UCD CENTRE FOR ECONOMIC RESEARCH

WORKING PAPER SERIES

2022

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

Doina Caragea and Mihai Dobri, Kansas State University
Theodor Cojoianu, University of Edinburgh Business School
Andreas Hoepner and Oana Peia, University College Dublin

WP22/01

January 2022

UCD SCHOOL OF ECONOMICS
UNIVERSITY COLLEGE DUBLIN
BELFIELD DUBLIN 4

Competition and innovation in the financial sector: Evidence from the rise of FinTech start-ups*

Doina Caragea
Kansas State University

Theodor Cojoianu
University of Edinburgh Business School

Mihai Dobri
Kansas State University

Andreas Hoepner
University College Dublin

Oana Peia†
University College Dublin

December 2021

Abstract

We provide new evidence on the effects of competition on incumbents' innovative behavior by examining the rise of FinTech start-ups over the period 2000-2016. We employ machine learning techniques to classify a large global sample of patent applications into five FinTech categories. We exploit the variation in the share of FinTech patent applications by non-financial startups to incumbent financial firms to measure competitive pressures from outside the financial industry. We show that higher competitive pressures from non-financial start-ups increases the probability that financial incumbents innovate. Moreover, competition from start-ups results in a higher number of FinTech patent applications by financial incumbents as compared to non-financial firms, especially when the innovations of FinTech start-ups are more important, as proxied by future patent citations count.

Keywords: FinTech, patents, machine learning, financial incumbents, innovation.
JEL Classification: G20, G21, O31.

*We would like to thank Mark Chen, Luis Cabral and Davide Romelli for providing valuable feedback. Colin Woodroofe provided excellent research assistance. Peia thanks UCD Seed Funding for financial support.

†Contacts: Doina Caragea, Department of Computer Science, Kansas State University, email: dcaragea@ksu.edu; Theodor Cojoianu, University of Edinburgh Business School, email: theodor.cojoianu@ed.ac.uk; Mihai Dobri, Department of Computer Science, Kansas State University, email: mihaidobri@ksu.edu; Andreas Hoepner, University College Dublin, College of Business, email: andreas.hoepner@ucd.ie; Oana Peia, University College Dublin, School of Economics, email: oana.peia@ucd.ie.

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. From 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 happening outside the financial sector, as young start-ups and technology companies attempt to disrupt incumbent financial institutions.

In this paper, we study whether this increased competition from FinTech start-ups affects the innovative behavior of incumbent financial firms. The link between competition and innovation is complex. On the one hand, an increase in competition can discourage innovation by reducing the rents for innovators and therefore lowering incentives to innovate (Romer 1990, Aghion & Howitt 1992). On the other hand, when competitive pressure is low and pre-innovation rents are high, firms may have less incentives to innovate (Arrow 1962, Gilbert & 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 where competition is already high, 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" suggests that, in industries where competition is low, incumbents will try to innovate in order to escape from a situation in which competition constrains profits.

Given that the lack of entry of competition have been an endemic problem in the financial sector, the recent FinTech revolution offer a unique experiment to test these theories (Philippon 2016). As such, our main testable hypothesis is that the threat of competition from non-financial innovators leads to an increase in innovation by incumbent financial institutions. To test this hypothesis, we focus on patent filling as a way of measuring the threat of entry by firms outside the finance industry.

One empirical challenge in studying the impact of FinTech innovation on the finance industry 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 documents.

[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 construct a novel dataset of FinTech patent applications based on patents published in the PATSTAT Global dataset over the period 2000-2016. We apply machine learning algorithms to the textual data of patent abstracts to classify patents into five categories reflecting key technological characteristics of FinTech innovations: (1) software and techniques for data analytics, (2) fraud detection and cyber security infrastructure, (3) insurance software analytics, (4) investments, lending and portfolio management tools and (5) mobile transfers and digital wallets.

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

Employing this sample of FinTech innovators, we examine whether the rise of non-financial FinTech start-up 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 exploiting the variation in the ratio of patent applications by startups to incumbents financial firms over the past three to five years. We interpret a change in this ratio as a variation in outside competitive pressures that is unexplained by technology shocks (which would increase innovation by both incumbents and outsiders to relatively the same extent keeping the ratio stable over time).

We find that competitive pressures by non-financial startups, measured as a relative larger number of patent applications as compared to incumbent firms, increases the probability that financial incumbents innovate. These results are stronger when we weight the number of patent applications by their relative importance, measured by the forward citations count. Specifically, we find that a one standard deviation increase in our measure of competitive pressure increases the probability that a financial incumbent applies for a FinTech patent by 0.6%. The results are robust to controlling for time-varying firms characteristics such as past patenting activity, size and profitability. We also control for financial incumbents' propensity to acquire FinTech innovators, as opposed to innovating themselves. Not surprisingly, we find that financial firms who have acquired FinTech innovators in the past are less likely to apply for FinTech patents.

We also show that the total number of FinTech patent applications is higher among financial firms relative to non-financial firms when competitive pressures by start-ups is higher. This suggests that the rise of FinTech startups has disproportionally driven financial incumbents to innovate. Moreover, we show that competitive pressures from incumbent technology firms do not have the same effect and are not followed by an increase in FinTech innovation by financial incumbents. This might be due to the fact that the majority of FinTech patent applications by mature technology firms are in the category of payments applications, while financial incumbents mainly innovate in areas such as investments and lending or insurance analytics.

Our estimations are robust to a wide array of model specification 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.

Our work is related to several branches of literature. First, we contribute to a growing literature that employs taxonomies to automatically classify patents across different categories of innovations (see, among, others Fall et al. 2003, Benzineb & Guyot 2011, Gomez & Moens 2014, Grawe et al. 2017, Li et al. 2018, Shalaby et al. 2018, Hu et al. 2018, Abdelgawad et al. 2019, Lee & Hsiang 2019, Mann & Püttmann 2021). Chen et al. (2019) were the first to propose a typology to classify a sample of 2 million patents filled with the U.S. Patent and Trademark Office (USPTO) into seven categories of FinTech innovations: cybersecurity, mobile transactions, data analytics, blockchain, peer-to-peer (P2P), robo-advising, and Internet of Things. They train and employ several families of machine learning algorithms such as support vector machines (SVM) and neural networks to identify a sample of FinTech innovations by U.S.-based firms or individuals. Chen et al. (2019) then investigate the value of these FinTech innovations by studying the 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 (RF) classifiers to identify FinTech patents in a sample of patent applications to the USPTO over 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 (Bidirectional Encoder Representations from Transformers) models (Devlin et al. 2018). Specifically, while previous work studying patent text has employed 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 employing state-of-the art deep learning methodology.

Second, we employ our newly build classification of FinTech patents to investigate how competition drives innovating behavior in the financial industry. The relationship between innovation and competition is of long-standing theoretical and empirical interest (Aghion et al. 2001, 2005, Aghion & 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 reduce the probability of surviving entry (*Schumpeterian effect*). In industries with low levels of competition, where leaders compete *neck-to-neck*, a 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 work has documented these two effects across different countries, time periods or 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 activity (Tang 2006). A key empirical challenge in this literature is that entry is endogenous to the innovative activity of incumbents. The FinTech revolution offers an interesting setting to testing 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 FinTech start-up emergence across 21 OECD countries and find that FinTech venture creation is positively related to the regional productivity and new knowledge created in the IT sector, but not that in the financial sector.

Finally, our work relates to a growing literature that looks at how the dramatic growth of FinTech start-ups has shaped the financial landscape (see Allen et al. 2020, and references within). For example, Hornuf et al. (2021) study how banks interact with FinTech start-ups using detailed information on strategic alliance 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.

The remainder of the paper is organized as follows. Section 2 presents the patent data and describes the machine learning techniques employed to classify FinTech patents, while section 3 discusses the identification strategy and methodology. 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 employed to classify FinTech patents. We also provide some descriptive statistics of the resulting FinTech patents dataset.

2.1 Patent data

Patent data is obtained from the BvD Orbis intellectual property database, which sources information from PATSTAT Global[2]. We retrieve all patents applications filled between 2000 and 2016 belonging to the International Patent Classification (IPC) Classes G and H, which cover areas related to digital computing that underline the FinTech technologies classified in this paper. This results in a sample to 6.8 million patents. To narrow down the search, we employ 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 38,299 patents that are related to financial services.[3]

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 analyse this sector. Hence, we build a FinTech taxonomy by corroborating taxonomies from several academic works, industry reports and market maps (such as Mellon 2015, Levy 2015, Amalia 2016, Young &

[2]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.

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

Figure 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.

