Research Policy 51 (2022) 104373

Contents lists available at ScienceDirect

Research Policy

journal homepage: www.elsevier.com/locate/respol

Innovative events: product launches, innovation and firm performance

Max Nathan [a,#], Anna Rosso [b,$,*]

[a] UCL, CEP and IZA
[b] Università degli Studi di Milano, CEP and LdA

ARTICLE INFO

JEL:
C55
L86
O40

Keywords:
innovation
productivity
ICT
data science

ABSTRACT

In this paper, we shed new light on the links between firm-level innovation and growth. We introduce data that capture a difficult-to-observe aspect of firms' innovative activity – new product/service launches – at scale. We show that our novel measures complement existing innovation metrics. We build a simple framework covering firm-level innovation, launches and revenue productivity. Then, we show positive linkages between past patenting and launches and between launches and performance for a large panel of small and medium-sized enterprises (SMEs) in the UK. We go on to explore the roles of age, size, industry and product/service quality in these relationships. A subset of SMEs with high-quality launches explains our results.

1. Introduction

There is decades' worth of research exploring the role of innovation in explaining economic performance. Endogenous growth theory helps explain country-level innovation-productivity linkages (Lucas, 1988; Romer, 1990). Schumpeterian frameworks highlight the roles of entrepreneurial entry, competition and factor reallocation (Schumpeter, 1962; Aghion and Howitt, 1992). They also provide firm-level microfoundations linking R&D activity, innovation and new products and services (see Akcigit (2017) for a review). Evolutionary perspectives emphasise that firms' capabilities vary greatly. Firms' resources accumulate over time, and it is difficult to shift them (Nelson and Winter 1982, Dosi et al 2000). This variation in firms' capabilities makes predictors of average firm performance difficult to identify (Nightingale and Coad, 2013).

The empirical literature on firm-level R&D, innovation and productivity dates back to Griliches (1979; 1986) and Crepon, Duguet and Mairesse (1998). Most subsequent studies use R&D, patent or innovation survey data to identify cross-sectional links between these factors and firm performance (see Hall, 2011; Syverson, 2011; Mohnen and Hall, 2013 and Audretsch et al, 2014 for reviews). More recent contributions use panel data and more sophisticated estimators (for example Fernandes and Paunov, 2015; Howell, 2015; Coad et al., 2016a,b; Baumann and Kritikos, 2016; Bianchini et al, 2018; Morris, 2018; Grillitsch et al, 2019; Spescha and Woerter, 2019; Audretsch and Belitski, 2020; Audretsch et al, 2020).

Consistent with evolutionary theory, these studies broadly confirm a positive linkage between innovation and firm performance – but with much heterogeneity across firm characteristics, behaviours and macro factors. The overall relationship is thus hard to pin down. For example, Mohnen (2019) shows that innovation has long-term effects on economic growth as measured by TFP, both at the firm level and the aggregate level, confirming Schumpeter's view of innovation. Conversely, Guarascio and Tamagni (2019) suggest that innovation-sales links are largely random, consistent with Gibrat's Law (Coad, 2009).

Starting with Mendonca et al. (2004), a more recent stream of work finds positive links between trademarks and innovation (see Taques et al. (2021), Castaldi et al. (2020), and Schautschick and Greenhalgh (2016) for reviews). These studies concentrate on start-ups, SMEs, and service industry settings (for example, Crass, 2020; Flikkema et al, 2019; De Vries et al, 2017; Block et al, 2015; Flikkema et al, 2014; Gotsch and Hipp, 2012).

One limitation of both bodies of work is that much innovative activity is informal and unobserved. Most companies rely on tools such as lead time, design complexity, or less often, NDAs or other forms of secrecy (Hall et al 2014). Only 1.6% of UK businesses file patents (Hall et al, 2013), and some firms employ patents defensively (Noel and

* Corresponding author:
E-mail addresses: max.nathan@ucl.ac.uk (M. Nathan), anna.rosso@unimi.it (A. Rosso).
Present Address: CASA, 90 Tottenham Court Road, London, W1T 4TJ UK.
$ Present Address: Università Statale di Milano, Dipartimento di Economia, Management e Metodi Quantitativi, Via Conservatorio 7, 20122, Milano Italy.

https://doi.org/10.1016/j.respol.2021.104373
Received 26 February 2020; Received in revised form 13 September 2021; Accepted 13 September 2021
Available online 30 September 2021
0048-7333/

Sample fragment	**Masterwork** goes large with new die cutter. Postpress equipment manufacturer Masterwork Graphic Equipment has expanded its range of products with the addition of the MK1450ER large-format die cutter with stripping and blanking facilities.…
doc_title	Masterwork goes large with new die cutter
url	http://www.XXX/NewsStory.aspx?i=2296
event_date	2014-03
source_name	xxx
company_id	13724
event_type_id	product_launch

Sample fragment	**Hammond Electronics** has launched a range of design specific moulded enclosures to support the new types of credit card sized, low cost bare board computers, which, typically running Linux, provide basic functionality across a wide range of applications…
doc_title	Enclosures for credit-card sized computers
url	http://www.XXXX/content/enclosures-credit-card-sized-computers
source_name	xxx
event_date	2013-12
company_id	1542955
event_type_id	product_launch

Sample fragment	New social housing energy switching service has been launched this week with Wigan and Leigh Housing, **Salix Homes** in Salford and Blackburn-based Twin Valley Homes, who together are responsible for a total of more…
doc_title	New social housing energy switching service launches
url	http://www.XXXX/energy-news/17541-new-social-housing-energy-switching-service-launches
source_name	xxx
event_date	2013-12
company_id	625998
event_type_id	product_launch

Figure 1. Example 'events', showing raw text and classification.
Source: GI. Each example shows the workflow from raw data to modelled variable. GI start with the raw text. We show a sample text fragment here with the company subject in bold. Title, URL, date and source name provide further information. As agreed with the data provider we cannot report the source name or the full text. The company ID field shows the match to Companies House data. Event type ID is the eventual classification into an event type: in both cases, these are new product launches.

Schankerman, 2013). While companies use trademarks more widely, including for non-technological and service innovations, it is still unclear which trademarks relate to innovation (Castaldi et al, 2020; Flikkema et al, 2019).[1] Low response rates and small samples can limit the usefulness of innovation surveys, as do the widely varying answers given by respondents (Mairesse and Mohnen, 2010).[2]

One newer strand of empirical work seeks to close these gaps with novel innovation metrics derived from firms' website text (Axenbeck and Breithaupt, 2019; Kinne and Lenz, 2019; Lenz and Winker, 2020), patents (Arts et al, 2021; Kelly et al, 2018) or regulatory filings (Saunders and Tambe, 2015; Hoberg and Philips, 2016; Kogan et al, 2017). These studies typically involve larger and/or listed firms rather than the SMEs that make up the bulk of the economies of more developed countries. Another strand of research uses product-level data to examine innovation-growth links (Bottazzi et al, 2001; Stam and Wennberg, 2009; Corsino and Gabriele, 2010; Cucculelli and Ermini, 2012; Argente et al 2018; 2019, Bokhari et al, 2020) or, relatedly, the timing of new product introductions (Ortega and García-Villaverde, 2011; Rodríguez-Pinto et al., 2012; Hsiao et al, 2017). Data constraints mean that this literature usually involves single or restricted industry cases: many studies use bespoke, small-*n* surveys.[3]

This paper makes three linked contributions to these debates. First, we use a novel mix of UK administrative microdata and media content to develop novel measures of innovative activity at the firm level across all industries and firm types. We have substantially broader coverage than both the abovementioned product data analyses and pioneering studies on media coverage and innovation.[4] To do this, we exploit a cutting-edge dataset developed by the data science firm Growth Intelligence (GI). This firm uses machine-learning routines on company website and media content to model firms' lifecycle 'events'. We focus on one of

[1] Castaldi et al. (2020) show that firms use trademarks for multiple purposes, including securing market position (allowing markups or deterring entry) and attracting resources (from venture capitalists and other investors).

[2] For example, the response rate of the 2017 UK Innovation Survey was 43 percent. This rather modest response rate mainly affects the creation of a balanced panel of firms over consecutive CIS years (Mairesse and Mohnen, 2010).

[3] For example, Bokhari et al, Bottazzi et al, Corsino and Gabriele, and Ortega and Garcia-Villaverde all examine single industries; Rodriguez-Pinto et al survey 136 manufacturing firms, Cucculelli and Ermini investigate 204 SMEs, and Stam and Wennberg analyse approximately 2,000 start-ups. Focusing on publicly listed firms, Argente et al (2018, 2019) use barcode data with much broader, fine-grained coverage closer in spirit to our data.

[4] Katila and Ahuja (2002) and Fosfuri et al (2008), for example, are restricted to a few hundred firms in single sectors. These studies use the 'counting innovation' method, to examine a selection of innovations introduced in a given year and reported in trade journals. See Kleinknecht et al (1993), Coombs et al (1996) and Santarelli and Piergiovanni (1996).

these variables – new products/services reported in the news. We also exploit overall event exposure to aid our research design. We clean and refine these data using structural topic modelling to better align reported and real-world activity. We also advance innovation-output studies, developing measures of launch quality analogous to patent citations.

Second, we show that these reported launches complement measures of formal IP. This includes frequency and industry/geographic coverage. For example, for single-plant SMEs, we find 24,720 UK launches in 2014/2015, versus 4,194 patent applications and 4,510 trademarks filed. We then show positive links between past IP activity and current launches. We find cross-industry and trademark-type variation consistent with prior literature and our framework's predictions.

Third, we develop a simple framework linking firm-level IP, launches and performance and test its predictions on a panel of UK SMEs. In particular, we test whether a company grows by transforming new products/services into higher revenue per worker and explore the roles of age, size, industry and product/service quality in explaining our results. We pay careful attention to the fact that event exposure is not random and that we are working with reported rather than observed activity. We find that launch activity is associated with higher SME revenue per worker, especially in the service sector, among medium-sized firms and among firms with specialised trademarks. Consistent with a world of heterogeneous firm capabilities, a small subset of high-quality launches helps drive the main result. Robustness checks on all firms with event exposure find relatively weak links – consistent with larger, multi-plant firms having multiple sources of revenue growth, such as advertising existing products.

