Patents that Match your Standards: Firm-level Evidence on Competition and Innovation

A. Bergeaud[1], J. Schmidt[2] & R. Zago[2]

[1]HEC
[2]Banque de France

ESCB cluster 2 - Athens

September 2022

[The views expressed herein are ours and should not be attributed to BdF]

Introduction

- **Standard Organizations (SSOs)** are independent institutions empowered to select the **set of technologies** to be **adopted** for specific production/industry

Introduction

- **Standard Organizations (SSOs)** are independent institutions empowered to select the **set of technologies** to be **adopted** for specific production/industry
- SSOs choose these technologies since
  - they are the best among competing ones or they surpass old ones, i.e. they are **frontier technologies**

Introduction

- **Standard Organizations (SSOs)** are independent institutions empowered to select the **set of technologies** to be **adopted** for specific production/industry
- SSOs choose these technologies since
  - they are the best among competing ones or they surpass old ones, i.e. they are **frontier technologies**
  - they lead to **efficiency** in production and/or **welfare gains**

Introduction

- **Standard Organizations (SSOs)** are independent institutions empowered to select the **set of technologies** to be **adopted** for specific production/industry

- SSOs choose these technologies since
  - they are the best among competing ones or they surpass old ones, i.e. they are **frontier technologies**
  - they lead to **efficiency** in production and/or **welfare gains**
  - they (should) ultimately foster further **innovation and growth**

Introduction

- **Standard Organizations (SSOs)** are independent institutions empowered to select the **set of technologies** to be **adopted** for specific production/industry
- SSOs choose these technologies since
  - they are the best among competing ones or they surpass old ones, i.e. they are **frontier technologies**
  - they lead to **efficiency** in production and/or **welfare gains**
  - they (should) ultimately foster further **innovation and growth**
- The final output of a SSO is a **standard**, a document describing in details the frontier technology (e.g. 5G protocol)

Introduction

- **Standard Organizations (SSOs)** are independent institutions empowered to select the **set of technologies** to be **adopted** for specific production/industry
- SSOs choose these technologies since
  - they are the best among competing ones or they surpass old ones, i.e. they are **frontier technologies**
  - they lead to **efficiency** in production and/or **welfare gains**
  - they (should) ultimately foster further **innovation and growth**
- The final output of a SSO is a **standard**, a document describing in details the frontier technology (e.g. 5G protocol)
- Once published, **firms have to comply to the standard** in order to produce a good or deliver a service with specific characteristics

Introduction

- **Standard Organizations (SSOs)** are independent institutions empowered to select the **set of technologies** to be **adopted** for specific production/industry
- SSOs choose these technologies since
  - they are the best among competing ones or they surpass old ones, i.e. they are **frontier technologies**
  - they lead to **efficiency** in production and/or **welfare gains**
  - they (should) ultimately foster further **innovation and growth**
- The final output of a SSO is a **standard**, a document describing in details the frontier technology (e.g. 5G protocol)
- Once published, **firms have to comply to the standard** in order to produce a good or deliver a service with specific characteristics
- The **ability to adapt** to the new technology frontier is **heterogeneous across firms** and depends on **past innovation choices**

Introduction

- **Standard Organizations (SSOs)** are independent institutions empowered to select the **set of technologies** to be **adopted** for specific production/industry
- SSOs choose these technologies since
  - they are the best among competing ones or they surpass old ones, i.e. they are **frontier technologies**
  - they lead to **efficiency** in production and/or **welfare gains**
  - they (should) ultimately foster further **innovation and growth**
- The final output of a SSO is a **standard**, a document describing in details the frontier technology (e.g. 5G protocol)
- Once published, **firms have to comply to the standard** in order to produce a good or deliver a service with specific characteristics
- The **ability to adapt** to the new technology frontier is **heterogeneous across firms** and depends on **past innovation choices**
- This has implications for **market structure**, **competition** and **(future) innovation**

Example: ITU Standard 2015, IMT-2020 5G protocol

- Date of release: 9 June 2015
- Key minimum specifications:
  - bandwidth to be at least 100 MHz
  - bandwidths up to 1 GHz are required for higher frequencies (above 6 GHz)
  - connection density is 1 million devices per square kilometer
  - downlink peak data rate of 20 Gb/s
  - uplink peak data rate of 10 Gb/s
  - target downlink user experienced data rate of 100 Mb/s
  - target uplink user experienced data rate of 50 Mb/s
- Which firms were ready to adapt and comply to the new standard?