Treasury n.d., CBInsights 2017, Eckenrode & Friedman 2017, Chen et al. 2019, Haddad & 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 filling abstract in each category are shown in Table 1. Our FinTech taxonomy is aligned with that of Chen et al. (2019) although we include a broader range of financial services, such as insurance. We also exclude some applications which are not necessarily specific to the financial sector such as *Blockchain* or the *Internet-of-Things*, as our main goal is to study the innovative behavior of financial incumbents.

2.3 Machine learning algorithms to classify FinTech patents

To be able to train machine learning and deep learning models for FinTech patent identification, we manually annotated/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 Data Analytics category). 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 classifying FinTech patents using BERT-type models, which stands for Bidirec-

Table 1: FinTech Taxonomy

Fintech category	Applications	Examples of patent filing abstracts
Data Analytics	Software, data infrastructure and analytics for financial services	"computer programs product 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."

tional Encoder Representations from Transformers (Devlin et al. 2018). 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). We focus on BERT models as our main text classifier, given the success of the deep learning approaches on general patent classification tasks (Li et al. 2018, Shalaby et al. 2018, Sun et al. 2019, Lee & Hsiang 2019).

BERT is a transformer-based machine learning technique for natural language processing pre-trained on a large corpus of un-annotated text (e.g., Wikipedia) using two self-supervised learning tasks: masked word prediction and next sentence prediction. A generic BERT model

Table 2: Performance results of traditional and deep learning models

Category	Metric	L-SVM	G-SVM	NB	kNN	RF	GB	MLP	ML-V	CNN	RNN	BERT
		Traditional machine learning models								Deep learning models		
Insurance	Pr	0.779	0.857	**0.878**	0.689	0.824	0.864	0.848	0.854	0.904	**0.922***	0.891
	Re	0.844	0.812	0.750	0.875	0.635	0.792	**0.927**	0.792	0.979	**0.990***	**0.990***
	F1	0.810	0.834	0.809	0.771	0.718	0.826	**0.886**	0.822	0.940	**0.955***	0.938
Payments	Pr	0.827	**0.852**	0.645	0.636	0.802	0.800	0.835	0.851	0.884	0.884	**0.906***
	Re	0.910	0.920	0.800	0.840	0.650	0.840	**0.960**	0.860	**0.990***	**0.990***	0.970
	F1	0.867	0.885	0.714	0.724	0.718	0.820	**0.893**	0.856	0.934	0.934	**0.937***
Investment	Pr	0.837	0.850	0.765	0.725	0.778	0.796	**0.870**	0.861	**0.943***	0.934	0.924
	Re	0.870	**0.960**	0.910	0.870	0.770	0.860	0.940	0.930	**0.990***	**0.990**	0.980
	F1	0.853	0.901	0.831	0.791	0.774	0.827	**0.904**	0.894	**0.966***	0.961	0.918
Fraud	Pr	0.788	0.790	0.676	0.617	0.813	0.815	0.849	**0.884**	0.951	**0.961***	0.960
	Re	0.780	**0.830**	0.750	0.740	0.610	0.750	0.790	0.760	**0.980***	**0.980***	**0.980***
	F1	0.784	0.810	0.711	0.673	0.697	0.781	**0.819**	0.817	0.966	**0.970***	**0.970***
Data Analytics	Pr	0.530	0.581	0.351	0.420	**0.700***	0.565	0.581	0.513	0.635	0.635	**0.652**
	Re	0.700	0.720	0.660	0.420	0.280	0.700	0.720	**0.800**	0.800	0.800	**0.900***
	F1	0.603	0.643	0.458	0.420	0.400	0.625	**0.643**	0.625	0.708	0.708	**0.753***
Non-FinTech	Pr	0.939	0.955	0.968	**0.980**	0.700	0.880	0.962	0.923	0.992	0.996	**1.000***
	Re	0.823	0.850	0.697	0.660	**0.927***	0.853	0.843	0.873	0.857	**0.863**	0.850
	F1	0.877	**0.899**	0.810	0.789	0.798	0.866	**0.899**	0.897	0.919	**0.925***	0.918
Average	Pr	0.783	0.814	0.714	0.678	0.770	0.787	**0.824**	0.814	0.885	**0.889***	**0.889***
	Re	0.821	0.849	0.761	0.734	0.645	0.799	**0.863**	0.836	0.933	0.935	**0.945***
	F1	0.799	0.829	0.722	0.695	0.684	0.791	**0.840**	0.819	0.905	0.909	**0.912***
	Acc	0.830	0.858	0.751	0.735	0.745	0.820	**0.867**	0.849	0.921	**0.924***	0.923

Table shows the precision (Pr), recall (Re), F1-score (F1) and overall accuracy (Acc) for the machine learning models employed. The traditional machine learning models are: Linear Support Vector Machines (LSVM), Gaussian Support Vector Machines (G-SVM), Naïve Bayes (NB), k-nearest neighbour (kNN), random forest (RF), gradient boosting (GB), Multilayer Perceptron (MLP). We also show the results of a voting classifier (ML-V), which consists of L-SVM, G-SVM and MLP, similar to the voting classifier in Chen et al. (2019), 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. * marks the best result in a row.

can then be further pre-trained and/or fine-tuned for specific natural language processing tasks. Particularly, the BERT model for the FinTech patent classification task is initialized with the parameters of a generic pre-trained BERT model, further pre-trained using the manually classified FinTech patents, and subsequently fine-tuned for patent classification. We also compare the results obtained from BERT models with text classifiers such as CNN (Convolutional Neural Networks) and RNN (Recurrent Neural Networks), which take 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, as well as Naïve Bayes (NB), k-nearest neighbour (kNN), random forest, and gradient boosting. Appendix 6.2 details the machine learning methodology employed.

To evaluate the performance of the various models that we train, we use several standard metrics, including the overall accuracy (Acc), precision (Pr), recall (Re) and F1-score (F1). We also report the precision, recall, and F1 scores for each of our five FinTech categories to determine what categories might be easier or harder to identify.

Table 3: Filtering steps in the construction of the patent applications dataset

(1) Source of patents	PATSTAT
(2) Years covered	2000-2016
(3) Legal jurisdiction of patents	US + Europe
(4) IPC classes used	G&H
(5) Initial number of patents based on criteria (1)-(4) above	6.8M
(6) Financial terms for filtering financial patents	516
(7) Number of patent applications after filtering with financial terms (6)	38,228
(8) Number of FinTech categories considered	5
(9) Number of manually annotated patents used for training	3,850
Total number of FinTech patents identified out of (7)	19,055

Performance results on the test data are presented in Table 2. The first eight columns contain traditional machine learning models, while the last three correspond to 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 employ the BERT model to classify our sample of patents. Table 3 summarizes the number of patent fillings removed at each step of our methodology. The deep learning algorithms employed identify 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 our machine learning algorithm described in the previous section, 3,166 (20%) were filled by individuals. This leaves a sample of 15,889 patent applications by companies that will be the focus of our analysis. Among these, private firms are the most important group of FinTech innovators accounting for 68% of the applications. Non-financial firms are also responsible for the largest share of innovations in financial technologies, while startups account for 13% of the applications.

Furthermore, while the 15,889 patents are filled at the European and US Patent offices by companies across 58 countries, the large majority of fillings (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 has been filled), as opposed to the headquarters of the inventor company, 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 4: FinTech innovation activity by innovator type

	Number of FinTech patent fillings
Innovator type	
Individual	3,166
Public company	5,028
Private company	10,861
Financial services firms	3,236
Non-financial firms	12,653
Startups	2,012
Geographical location of company	
US	12,322
Europe	1,636
Rest of the world	1,931

Table shows the frequencies of FinTech patent fillings 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 8 years prior to the patent filling date.

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.

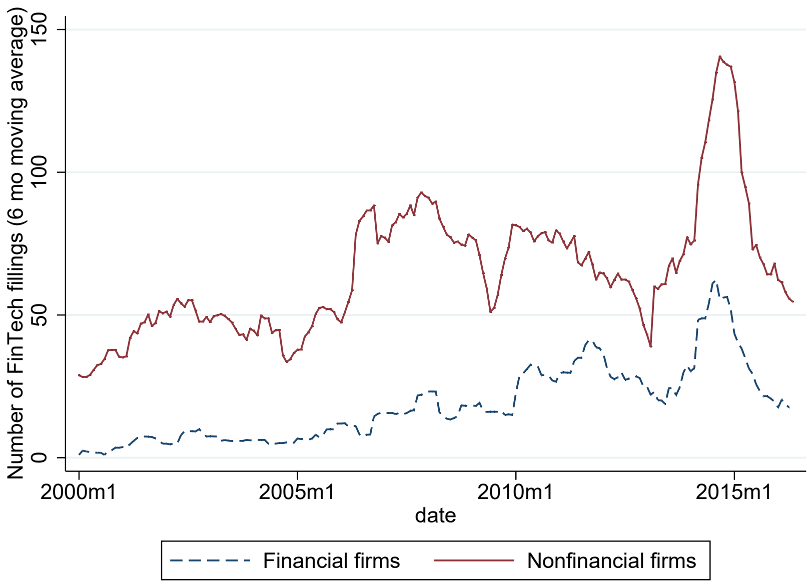
Table 5: FinTech innovation activity by technology category and innovator type

		Innovator type				
FinTech Classification	Number of applications	Private firms (%)	Public firms (%)	Financial firms (%)	Non financial firms (%)	Startups (%)
Data analytics	4790	0.64	0.36	0.18	0.82	0.11
Fraud	1075	0.65	0.35	0.09	0.91	0.13
Insurance	1343	0.81	0.19	0.35	0.65	0.11
Investment	3657	0.76	0.24	0.30	0.70	0.15
Payments	5024	0.64	0.36	0.14	0.86	0.13

The table shows the frequencies of FinTech patent fillings 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 8 years prior to the patent filling date.

