE Rev 2 codes: 26, 47, 49, 58, 61, 62 and 63, represent a large fraction of that. 21% of the FinTech patent application come from technology firms that are not start-ups (older than

Table 11 FinTech patent application and competition from tech incumbents. The dependent variable in columns (1)-(3) is an indicator variable equal to one if a financial firm applied for a FinTech patent in year t, while in columns (4)-(6) it is the total number of FinTech patent applications a firm in year t. Financial firms correspond to the 2-digit NACE Rev 2 codes: 64, 65 and 66. Competition Tech is the ratio of the sum of FinTech patent applications by Technology firms relative to that of financial incumbents during the previous four years. Technology firms correspond to the 2-digit NACE Rev 2 codes: 26, 47, 58, 61, 62, are 63, respectively. Competition Tech (Citations) is the Competition Tech ratio scaled by the number of forward citations of Technology firms. Number Tech is the ratio of the total number of technology firms applying for a FinTech patent over the total number of incumbents in the previous four years. Finance is a dummy variable for financial firms. Past Patents is the sum of FinTech patent applications by firm i over $[t - 4, t - 1]$. FinTech Acquisitions is the sum of acquisitions, mergers, follow-on funding, MBO, or joint ventures by financial incubator i of a FinTech innovator over the period $[t - 4, t - 1]$. Standard errors clustered at the country level are reported in parentheses.

Dependent variable:	FinTech patent application			Total number of patents		
	(1)	(2)	(3)	(4)	(5)	(6)
Competition Tech	0.003			-0.001		
	(0.003)			(0.006)		
Competition Tech (Citations)		0.007*			0.018**	
		(0.004)			(0.006)	
Number Tech			0.001			-0.005
			(0.003)			(0.007)
Competition Tech × Finance				0.001		
				(0.019)		
Competition Tech (Citations) × Finance					0.086***	
					(0.002)	
Number Tech × Finance						0.004
						(0.017)
Past Patents	0.000***	0.000***	0.000***	0.006***	0.006***	0.006***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
FinTech Acquisitions	-0.041***	-0.043***	-0.042***	-0.150***	-0.154***	-0.151***
	(0.013)	(0.014)	(0.013)	(0.014)	(0.013)	(0.014)
Total Assets	0.000	0.000	0.000	0.003	0.003	0.003
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)

Table 11 continued

Dependent variable:	FinTech patent application			Total number of patents		
	(1)	(2)	(3)	(4)	(5)	(6)
Operating Revenue	0.002	0.002	0.002	0.013***	0.013***	0.013***
	(0.002)	(0.002)	(0.002)	(0.004)	(0.004)	(0.004)
Constant	0.008	0.009	0.008	-0.007	-0.006	-0.006
	(0.006)	(0.006)	(0.006)	(0.015)	(0.014)	(0.015)
Observations	18,507	18,507	18,507	18,507	18,507	18,507
R-squared	0.469	0.469	0.469	0.528	0.530	0.528
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

*/**/*** represent significance at 10, 5 and 1% level, respectively

eight years at the moment of application). In this subsection we check whether the effect of competition on financial incumbents' innovation is present when competition is coming from these incumbent technology firms and not start-ups.

As such, we re-construct the measures of competition, but considering incumbent technology companies in the numerator of the ratio in Eq. (2).

The results are presented in Table 11. Overall, we find little evidence that competition from technology incumbents spurs innovation in the financial sector. The results are overall less precise and not always positive. We only find a statistically significant effect when we consider the interaction of Competition Tech (Citations) × Finance in column (5). In the Appendix we estimate the alternative specifications using this definition of competition from technology incumbents. Specifically, Appendix Fig. 8 considers alternative time frames, Appendix Fig. 10 looks at the type of FinTech technology, while Appendix Table 13 measures competition based on priority country. Overall, these results show weak evidence that the competitive pressures from mature technology firms are important in driving innovation in the financial industry.

5 Conclusion

Technological innovations have long shaped the financial sector and have had large effects on long-term profitability (Fuentelsaz et al. 2009; Haynes and Thompson 2000; Scott et al. 2017). Yet, the rise of FinTech innovators over the past few decades has the potential to represent the largest disruption in the financial sector so far (Gomber et al. 2017).