To our knowledge, this study is the first to use such data, at scale, to study innovation and firm performance. More broadly, it contributes to the literature on firm growth (Gilbert and Newbery, 1982; Coad, 2009; Audretsch et al, 2014; Castaldi et al, 2020), IP choices (Hall et al, 2014), determinants of firm productivity (Syverson, 2011), high-growth businesses (Coad et al, 2014), and economic applications of natural language processing (Gentzkow et al, 2019). Crucially, we combine text-based measures of innovative activity with high-quality administrative microdata. This gives us a clear sampling frame, aiding inference and interpretation (Einav and Levin, 2014). We also advance on rich data papers such as those of Hall et al. (2013) and Coad et al., 2016a, who combine conventional administrative data, patents, trademarks and innovation surveys to examine smaller groups of firms. Overall, we focus on demonstrating data use cases and estimating clean associations rather than causal effects. Our dataset is replicable and extendable for future research.[5]

2. Data

We use modelled company 'events' to develop new measures of innovative activity. Our method extends the innovative outputs approach (Corsino and Gabriele, 2010; Cucculelli and Ermini, 2012; Bokhari et al, 2020). Each 'event' derives from article text taken from 3, 740 online news sources (including major sources such as Reuters and Yahoo news and industry sources such as IT Briefing and PRWeb).[6] Our raw data consist of 318,899 observations corresponding to 30,205 companies during financial years 2014 and 2015. Growth Intel matches the text of each article to the UK company register (Companies House) using firm names and contextual information, then uses supervised learning to classify the text as one of several event types. Nathan and Rosso (2015) provide more details on the data sources and workflow.[7] We focus on one event type: 'product/service launch'. Fig. 1 provides two examples of product/service launches, showing both raw inputs and modelled outputs.

In theory, each launch event represents a new product/service that a given firm releases into the world and that is then covered by at least one of our media sources. In practice, we need to deal with three ascription challenges. First, our coverage may be uneven, analogous to the well-discussed limitations of patents (Hall and Harhoff, 2012). Second, media exposure is determined by a combination of firm decisions, firm media-generating capacity and reporting norms, some of which are difficult to observe. Third, and relatedly, reported launch content might reflect firms' advertising or PR rather than 'true' innovation.

We deal with these issues as follows. First, we carefully clean the raw data, as detailed below. Additionally, we show that launch coverage is substantially more frequent and even across industries and regions than patents and trademarks.

Second, to address selection issues, we use firms' overall 'event exposure' or 'coverage'; that is, whether a firm has *any* reported events of *any* kind. We show that event exposure is not random and that it is correlated with a range of observables, while launch exposure is more balanced. We focus on SMEs with event exposure and then show that our main results hold for larger samples.

The third issue is harder to disentangle. Our input data are news articles, not advertising copy. However, we may be (a) reporting trivial innovations, (b) missing important innovations or both. To address (a), we measure the number of raw mentions of each launch and use this to construct a proxy measure of launch 'importance' analogous to patent citations. Existing literature provides some reassurance on (b). Firms' predominant use of informal IP protection means that *all* innovation measures undercount the true level of innovative activity (Hall et al, 2014). However, coverage of variables based on observations of products and services in the marketplace is then more affected while that of measures based on formal IP practices is less affected.

2.1. Build

Our data cleaning is summarised here and detailed in Appendix A1.

[5] A sample of the GI data, plus cleaning, topic modelling and matching code is available at https://osf.io/bjykc/. Our data is part of a growing body of similar resources. Existing datasets on news events such as GDELT and Events Registry are designed for country-level analyses, especially those on politics/current affairs. Other proprietary firm-level datasets such as Mattermark (US) and Beauhurst (UK) provide rich information from a range of sources but are restricted to small numbers of 'high-potential' businesses. Crunchbase (US) is a global wiki-type dataset regarding the tech sector with good US coverage but limited coverage for other countries, as well as significant quality concerns due to the self-reported nature of its data (Motoyama and Bell-Masterson, 2014). Glass AI (UK) draws data from firms' websites with some observations now linked back to administrative data; see Siepel et al (2020) for one recent use case. SpazioDati (Italy) provides similar website-sourced data, as do the German datasets used in Axenbeck and Breithaupt (2019) and Kinne and Lenz (2019) to measure firm-level innovation.

[6] We cannot split the sample of product launches based on sources. In future research, it would be interesting to compare events reported via online news outlets (as we have here) with events directly reported on company websites (which we do not include). Similarly, we lack detailed information on local vs. national online news sources.

[7] Text fragment for illustration. GI uses a full page of content to assign text to a subject company and to classify the related activity. Where a text describes more than one subject company, as in mergers or joint ventures, GI assigns the event to a pair of companies or *n*-groups. GI also filters the data to remove results from irrelevant domains (for example, mentions of companies in celebrity magazines or results from sites that largely or wholly deal with markets outside the UK).

We first remove duplicates and control for 'farmed' content.[8] Next, we run two quality checks on GI's ascription routines. We then improve the realism of the data using structural topic modelling (STM). A major product launch is likely to be reported hundreds of times; in the raw data, each launch is reported as a distinct event. STM is used to cluster text fragments that refer to the same topic into single observations representing the underlying real-world launch event (Roberts et al, 2016). Overall, our cleaning steps substantially reduce the number of event observations from 318,899 observations for 30,205 firms, to 257, 056 observations for 30,187 firms. STM accounts for the bulk of this reduction.

We combine this cleaned event data with data from other sources. To do this, we link Companies House identifiers to the Business Structure Database (BSD) (Office of National Statistics, 2017). The high-quality administrative microdata in this database cover 99% of UK enterprises and provide a clearly defined sampling frame. We then use various matching routines, detailed in Appendix A2,[9] to link US, European and other patent data (from Orbis, application years 1900-2015) and UK trademark data (from the UK Intellectual Property Office, 2012 - 2015).[10] The resulting dataset includes 1,399,146 firms and is an unbalanced panel of 5,039,811 firm-year observations from 2014-2017. Within this set, there are 22,497 firms with event exposure and 212, 426 unique events.

In our descriptive analysis (Section 3), we use data for the subset of single-plant SMEs from 2014-2015 – the years in which events are observed – to compare the characteristics of GI product launches with those of traditional IP measures. Crucially, using single-plant SMEs allows us to cleanly ascribe a launch to a given firm and location. Removing the largest firms also reduces ascription error (see Appendix A1). Over 95% of firms in the BSD are single-plant SMEs.[11]

In our regression analysis (Sections 4-6), we further focus on the subset of single-plant SMEs with 'event exposure', which we define as firms that had an event (of any kind) during 2014, 2015 or both. This allows us to work with the variation in launch activity across firms and time, conditional on non-random event exposure. Given that these restrictions remove much of the firm-level variation in event activity – see Fig. 2 below – for robustness, we rerun our analysis for a) all firms with event exposure, including large and multi-plant businesses, and b) all SMEs.

Table 1 provides a comparison of our full dataset and subsamples. For each, the first panel gives the number of firm-year observations for each year, the years during which we observe events, and the number of unique firms. The next panel gives the number of firms for which we observe events of any kind and the total number of events for each year. We repeat this for launches and then for patents and trademarks. Overall, 1.6% of all the firms and 1.3% of the SMEs have event exposure; in both cases, over 37% of the firms also have launch exposure. Events and launches are well balanced across all the years. There are rather smaller shares of firms with patents and trademarks, with more uneven coverage over time.

3. Descriptive analysis

We now explore how modelled launches compare with other innovation measures. We make a launch dummy that takes the value of one when a firm has at least one launch during a given year. We also count each firm's launches in that year. We then use the number of raw observations per modelled event to create measures of launch 'importance'. For each firm-year cell, we make a count of mentions, a dummy for whether a firm has an 'important' launch, and a count of such launches. Further details are given in Appendix A1.

Single-plant SMEs with event coverage have approximately 2.2 events, of which approximately 0.7 are product/service launches. The firms have fewer patents and trademarks. While the average launch has over 250 raw media reports, this is driven by a small number of high-profile events; only 2% of the SMEs with events have 'important' launches with more than one underlying media report. Appendix Table B1 provides details.

As anticipated, event exposure is not random, with firms selected on a range of observables. Table 2 shows the mean characteristics of SMEs with and without events and launches for 2014-2015, the years during which we observe event activity. We can see that the mean differences between firms with and without event exposure are large; rank-sum tests confirm significant mean differences for all observables. In contrast, for firms with events, differences between those with and without launch activity are rather small and often nonsignificant.[12] Companies with launches are more likely to obtain patents and trademarks. They are older, have significantly lower revenue productivity, are less likely to have high revenue and growth episodes, are more likely to be foreign-owned, and are more likely to be listed companies than partnerships. However, the balance on shares of start-ups, sole proprietors and small firms, business group structure, number of employees, urban location, revenue, revenue per worker growth, employment level, growth and high-growth episodes. Therefore, we focus the second part of our analysis on SMEs with events.

Table 3 compares coverage of patents, trademarks and reported launches at the industry level for 2014-2015, the years during which we observe events. We show launch, patent, and trademark coverage across SIC1 bins for firms with event exposure (Table B3 repeats this analysis for all SMEs, showing similar results). Overall, launches have a wider industry spread than patents or trademarks; for a minority of firms, we also find correlations between the three measures.[13] Patenting is concentrated in manufacturing, but it is also present in services, notably

[8] Recent structural changes to the media industry – notably, the rise of online platforms – may be the reducing levels of data quality and scrutiny in this industry, for example through 'content farming' and 'churnalism' (Viner, 2016; Gentzkow and Shapiro, 2010; Davies, 2009). The first leads to duplicate reported events, while the second alters the distribution of event activity. Both may be particularly prevalent in the information and communications technology (ICT) sector (Lafrance, 2016). We identify duplicate observations of events using all available variables except the source and time. Within each group, we keep only one event; thus, we are not selecting events on the basis of source quality.

[9] Bureau Van Dijk identifiers or firm name and full postcode. An alternative approach is the automated method developed by Autor et al (2020), which exploits internet search results.