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

Figure 2: FinTech innovation over time by financial versus non-financial firms



The figure shows the 6 months moving average of patent applications from June 2000 to April 2016 by financial firms (corresponding to NACE 2 digits codes 64, 65 and 66) versus non-financial firms.

Figure 3 plots the evolution by company age and sector. It shows that while non-financial start-ups and financial incumbents have similar patent filling rates prior to 2007, incumbents dominate filling in the second half of the sample. Our empirical strategy will exploits this time variation in patent filling across these two sectors.

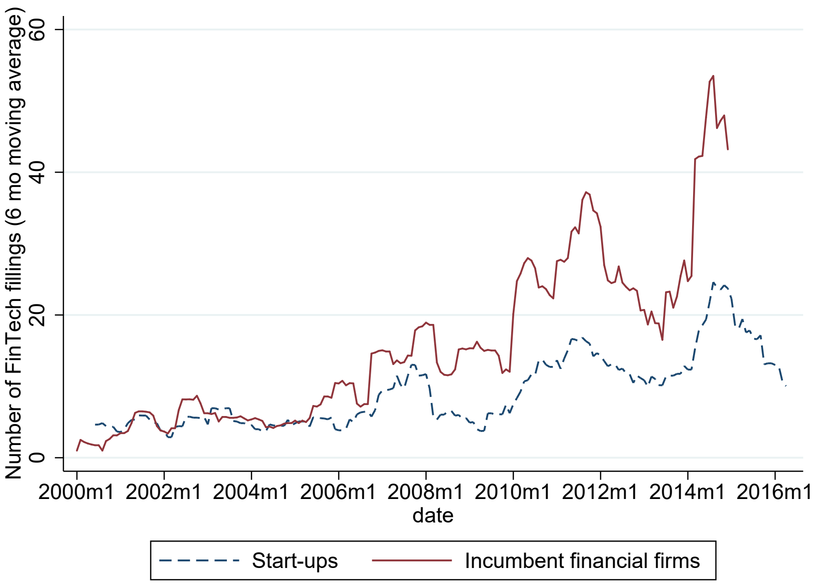
Finally, Figure 4 shows the evolution of patent applications by type of FinTech technology. It is apparent that the increase in FinTech innovations occurred across all types of technologies.

3 Identification strategy

Our empirical approach aims to identify how competition from FinTech startups affects innovation efforts by incumbent financial firms. This strategy face two main identification challenges. The first is the fact that entry can be endogenous to innovation. In order words, innovations in financial technologies by entrants are driven by the innovations and potential productivity gains in the finance industry. 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

Figure 3: FinTech innovation over time by start-ups versus incumbent financial firms

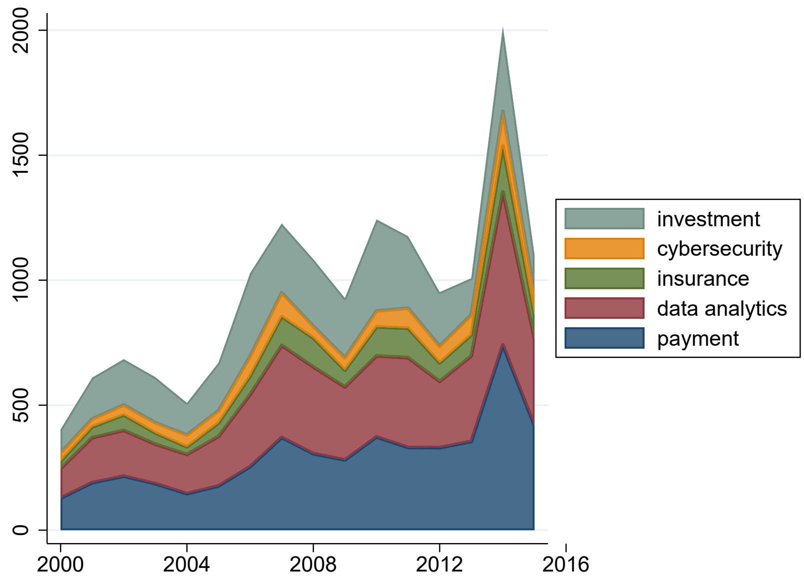


The figure shows the 6 months moving average of patent applications from June 2000 to April 2016 by incumbent financial firms (corresponding to NACE 2 digits codes 64, 65 and 66) versus start-ups outside the finance industry. Start-ups are defined as companies less that 8 years old at the time of patent application.

cost of financial intermediation in the U.S. 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. It is thus plausible to argue that the most recent FinTech revolution is not driven by innovations already happening in the financial sector.

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 employ an instrument or exogenous shock to identify entry by non-financial firms in the finance industry. However, since such instruments are not easily available, 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 incumbent and outsiders, then the rate of innovations 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 these 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

Figure 4: FinTech innovation over time by region



The figure shows the total number of applications by year and type of FinTech innovation.

than 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 the competitive pressure 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 citations count (Akcigit & Kerr 2018). This is done by scaling our measure of competitive pressure by the forward citations count of the start-up patents at the end of the sample. Competitive pressures should be larger when the innovation by outsiders is more important, as proxied by forward citations count.

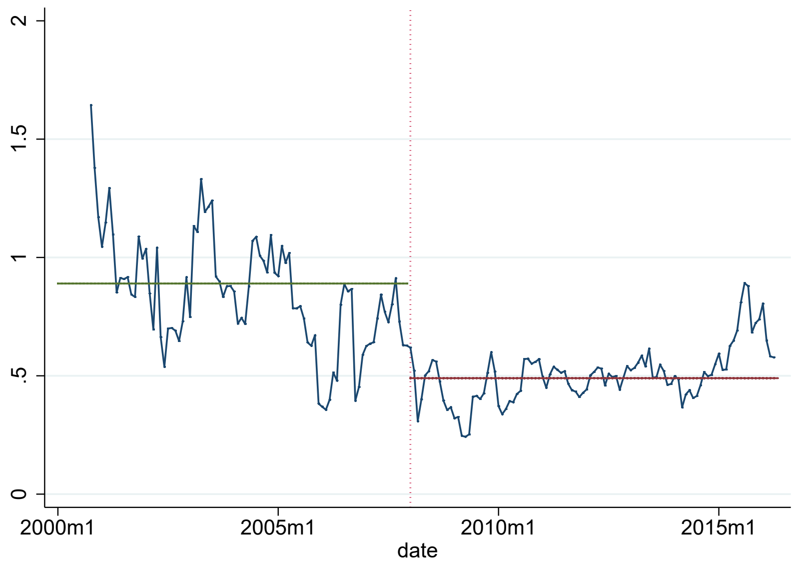
Our baseline strategy is as follows:

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

where FinTech Innovation$_{i,t}$ is an indicator variable equal to 1 if company i is an incumbent financial institution and has applied for at least one FinTech patent in year t. We define incumbent financial firms as those operating in the 2-digit NACE Rev 2 codes: 64, 65 and 66 and are not start-ups, i.e., were founded more than 8 years prior to the patent application date.

Our main independent variable is the ratio of FinTech applications by startups relative to

Figure 5: Evolution of patent applications (6 months moving average)



financial incumbents, computed as:

$$\text{Competition Startups}_{\,c,t-4\to t-1} = \frac{\sum_{\tau=t-4}^{t-1} \text{Startup Patents}_c}{\sum_{\tau=t-4}^{t-1} \text{Financial incumbents}_{-i,c}}. \tag{2}$$

In the baseline model, we consider the sum of patents in the previous 4 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 include as a separate explanatory variable.[4]

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 past innovation and financial performance of the firm. We control for the lagged sum of patent applications by firm i over $[t-5, t-1]$ to account for the propensity of a firm to produce FinTech innovations. We also include measures of firms size, such as log of Total Assets and Operating revenue, as well as profitability, captured by net profits. 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 incubator i of 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.