This Paper

- **This paper studies the firm-level implications of large-scale technology adoption for competition, innovation and growth**

This Paper

- **This paper studies the firm-level implications of large-scale technology adoption for competition, innovation and growth**
- This research question implies **three measurement problems**:
  1. how to measure the new (adopted) technology frontier
  2. how to measure firms' past innovation choices
  3. how to measure the proximity of the firm technology to the new frontier

This Paper

- **This paper studies the firm-level implications of large-scale technology adoption for competition, innovation and growth**
- This research question implies **three measurement problems**:
  1. how to measure the new (adopted) technology frontier
  2. how to measure firms' past innovation choices
  3. how to measure the proximity of the firm technology to the new frontier
- We use:
  1. data on **published standard documents** ← the new tech. frontier
  2. data on **firm-level patents** ← firms past innovation outcomes
  3. **textual analysis** to measure the **semantic proximity** of the firm' stock of **patents** to the newly published **standards** [!NOVELTY!] ← proximity to the frontier
- We cross this new measure with firm-level data (Compustat, Crisp, IBES)

Findings

1. Patent-level results

- Our **measure of proximity** to the technology frontier is **economically meaningful** as patents semantically closer to the new standards have higher **economic value**, are **more cited** and **renewed**

Findings

1. Patent-level results

- Our **measure of proximity** to the technology frontier is **economically meaningful** as patents semantically closer to the new standards have higher **economic value**, are **more cited** and **renewed**

2. Firm-level results

- We exploit the exogeneity of the timing and content of the standard to show that **frontier firms**
  1. temporarily gain in terms of sales and market-shares
  2. temporarily increase R&D expenditure, in particular if they operate in a competitive market
- These results suggest that **the release of a new standard** can be interpreted as a **temporary competition shock** for frontier firms

Findings

1. Patent-level results

- Our **measure of proximity** to the technology frontier is **economically meaningful** as patents semantically closer to the new standards have higher **economic value**, are **more cited** and **renewed**

2. Firm-level results

- We exploit the exogeneity of the timing and content of the standard to show that **frontier firms**
  1. temporarily gain in terms of sales and market-shares
  2. temporarily increase R&D expenditure, in particular if they operate in a competitive market
- These results suggest that **the release of a new standard** can be interpreted as a **temporary competition shock** for frontier firms

3. Industry-level results

- **Growth is higher** in industries with initial larger gap between leaders and laggards
- Growth is driven by frontier firms in the **short-run**, but by laggards in the **long-run**

Literature

1. The pros and cons of standardization (IO literature):

   Katz and Shapiro 1985; Farrell and Saloner 1985; Leland 1979; Rysman and Simcoe 2008; Lerner and Tirole 2015; Schmalensee 2009; Llanes and Poblete 2014; Spulber 2019; Bekkers et al. 2017; Baron and Pohlmann 2018

2. Innovation literature:

   - Real Effects: Arrow 1962; Romer 1990; Aghion and Howitt, 1992; Grossman and Helpman 1991; Aghion et al. 2005; Aghion and Griffith 2005
   - Financial Effects: Kogan et al. 2017, Pakes 1985, Nicholas 2008; Daniel et al. 1998; Mitchell and Stafford 2000

3. Text mining applied to semantic analysis of patents and standards

   Arts et al. 2018; Kuhn et al. 2020; Kelly et al. 2021; Argente et al. 2020; Bergeaud et al. 2017; Webb et al. 2018; Bloom et al. 2021; Brachtendorf et al. 2020

1. Data

1.1 Semantic Matching of Standards to Patents

- Goal: Calculation of a score between each patent and each standard that measures to what extent their content overlaps

1.1 Semantic Matching of Standards to Patents

- Goal: Calculation of a score between each patent and each standard that measures to what extent their content overlaps