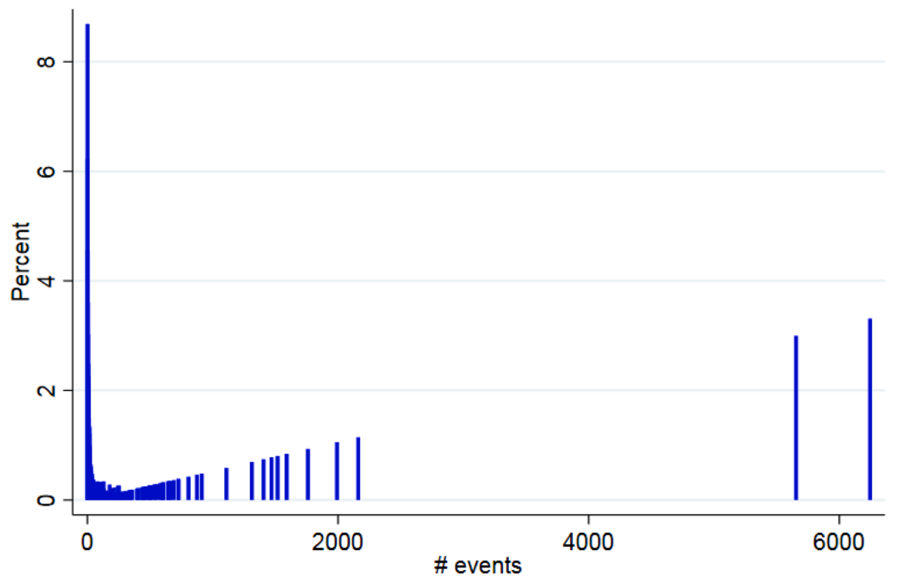
[10] We also link our data with data from the UK Innovation Survey, although this is less successful. Specifically, we link our data with Waves 4-9 of the UK Innovation Survey (UKIS), covering the years 2002-2014. We match 26,708 firms, of which 1,173 are single-plant SMEs. We run Hotelling tests to determine whether this set of firms is systematically different from the rest of the sample. The results, which are all significant at the 1% level, show that the UKIS subsample differs on a large set of observable characteristics (Hotelling's $T2 = 24866.03$, $F(29,2468834) = 857.440^{***}$). In particular, the UKIS firms have a substantially higher probability of event exposure (2.91% vs 0.86%) and launch activity (1.03% vs 0.31%). The firms with event exposure that are in the UKIS subsample are also systematically different from the other firms with event exposure (Hotelling's $T2 = 487.983$, $F(21868) = 17.407^{***}$). Given these substantial differences, we do not use UKIS data in the subsequent analysis.

[11] We also remove outliers: for each year, we remove observations with an event count higher than 1 standard deviation above the mean event count. This eliminates 84 observations.

[12] Rank-sum tests are preferred, as we do not know the underlying distribution of events. T-tests give virtually identical results and are available on request.

[13] Specifically, Table B2 further checks for the dispersion of launch activity. It reports the number of launches for firms with one patent or one trademark in either 2014 or 2015 for all SMEs, and for those with event exposure. While the majority have zero launches to show for the IP, a small number have two or more reported launches per patent or trademark. Table B3 reports the industry coverage for all single plant SMEs.

A. Raw sample, all firms with events exposure.



B. Single-plant SMEs with events exposure. Disclosive cell counts suppressed.

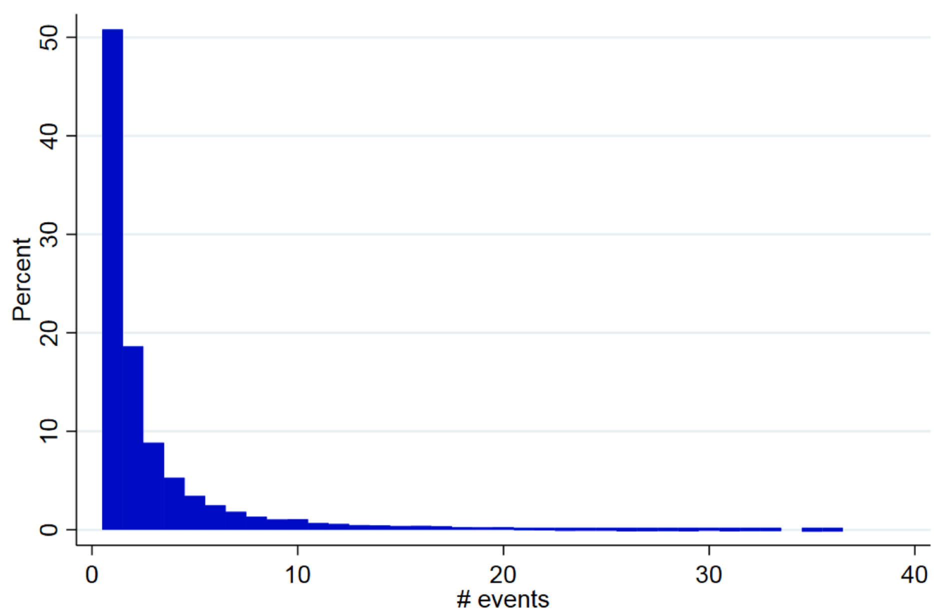


Figure 2. Histogram of events activity, 2014-2015. Raw sample (top), single-plant SMEs (bottom). A. Raw sample, all firms with events exposure. B. Single-plant SMEs with events exposure. Disclosive cell counts suppressed.. Source: GI data, years 2014-2015.

business services (including software and other 'knowledge-intensive' activities (Castellacci, 2008)). Given their broader functionality, trademarks are more evenly distributed, with most related activity in manufacturing, wholesale/retail/repair and social/personal services.

Finally we turn to spatial variation. Appendix Figure B1 shows geographical coverage of events, launches, patents and trademarks across urban travel to work areas (TTWAs), which approximate local spatial economies. Panel A provides a simple scatterplot of launches, patents and trademarks across TTWAs. Based on raw counts, coverage across spatial economies appears even, although launch counts are substantially higher than patent or trademark counts. London is a major outlier in terms of counts, even for single-plant SMEs.[14] To correct for this, Panel B plots TTWA counts weighted by the number of firms in each TTWA. We can see that when local economic conditions are taken into account, launches have a far more even geographical distribution than either patents or trademarks.

[14] Table B4 shows the urban/non-urban and London/non-London shares of launches, patents and trademarks for single-plant SMEs. Launches and other innovation metrics are highly urbanised.

Table 1
Panel characteristics, 2014-2017.

	All firms	Single plant SMEs	SMEs with events	All firms with events
Observations, all years	5,039,811	4,878,532	67,739	87,390
Observations, 2014-15	*2,723,875*	*2,643,043*	*35,289*	*44,763*
Unique firms	1,399,146	1,364,624	17,905	22,497
Firms with events	22,497	17,905	17,905	22,497
#events	212,426	78,090	78,090	212,426
of which 2014	*113,423*	*42,225*	*42,225*	*113,423*
of which 2015	*99,003*	*35,865*	*35,865*	*99,003*
Firms with product launches	8,435	6,640	6,640	8,435
#launches	89,027	24,720	24,720	89,027
of which 2014	*47,236*	*12,527*	*12,527*	*47,236*
of which 2015	*41,791*	*12,193*	*12,193*	*41,791*
Firms with patents	2,355	1,795	295	645
#patents	9,064	4,194	1,141	4,474
of which 2014	*8,235*	*3,892*	*1,055*	*3,393*
of which 2015	*829*	*302*	*86*	*481*
Firms with trademarks	3,961	3,164	280	589
#TMs	6,407	4,510	491	1,322
of which 2014	*6,407*	*4,510*	*491*	*1,322*
of which 2015	*0*	*0*	*0*	*0*

Source: BSD / GI / Companies House / Orbis / UKIPO.

4. Research design

4.1. Theoretical framework

We develop a simple framework to formally explore links between launches, other innovation measures and firm performance, which we take to data in the rest of this paper. We start with the quality ladder model developed by Aghion and Howitt (1992), in which product innovation stems from a combination of entrants seeking market share (and thus temporary revenue markups) and incumbents seeking to protect it. Innovation at the firm level leads to increased average product quality, knowledge spillovers across firms, and the reallocation of capital, workers and products across and within firms as products and businesses enter and exit the market (Klette and Kortum, 2004; Lentz and Mortensen, 2008).[15] In turn, these three forces drive aggregate revenue and productivity growth. In practice, any firm's growth path is further shaped by its individual endowments, including capabilities built up over time (Nelson and Winter, 1982; Cohen and Levinthal, 1990; Dosi et al, 2000). Firms and their managers also operate with bounded rationality and may individually maximise profits, growth or some combination of the two (Marris, 1963). These factors skew growth distributions so that a small number of high-growth/superstar firms typically has a disproportionate impact on aggregate outcomes (Nightingale and Coad, 2013).

Given our short timeframe, we say a firm has fixed capability endowments and grows through developing products and services – either new to the firm or the market. Specifically, innovation is reflected in IP and is a function of costly past and current R&D and managerial and organisational capabilities. This can come from either inside a firm or external sources via spillovers (Akcigit and Kerr, 2018). Incumbents with valuable IP may be less likely to source innovations externally (Bei, 2019). A firm's probability of conducting a successful launch then depends on existing capabilities and knowledge stocks accumulated through past product innovations. Higher-quality products/services, produced by a subset of more capable and/or experienced firms, allow for higher markups but are also costlier to develop.

Firms also need to decide how to protect innovations and to what ends. As noted above, the majority of firms avoid using formal IP tools or combine formal and informal tools (Hall et al, 2014). In principle, patents indicate 'upstream' inventions and trademarks denote 'downstream' commercialisation (Castaldi et al., 2020; Flikkema et al., 2019). In practice, they are used as complements or substitutes (Llerena and

[15] Most of the earlier literature on reallocation has focused on input markets, combining labour market and establishment data (see, for example, Foster et al, 2016). In contrast, Argente et al (2018) focus on output markets.

Table 2
Comparing observable characteristics across samples, 2014-2015.