[4]One concern with the specification in (2) is that the amount of innovations of incumbent i is positively correlated with the sum of innovations of other financial incumbents. This, however, puts a downward bias on our estimations lessening the effect of competition by startups.

Table 6: Descriptive statistics

	Full Sample		Financial incumbents		No financial startups	
	Mean	St dev	Mean	St dev	Mean	St dev
ln(Total Assets)	7.767	4.631	11.282	5.198	4.855	4.134
Net profit/Total assets	0.053	0.183	0.083	0.173	-0.013	0.242
ln(Operating Revenue)	7.141	3.885	9.17	4.353	4.707	3.858
Past Patents	1.245	8.9	2.871	16.107	0.612	2.301
FinTech Acquisitions	0.001	0.028	0.006	0.075	0	0

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) which 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 filling from a financial incumbent, which yields a final dataset of 15,879 patent applications from 3,446 unique firms over the period 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 of one the members of the deal is a FinTech innovator as identified by our machine learning algorithm. We obtained this data from Zephyr for the period 2000-2016.

A summary of the financial information by types of FinTech innovators is presented in Table 6. Financial incumbents tend to be larger and more profitable than the average firm in the overall sample. They also have, on average, more patents and acquire more FinTech innovators.

4 Results

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

Our main explanatory variable is competition by startups (Column 1) measured by the ratio of the sum of patent applications by startups to financial incumbents over the period $[t-4, t-1]$. We find evidence that following periods with relatively higher FinTech innovations by non-financial startups, financial incumbents are more likely to innovate themselves. To ease interpretation of the point estimates, we have standardised the main independent variable. As such, the point estimate in Columns (1) suggests that a one standard deviation increase in the Competition Startups ratio increases financial incumbents' probability of innovating by

Table 7: FinTech patent applications and competition from startups

Dependent variable:	FinTech patent application			Total number of patents		
	(1)	(2)	(3)	(4)	(5)	(6)
Competition Startups	0.001**			-0.004*		
	(0.001)			(0.002)		
Competition Startups (Citations)		0.006**			-0.007	
		(0.002)			(0.008)	
Number Startups			0.001**			-0.002
			(0.001)			(0.002)
Competition Startups × Finance				0.030**		
				(0.014)		
Competition Startups (Citations) × Finance					0.071***	
					(0.004)	
Number Startups × Finance						0.025*
						(0.013)
Past Patents	0.000***	0.000***	0.000***	0.006***	0.007***	0.006***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
FinTech Acquisitions	-0.069**	-0.069**	-0.069**	-0.170***	-0.146***	-0.170***
	(0.026)	(0.026)	(0.026)	(0.013)	(0.016)	(0.014)
Total Assets	-0.002**	-0.002**	-0.002**	-0.000	0.001	-0.000
	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)	(0.003)
Net Profit	-0.000	-0.000	-0.000	-0.000*	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Operating Revenue	0.001*	0.002*	0.001*	0.013***	0.013***	0.013***
	(0.001)	(0.001)	(0.001)	(0.004)	(0.004)	(0.004)
Observations	16,997	16,997	16,997	16,997	16,997	16,997
R-squared	0.392	0.392	0.392	0.473	0.475	0.473
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

The dependent variable in columns (1)-(3) is an indicator variable equal to 1 if a financial firm applied for a FinTech patent in year t, while is 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 Startups is the ratio of the sum of FinTech patent applications by non-financial start-ups to that of financial incumbents over the period $[t-4, t-1]$. Competition Startups (Citations) is the Competition Startups ratio scaled by the number of forward citations of non-financial startups. Number startups is the ratio of the total number of startups applying for a FinTech patent over the total number of incumbents during the previous 4 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 a firm i of a FinTech innovator over the period $[t-4, t-1]$. Standard errors are clustered at the country level. */**/*** represent significance at 10, 5 and 1% level, respectively.

0.1%.

We consider next the importance of FinTech innovations by startups 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) suggests a 0.6% higher likelihood of financial incumbent innovation for a one standard deviation increase in the *Competition Startups (Citations)* variable.

Our proxy for outside competitive pressures counts the number of patent applications. This approach, however, does not take into account the fact that these applications might come from a small number of very successful innovators. Consequently, in column (3), we count the number of FinTech startup firms as opposed to their total number of applications. The results are very similar, suggesting that the applications counted in the measure "Competition Startups" 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 considering the total number of FinTech patent applications by a firm in year t. 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 startups on financial firms relative to non-financial firms. Specifically, the model tested is as follows:

$$\begin{aligned}\text{FinTech Innovation}_{i,t} \quad &= \quad \alpha_i + \gamma_t + \beta_1 \text{Competition Startups}_{c,t-4\to t-1} \\ &\beta_2 \text{Competition Startups}_{c,t-4\to t-1} \times \text{Finance} + \theta' X_{i,t} + \epsilon_{i,t}, \quad (3)\end{aligned}$$

where FinTech Innovation$_{i,t}$ is now the total number of patents applications by firm i in year t, measured as $log(1+\text{Total Patents}_{i,t})$.[5] The coefficient of interest, β_2, is that of the interaction between the Competition Startups$_{c,t-4\to t-1}$ 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 Startups$_{c,t-4\to t-1}$, 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 startups results in a relatively higher increase in the number of FinTech patent applications by financial firms as compared to non-financial

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

ones. We repeat next the same exercise as in columns (2)-(3) and replace the number of patents by startups 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 suggest that competitive pressures from startups 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 firm i over $[t-4, t-1]$. As expected, firms who produce FinTech innovations in the past are more likely to continue innovating. We also control for whether a financial firm has acquired, merged, or bought a FinTech innovator over the past 4 years. Again, as expected, firms that purchase FinTech innovators are less likely to apply for new patent themselves. Firm size is also negatively related to the probability of innovating, in line with a large literature that suggests that smaller firms are responsible for more radical innovations (see Akcigit & Kerr 2018, and references within). At the same time, a higher operating revenue is positively correlated with higher FinTech innovation, while the effect of net profit is negligible. Finally, all estimations include firm fixed effects, which means that we obtain identification from 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.[6]

Overall, the results in Table 7, point to evidence of an *escape competition* effect, whereby financial incumbents innovate when the competitive pressure from non-financial startups relative to other financial incumbents is higher. In Appendix Figure 8, 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 8 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) (ii) zero, which

[6]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.

Table 8: FinTech patents and competition: alternative model specifications

	Logistic model			Poisson model		
	(1)	(2)	(3)	(4)	(5)	(6)
Competition Startups	0.209**			-0.049*		
	(0.098)			(0.028)		
Competition Startups (Citations)		0.801***			-0.034	
		(0.014)			(0.087)	
Number Startups			0.178**			-0.030*
			(0.089)			(0.016)
Competition Startups × Finance				0.292***		
				(0.087)		
Competition Startups (Citations) × Finance					0.031**	
					(0.013)	
Number Startups × Finance						0.267***
						(0.089)
Observations	2,240	2,240	2,240	9,428	9,428	9,428
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Columns (1)-(3) estimate a logit model where the dependent variable is an indicator variable whether a financial firm applied for a FinTech patent in year t. Columns (4)-(6) show Poisson estimates where the dependent variable is the total number of FinTech patent applications a firm in year t. Financial firms correspond to the 2-digit NACE Rev 2 codes: 64-66. Competition Startups is the ratio of the sum of FinTech patent applications by non-financial start-ups to that of financial incumbents during the previous 4 years. Competition Startups (Citations) is the Competition Startups ratio scaled by the number of forward citations of non-financial startups. Number startups is the ratio of the total number of startups applying for a FinTech patent over the total number of incumbent the previous 4 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, Net profit and Operating revenue. Standard errors are clustered at the country level. */**/*** represent significance at 10, 5 and 1% level, respectively.

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 at different points in time.

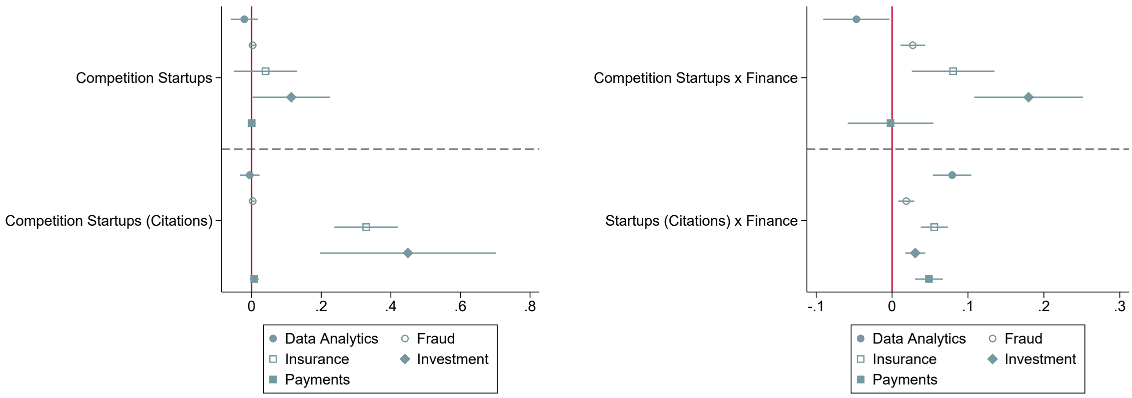
Furthermore, the dependent variable in model (3) is the log-transformation of patent count. However, while widely used, this linear transformation can produce biased results (see Cohn et al. 2021, Wooldridge 2010, page 726). We thus re-estimate Eq.(3) using a Poisson model, which admits separable fixed effects, as in our baseline linear model.[7]

The results using these alternative models are presented in Table 8. In all specifications, we control for time-varying firm characteristics, as well as time and firm fixed effects. We obtain consistent estimations both employing the logistic model, as well as the Poisson 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.