In this study, we provide the first test of how financial incumbents responded to the competition from non-financial FinTech start-ups. To this end, we classify a large sample of patents into five FinTech categories using machine learning classifiers. We then exploit the fluctuations in the ratio of FinTech patent applications by non-financial start-ups versus incumbents to capture competitive pressures coming from outside the financial sector.

We show that financial incumbents are more likely to innovate when they face greater competitive pressure from non-financial start-ups. This competitive pressure is even greater if the patent applications turn out to be radical innovations, as captured by their forward citation count.

These results provide support for an escape competition effect, whereby in markets with a low level of competition, the threat of entry pushes incumbents to innovate. They also hold several policy implications. In particular, given the documented inefficiencies in the financial sector (Philippon 2016), our results suggest that an increase in competition from non-financial start-ups can result in important technological innovations in the financial sector. This is important because theoretical models such as Laeven et al. (2015) show that, as economies become more technologically advanced, the well-documented link between financial development and economic growth (Peia and Roszbach 2015; Valickova et al. 2015) disappears unless financiers themselves innovate and become more efficient. Our results suggest that policies that encourage entry from technologically advanced FinTech start-ups can lead to such efficiency gains in the financial sector. As such, policies that encourage the funding of FinTech start-ups, though, for example, venture capital, can lead to technological innovations in the financial sector as well.

Declarations

Conflicts of interests/Competing interests The authors certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

Appendix

A. Database construction

This section provides further details on the construction of our FinTech patent dataset. It also compares our dataset with that of Chen et al. (2019) and Xu et al. (2020), which are, to the best of our knowledge, the closest related studies that employ machine learning approaches to classifying FinTech innovations.

A summary of the dataset characteristics and filtering steps employed is presented in Table 12. Several key differences between our approach and that in previous work should be noted. First, while the datasets in Chen et al. (2019) and Xu et al. (2020) employ the universe of patents submitted to the USPTO, our initial dataset covers also patents submitted to the EPO, as well as an extended sample of countries covered by the PATSTAT Global database.

Similar to Chen et al. (2019) and Xu et al. (2020), we filter patents belonging to International Patent Classification classes G and H, which yields a significantly larger sample of potential FinTech patents.

Next, both Chen et al. (2019) and Xu et al. (2020) used a similar lists of financial terms consisting of 487 and 478 terms, respectively, to filter patents potentially related to financial services. We complement the list of financial terms from Chen et al. (2019) with additional FinTech identification terms which takes the total number of initial filtering keywords to 516.

Chen et al. (2019) identified seven FinTech categories based on insights from a general reading of FinTech reports and articles. Xu et al. (2020) selected their seven FinTech categories based on a Financial Stability Board (FSB) report from 2017. We employ a more

Table 12 Comparison with FinTech datasets in Chen et al. (2019) and Xu et al. (2020)

	Dataset Characteristics	Chen et al. (2019)	Xu et al. (2020)	Our dataset
(1)	Source of patents	BDSS	Lens	**Orbis/PATSTAT**
(2)	Years covered	2003-2016	2014-2018	**2000-2016**
(3)	Legal jurisdiction of patents	US	US	**US + Europe**
(4)	IPC classes used	G and H	G and H	**G and H**
(5)	Initial number of patents based on criteria (1)-(4) above	1,181,162	1,328,623	**6.8M**
(6)	Financial terms for filtering financial patents	487	478	**516**
(7)	Number of patents after filtering with financial terms (6)	67,948	37,156	**38,228**
(8)	Number of FinTech categories considered	7	7	**5**
(9)	Number of manually annotated patents used for training	1,800	1,800	**3850**
(10)	Total number of FinTech patents identified out of (7)	6,511	3,602	**19,055**

Table compares the dataset characteristics and filtering steps employed for the contraction of our dataset with those in Chen et al. (2019) and Xu et al. (2020)

stringent taxonomy of 5 broad FinTech-related categories to avoid overlapping our patents to what are traditionally software innovations applicable across industries.