Variables	A. Single plant SMEs		B. SMEs with events		
	No events	Events	No launches	Launches	
	(1)	(2)	(3)	(4)	(3)-(4)
Patent count	0.001	0.042	0.023	0.062	***
Weighted patent count	0.001	0.041	0.023	0.062	***
EPO/US/PCT patents	0.001	0.023	0.013	0.037	***
Weighted EPO/US/PCT patents	0.001	0.023	0.013	0.037	***
TM count	0.002	0.015	0.011	0.022	***
Rev per worker two-year average	146.55	781.3	900.341	461.6825	***
Annual % rev per worker growth	-0.006	0.017	0.019	0.0132	
High rev per worker growth firm	0.129	0.148	0.151	0.139	**
Revenue two-year average	811	13752	12932.89	13264.97	***
Annual % revenue growth	0.011	0.049	0.05	0.048	
High revenue growth firm	0.15	0.215	0.22	0.208	**
Employment two-year average	5.1	21.2	21.254	21.352	
Annual % employment growth	0.017	0.032	0.03	0.036	
High jobs growth firm	0.014	0.06	0.059	0.059	
Age entered BSD / incorporated	12.4	17.9	17.421	17.943	***
Startup	0.142	0.028	0.028	0.028	
Firm has 1-9 staff	0.892	0.571	0.592	0.568	***
Firm has 10-49 staff	0.086	0.284	0.278	0.289	*
Firm has 50-249 staff	0.013	0.124	0.11	0.125	***
Immediate foreign ownership	0.165	0.328	0.269	0.44	***
Firm is in a group of enterprises	0.003	0.055	0.048	0.047	
Number of companies in the group	0.008	0.187	0.168	0.119	
Firm is a company	0.942	0.903	0.899	0.931	***
Firm is a sole proprietor	0.021	0.004	0.004	0.003	
Firm is a partnership	0.014	0.004	0.005	0.002	***
Firm is a public company	0	0.001	0.001	0.001	
Firm is a non-profit / social enterprise	0.023	0.088	0.103	0.063	***
Services sector	0.909	0.883	0.891	0.858	***
Urban TTWA	0.788	0.838	0.832	0.836	
Greater London	0.228	0.303	0.292	0.291	
Observations	*2,643,043*		*35,289*		
Unique firms	*1,346,719*	*17,905*	*11,265*	*6,640*	

Source: BSD / CH / GI / Orbis / UKIPO. The table shows mean differences between all single-plant SMEs with and without events exposure (Panel A) and for SMEs with and without launches (Panel B). For Panel B, the stars in the last colum give the results of rank-sum tests for each variable between columns 3 and 4. *** denotes 1% significance
** denotes 5% significance.

Table 3
Coverage by SIC1 sectors for product launch, patents and trademarks, 2014-15.

SIC1	Section Name	SMEs with events exposure. % of firms with coverage, within sectors Launch	Patent	TM	*N*
A	Agriculture, hunting and forestry	19.5			*200*
B	Fishing				*12*
C	Mining and quarrying	15.71	.	.	*70*
D	Manufacturing	29.24	2.66	1.33	*3,837*
E	Electricity, Gas and Water Supply				*92*
F	Construction	14.11			*1,680*
G	Wholesale and retail trade, etc	33.92	0.74	0.85	*4,593*
H	Hotels and restaurants	19.34			*543*
I	Transport, storage and communications	23.2			*1,319*
J	Financial intermediation	12.43		0.87	*1,601*
K	Real estate, renting and business activities	22.3	1.05	0.73	*16,814*
L	Public administration and defence, etc				
M	Education	14.62			*643*
N	Health and social work	15.35			*951*
O	Other community, social and personal services	23.93		1.02	*2,934*
P	Household domestic employment				
Q	Extra-terrestrial organisations, bodies				
	Average coverage, %	*23.45*	*0.96*	*0.79*	
	Observations	*8,275*	*339*	*280*	*35,289*
	Unique firms	*6,640*	*295*	*280*	

Source: BSD / CH / GI / Orbis / UKIPO. For each sector, the table shows the share of firms with coverage reported in the column (event, launch, patent, tm). N is the total number of firms in each sector. Panel A reports all single plant SMEs, Panel B reports single plant SMEs with at least one event in some year. Observations are instances of a firm having at least one event, launch, patent or TM in that year. Trademark data is only available to 2014, so that observations are the same as unique firms. Cells with under 10 observations are suppressed to avoid disclosure.

Millot, 2020; De Vries et al, 2017) at varying points in the innovation process (Seip et al, 2018), and varying across industry and market contexts (Jensen and Webster, 2009). In particular, brand creation trademarks and trademarks with a narrow scope may be important indicators of innovation in start-ups and young firms (Flikkema et al, 2019).

Formally, Argente et al. (2019) distinguish between 'productive' IP, which creates new products and protects revenue markups, and 'protective' IP, which creates future revenue sources. Large firms worry about cannibalizing their existing products; thus, returns to new products decrease with firm size while returns to patenting increase. Large firms are also better able than small firms to bear the costs of formal IP protection. This implies that small firms below a certain size cut-off may not patent or trademark at all, while large businesses have multiple filings per new product; the cost differentials between patents and trademarks also influence this (Castaldi et al, 2020). Additionally, firms can choose between engaging in product/service innovation and advertising/marketing their existing offerings. Advertising may complement innovation by increasing markups on new products or substitute for it by increasing revenues on existing products (Bokhari et al, 2020; Cavenaile and Roldan 2020). Larger (and older) firms tend to prefer advertising their existing products over product innovation due to spillovers from umbrella branding.

4.1. Empirical specification

This framework generates a number of predictions that we can explore in our data – both on links between IP activity and launches and from launches to firm growth (proxied by revenue per worker, as is usual in this literature). First, we should expect a positive link between past IP (patents and trademarks) and launches. The timing of IP activity is ambiguous. Knowledge stock decay implies that past activities have weaker links, but to the extent that past IP proxies for individual firm capabilities, it positively predicts launch activity. Per Flikkema et al. (2019), we should expect stronger links for more narrowly focused trademarks and for trademarks related to services. Per Argente et al. (2018), larger firms should produce more IP per launch than smaller firms; the exact size cut-off is an empirical question. Relatedly, younger firms should have higher probabilities of launching and lower returns to formal IP. Sectors where the cost of R&D is lower should exhibit stronger IP-launch links – for example, services should have stronger links than manufacturing (Audretsch et al, 2020).

Second, we should expect a positive relationship between firm launch activity and levels of revenue productivity, as innovations generate temporary monopolies for their producers. To the extent that they reflect higher-quality products/services, more 'important' launches should generate stronger revenue/worker effects. In contrast, the link to revenue productivity *growth* is ambiguous, as we do not observe different levels of overall competition in our data.

In theory, to estimate the link between IP and launches, we can estimate for firm i in year t, TTWA a and sector s:

$$\mathrm{L}_{itas} = \mathrm{F}(\mathbf{IP}_{\mathrm{it-n}},\ \mathbf{X}_{\mathrm{it-n}},\ \mathrm{T}_t,\ \mathrm{A}_a,\ \mathrm{S}_s,\ \mathrm{e}_{itas}) \tag{1}$$

where L is a measure of launch activity, including proxies for launch importance/quality; **IP** is a vector of past patenting, trademarking and self-reported innovation; **X** is a vector of time-varying controls; and T, A and S are year, area and industry fixed effects, respectively. Similarly, to estimate the launch-growth link, we can estimate:

$$\mathrm{Y}_{itas} = \mathrm{F}(\mathrm{L}_{it-n},\ \mathbf{IP}_{it-n},\ \mathbf{X}_{it-n},\ \mathrm{T}_t,\ \mathrm{A}_a,\ \mathrm{S}_s,\ \mathrm{u}_{itas}) \tag{2}$$

where Y is a measure of revenue productivity, and other terms are defined as above.

This design leaves us with three main challenges. First, in our framework, firms have fixed, individual capability endowments; additionally, the decision to innovate varies at the firm level. Only some relevant determinants are observable, and our short panel makes it challenging to fit firm fixed effects. Blundell et al. (1995) propose using firm-specific 'level effects' based on historic patenting activity, and we follow this alternative to capture firm-level heterogeneity. Thus we estimate the following:

$$\mathrm{L}_{itas} = \mathrm{F}(\mathbf{IP}_{it-n},\ \mathbf{X}_{it-n},\ \mathbf{HP}_i,\ \mathrm{T}_t,\ \mathrm{A}_a,\ \mathrm{S}_s,\ \mathrm{e}_{itas}) \tag{3}$$

$$Y_{itas} = \mathrm{F}(L_{it-n},\ \mathbf{IP}_{it-n},\ \mathbf{X}'_{it-n},\ \mathbf{HP}_i,\ \mathrm{T}_t,\ \mathrm{A}_a,\ \mathrm{S}_s,\ \mathrm{u}_{itas}) \tag{4}$$

Second, launches (and events, more broadly) are media-reported rather than directly observed. For a given firm, event exposure is determined by a) a firm's decision to seek coverage, b) its capacity to do so, and c) media interest in reporting the firm's activity. The value of media coverage varies across firms and is a function of management strategy. The capacity to achieve is a function of management quality (Cohen and Levinthal, 1990), resources and other characteristics (such as age, size, legal and corporate structure) (Teece et al, 1997). Both firm choices and capacity are also shaped by industry characteristics, trends and macro forces, such as national/international policy regimes, trade frictions and changes in these factors (Cockburn et al, 2016). Media interest may vary across industries (for instance, on levels of newsworthy content) and locations (physical proximity to media producers), and is affected by media industry trends related to reporting capacity and coverage (Davies, 2009; Viner, 2016).

Much of this can be addressed with controls and fixed effects, while

Table 4
Linking past IP activity to product launches. Stepwise regressions, SMEs with events, 2014-2017.