- Main ingredient: text describing the respective patent and standard

1.1 Semantic Matching of Standards to Patents

- Goal: Calculation of a score between each patent and each standard that measures to what extent their content overlaps

- Main ingredient: text describing the respective patent and standard

- Link between patent and standard via keyword matching (common words in both texts conditional on length, repetition, diffusion) Appendix

1.1 Semantic Matching of Standards to Patents

- Goal: Calculation of a score between each patent and each standard that measures to what extent their content overlaps
- Main ingredient: text describing the respective patent and standard
- Link between patent and standard via keyword matching (common words in both texts conditional on length, repetition, diffusion) Appendix
- Challenge: Large data volume and string processing
  - 24 mio patents (Google Patents Data)
  - 800,000 standard documents (Perinorm Data)

1.1 Semantic Matching of Standards to Patents

- Goal: Calculation of a score between each patent and each standard that measures to what extent their content overlaps
- Main ingredient: text describing the respective patent and standard
- Link between patent and standard via keyword matching (common words in both texts conditional on length, repetition, diffusion) Appendix
- Challenge: Large data volume and string processing
  - 24 mio patents (Google Patents Data)
  - 800,000 standard documents (Perinorm Data)
- $Score(p, s)$ measures the proximity of patent p to standard s

	Mean	*SD*	*p1*	*p5*	*p25*	*p50*	*p75*	*p95*	*p99*	*N*
(A) Keyword matching sample										
Score	715.7	1,766.6	141.3	151.8	211.8	315.2	638.7	2,345.7	6,289.0	100M

1.2 Firm-level Data

- CRISP for stock market abnormal returns
- IBES for market expectation on future EPS
- COMPUSTAT for firms fundamentals: sales, market shares (defined at NAICS 3-digit level), R&D, CapX, Tobin's Q, market capitalization, leverage, age, industry markup
- Aggregation of **patent-to-standard scores** at the firm level:

$$Shock_{i,t} = \sum_{s \in S} \sum_{p \in P} Score(p, s)_{i,t}$$

	Mean	*SD*	*p1*	*p5*	*p25*	*p50*	*p75*	*p95*	*p99*	*N*
(A) Standardization Shock										
Shock	0.34	2.02	0.00	0.00	0.00	0.00	0.10	1.27	6.24	24,162
$\mathbb{I}[Shock > 0]$	0.48	0.49	0.00	0.00	0.00	0.00	1.00	1.00	1.00	24,162

2. Results

2.1 Patent-level Results

Is the score economically meaningful?

- For scores based on patents filed 1 year before the publication of each standard, consider this:

$$\log(\text{value}_i) = c + \alpha \log(1 + \text{score}_i) + \beta \log(1 + \text{cit}_i) + \gamma Z_i + \varepsilon_i$$

where *value$_i$* is the economic value of the patent from Kogan et al. 2017

	(1)	(2)	(3)	(4)	(5)	(6)
Score	0.0088***	0.0062***	0.0064***	0.0051***	0.0062***	0.0050***
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Observations	1,165,487	1,165,487	1,165,462	1,165,462	1,163,913	1,163,913
R^2	0.05	0.06	0.10	0.11	0.13	0.14
Time	Yes	Yes	Yes	Yes	No	No
IPC	No	No	Yes	Yes	No	No
Time × IPC	No	No	No	No	Yes	Yes

- Since the mean patent is valued at 16.3 mio USD, patents with a positive score raise their value by 310,000-592,000 USD.

2.2 Firm-level Results

Is the release of new standard exogenous to the firm?
What are its implications for competition and innovation?