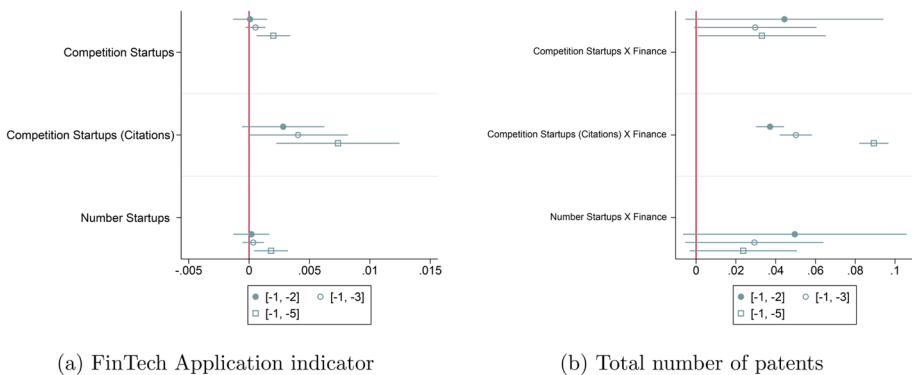Finally, we manually labeled a much larger subset of patents as compared to previous work.

(a) FinTech Application indicator   (b) Total number of patents

Fig. 7 Alternative time frames for measuring startup competition

(a) FinTech Application indicator   (b) Total number of patents
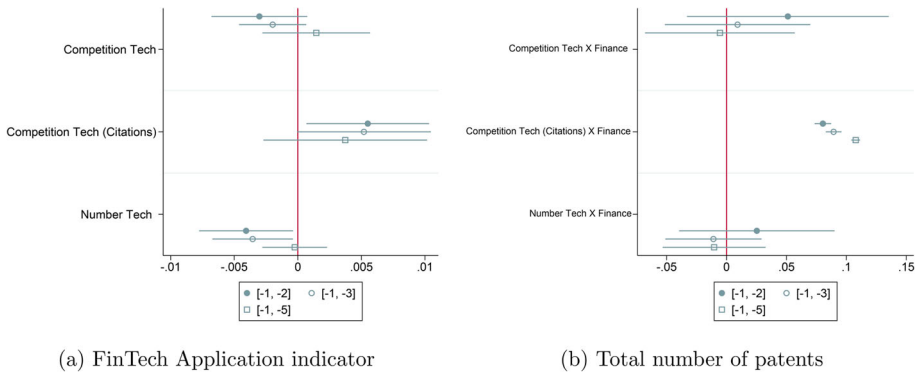
Fig. 8 Alternative time frames for measuring Tech competition

B. Other robustness tests

Table 13 Alternative measures of competition (based on Priority Country).The dependent variable in columns (1)-(3) is an indicator variable equal to one if a financial firm applied for a FinTech patent in year t, while in columns (3)-(4) it is the total number of FinTech patent applications a firm in year t. Financial firms correspond to the 2-digit NACE Rev 2 codes: 64, 65 and 66. Competition Start-ups is the ratio of the sum of FinTech patent applications by non-financial start-ups relative to that of financial incumbents during the previous four years. Competition Start-ups (Citations) is the Competition Start-ups ratio scaled by the number of forward citations of non-financial start-ups. Number Start-ups is the ratio of the total number of start-ups applying for a FinTech patent over the total number of incumbents in the previous four years. Finance is a dummy variable for financial firms. Past Patents is the sum of FinTech patent applications by firm i over $[t-4, t-1]$. FinTech Acquisitions is the sum sum of acquisitions, mergers, follow-on funding, MBO, or joint ventures by financial incubator i of a FinTech innovator over the period $[t-4, t-1]$. . Standard errors clustered at the country level are reported in parentheses

Dependent variable:	FinTech patent application			Total number of patents		
	(1)	(2)	(3)	(4)	(5)	(6)
Competition Tech	0.016			0.107***		
	(0.011)			(0.028)		
Competition Tech (Citations)		0.018			0.077**	
		(0.010)			(0.033)	
Number Tech			0.015			0.104***
			(0.010)			(0.026)
Competition Tech × Finance				0.140		
				(0.096)		
Competition Tech (Citations) × Finance					0.121	
					(0.080)	

Table 13 continued

Dependent variable:	FinTech patent application			Total number of patents		
	(1)	(2)	(3)	(4)	(5)	(6)
Number Tech × Finance						0.150
						(0.101)
Observations	16,997	16,997	16,997	16,997	16,997	16,997
R-squared	0.408	0.410	0.406	0.570	0.531	0.565
Firm controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