	(1)	(2)	(3)	(4)
	A. Probability to Launch			
L1.15% depreciated PCT / EPO /	0.011***	0.010***	0.008***	0.007***
US patent count	(0.002)	(0.002)	(0.002)	(0.001)
L1.15% depreciated TM count	0.009	0.006	0.006	-0.001
	(0.010)	(0.010)	(0.010)	(0.008)
Ave pre-2009 patenting			-0.026***	-0.022***
			(0.007)	(0.007)
Firm patents pre-2009			0.138***	0.097***
			(0.036)	(0.035)
Observations	29528	29528	29528	29189
R^2	0.0012	0.0059	0.0070	0.0614
B. Launch counts				
L1.15% depreciated PCT / EPO	0.063***	0.056***	0.059***	0.055***
/ US patent count	(0.009)	(0.010)	(0.010)	(0.011)
L1.15% depreciated TM count	0.010	-0.009	-0.011	-0.032
	(0.030)	(0.030)	(0.030)	(0.029)
Ave pre-2009 patenting			-0.138**	-0.120*
			(0.053)	(0.067)
Firm patents pre-2009			0.371	0.196
			(0.250)	(0.268)
Observations	29528	29528	29528	29189
R^2	0.0008	0.004	0.004	0.031
Controls	N	Y	Y	Y
Pre-sample patenting	N	N	Y	Y
Year, area and industry dummies	N	N	N	Y

Source: BSD/CH/GI/Orbis/UKIPO. The dependent variable is a dummy for whether the firm has a product launch in a given year (Panel A) and the count of a firm's product launches in that year (Panel B). We control for log mean turnover and employment, age, firm size dummies, company legal status and structure dummies, and an urban TTWA dummy. Controls are lagged one year except age. Pre-sample patenting levels effects are detailed in the main text. Standard errors are clustered on 2-digit SIC.

*** Denotes a result significant at 1%.

** Significant at 5%.

* Significant at 10%. Constant not shown.

our single-country setting eliminates cross-country differences.[16] Nevertheless, unobservables that affect event exposure may also condition both sides of Eqs. (3) and (4). Because launches are observed only conditional on event exposure, we cannot directly control for the latter. While we could in principle use a Heckman or IV estimator to handle selection, in this case there is no obvious instrument. Thus, our preferred approach is to estimate Eqs. (3) and (4) for the sample of SMEs with events so that we can estimate the examined linkages conditional on all drivers of event exposure. We also run diagnostic/falsification tests on this sample, showing that past IP is linked to launches and not other event counts, while other events are, as expected, correlated with our level effect (past patenting). This provides further support that we are both estimating a true IP-launch relationship and controlling for unobservables that drive both selection and outcomes. In robustness checks, we quantify these sources of bias by re-estimating (3) and (4) for all SMEs and for all firms with events.

Third, per Cavenaile and Roldan (2020), launches have measurement error to the extent that media coverage functions as a form of advertising for firms. It is extremely challenging to distinguish the direct revenue effect of a new product/service from that of the launch process. Nevertheless, an observation of the predicted significant positive relationship between past IP and launches would support the idea that reported launches are linked to innovation rather than being purely a form of advertising. Relatedly, we lack data on advertising spending at the firm level.[17] However, if advertising spending is equal within firm size and industry bins, our specification eliminates spending as an unobservable.

5. Results

First, we examine the link between patenting/trademarks and the extensive and intensive margins of product launches. We then move to the gains of innovation, estimating the link between launches and revenue productivity.

5.1. Linking past IP with launch activity

Table 4 gives the results of Eq. (3): for SMEs with events, we regress launch activity on past IP stocks, controlling for a range of firm characteristics, local and sectoral conditions. In **IP**, patents and trademark stocks are depreciated with the standard 15% depreciation rate (Hall and Harhoff, 2012).[18] Trademark stocks are constructed in the same way. We define 'recent' patenting as that occurring in a given five-year period such that n takes the value 0, 1... 5 for patents and for EPO/US/PCT filings in any given year since 2009.[19] For trademarks, n takes the value 0, 1 or 2 based on the available data. As discussed above, following Blundell et al. (1995), we use individual firms' historic patent stocks as proxies for firm-level experience, absorptive capacity and other unobservables.[20] We define 'historic' patenting as that taking place before 2009. Specifically, in **HP,** we include a dummy taking the value 1 if a firm has patented before this date and an average of pre-2009 patenting activity taking the values $p = 0.... p$.[21]

Table 4 gives results for both the product launch dummy (Panel A) and the count of product launches (Panel B), fitting progressively more demanding specifications.[22] Overall, we find the positive link between patents and launches predicted in our framework. For our preferred linear probability model (Panel A, column 4), past patenting increases launch probabilities by 0.7% points the following year. For trademarks,

[16] Additionally, we assume that media interest in any given firm is equal, conditional on sector, year and individual level effects. While we might worry that individual firms could influence media interest through their market position or by buying advertising, this is less plausible in our main sample of single-plant SMEs.

[17] While the UK Innovation Survey asks many questions about firm spending, it does not cover advertising.

[18] This 15% rate is varied in sensitivity tests.

[19] We use filings to these offices as a proxy for invention quality: inventions filed in international domains rather than to a single country are 'worth' more to applicants (Helmers and Rogers, 2010). Alternatives are triadic patent family constructs as an ex-ante measure of quality or patent citations as an ex-post measure.

[20] Many of the cited approaches normally include R&D and advertising expenditures. Our data makes this challenging. We do not observe firm-level advertising spending. The UKIS data contain R&D spend information, and we match this to our panel, but the sample is small and highly selective. Commercial sources such as Orbis have limited direct coverage (7,600 'industrial companies' in the UK with R&D expenditures in their annual accounts); UK SMEs file minimal returns with Companies House, so it is difficult to reconstruct standard proxies. As an alternative, we follow Audretsch et al. (2020) and infer the role of R&D activity by subsetting it across industry bins.

[21] We estimate using OLS because nonlinear estimates converge to OLS results once converted to marginal effects (Angrist and Pischke, 2009). OLS is also more efficient given the very large number of fixed effects in our data. The functionally 'correct' estimation methods are the Zero-inflated Poisson or Zero-inflated Negative Binomial methods. Angrist and Pischke (2009) convincingly show that once the raw coefficients produced by these estimators are converted to marginal effects, the results are essentially identical to those of OLS.

[22] Sample size changes drive results in different columns. To make sure the small differences in the results are driven by sample selection, we run the same regressions, keeping the sample size constant. Results are qualitatively the same, with very minor changes to the coefficients.

Table 5
Linking launch dummies and firm revenue productivity, SMEs with events, 2014-2017.

	Log revenue/worker		Rev/worker growth		High-growth episodes	
	(1)	(2)	(3)	(4)	(5)	(6)
L.new product launch	0.064***		0.000		-0.006	
	(0.019)		(0.007)		(0.005)	
L2.15% depreciated PCT /EPO/US patent count	0.004	0.005	-0.006*	-0.006*	0.002	0.002
	(0.007)	(0.007)	(0.003)	(0.003)	(0.002)	(0.002)
L2.15% depreciated TM count	0.081***	0.081***	0.002	0.002	0.003	0.003
	(0.024)	(0.024)	(0.005)	(0.005)	(0.005)	(0.005)
Ave pre-2009 patenting	0.072	0.070	0.009	0.009	0.029*	0.029*
	(0.057)	(0.058)	(0.022)	(0.022)	(0.016)	(0.016)
Firm patents pre-2009	-0.223*	-0.217*	0.014	0.014	-0.019	-0.020
	(0.116)	(0.116)	(0.043)	(0.043)	(0.029)	(0.029)
Observations	27019	27019	27019	27019	27019	27019
R^2	0.166	0.165	0.010	0.010	0.023	0.023

Source: BSD / CH / GI / Orbis / UKIPO. The dependent variables are log revenue per worker, annual growth in revenue per worker, and a dummy for whether a firm has a high-growth episode, per the OECD definition. L2 is the stock of patents or trademarks two years before. All models fit controls for log turnover and employment, age, firm size dummies, company legal status and structure dummies and an urban TTWA dummy. Controls are lagged one year except age. All models also fit TTWA, 2-digit industry and year dummies. Pre-sample patenting levels effects detailed in the main text. Standard errors are clustered on firms.

*** Denotes a result significant at 1%

**Significant at 5%.

* Significant at 10%. Constant not shown.

conversely, the overall relationship is close to zero and non-significant: this varies when we decompose trademarks by type and scope, as seen below. Consistent with our framework, historical patenting predicts current launch activity: firms with some historical IP are 9.7 percentage points more likely to have a launch in any sample year. The *number* of historical patents is a significant negative predictor, however, consistent with depreciation from bigger stocks of older patents. We see similar patterns for the launch count model (Panel B): 10 additional patents in a given year are linked to over 0.5 extra launch events the following year, while trademarks have no effect. Here, historical patenting has no significant link with the intensive margin of launches.[23]

Then, we focus on other predictions. We re-estimate (3) separately for product, service and specialised trademarks – defined using NICE codes (see Appendix 2 for details). Table B8 shows the results. We find larger IP-launch dummy coefficients for service trademarks than for product trademarks, although we have significantly fewer of the former.[24] In line with Flikkema et al. (2019), for specialised trademarks, the links are larger still and are marginally significant.

We also divide the panel into manufacturing and services subsamples (based on SIC1 classifications) and subgroups based on firm size (following OECD definitions of sole traders, micro, small and medium-sized firms) and age (those under 10 years old, the youngest 25% of the sample, and the remainder). The results are given in Tables B9 and B10. Although we do not observe R&D costs directly, we can infer them from sectoral information (Audretsch et al 2020): as expected, we find stronger links between patents and launches in the services sector, where R&D capital costs are lower than they are in the manufacturing sector. Consistent with Argente et al. (2019), since larger firms file more patents per launch, the coefficients of patents on launch probabilities and launch counts are overall increasing with firm size. There is one exception: micro firms (with 1-9 staff) are more likely than medium-sized firms (with 25-249 staff) to launch with a prior patent, and they generate more launches per patent. Consistent with our framework, we also find that the youngest 25% of firms have higher probabilities of launching (per their past IP) and lower returns to patenting than more established businesses (nonsignificant coefficients of IP on launch count, versus a link that is significant at the 1% level for the oldest 75%).[25]

There are three main caveats to these exercises. First, although our results are robust to varying the lag, the true time decay function between IP and launches is unclear. Second, measurement error on both sides of Eq. (3) affects our estimates. The majority of UK innovations are not protected with formal IP (Hall et al, 2013). Many new products/services involve multiple patents; for instance, the iPhone reportedly has over 100 (Mazzucato, 2013). We also test aggregate links for each firm using many years of patents and trademarks, but only two financial years' worth of reported launches. While measurement error related to patents and trademarks may downward bias the estimates, we can consider the error in product launches to be as good as random conditional on observables.