- For firms in the COMPUSTAT universe, consider the following:

$$Y_{i,t} = \alpha_i + \phi_{s,t} + \sum_{n=-12}^{N=16} \beta_n Shock_{i,t+n} + X'_{i,t-1}\eta + \varepsilon_{i,t}$$

where

- α_i is the firm FE
- $\phi_{s,t}$ is the Naics 3-digit FE interacted with a time FE
- $X_{i,t-1}$ controls for mkt cap, leverage, Q, dummy for tech. industries, age

- The distributed lead-lag model takes into account that different standards can be released in subsequent periods such that firms can receive multiple shocks throughout time

Exogeneity of the Shock

- Each standard is voted by an ad-hoc committee. If the committee approves it, the first version of the standard is circulated to sub-committees for comments

Exogeneity of the Shock

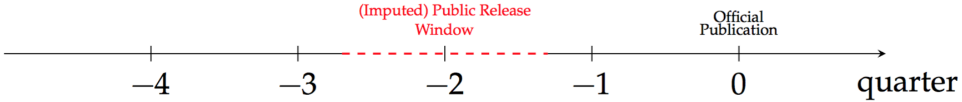
- Each standard is voted by an ad-hoc committee. If the committee approves it, the first version of the standard is circulated to sub-committees for comments
- If no adjustment is required from sub-commitees, the standard is published

Exogeneity of the Shock

- Each standard is voted by an ad-hoc committee. If the committee approves it, the first version of the standard is circulated to sub-committees for comments
- If no adjustment is required from sub-commitees, the standard is published
- Otherwise, the standard is revised and voted again

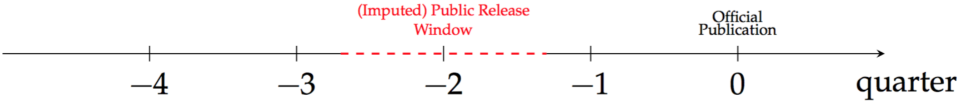
Exogeneity of the Shock

- Each standard is voted by an ad-hoc committee. If the committee approves it, the first version of the standard is circulated to sub-committees for comments
- If no adjustment is required from sub-commitees, the standard is published
- Otherwise, the standard is revised and voted again
- Depending on the administrative procedure, the content of the new standard becomes public between 1.5 and 2.5 quarters before its official publication



Exogeneity of the Shock

- Each standard is voted by an ad-hoc committee. If the committee approves it, the first version of the standard is circulated to sub-committees for comments
- If no adjustment is required from sub-commitees, the standard is published
- Otherwise, the standard is revised and voted again
- Depending on the administrative procedure, the content of the new standard becomes public between 1.5 and 2.5 quarters before its official publication

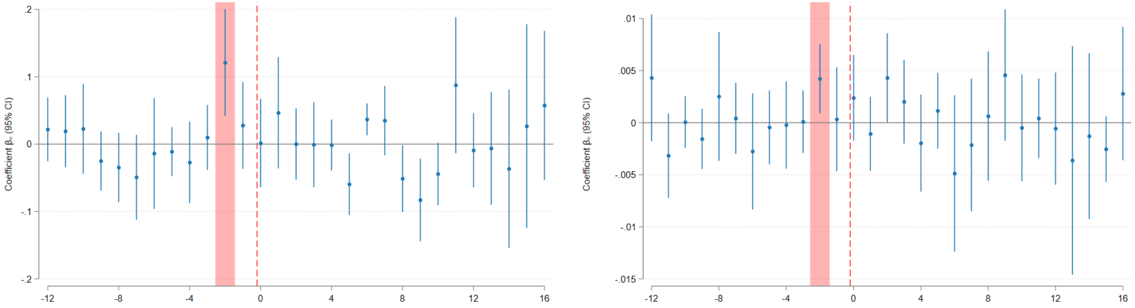


- **Market Efficiency Hypothesis:** if the timing and content of the standard are unexpected, financial markets should react at the moment of the information release

Exogeneity of the Shock

- Consider the following dependent variables:
  - the abnormal return over a NAICS3-industry portfolio, i.e. $ar_{i,t}^{NAICS3}$
  - the change in the 1-year EPS forecast from professional agencies, i.e. $\Delta\mathbb{E}[EPS_{i,t+4}] = \mathbb{E}[EPS_{i,t+4}|\mathbb{I}_t] - \mathbb{E}[EPS_{i,t+4}|\mathbb{I}_{t-1}]$