*/**/*** represent significance at 10, 5 and 1% level, respectively
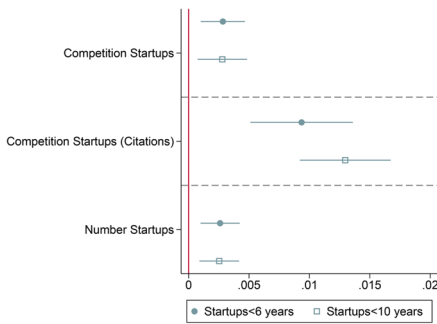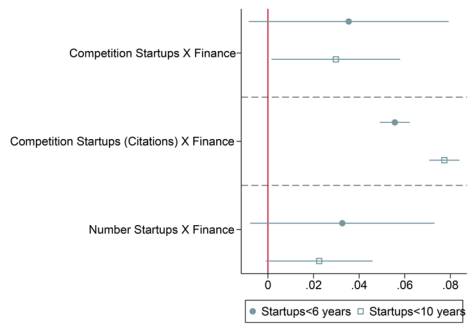
(a) FinTech Application indicator   (b) Total number of patents

Fig. 9 Alternative age definitions of start-ups
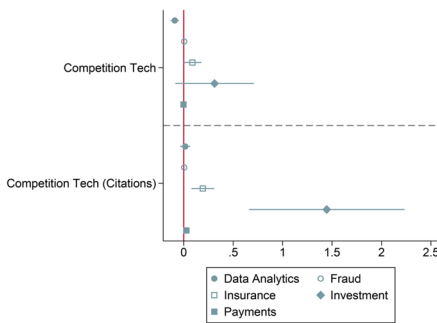
(a) FinTech Applications by technology   (b) Number of Patents by technology
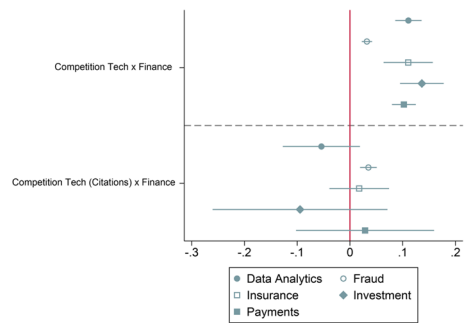
Fig. 10 Competition Tech and type of FinTech technologies

References

Aaron M, Rivadeneyra F, Sohal S (2017) Fintech: Is this time different? a framework for assessing risks and opportunities for central banks. Tech. rep, Bank of Canada Staff Discussion Paper

Abdelgawad L, Kluegl P, Genc E, Falkner S, Hutter F (2019) Optimizing neural networks for patent classification. In: Proc. of ECML-PKDD, vol 16

Aghion P, Griffith R (2008) Competition and growth: reconciling theory and evidence. MIT press

Aghion P, Howitt P (1992) A model of growth through creative destruction. Econometrica 60(2):323–351

Aghion P, Harris C, Howitt P, Vickers J (2001) Competition, imitation and growth with step-by-step innovation. Rev Econ Stud 68(3):467–492

Aghion P, Bloom N, Blundell R, Griffith R, Howitt P (2005) Competition and innovation: An inverted-u relationship. Q J Econ 120(2):701–728

Aghion P, Blundell R, Griffith R, Howitt P, Prantl S (2009) The effects of entry on incumbent innovation and productivity. Rev Econ Stat 91(1):20–32

Aghion P, Akcigit U, Howitt P (2014) What do we learn from schumpeterian growth theory? In: Handbook of economic growth, vol 2, Elsevier, pp 515–563

Akcigit U, Kerr WR (2018) Growth through heterogeneous innovations. J Polit Econ 126(4):1374–1443

Allen F, Gu X, Jagtiani J (2020) A survey of fintech research and policy discussion

Alt R, Beck R, Smits MT (2018) Fintech and the transformation of the financial industry

Angrist JD, Pischke JS (2009) Mostly harmless econometrics: An empiricist's companion. Princeton university press

Arrow KJ (1962) Economic Welfare and the Allocation of Resources for Invention

Ayyagari M, Demirgüç-Kunt A, Maksimovic V (2011) Firm innovation in emerging markets: The role of finance, governance, and competition. Journal of Financial and Quantitative Analysis 46(6):1545–1580

Beck N (2020) Estimating grouped data models with a binary-dependent variable and fixed effects via a logit versus a linear probability model: The impact of dropped units. Polit Anal 28(1):139–145