[7]Poisson estimates lose efficiency 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).

Figure 6: Competition and type of FinTech technologies



(a) FinTech Applications by technology

(b) Number of Patents by 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. As such, we also recompute the measures of competition to count the number of patent applications for a particular technology.

The results are presented in Figure 6, where figure (a) corresponds to the point estimates of β_1 in Eq. (1) , while figure (b) to those of coefficient β_2 in Eq. (3). When considering the effect of the Competition Startups ratio, 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 startups in this case. Figure 6(b) shows the impact of competition on the total number of applications of financial incumbents as compared to other FinTech innovators. 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 startups push financial incumbents to innovate to escape competition.

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

Table 9: Alternative measures of competition (based on Priority Country)

Dependent variable:	FinTech patent application			Total number of patents		
	(1)	(2)	(3)	(4)	(5)	(6)
Competition Startups	0.027*			0.143***		
	(0.013)			(0.028)		
Competition Startups (Citations)		0.038***			0.145***	
		(0.002)			(0.006)	
Competition Number Startups			0.028**			0.145***
			(0.013)			(0.029)
Competition Startups × Finance				0.193***		
				(0.028)		
Competition Startups (Citations) × Finance					0.140***	
					(0.006)	
Number Startups × Finance						0.188***
						(0.027)
Observations	16,997	16,997	16,997	16,997	16,997	16,997
R-squared	0.428	0.452	0.431	0.631	0.606	0.630
Firm controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

The dependent variable in columns (1)-(3) is a indicator variable whether a financial firm applied for a FinTech patent in year t, while is 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: 64-66. Competition Startups is the ratio of the sum of FinTech patent applications by non-financial start-ups to that of financial incumbents during the previous 4 years. Competition Startups (Citations) is the Competition Startups ratio scaled by the number of forward citations of non-financial startups. Number startups is the ratio of the total number of startups applying for a FinTech patent over the total number of incumbent the previous 4 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, Net profit and Operating revenue. Standard errors are clustered at the country level. */**/*** represent significance at 10, 5 and 1% level, respectively.

startups 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 is 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 a 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 startups domiciled abroad, but whose FinTech innovation is patented in the financial incumbent's domicile country.

The results using this alternative measure are presented in Table 9 and are robustly estimated 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 consider next 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

Table 10: Alternative measures of competition (based on patents granted)

	(1)	(2)	(3)	(4)
Competition Startups	-0.000		-0.000	
	(0.002)		(0.006)	
Competition Startups (Citations)		0.006**		-0.006
		(0.002)		(0.009)
Competition Startups × Finance			0.022	
			(0.042)	
Competition Startups (Citations) × Finance				0.070***
				(0.004)
Observations	16,997	16,997	16,997	16,997
R-squared	0.392	0.392	0.473	0.475
Firm controls	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Yes				
Year FE	Yes	Yes	Yes	Yes
Yes				

The dependent variable in columns (1)-(3) is a indicator variable whether a financial firm applied for a FinTech patent in year t, while is 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: 64-66. Competition Startups is the ratio of the sum of FinTech patent applications by non-financial start-ups to that of financial incumbents during the previous 4 years. Competition Startups (Citations) is the Competition Startups ratio scaled by the number of forward citations of non-financial startups. Number startups is the ratio of the total number of startups applying for a FinTech patent over the total number of incumbent the previous 4 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, Net profit and Operating revenue. Standard errors are clustered at the country level. */**/*** represent significance at 10, 5 and 1% level, respectively.

competitive pressures. 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 Table 10 and support our main hypothesis although are robustly estimated only when we consider the measure weighted by citations count. This suggests that the importance of the applications, measured by the citation count, matters more than whether the applications are merely granted.

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 8 years at the moment of application). In this section 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 Competition Tech (Citations) $\times$ Finance in column (5). In the Appendix we estimate the alternative specifications using this definition of competition by technology incumbents. Specifically, Figure 9 considers alternative time frames, Figure 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 & Thompson 2000, Scott et al. 2017). Yet, the rise of FinTech innovators over the past decades has the potential to represent the largest disruption in the financial sector so far (Gomber et al. 2017).

In this paper, we provide the first test of how financial incumbents responded to the competition from non-financial startups. To this end, we classify a large sample of patents into 5 FinTech categories using novel machine learning techniques. 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 to an escape competition effect, whereby in markets with a low level of competition, the threat of entry pushes incumbents to innovate.

Table 11: FinTech patent application and competition from tech incumbents

Dependent variable:	FinTech patent application			Total number of patents		
	(1)	(2)	(3)	(4)	(5)	(6)
Competition Tech	0.000			-0.009*		
	(0.001)			(0.005)		
Competition Tech (Citations)		0.004			0.012	
		(0.003)			(0.010)	
Competition Number Tech			-0.001			-0.011**
			(0.001)			(0.005)
Competition Tech × Finance				-0.000		
				(0.022)		
Competition Tech (Citations) × Finance					0.098***	
					(0.002)	
Competition Number Tech × Finance						-0.010
						(0.017)
Past Patents	0.000***	0.000***	0.000***	0.006***	0.006***	0.006***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
FinTech Acquisitions	-0.069**	-0.069**	-0.069**	-0.177***	-0.153***	-0.180***
	(0.026)	(0.026)	(0.026)	(0.018)	(0.016)	(0.016)
Total Assets	-0.002**	-0.002**	-0.002**	-0.000	0.001	-0.000
	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)	(0.003)
Net Profit	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Operating Revenue	0.001*	0.001*	0.001*	0.013***	0.013***	0.013***
	(0.001)	(0.001)	(0.001)	(0.004)	(0.004)	(0.004)
Observations	16,997	16,997	16,997	16,997	16,997	16,997
R-squared	0.392	0.392	0.392	0.473	0.477	0.473
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

The dependent variable in columns (1)-(3) is a indicator variable whether a financial firm applied for a FinTech patent in year t, while is 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: 64-66. Competition Tech is the ratio of the sum of FinTech patent applications by Technology firms to that of financial incumbents during the previous 4 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 Startups 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 incumbent the previous 4 years. Finance is a dummy variable for financial firms. Past Patents is the sum of FinTech patent applications by firm i over $[t-5, t-1]$. 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 are clustered at the country level. */**/*** represent significance at 10, 5 and 1% level, respectively.

References

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., Akcigit, U. & Howitt, P. (2014), What do we learn from schumpeterian growth theory?, *in* 'Handbook of economic growth', Vol. 2, Elsevier, pp. 515–563.

Aghion, P., Bloom, N., Blundell, R., Griffith, R. & Howitt, P. (2005), 'Competition and innovation: An inverted-u relationship', *The quarterly journal of economics* **120**(2), 701–728.

Aghion, P., Blundell, R., Griffith, R., Howitt, P. & Prantl, S. (2009), 'The effects of entry on incumbent innovation and productivity', *The Review of Economics and Statistics* **91**(1), 20–32.

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

Aghion, P., Harris, C., Howitt, P. & Vickers, J. (2001), 'Competition, imitation and growth with step-by-step innovation', *The Review of Economic Studies* **68**(3), 467–492.

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

Akcigit, U. & Kerr, W. R. (2018), 'Growth through heterogeneous innovations', *Journal of Political Economy* **126**(4), 1374–1443.

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

Amalia, F. (2016), 'The fintech book: The financial technology handbook for investors, entrepreneurs and visionaries', *Journal of Indonesian Economy and Business* **31**(3), 345–348.

Arrow, K. J. (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', *Political Analysis* **28**(1), 139–145.

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

Bloom, N., Draca, M. & Van Reenen, J. (2016), 'Trade induced technical change? the impact of chinese imports on innovation, it and productivity', *The review of economic studies* **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.

CBInsights (2017), 'Wealth tech market map'.

Chen, M. A., Wu, Q. & Yang, B. (2019), 'How valuable is fintech innovation?', *The Review of Financial Studies* **32**(5), 2062–2106.