Figure: Standardization Shock and Financial Markets' Reaction

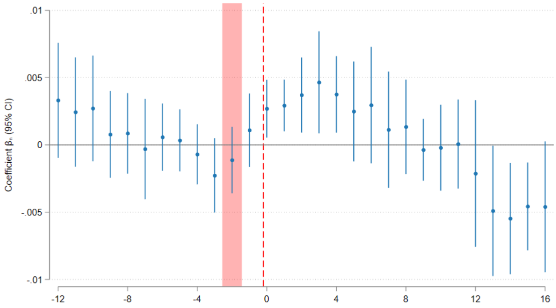


(a) ar^{NAICS3} (b) $\Delta\mathbb{E}(EPS)$

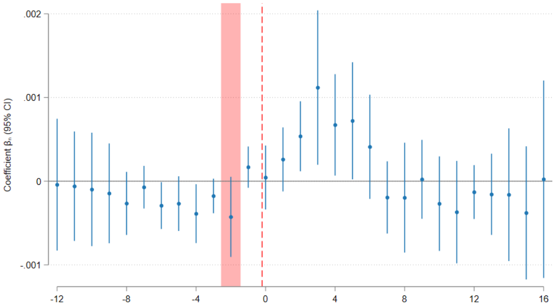
Sales & Market Shares

- Consider sales and market-shares (defined at the Naics 3-digit level) as dependent variables

Figure: Standardization Shock, Sales and Market Shares



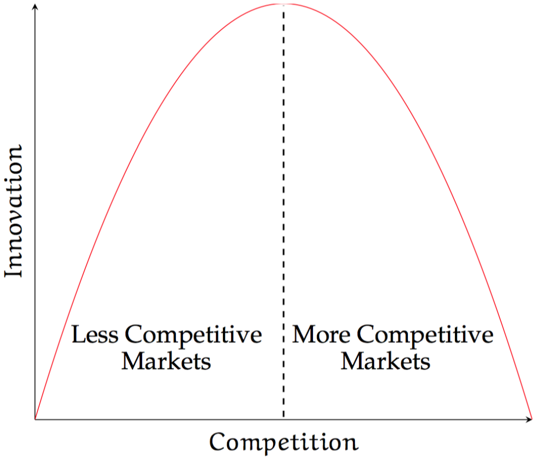
(a) Sales



(b) Market Share

Is it a competition shock?

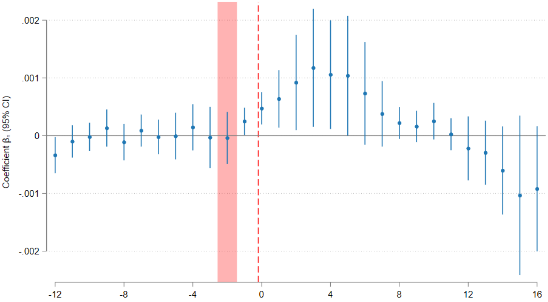
Figure: Aghion et al. 2005, Innovation vs. Competition



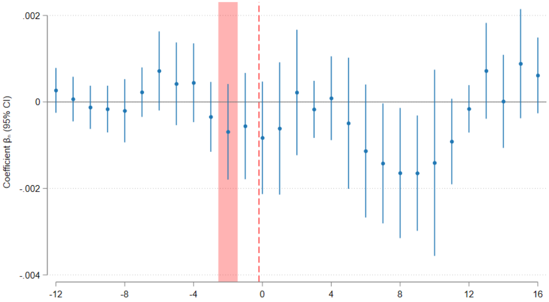
Is it a competition shock?

- Split Naics 3-digit industries in competitive and non-competitive according to their historical markup (De Loecker et al. 2020) and look at results for R&D

Figure: Standardization Shock and R&D



(a) R&D (Competitive Ind)



(b) R&D (Non-Competitive Ind)

Robustness Checks

- These results are robust
  - to different clustering methods:
    - firm-level clustering
    - industry-time double clustering
  - to sample selection:
    - results are not driven by frontier firms
    - results are not driven by listed firms only
  - when considering the intensive/extensive margin of the variable $Shock_{i,t}$
  - to different definitions of the variable $Shock_{i,t}$, i.e. they do not depend on our text analysis methodology

2.3 Industry-level Results

What are the sectoral implications of technology adoption in terms of growth? Is growth driven by leaders or followers in the long-run?