Benzineb K, Guyot J (2011) Automated patent classification. In: Current challenges in patent information retrieval, Springer, pp 239–261

Bijlsma MJ, Zwart GT (2013) The changing landscape of financial markets in Europe, the United States and Japan. Tech. rep., Bruegel working paper

Bloom N, Draca M, Van Reenen J (2016) Trade induced technical change? the impact of chinese imports on innovation, it and productivity. Rev Econ Stud 83(1):87–117

Bojanowski P, Grave E, Joulin A, Mikolov T (2017) Enriching word vectors with subword information. Trans of the ACL 5:135–146

Caragea D, Chen M, Cojoianu T, Dobri M, Glandt K, Mihaila G (2020) Identifying fintech innovations using bert. In: 2020 IEEE International Conference on Big Data (Big Data), pp 1117–1126

CBInsights (2017) Wealth tech market map

Chen MA, Wu Q, Yang B (2019) How valuable is fintech innovation? The Review of Financial Studies 32(5):2062–2106

Cockburn IM, MacGarvie MJ (2011) Entry and patenting in the software industry. Manage Sci 57(5):915–933

Cohn JB, Liu Z, Wardlaw M (2021) Count data in finance. Available at SSRN

Cojoianu TF, Clark GL, Hoepner AG, Pazitka V, Wojcik D (2020) Fin vs. tech: Determinants of fintech start-up emergence and innovation in the financial services incumbent sector. Small Business Economics (forthcoming)

Czarnitzki D, Etro F, Kraft K (2014) Endogenous market structures and innovation by leaders: an empirical test. Economica 81(321):117–139

Devlin J, Chang MW, Lee K, Toutanova K (2018) Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805

Dorn D, Hanson GH, Pisano G, Shu P et al (2020) Foreign competition and domestic innovation: Evidence from us patents. American Economic Review: Insights 2(3):357–74

Eckenrode J, Friedman S (2017) Fintech by the numbers. Deloitte Services LP

Etro F (2004) Innovation by leaders. Econ J 114(495):281–303

Fall CJ, Törcsvári A, Benzineb K, Karetka G (2003) Automated categorization in the international patent classification. Acm Sigir Forum, ACM New York, NY, USA 37:10–25

Fuentelsaz L, Gómez J, Palomas S (2009) The effects of new technologies on productivity: An intrafirm diffusion-based assessment. Res Policy 38(7):1172–1180

Gilbert RJ, Newbery DM (1982) Preemptive patenting and the persistence of monopoly. Am Econ Rev 514–526

Gomber P, Koch JA, Siering M (2017) Digital finance and fintech: current research and future research directions. J Bus Econ 87(5):537–580

Gomez JC, Moens MF (2014) A survey of automated hierarchical classification of patents. In: Prof. search in the modern world, Springer, pp 215–249

Grawe MF, Martins CA, Bonfante AG (2017) Automated patent classification using word embedding. In: 16th IEEE Int. Conference on Machine Learning and Applications (ICMLA), IEEE, pp 408–411

Haddad C, Hornuf L (2019) The emergence of the global fintech market: Economic and technological determinants. Small Bus Econ 53(1):81–105

Haynes M, Thompson S (2000) The productivity impact of it deployment: An empirical evaluation of atm introduction. Oxford Bull Econ Stat 62(5):607–619

Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comp 9(8)

Hornuf L, Klus MF, Lohwasser TS, Schwienbacher A (2021) How do banks interact with fintech startups? Small Bus Econ 57(3):1505–1526

Horrace WC, Oaxaca RL (2006) Results on the bias and inconsistency of ordinary least squares for the linear probability model. Econ Lett 90(3):321–327

Hu J, Li S, Hu J, Yang G (2018) A hierarchical feature extraction model for multi-label mechanical patent classification. Sustainability 10(1):219

Janeway WH, Nanda R, Rhodes-Kropf M (2021) Venture capital booms and start-up financing. Annu Rev Financ Econ 13:111–127

Kim Y (2014) Convolutional neural networks for sentence classification. arXiv preprint arXiv:1408.5882

Kowalewski O, Pisany P (2020) The rise of fintech: a cross-country perspective

Laeven L, Levine R, Michalopoulos S (2015) Financial innovation and endogenous growth. Journal of Financial Intermediation 24(1):1–24