Third, even conditioning on event exposure may not fully control for unobservables. Per our framework, past IP should have a stronger link with launch activity than with other types of events, such as mergers or staff changes. However, if *any* kind of event exposure is a proxy for underlying knowledge capabilities (Klette and Kortum, 2004), it may affect both IP and launch activity differently across individual firms. We test this in Table B13, restricting to the set of SMEs with event exposure but no launches, then regressing this alternative event count on **IP**. Reassuringly, we find that recent patenting and trademarks are not associated with non-launch event counts. Additionally, we find positive, significant links between firms' historic patenting and other event exposure. This provides further evidence that we are estimating a link

[23] These results survive an extensive set of sensitivity checks and re-running on different samples. In Table B5 we vary the lags for patents and trademarks; in Tables B6-B7 we add controls for past high-growth episodes; add technology field fixed effects; re-specify patents using cumulative patent counts; and use 40% depreciation rates, following Li and Hall (2020).

[24] See Table B1. We have 8,493 trademarks across 5,189 firms, of which 4,744 have only product NICE codes and 1,969 have only service NICE codes. When we re-run this test for all SMEs, we find the same pattern of results, but effect sizes are larger and all significant. Launch activity is unaltered, so this is driven by the larger sample plus more variation in trademarking activity. Results are available on request.

[25] In Table B11, we examine links between past IP activity and launch quality/importance measures, but we find no significant linkages, and these factors are instead linked to revenue per worker (Section 6). In Table B12, we rerun our main analysis on all single-plant SMEs (Panel A) and on all firms with event exposure (Panel B). For all SMEs, we find slightly smaller coefficients, with a significant link to both patenting and trademarking; for all firms with events, we find similar results on the extensive margin, but nonsignificant links to launch counts. This implies that in explaining the IP-launch relationships across the examined population of firms, event exposure may be less salient than the role of firm size, although as explained in Section 2, in our all-firm sample, we risk error in ascribing launches to specific plants and locations.

Table 6
Linking launch counts and firm revenue productivity, SMEs with events, 2014-2017.

	Log revenue/worker		Rev/worker growth		High-growth episodes	
	(1)	(2)	(3)	(4)	(5)	(6)
L.new product launch count	0.017***		0.000		0.001	
	(0.005)		(0.001)		(0.001)	
L2.15% depreciated PCT	0.004	0.005	-0.006*	-0.006*	0.002	0.002
/ EPO / US patent count	(0.007)	(0.007)	(0.003)	(0.003)	(0.002)	(0.002)
L2.15% depreciated TM	0.081***	0.081***	0.002	0.002	0.003	0.003
count	(0.024)	(0.024)	(0.005)	(0.005)	(0.005)	(0.005)
Ave pre-2009 patenting	0.073	0.070	0.009	0.009	0.029*	0.029*
	(0.058)	(0.058)	(0.022)	(0.022)	(0.016)	(0.016)
Firm patents pre-2009	-0.220*	-0.217*	0.014	0.014	-0.020	-0.020
	(0.116)	(0.116)	(0.043)	(0.043)	(0.029)	(0.029)
Observations	27019	27019	27019	27019	27019	27019
R^2	0.167	0.165	0.010	0.010	0.023	0.023

Source: BSD / CH / GI / Orbis / UKIPO. Notes as in Table 5.

between past IP and launches specifically, which is not affected by individual variation in other types of event exposure, and that our individual-level effects capture relevant firm-level heterogeneity.

5.2. Linking IP, launches and firm performance

To explore links between IP/launch activity and firm performance, for SMEs with events, we estimate:

$$Y_{itas} = a + bL_{it-1} + c\mathrm{PATS}_{it-2} + d\mathrm{TM}_{it-2} + \mathbf{HP}e_i + \mathbf{X}f_{it-n} + \mathrm{T}_t + \mathrm{A}_a + \mathrm{S}_s + \mathrm{u}_{itas} \tag{5}$$

This function allows us to study the link between launch activity and subsequent performance changes at the firm level, conditional on previous patenting and trademarking. As before, we then examine subsamples to explore heterogeneity in the innovation-launch-performance relationship. As in existing studies, we specify Y as revenue per worker (Mohnen and Hall, 2013; Klette and Kortum, 2004). We fit Y in both levels (log revenue/worker) and changes (% revenue worker growth/year). Alternately, we specify Y as a dummy indicating whether a firm has at least one ‘high-growth’ revenue growth episode – per the OECD definition – during the sample period. This last specification focuses on the most dynamic firms in the sample. Given the short panel, levels is likely to be more informative than changes or growth episodes.

Controls, lagged launches,[26] patents and trademarks are specified as in the previous subsection: we lag the latter two periods to allow ‘upstream’ IP to influence ‘downstream’ launches. Given the short panel, we use each firm’s pre-2009 patenting activity as a proxy for firm-level heterogeneity.

Tables 5 and 6 give results for the subsample of firms with event exposure. For each outcome, we fit the model with launches (columns 1, 3, and 5) and without launches (columns 2, 4 and 6). We interpret coefficients of *b* as expressing the association between launches and revenue productivity, conditional on media exposure.

As suggested by our framework, Table 5 shows that product launches have a positive, significant relationship with log revenue productivity. Specifically, SMEs with launches have 6.4% more revenue per worker than firms with other types of event exposure (column 1): a 1 standard deviation (0.42) increase in the average launch probability is associated with a 2.7% increase in revenue productivity. While recent patenting has no relationship with revenue per worker, recent trademarking has a positive, significant association, as each additional trademark is linked to an 8.1% increase in revenue productivity (column 1). This result is in opposition to our earlier IP-launch results, suggesting both that patents affect revenue per worker through launches and that patents and trademarks are complements, consistent with De Vries et al (2017). We speculate that the two-year lagged trademark results may partly reflect revenue markups from launches prior to 2014, which we do not observe in our data. We find no link between launch activity and revenue productivity growth or high-growth episodes; patent and trademark links are also weak or non-significant here (columns 3 and 5). Given our short panel, this is, perhaps, not surprising.

Table 6 gives results for launch counts. Each additional launch is linked to a 1.7% increase in revenue productivity (column 1), and this significant at the 1% level. As in Table 5, we find no link with recent patenting, but we see a clear, positive link with recent trademarks, which is larger than that with launches. Removing launches from the model (column 2) reduces the model fit, as before. As before, we find that launch counts do not predict revenue per worker growth or high-growth episodes.

These results are robust to a battery of robustness checks. Tables B14-B16 in the Appendix give results for our three dependent variables. Table B17 reruns the levels result for alternate samples. For all SMEs (Panel A), selection into the events sample drives many of the major associations in our main results. Firms with product launches have 45% higher revenue productivity than those without, and this is significant at the 1% level (column 1). We find similar results regarding the launch count, which has a positive, significant relationship with log revenue productivity. As expected, not controlling for underlying media exposure substantially strengthens the launch count-performance link. Specifically, each additional launch increases revenue productivity by 4.7%, though underlying media exposure is uncontrolled. In contrast, adding in larger, multiplant firms with events (Panel B) only increases the launch-performance link to 8.5%, from 6.4% in our main results, and there is now a nonsignificant link between revenue productivity and the number of launches. Per our framework and Cavenaile and Roldan (2020), this is consistent with larger firms being more likely to have other sources of revenue such as advertising and with the presence of spillover effects from existing products and services.

6. Extensions

We extend our main results in four ways. First, we decompose trademarks into product, service and specialised categories, as before (Table B18). We find that coefficients of launch activity are essentially unchanged; however, consistent with Castaldi et al. (2020), counts of

[26] We can only lag launches by one year. We are aware that any estimated correlations may be industry-specific. We do not run each regression separately by sector, although we control for industry fixed effects (and in some specifications, for industry-by-year fixed effects) to account for average sectoral heterogeneity.

Table 7
Launch quality, launch importance and firm revenue productivity, SMEs with events, 2014-2017.

	Log rev/ worker (1)	Rev/worker growth (2)	High growth episodes (3)
A. L.total launch reports, main topic	0.000***	-0.000**	-0.000
	(0.000)	(0.000)	(0.000)
R^2	*0.166*	*0.010*	*0.023*
B. L.weighted launch reports, main topic	0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
R^2	*0.165*	*0.010*	*0.023*
C. L.firm has important launch, main topic	0.168***	-0.027	-0.007
	(0.052)	(0.018)	(0.013)
R^2	*0.166*	*0.010*	*0.023*
D. L.count of important launches, main topic	0.218***	-0.014	-0.002
	(0.069)	(0.016)	(0.014)
R^2	*0.166*	*0.010*	*0.023*
Observations	27019	27019	27019

Source: BSD/CH/GI/Orbis/UKIPO. Each panel of the table represents a different regression for Eq. (2), with dependent variables specified A-D. For each panel, each cell is a different specification showing the coefficient of *b* in Eq. (2), with standard errors in parentheses and R^2 in italics. All models fit controls for log turnover and employment, age, firm size dummies, company legal status and structure dummies and an urban TTWA dummy. Controls are lagged one year except age. All models also fit TTWA, 2-digit industry and year dummies, plus pre-sample patenting levels effects detailed in the main text. Standard errors are clustered on firms.

*** Denotes a result significant at 1%.

** Significant at 5%.

*Significant at 10%. Constant not shown.

specialised trademarks are more strongly linked to revenue per worker than the simple trademark count is.[27] Second, we split the events subsample to examine the launch-revenue productivity links in manufacturing and services industries. Tables B19 and B20 give results for the linear probability and count models, respectively. In both cases, overall, positive links are driven by firms in the services sector. Service firms also drive the trademark results, consistent with previous studies (Castaldi et al., 2020; Flikkema et al., 2019). For manufacturing firms, recent patenting is linked to lower revenue productivity growth, but historical patenting is correlated with higher revenue per worker growth. Overall, these results are consistent with Audretsch et al. (2020), who suggest that barriers to (reported) innovation are lower for service firms than manufacturing firms.

Third, we investigate the role of firm size and age in explaining our results (Table B21). As before, we group firms into size bins using OECD definitions, and define young firms as the youngest 25% of firms in the events sample. For launch dummies, we first fit our main regression with age and size dummies (column 1), then add size and age group interactions (columns 2 and 3, respectively). In columns 4-6, this is repeated for launch counts. Overall, the results are driven by medium-sized firms, while there is no effect of age. Specifically, for the extensive margin, we find a positive revenue/worker link for small firms, but this is half as strong as that found for medium-sized firms.