Cohn, J. B., Liu, Z. & Wardlaw, M. (2021), 'Count data in finance', *Available at SSRN* .

Cojoianu, T. F., Clark, G. L., Hoepner, A. G., 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)* .

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

Dorn, D., Hanson, G. H., 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.* .

Fall, C. J., Törcsvári, A., Benzineb, K. & Karetka, G. (2003), Automated categorization in the international patent classification, *in* 'Acm Sigir Forum', Vol. 37, ACM New York, NY, USA, pp. 10–25.

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

Gilbert, R. J. & Newbery, D. M. (1982), 'Preemptive patenting and the persistence of monopoly', *The American Economic Review* pp. 514–526.

Gomber, P., Koch, J.-A. & Siering, M. (2017), 'Digital finance and fintech: current research and future research directions', *Journal of Business Economics* **87**(5), 537–580.

Gomez, J. C. & Moens, M.-F. (2014), A survey of automated hierarchical classification of patents, *in* 'Prof. search in the modern world', Springer, pp. 215–249.

Grawe, M. F., Martins, C. A. & Bonfante, A. G. (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 Business Economics* **53**(1), 81–105.

Haynes, M. & Thompson, S. (2000), 'The productivity impact of it deployment: An empirical evaluation of atm introduction', *Oxford Bulletin of Economics and Statistics* **62**(5), 607–619.

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

Hornuf, L., Klus, M. F., Lohwasser, T. S. & Schwienbacher, A. (2021), 'How do banks interact with fintech startups?', *Small Business Economics* **57**(3), 1505–1526.

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

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

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, J.-S. & 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.

Liu, X., Hodgkinson, I. R. & Chuang, F.-M. (2014), 'Foreign competition, domestic knowledge base and innovation activities: Evidence from chinese high-tech industries', *Research 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', *Review of Economics and Statistics (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* .

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, Technical report, National Bureau of Economic Research.

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

Romer, P. M. (1990), 'Endogenous technological change', *Journal of political Economy* **98**(5, Part 2), S71–S102.

Scott, S. V., 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', *Research 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.

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

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

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

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., 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', *Annual Review of Financial Economics* **11**, 243–272.

Wooldridge, J. M. (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, R. R. & Le, Q. V. (2019), Xlnet: Generalized autoregressive pretraining for language understanding, *in* 'Advances in NIPS', pp. 5754–5764.

Young, E. . & Treasury, G. B. (n.d.), *UK Fintech on the Cutting Edge: An Evaluation of the International Fintech Sector*, Ernst & Young.

6 Appendix

6.1 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.

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

	Dataset Characteristics	**Chen** *et al.* Chen et al. (2019)	**Xu et al. (2020)**	**Our dataset**
(1)	Source of patents	BDSS	Lens	**Orbis/Patsat**
(2)	Years covered	2003-2016	2014-2018	**2000-2016**
(3)	Legal jurisdiction of patents	US	US	**US + Europe**
(4)	IPC classes used	G&H	G&H	**G&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	**2350+1500**
(10)	Total number of FinTech patents identified out of (7)	6,511	3,602	**19,055**

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 in our study.

Chen et al. (2019) identified the 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 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.

6.2 Machine-learning classifiers

We detail in this section the main classification algorithms employed.

Prior work on FinTech patents has employed traditional machine learning approaches such as ensemble classifiers Chen et al. (2019) and random forests Xu et al. (2020), and has found that ensemble-type approaches show promising results. However, in the light of the growing success that deep learning approaches have seen in recent years, several works have used such approaches to automatically classify general patents according to standard categories in the International Patent Classification (IPC) or the Cooperative Patent Classification (CPC) taxonomies (Li et al. 2018, Hu et al. 2018, Shalaby et al. 2018, Abdelgawad et al. 2019, Lee & Hsiang 2019).[8]

The main classification method we employ in this paper is BERT, which stands for Bidirectional Encoder Representations from Transformers (Devlin et al. 2018). BERT is a language model that uses a deep bidirectional transformer encoder architecture Vaswani et al. (2017) to encode sentences and their tokens into dense vector representations. A generic model is pre-trained on a large corpus of un-annotated text (e.g., Wikipedia) using two self-supervised learning tasks: masked word prediction (a.k.a., masked language modeling, or MLM) and next sentence prediction (NSP). BERT takes as input a sequence of word tokens, where the first token is a special token denoted by [CLS] (the output representation of the [CLS] token can be seen as a semantic representation for the whole input sequence). A BERT input sequence consists of one or two sentences. For an input sequence consisting of two sentences, the two sentences are separated by another special token, denoted by [SEP]. Embeddings of the input tokens are provided to a multi-layer bidirectional transformer encoder, which transforms the original input embeddings into contextual output embeddings.

A generic BERT model, pre-trained on a large corpus, can be further *pre-trained* and/or *fine-tuned* for specific NLP tasks (Devlin et al. 2018). Particularly, a BERT model for the FinTech patent classification task can be initialized with the parameters of a generic pre-trained BERT

[8]For example, Grawe et al. (2017), Shalaby et al. (2018) used recurrent neural networks, specifically, long short-term neural networks (LSTM) Hochreiter & Schmidhuber (1997), together with word2vec embeddings Mikolov et al. (2013), to classify patents into IPC categories, while Risch & Krestel (2019) used gated recurrent unit (GRU) networks, together with fastText embeddings Bojanowski et al. (2017) for the same task. Similarly, Abdelgawad et al. (2019), Li et al. (2018) used word embeddings, including Word2vec Mikolov et al. (2013) and GloVe Pennington et al. (2014), together with convolutional neural networks (CNN) for text classification Kim (2014). Hu et al. (2018) build a hierarchical feature model that combined CNN and bidirectional LSTM (bi-LSTM) networks to capture both local lexical-level features and global sequential dependencies. The authors showed that the combined model achieved better performance than the independent CNN and LSTM/Bi-LSTM models on mechanical patent documents. Finally, Lee & Hsiang (2019) obtained state-of-the-art results with BERT models Devlin et al. (2018) on the task of classifying patent documents according to the IPC or CPC taxonomies. Specifically, Lee & Hsiang (2019) used large datasets of patent documents to fine-tune a pre-trained BERT-base model on the general patent classification task.

model, further pre-trained using FinTech patent data, and subsequently fine-tuned for patent classification.

6.2.1 BERT Variants

The success of the initial BERT-based models has resulted in an unparalleled suite of variants that can be used with the pre-training/fine-tuning framework proposed in Devlin et al. (2018). We used three variants in our analysis, specifically, RoBERTa (Liu et al. 2019), ALBERT (Lan et al. 2020) and XLNet (Yang et al. 2019). RoBERTa (Robustly Optimized BERT Approach) uses a larger dataset and an improved procedure to pre-train the BERT architecture. Among others, the next sentence prediction is removed and the masking applied to the training data is changed dynamically. ALBERT (A Lite BERT) is focused on decreasing BERT's size (i.e., the number of parameters that need to be learned), while not hurting its performance. It achieves a reduction in the number of parameters by factorizing the embedding parameterization and sharing parameters across all layers. To improve the training, it replaces the next sentence prediction task with a sentence order prediction task that better captures the inter-sentence cohesion. XLNet is a large bidirectional transformer, whose authors argue against the masked language modeling task and introduce an autoregressive permutation language modeling task for training (specifically, prediction of the next token in a sequence using some random order of the sequence). This improvement in the training procedure enables XLNet to capture better bidirectional dependencies among tokens in a sequence. In total, we experiment with 22 pre-trained BERT models and variants, by fine-tuning the models using labeled data for our specific FinTech classification task. We train each model for 6 epochs, using the AdamW optimizer with a learning rate of $2e^{-}5$. We use default values for other hyper-parameters.

6.2.2 CNN and RNN Models

We compare the best BERT model for the FinTech patent classification task with models that train a CNN (Convolutional Neural Networks) or an RNN (recurrent neural networks) on top of the patent pre-trained BERT model.

Convolutional Neural Networks (CNNs), originally proposed in computer vision, have been successfully adapted to text classification (LeCun et al. 2015, Kim 2014). In general, a CNN consists of convolutional layers followed by non-linear activations, pooling layers and fully connected layers. A convolutional layer employs a sliding window approach to apply a set of filters (low-dimensional tensors) to its input. The convolution operation captures local dependencies in the original input, and it produces feature maps. The pooling operation is used to reduce the dimensionality of the feature maps. Following convolutional layers (used together with non-linear activations), and pooling layers, a CNN has one or more fully connected layers, that capture non-linear dependencies among features (LeCun et al. 2015). The last fully connected layer uses a softmax activation function, and has as many output neurons as the number of targeted classes. In this work, we adopt the simple CNN architecture for text classification proposed in Kim (2014), which consists of one convolution layer with multiple filter widths and feature maps (and non-linear activations), followed by a max-pooling layer, and finally a fully connected layer with softmax output. The input to our CNN models consists of vector embeddings of patent word tokens. We use a BERT model pre-trained on FinTech patent documents to represent input tokens.