Lan Z, Chen M, Goodman S, Gimpel K, Sharma P, Soricut R (2020) Albert: A lite bert for self-supervised learning of language representations. In: ICLR

LeCun Y, Bengio Y, Hinton G (2015) Deep learning. nature 521(7553):436

Lee JS, Hsiang J (2019) Patentbert: Patent classification with fine-tuning a pre-trained bert model. arXiv preprint arXiv:1906.02124

Levy R (2015) Fintech market map. Retrieved 11(08):2016

Li S, Hu J, Cui Y, Hu J (2018) Deeppatent: patent classification with convolutional neural networks and word embedding. Scientometrics 117(2):721–744

Li Y, Spigt R, Swinkels L (2017) The impact of fintech start-ups on incumbent retail banks' share prices. Financial Innovation 3(1):1–16

Liu X, Hodgkinson IR, Chuang FM (2014) Foreign competition, domestic knowledge base and innovation activities: Evidence from chinese high-tech industries. Res Policy 43(2):414–422

Liu Y, Ott M, Goyal N, Du J, Joshi M, Chen D, Levy O, Lewis M, Zettlemoyer L, Stoyanov V (2019) Roberta: A robustly optimized bert pretraining approach. arXiv preprint arXiv:1907.11692

Mann K, Püttmann L (2021) Benign effects of automation: New evidence from patent texts. Rev Econ Stat (forthcoming)

Mellon B (2015) Innovation in payments: The future is fintech. The Bank of Newyork

Mikolov T, Chen K, Corrado G, Dean J (2013) Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781

Peia O, Roszbach K (2015) Finance and growth: time series evidence on causality. J Financ Stab 19:105–118

Pennington J, Socher R, Manning C (2014) Glove: Global vectors for word representation. In: EMNLP Vol. 14, pp. 1532–1543

Philippon T (2015) Has the us finance industry become less efficient? on the theory and measurement of financial intermediation. American Economic Review 105(4):1408–38

Philippon T (2016) The fintech opportunity. Tech. rep, National Bureau of Economic Research

Risch J, Krestel R (2019) Domain-specific word embeddings for patent classification. Data Technologies and Applications

Romer PM (1990) Endogenous technological change. J Polit Econ 98(5, Part 2):S71–S102

Scott SV, Van Reenen J, Zachariadis M (2017) The long-term effect of digital innovation on bank performance: An empirical study of swift adoption in financial services. Res Policy 46(5):984–1004

Shalaby M, Stutzki J, Schubert M, Günnemann S (2018) An lstm approach to patent classification based on fixed hierarchy vectors. In: Proceedings of the 2018 SIAM Int. Conference on Data Mining, SIAM, pp 495–503

Stulz RM (2019) Fintech, bigtech, and the future of banks. J Appl Corp Financ 31(4):86–97

Sun C, Qiu X, Xu Y, Huang X (2019) How to fine-tune bert for text classification? Nat. Springer, Conf. on Chinese Comp. Linguistics, pp 194–206

Tang J (2006) Competition and innovation behaviour. Research policy 35(1):68–82

Temelkov Z (2018) Fintech firms opportunity or threat for banks? International journal of information, Business and Management 10(1):137–143

Thakor AV (2020) Fintech and banking: What do we know? Journal of Financial Intermediation 41:100833

Tseng PL, Guo WC (2022) Fintech, credit market competition, and bank asset quality. Journal of Financial Services Research 61(3):285–318

Valickova P, Havranek T, Horvath R (2015) Financial development and economic growth: A meta-analysis. Journal of economic surveys 29(3):506–526

Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I (2017) Attention is all you need. In: Advances in neural information processing systems, pp 5998–6008

Vives X (2019) Digital disruption in banking. Annu Rev Financ Econ 11:243–272

Wooldridge JM (2010) Econometric analysis of cross section and panel data. MIT press

Xu L, Lu X, Yang G, Shi B (2020) Identifying fintech innovations with patent data: A combination of textual analysis and machine-learning techniques. In: Int. Conference on Information, Springer, pp 835–843

Yang Z, Dai Z, Yang Y, Carbonell J, Salakhutdinov RR, Le QV (2019) Xlnet: Generalized autoregressive pretraining for language understanding. In: Advances in NIPS, pp 5754–5764

Young E, Treasury GB (2016) UK Fintech on the Cutting Edge: An Evaluation of the International Fintech Sector. Ernst & Young