Finally, we examine the link between launch quality and revenue productivity. We use the number of media reports per event as a proxy for quality, as detailed in Section 2. We re-estimate Eq. (5) using four *alternative* quality measures in separate regressions: 1) a simple count of the number of reports across each firm's launches per year; 2) firm-year counts weighted by the number of launches; 3) a dummy for whether a firm has an 'important' launch with many mentions; and 4) the number of important launches per firm per year. Table 7 gives the results when we look at counts for the main event topic (using counts across all topics and counts weighted by topics give identical findings). We find very small positive links between the report counts and levels of revenue productivity and very small negative links to revenue productivity growth. We do not find links for weighted report counts. We find large, significant associations between having an important launch and revenue per worker and between having an important launch and the count of important launches. Specifically, SMEs with at least one important launch have approximately 17% higher revenue productivity than other SMEs with media exposure; each additional important launch increases revenue per worker by nearly 22%. This suggests that our main results, which link launch activity to SME revenue productivity, are significantly driven by a small set of high-profile, important product and service launches.

7. Conclusions

A vast field of literature explores the links between innovation and economic performance (Romer, 1990; Aghion and Howitt, 1992; Coad, 2009; Akcigit, 2017). Four streams of empirical work unpack connections at the firm level. An established set of studies uses R&D, patents and innovation surveys; newer analyses use trademarks, text-based measures or product-level data. However, this body of work has two constraints: the informality of much innovative activity (Hall et al, 2014) and practical limits in processing richer text or product-level information. Our paper makes three practical and empirical advances on this literature. First, we develop novel product-level innovation metrics that extend existing studies. Second, we show that our new measures complement existing, formal IP metrics. Third, we find positive links between SME launch activity and revenue per worker. Importantly, we also find that industry, size, IP strategy and launch quality differences moderate our main results.

We highlight three main lessons for existing research, and for practice. First, our results further confirm the overall positive links between firm-level innovation and growth found in many previous studies (Audretsch et al, 2014). Our findings are also consistent with more recent work on trademarks and innovation (Castaldi et al, 2020). Second, however, and consistent with extreme heterogeneity, we show that a subset of high-growth firms can drive overall innovation and growth outcomes (Nightingale and Coad, 2013). Developing policy tools to identify and support such firms is both important and highly challenging. Third, we show the value of monitoring innovations not captured by formal IP or surveys – and the rich potential of text-based sources to achieve this (Gentzkow et al, 2019).

Four limitations of our work may inform future research. First, we explore heterogeneity mainly via subsamples. One could instead use data-driven approaches to identify high-growth businesses, as in Coad et al. (2016a), or richer firm-level information that covers management strategy, as in Grillitsch et al. (2019). Second, it would be valuable to link our data to information on firm advertising, capital intensity and R&D, as in Hall et al. (2013) and Cavenaile and Roldan (forthcoming). Third, we utilize a short timeframe; a longer time series would allow an analysis of macro conditions, as in Spescha and Woerter (2019). Finally, our analysis is not causal. Future work could improve on this issue by exploiting policy evaluation settings or finding viable instruments.

CRediT authorship contribution statement

Max Nathan: Conceptualization, Methodology, Software, Formal analysis, Validation, Visualization, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Anna Rosso:** Conceptualization, Methodology, Software, Formal analysis, Validation, Visualization, Writing – original draft, Writing –

[27] As before, when we rerun this test for all SMEs, we find the same pattern of results but with larger, more robust effect sizes. These results are available on request.

review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Thanks to seminar/conference participants at ZEW Economics of ICT workshop 2017, 4th VPDE-BRICK workshop, SpazioDati, CTCS 2016, Uddevalla 2016, AAG 2016, University of Birmingham, CIRCLE and NIESR, and to Cathy Atkinson, Steve Dempsey, Rob Elliott, Bill Kerr, Francesco Rentocchini and Brandon Stewart for helpful comments. Many thanks to Tom Gatten, Prash Majmudar and Francois Bouet at Growth Intelligence for data. Camilla Andretta, Neeraj Baruah, Aslam Ghumra, and Christoph Stich provided outstanding technical and research assistance. This work includes analysis based on data from the Business Structure Database and the UK Innovation Survey, produced by the Office for National Statistics (ONS) and supplied by the Secure Data Service at the UK Data Archive. The data is Crown copyright and reproduced with the permission of the controller of HMSO and Queen's Printer for Scotland. The use of the ONS statistical data in this work does not imply the endorsement of the ONS or the Secure Data Service at the UK Data Archive in relation to the interpretation or analysis of the data. This work uses research datasets that may not exactly reproduce National Statistics aggregates. All the outputs have been granted final clearance by the staff of the SDS-UKDA. This research builds on projects funded by NESTA and by Google UK. It represents our views not theirs.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.respol.2021.104373.

References

Akcigit, U, 2017. "Economic Growth: The Past, the Present, and the Future." - Commemorative essay in "The Past, Present, and Future of Economics: A Celebration of the 125 Year Anniversary of the JPE and of Chicago Economics," edited by John List and Harald Uhlig. Journal of Political Economy 125 (6), 1736–1747.
Akcigit, U., Kerr, W, 2018. Growth through Heterogenous Innovations. Journal of Political Economy 126 (4), 1347–1443.
Aghion, P., Howitt, P., 1992. A Model of Growth through Creative Destruction. Econometrica 60 (2), 323–351.
Angrist, J, Pischke, J-S, 2009. Mostly Harmless Econometrics. Princeton University Press, Princeton.
Argente, D., Lee, M., Moreira, S., 2018. Innovation and product reallocation in the great recession. Journal of Monetary Economics 93, 1–20.
Argente, D, Baslandze, S, Hanley, D, Moreira, S., 2019. Patents to Products: Innovation, Product Creation and Firm Growth. Mimeo.
Arts, S., Hou, J., Gomez, J.C., 2021. Natural language processing to identify the creation and impact of new technologies in patent text: Code, data, and new measures. Research Policy 50, 104144.
Audretsch, D.B., Belitski, M., 2020. The role of R&D and knowledge spillovers in innovation and productivity. European Economic Review 123, 103391.
Audretsch, D.B., Coad, A., Segarra, A., 2014. Firm growth and innovation. Small Business Economics 43, 743–749.
Audretsch, D.B, Hafenstein, M., Kritikos, A.S., 2020. Microfirms and innovation in the service sector. Small Business Economics 55, 997–1018.
Autor, D., Dorn, D., Hanson, G., Pisano, G., Shu, P., 2020. Foreign Competition and Domestic Innovation: Evidence from US Patents. American Economic Review: Insights 2 (3), 357–374.
Axenbeck, J. and Breithaupt, P. (2019), Web-Based Innovation Indicators - Which Firm Website Characteristics Relate to Firm-Level Innovation Activity?, ZEW Discussion Paper No. 19-063.
Baumann, J., Kritikos, A.S., 2016. The link between R&D, innovation and productivity: Are micro firms different? Research Policy 45, 1263–1274.
Bei, X., 2019. Trademarks, specialized complementary assets, and the external sourcing of innovation. Research Policy 48, 103709.
Bianchini, S., Pellegrino, G., Tamagni, F., 2018. Innovation complementarities and firm growth. Industrial and Corporate Change 27, 657–676.
Block, JH., Fisch, CO., Hahn, A., Sandner, PG., 2015. Why do SMEs file trademarks? Insights from firms in innovative industries. Research Policy 44, 1915–1930.
Blundell, R., Griffith, R., Van Reenen, J., 1995. Dynamic Count Data Models of Technological Innovation. The Economic Journal 105, 333–344.
Bokhari, F.A.S., Mariuzzo, F., Bennato, A.R., 2020. Innovation and growth in the UK pharmaceuticals: the case of product and marketing introductions. Small Business Economics.
Bottazzi, G., Dosi, G., Lippi, M., Pammolli, F., Riccaboni, M., 2001. Innovation and corporate growth in the evolution of the drug industry. International Journal of Industrial Organization 19, 1161–1187.
Castaldi, C., Block, J., Flikkema, M.J., 2020. Editorial: why and when do firms trademark? Bridging perspectives from industrial organisation, innovation and entrepreneurship. Industry and Innovation 27, 1–10.
Castellacci, F., 2008. Technological paradigms, regimes and trajectories: Manufacturing and service industries in a new taxonomy of sectoral patterns of innovation. Research Policy 37, 978–994.
Cavenaile, L. and Roldan, P., (2020) Advertising, Innovation and Economic Growth. *AEJ: Macroeconomics, forthcoming.*
Coad, A., 2009. The Growth of Firms: A Survey of Theories and Empirical Evidence. Edward Elgar, Cheltenham.
Coad, A., Pellegrino, G., Savona, M., 2016a. Barriers to innovation and firm productivity. Economics of Innovation and New Technology 25, 321–334.
Coad, A., Segarra, A., Teruel, M., 2016b. Innovation and firm growth: Does firm age play a role? Research Policy 45, 387–400.
Cockburn, I.M., Lanjouw, J.O., Schankerman, M., 2016. Patents and the Global Diffusion of New Drugs. The American Economic Review 106, 136–164.
Cohen, W.M., Levinthal, D.A., 1990. Absorptive Capacity: A New Perspective on Learning and Innovation. Administrative Science Quarterly 35, 128–152.
Coombs, R., Narandren, P., Richards, A., 1996. A literature-based innovation output indicator. Research Policy 25, 403–413.
Corsino, M., Gabriele, R., 2010. Product innovation and firm growth: evidence from the integrated circuit industry. Industrial and Corporate Change 20, 29–56.
Crass, D., 2020. Which firms use trademarks? Firm-level evidence from Germany on the role of distance, product quality and innovation. Industry and Innovation 27, 730–755.
Crepon, B., Duguet, E., Mairesse, J., 1998. Research, Innovation And Productivity: An Econometric Analysis At The Firm Level. Economics of Innovation and New Technology 7 (2), 115–158.
Cucculelli, M., Ermini, B., 2012. New product introduction and product tenure: What effects on firm growth? Research Policy 41, 808–821.
Davies, N., 2009. Flat Earth News. Vintage, London.
De Vries, G., Pennings, E., Block, J.H., Fisch, C., 2017. Trademark or patent? The effects of market concentration, customer type and venture capital financing on start-ups' initial IP applications. Industry and Innovation 24, 325–345.
Dosi, G., Nelson, R., Winter, S., 2000. The nature and dynamics of organizational capabilities. OUP, Oxford.
Einav, L., Levin, J.D., 2014. The Data Revolution and Economic Analysis. Innovation Policy and the Economy 14.
Fernandes, A.M., Paunov, C., 2015. The Risks of Innovation: Are Innovating Firms Less Likely to Die? The Review of Economics and Statistics 97, 638–653.
Flikkema, M., Castaldi, C., de Man, A.-P., Seip, M., 2019. Trademarks' relatedness to product and service innovation: A branding strategy approach. Research Policy 48, 1340–1353.
Flikkema, M., De Man, A.-P., Castaldi, C., 2014. Are Trademark Counts a Valid Indicator of Innovation? Results of an In-Depth Study of New Benelux Trademarks Filed by SMEs. Industry and Innovation 21, 310–331.
Fosfuri, A., Giarratana, M.S., Luzzi, A., 2008. The Penguin Has Entered the Building: The Commercialization of Open Source Software Products. Organization Science 19, 292–305.
Foster, L., Grim, C., Haltiwanger, J., 2016. Reallocation in the great recession: cleansing or not? Journal of Labor Economics 34 (S1), S293–S331.
Gentzkow, M., Shapiro, J.M., 2010. What Drives Media Slant? Evidence From U.S. Daily Newspapers. Econometrica 78, 35–71.
Gentzkow, M., Kelly, B.T., Taddy, M., 2019. Text as Data. Journal of Economic Literature 57 (3), 535–574.
Gilbert, R.J., Newbery, D.M.G., 1982. Preemptive Patenting and the Persistence of Monopoly. The American Economic Review 72, 514–526.
Gotsch, M., Hipp, C., 2012. Measurement of innovation activities in the knowledge-intensive services industry: a trademark approach. The Service Industries Journal 32, 2167–2184.
Griliches, Z., 1979. Issues in assessing the contribution of R&D to productivity growth. Bell Journal of Economics 10.
Griliches, Z., 1986. Productivity, R&D, and basic research at the firm level in the 1970s. American Economic Review 76.
Grillitsch, M., Schubert, T., Srholec, M., 2019. Knowledge base combinations and firm growth. Research Policy 48, 234–247.
Guarascio, D., Tamagni, F., 2019. Persistence of innovation and patterns of firm growth. Research Policy 48, 1493–1512.
Hall, B.H., 2011. Innovation and Productivity. NBER, Cambridge, MA. *NBER Working Paper 17178*.
Hall, B.H., Harhoff, D., 2012. Recent Research on the Economics of Patents. Annual Review of Economics 4 (1), 541–565.
Hall, B.H., Helmers, C., Rogers, M., Sena, V., 2013. The Importance (or not) of Patents to UK Firms. In: National Bureau of Economic Research Working Paper 19089. Cambridge, MA. NBER.