Among RNN models, we employ LSTM (Long Short-Term Memory) networks, which are a type of recurrent neural networks (RNN) that can be used to capture dependencies in sequence data, including long term dependencies (Hochreiter & Schmidhuber 1997). At the core of an RNN network, including LSTM networks, there is a recurrent cell represented as a hidden state, which enables the network to pass information from one time step to the next one. When unfolded, an RNN networks looks like a chain of repeated cells, which share the hidden state. General RNNs suffer from the gradient vanishing and exploding problem, when used with long sequences (Hochreiter & Schmidhuber 1997). LSTM networks avoid this problem by introducing a cell state, in addition to the internal hidden state, which carries information across the sequence. The information passed through the cell state is controlled by three gates, an *input gate*, a *forget gate* and an *output gate*. The gates help identify which information from the previous cell state needs to be forgotten, which information needs to be updated, and which information needs to be used as the output of the current cell state. A standard LSTM has one layer corresponding to one cell, and processes the sequence in the forward direction. A bidirectional LSTM (Bi-LSTM) includes a second layer/cell, which processes the sequence in reverse direction. In this paper, we use a Bi-LSTM network to capture sequence dependencies in both forward and reverse directions. As input to the Bi-LSTM, we use BERT embeddings pre-trained on FinTech data.

6.2.3 Traditional ML baselines

To compare our the performance of our deep learning classifiers with previous work, we also employ several traditional machine learning techniques. Specifically, we used Linear Support Vector Machines (L-SVM), Gaussian Support Vector Machines (G-SVM), Multilayer Perceptron (MLP), Naive Bayes (NB), Random Forest (RF), Gradient Boosting (GB), and also a voting classifier (ML-V). The voting classifier consists of L-SVM, G-SVM and MLP, similar to the voting classifier in Chen et al. (2019), and uses the prediction made by the MLP, in case of ties. All the classifiers, except for MLP, were implemented with `scikit-learn` library and used the default hyper-parameter values provided by the library (except for L-SVM, where $C = 0.0$, and G-SVM, where $C = 1$). The MLP network was implemented with PyTorch, and featured 4 layers. The first, second, and third layers had 256, 200, and 100 hidden neurons, respectively. During training, we used the standard cross-entropy as our optimization criterion and Adam as the optimizer. Further, we used a learning rate of $6e^{-4}$ for 30 epochs. One difference between our ML classifiers and the classifiers used in Chen et al. (2019) is that we used patent pre-trained BERT embeddings for the features as opposed to the bag-of-words representation.

6.2.4 Classifier Performance

In this section, we present the results of our performance benchmarking exercise that we conducted between the different state-of-the-art machine learning classifiers that we employed.

To train and evaluate our models, we split the dataset into training and test subsets, where the training subset contains 80% of the labeled data and the test dataset contains 20% of the data[9][10]

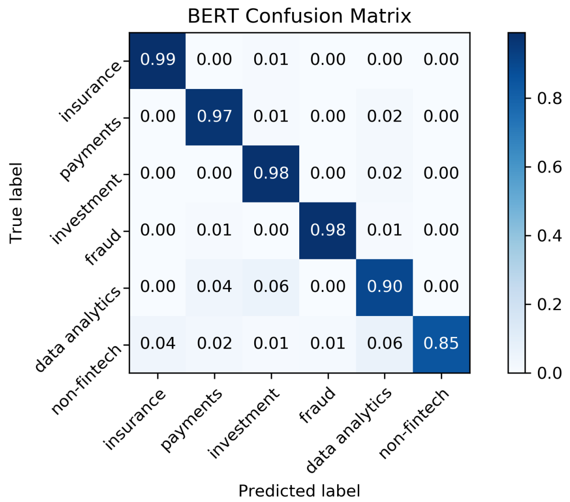
BERT models. We fine-tune 22 pre-trained BERT-based models on our labeled training dataset, and estimate their performance in terms of F1-scores on the test dataset. The F1-scores range from 89.39% (for "roberta-based") to 91.21% (for "bert-based-cased"), showing that the BERT-like models are generally good candidates for our FinTech patent classification problem. The best performing model, "bert-based-cased", has 12-layers, 768-hidden units, 12-heads and 110M parameters, and is pre-trained on cased English text. The next three best models (having similar scores to each other) are "roberta-large", "bert-large-cased" and "bert-large-cased-whole-word-masking", which all share the same architecture, specifically,

[9] The annotated dataset (in the form of patent identification number and label) is available from the authors upon request.

[10] To ensure that our deep learning models are capable of modeling more than one dataset split, we created five random splits that were used whenever we ran training experiments. Using these 5 random splits we are able to get a better idea of not only the raw performance of the model, but also its consistency when faced with different training data.

On top of these 5 random splits, we also manually curated a test set of 'hard' patent abstracts that we considered would be challenging even for a human to classify, informed by our own experience acquired through manually classifying well over 5000 patents throughout the development of the paper. This more challenging split has been constructed by recording the abstracts that the two researchers on the team manually classifying FinTech patents could not decide which FinTech subclass they belonged to, and needed to further discuss it with the wider team to ensure consistency. Finally, we run the classifiers on the 25580 patent dataset and obtain our shortlist of FinTech innovations. Throughout this process, we use the abstract text of each patent in this subset as the feature to be fed to the neural network.

Figure 7: Normalized confusion matrix corresponding to the best BERT model. The diagonal entries show the percentage of correctly classified instances in each category. Non-diagonal entries on a row show how the misclassified instances of a category are distributed among the other categories.



24-layers, 1024-hidden units, 16-heads, and 355M parameter, and are also trained on cased English text. As expected for patent documents, the uncased models have worse performance than the cased models overall.

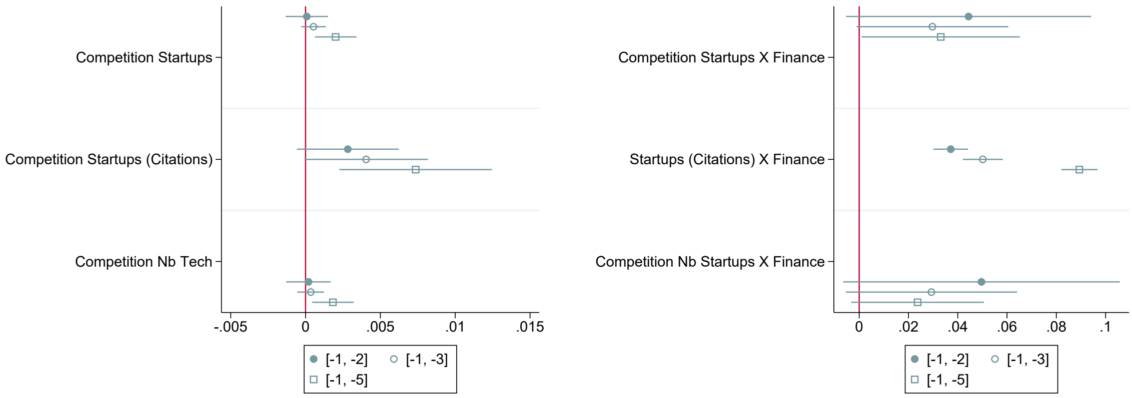
Fig. 7 shows the confusion matrix corresponding to the best BERT model, which is later used for the large-scale classification of FinTech patents. The diagonal entries show the percentage of correctly classified instances in each category, while the non-diagonal entries show how the misclassified instances for a particular category are distributed among the other categories. As can be seen, the *Insurance*, *Investment*, *Fraud* and *Payments* have a high percentage of instances correctly classified (0.99%, 0.98%, 0.98% and 0.97%, respectively). For *Data Analytics*, 90% of instances are classified correctly, while 10% are misclassified as *Payments* (4%) or *Investment* (6%). Finally, the *Non-FinTech* category has 85% correctly classified instances, and the misclassified instances are spread across all the FinTech categories. Together, these results show that our models are effective at identifying true positives for the FinTech categories, although some false positives (Non-FinTech patents classified as FinTech) will also be expected.

Convolutional and Recurrent Neural Networks. The result of the deep learning models, CNN and RNN, which take the patent pre-trained BERT embeddings as input, are shown on the right side of vertical double-line in Table 2, by comparison with the results of the best BERT model fine-tuned on the patent training data. The table shows the results for each category separately, and also the average over the 6 categories captured by our labeled data (including

5 FinTech categories and 1 Non-FinTech category), and the accuracy of the models. The three models have similar results, although BERT is the best model overall, with an accuracy of 92.30% and an average F1-score of 91.20% over the six categories. BERT is closely followed by the RNN model, while the CNN model has the worse performance among the three deep learning models. In terms of individual categories, the results show that the *Fraud* category has the best performance, with an F1-score of 97.00%, while the *Data Analytics* category has the worst performance, with an F1-score of 75.30%.