Hall, B.H., Helmers, C., Rogers, M., Sena, V., 2014. The Choice Between Formal and Informal Intellectual Property: A review. Journal of Economic Literature 52, 375–423.
Helmers, C., Rogers, M., 2010. Innovation and the Survival of New Firms in the UK. Review of Industrial Organization 36, 227–248.
Hoberg, G., Phillips, G., 2016. Text-Based Network Industries and Endogenous Product Differentiation. Journal of Political Economy 124, 1423–1465.
Howell, A., 2015. Indigenous' innovation with heterogeneous risk and new firm survival in a transitioning Chinese economy. Research Policy 44, 1866–1876.
Hsiao, Y.-C., Chen, C.-J., Guo, R.-S., Hu, K.-K., 2017. First-mover strategy, resource capacity alignment, and new product performance: a framework for mediation and moderation effects. R&D Management 47, 75–87.
Jensen, P.H., Webster, E., 2009. Another Look At The Relationship Between Innovation Proxies*. Australian Economic Papers 48, 252–269.
Katila, R., Ahuja, G., 2002. Something Old, Something New: A Longitudinal Study of Search Behavior and New Product Introduction. The Academy of Management Journal 45, 1183–1194.
Kelly, B., Papanikolaou, D., Seru, A., Taddy, M., 2018. Measuring Technological Innovation over the Long Run, NBER Working Paper 252266. NBER, Cambridge, MA.
Kinne, J. and Lenz, D. (2019), Predicting Innovative Firms Using Web Mining and Deep Learning, ZEW Discussion Paper No. 19-001, Mannheim.
Kleinknecht, A., Reijnen, J.O.N., Smits, W, 1993. Collecting literature-based innovation output indicators. The experience in the Netherlands. In: Kleinknecht, A., Bain, D. (Eds.), New Concepts in Innovation Output Measurement. Macmillan, London, pp. 42–84.
Klette, T.J., Kortum, S., 2004. Innovating Firms and Aggregate Innovation. Journal of Political Economy 112, 986–1018.
Kogan, L., Papanikolaou, D., Seru, A., Stoffman, N., 2017. Technological Innovation, Resource Allocation, and Growth. Quarterly Journal of Economics 132, 665–712.
Lafrance, A., 2016. Access, Accountability Reporting and Silicon Valley. Nieman Reports 17. August.
Lenz, D., Winker, P., 2020. Measuring the Diffusion of Innovations with Paragraph Vector Topic Models. PLOS ONE 15 (1), 1–18.
Lentz, R., Mortensen, D, 2008. An Empirical Model of Growth through Product Innovation. Econometrica 76 (6), 1317–1373.
Li, W., Hall, B.H., 2020. Depreciation of Business R&D Capital. Review of Income and Wealth 66 (1), 161–180.
Llerena, P., Millot, V., 2020. Are two better than one? Modelling the complementarity between patents and trademarks across industries. Industry and Innovation 27, 52–79.
Lucas, R., 1988. On the Mechanics of Economic Growth. Journal of Monetary Economics 22, 3–42.
Mazzucato, M., 2013. The Entrepreneurial State: debunking public vs. private sector myths. Anthem Press.
Mairesse, J., Mohnen, P., 2010. Using Innovation Surveys for Econometric Analysis. In: Bronwyn, HH, Nathan, R (Eds.), Handbook of the Economics of Innovation. North-Holland, pp. 1129–1155.
Marris, R., 1963. A Model of the "Managerial" Enterprise. The Quarterly Journal of Economics 77, 185–209.
Mendonca, S., Pereira, T.S., Godinho, M.M., 2004. Trademarks as an indicator of innovation and industrial change. Research Policy 33, 1385–1404.
Mohnen, P., Hall, B.H., 2013. Innovation and Productivity: An Update. Eurasian Business Review 3, 47–65.
Mohnen, P. (2019) R&D, Innovation and Productivity. In: ten Raa T., Greene W. (eds.), *The Palgrave Handbook of Economic Performance Analysis.*
Morris, D.M., 2018. Innovation and productivity among heterogeneous firms. Research Policy 47, 1918–1932.
Motoyama, Y. and Bell-Masterson, J. (2014) Beyond Metropolitan Start-Up Rates. Kansas City: Kauffman Foundation.
Nathan, M., Rosso, A., 2015. Mapping digital businesses with Big Data: some early findings from the UK. Research Policy 44, 1714–1733.
Nelson, R., Winter, S., 1982. An Evolutionary Theory of Economic Change. Harvard University Press, Cambridge, MA.
Nightingale, P., Coad, A., 2013. Muppets and gazelles: political and methodological biases in entrepreneurship research. Industrial and Corporate Change 23, 113–143.
Noel, M., Schankerman, M., 2013. Strategic Patenting and Software Innovation. The Journal of Industrial Economics 61, 481–520.
Office of National Statistics. (2017) Business Structure Database, 1997-2017 SN: 6697. *Secure Data Service Access [computer file].* Colchester: UK Data Archive.
Ortega, M.J.R., García-Villaverde, P.M., 2011. Pioneer orientation and new product performance of the firm: Internal contingency factors. Journal of Management & Organization 17, 474–497.
Roberts, M.E., Stewart, B.M. and Tingley, D. (2016) stm: R Package for Structural Topic Models. http://www.structuraltopicmodel.com.
Rodríguez-Pinto, J., Rodríguez-Escudero, A.I., Gutiérrez-Cillán, J., 2012. How market entry order mediates the influence of firm resources on new product performance. Journal of Engineering and Technology Management 29, 241–264.
Romer, P., 1990. Endogenous Technological Change. Journal of Political Economy 98, 71–102.
Santarelli, E., Piergiovanni, R., 1996. Analysing literature-based innovation output indicators: The Italian experience. Research Policy 25, 689-71.
Saunders, A. and Tambe, P. (2015) Data Assets and Industry Competition: Evidence from 10-K Filings. Available at SSRN: https://ssrn.com/abstract=2537089.
Schautschick, P., Greenhalgh, C., 2016. Empirical studies of trade marks – The existing economic literature. Economics of Innovation and New Technology 25, 358–390.
Schumpeter, J., 1962. The Theory of Economic Development. Springer, Berlin.
Siepel, J., Camerani, R., Masucci, M., Velez Ospina, J., Casadei, P. and Bloom M. (2020) Creative Industries Radar: Mapping the UK's creative clusters and microclusters. Available at https://www.pec.ac.uk/research-reports/creative-radar.
Seip, M., Castaldi, C., Flikkema, M., De Man, A.-P., 2018. The timing of trademark application in innovation processes. Technovation 72-73 34–45.
Spescha, A., Woerter, M., 2019. Innovation and firm growth over the business cycle. Industry and Innovation 26, 321–347.
Stam, E., Wennberg, K., 2009. The roles of R&D in new firm growth. Small Business Economics 33, 77–89.
Syverson, C., 2011. What Determines Productivity? Journal of Economic Literature 49, 326–365.
Taques, F.H., López, M.G., Basso, L.F., Areal, N., 2021. Indicators used to measure service innovation and manufacturing innovation. Journal of Innovation & Knowledge 6, 11–26.
Teece, D.J, Pisano, G., Shuen, A., 1997. Dynamic Capabilities and Strategic Management. Strategic Management Journal 18, 509–533.
Viner, K. (2016) How Technology Disrupted the Truth. *Guardian.*