The results of the traditional machine learning models, which also take the patent pre-trained BERT embeddings as input are shown on the left side of the vertical double-line in Table 2. As can be seen, the best model in this category is the MLP model. However, except for two cases when the RF model gives the best result in terms of precision for the *Data Analytics* category and the best result in terms of recall for the *Non-FinTech* category, respectively, the deep learning models are superior to the traditional machine learning models.

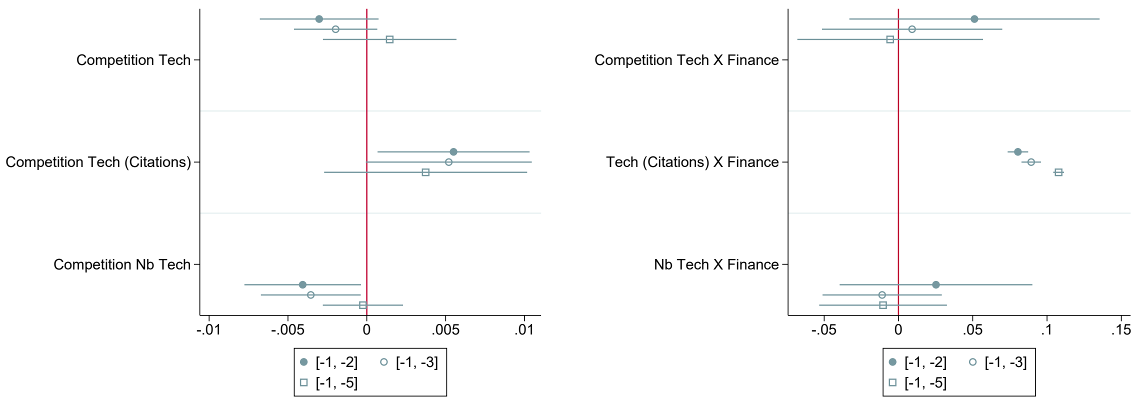
Figure 8: Alternative time frames for measuring startup competition



(a) FinTech Application indicator

(b) Total number of patents

Figure 9: Alternative time frames for measuring Tech competition

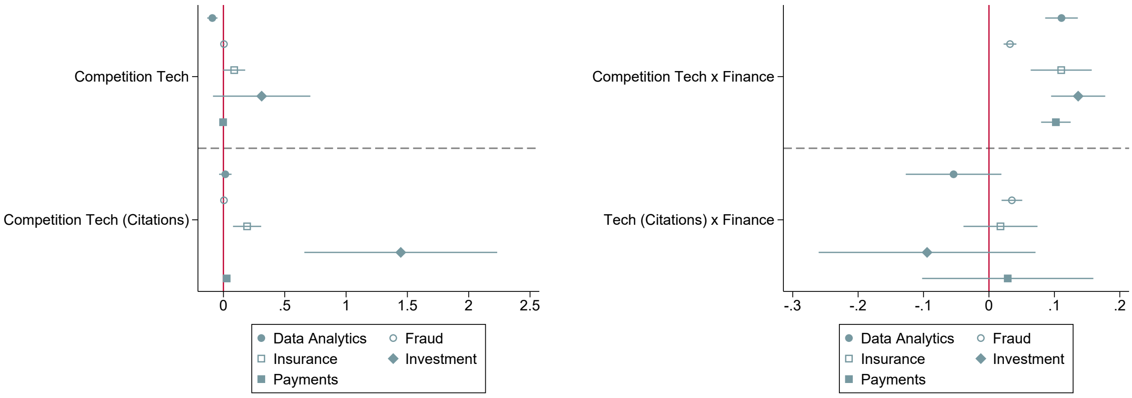


(a) FinTech Application indicator

(b) Total number of patents

6.3 Other robustness tests

Figure 10: Competition Tech and type of FinTech technologies



(a) FinTech Applications by technology (b) Number of Patents by technology

Table 13: Alternative measures of competition (based on Priority Country)

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)	
Competition 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)	
Competition 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

The dependent variable in columns (1)-(3) is a indicator variable whether a financial firm applied for a FinTech patent in year t, while is 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: 64-66. Competition Startups is the ratio of the sum of FinTech patent applications by non-financial start-ups to that of financial incumbents during the previous 4 years. Competition Startups (Citations) is the Competition Startups ratio scaled by the number of forward citations of non-financial startups. Number startups is the ratio of the total number of startups applying for a FinTech patent over the total number of incumbent the previous 4 years. Finance is a dummy variable for financial firms. Past Patents is the sum of FinTech patent applications by firm i over $[t-5, t-1]$. 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 $[t-4, t-1]$.

UCD CENTRE FOR ECONOMIC RESEARCH – RECENT WORKING PAPERS

WP21/04 Pierluigi Conzo, Laura K Taylor, Juan S Morales, Margaret Samahita and Andrea Gallice: 'Can ❤s Change Minds? Social Media Endorsements and Policy Preferences' February 2021
WP21/05 Diane Pelly, Michael Daly, Liam Delaney and Orla Doyle: 'Worker Well-being Before and During the COVID-19 Restrictions: A Longitudinal Study in the UK' February 2021
WP21/06 Margaret Samahita and Leonhard K Lades: 'The Unintended Side Effects of Regulating Charities: Donors Penalise Administrative Burden Almost as Much as Overheads' February 2021
WP21/07 Ellen Ryan and Karl Whelan: 'A Model of QE, Reserve Demand and the Money Multiplier' February 2021
WP21/08 Cormac Ó Gráda and Kevin Hjortshøj O!Rourke: 'The Irish Economy During the Century After Partition' April 2021
WP21/09 Ronald B Davies, Dieter F Kogler and Ryan Hynes: 'Patent Boxes and the Success Rate of Applications' April 2021
WP21/10 Benjamin Elsner, Ingo E Isphording and Ulf Zölitz: 'Achievement Rank Affects Performance and Major Choices in College' April 2021
WP21/11 Vincent Hogan and Patrick Massey: 'Soccer Clubs and Diminishing Returns: The Case of Paris Saint-Germain' April 2021
WP21/12 Demid Getik, Marco Islam and Margaret Samahita: 'The Inelastic Demand for Affirmative Action' May 2021
WP21/13 Emmanuel P de Albuquerque: 'The Creation and Diffusion of Knowledge - an Agent Based Modelling Approach' May 2021
WP21/14 Tyler Anbinder, Dylan Connor, Cormac Ó Gráda and Simone Wegge: 'The Problem of False Positives in Automated Census Linking: Evidence from Nineteenth-Century New York's Irish Immigrants' June 2021
WP21/15 Morgan Kelly: 'Devotion or Deprivation: Did Catholicism Retard French Development?' June 2021
WP21/16 Bénédicte Apouey and David Madden: 'Health Poverty' July 2021
WP21/17 David Madden: 'The Dynamics of Multidimensional Poverty in a Cohort of Irish Children' August 2021
WP21/18 Vessela Daskalova and Nicolaas J Vriend: 'Learning Frames' August 2021
WP21/19 Sanghamitra Chattopadhyay Mukherjee: 'A Framework to Measure Regional Disparities in Battery Electric Vehicle Diffusion in Ireland' August 2021
WP21/20 Karl Whelan: 'Central Banks and Inflation: Where Do We Stand and How Did We Get Here?' August 2021
WP21/21 Tyler Anbinder, Cormac Ó Gráda and Simone Wegge: '“The Best Country in the World”: The Surprising Social Mobility of New York’s Irish Famine Immigrants' August 2021
WP21/22 Jane Dooley and David Madden: 'Ireland’s Post Crisis Recovery, 2012-2019: Was It Pro-Poor?' September 2021
WP21/23 Matthew Shannon: 'The Impact of Victimisation on Subjective Well-Being' September 2021
WP21/24 Morgan Kelly: 'Persistence, Randomization, and Spatial Noise' October 2021 *(For revised version of this paper see WP21/25)*
WP21/25 Morgan Kelly: 'Persistence, Randomization, and Spatial Noise' November 2021
WP21/26 Eliane Badaoui and Frank Walsh: 'Productivity, Non-Compliance and the Minimum Wage' November 2021
WP21/27 Annette Broocks and Zuzanna Studnicka: 'Gravity and Trade in Video on Demand Services' December 2021
WP21/28 Linda Mastrandrea: 'Linking Retail Electricity Pricing and the Decarbonisation of the Energy Sector: a Microeconomic Approach' December 2021

UCD Centre for Economic Research
Email economics@ucd.ie