- Define as leaders (followers) those with $Shock_{i,t} > 0$ ($Shock_{i,t} = 0$)
- Aggregate variables at industry level and use same econometric model

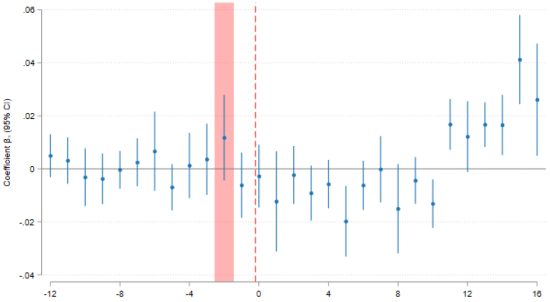
	Industry	*Leaders*	*Followers*
Mean Sectoral Growth Rate (%)	1.64	1.04	0.60
(1yr-Cumulative) Change in Growth (pp)	-0.03	0.08	-0.11
(4yr-Cumulative) Change in Growth (pp)	0.11	0.02	0.09

- Consistently with the step-by-step model of Aghion et al. (1997,2001)
  - **Short-run effect:** leaders try to **escape competition** by increasing R&D
  - **Long-run/Catching-up effect:** followers adapt, increase research effort and surpass leaders

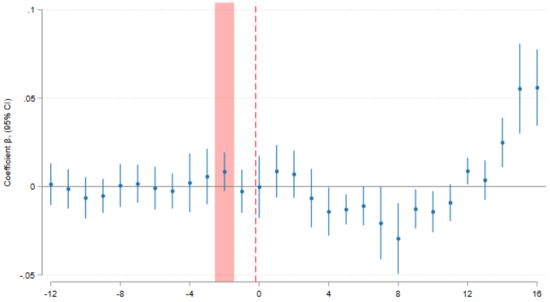
Followers explain sectoral growth in the long-run

- Consider the followers' share of industry aggregate sales and R&D expenditure

Figure: Sectoral Sales and R&D of Followers



(a) Market Share (Followers)



(b) R&D (Followers)

Conclusion

- This paper studies the implications of large-scale technology adoption at the firm level
- To do so, we introduce a new measure of proximity of firms to the technological frontier
- This measure is based on text analysis of patents and standard documents and expresses how much the content of the patents of a firm overlaps with the content of a new standard
- We find that when a new standard is released, firms closer to the new frontier
  - gain in sales and market shares
  - increase R&D expenditure, if they operate in a competitive industry
- We can interpret the release of a new standard as a competition shock in favor of frontier firms
- Effects are only temporary. Laggards catch up through higher innovation efforts. This explains sectoral growth in the long-run

Thank you!

Appendix

Linking patents to standards

- **Basic idea: look at keywords in the abstract of the patent and in the standard document**

1. Establish a list of all patent keywords with respective list of patent ids

keyword	patent id
air condit	41
aluminium	41,106
beef	2
electric	7,84,41,91,102
screw	3,9,17,38

2. Establish a list of all standard keywords with respective list of standard ids

keyword	standard id
air condit	B,M
aluminium	B
diameter	C
rubber	A,E,G,H,R,Z

Linking patents to standards

4. Find common keywords

keyword	patent id	standard id	idf
aluminium	41,106	B	0.2
air condit	41,65	B,M	0.5

5. For each keyword: Cartesian product of respective patent and standard ids

keyword	match	idf
aluminium	41-B	0.2
aluminium	106-B	0.2
air condit	41-B	0.5
air condit	41-M	0.5

6. Calculate score by adding up combinations of ids:

match	idf
41-B	0.7
41-M	0.5
106-B	0.2

Building up the score

Back

- Evaluation of importance of each keyword k via calculation of inverse document frequency (IDF) across all patents (N):

$$IDF\,(k) = 1 + \log\left(\frac{1+N}{1+N(k)}\right)$$

- For every patent $i \in \mathcal{P}$ and standard $j \in \mathcal{J}$, we extract
  - the set of k-grams $\mathcal{B}(i)$ for patent i
  - the set of k-grams $\mathcal{A}(j)$ for standard j
  - $n(k,i)$, i.e. the number of times k-gram k appears in $\mathcal{B}(i)$
  - $s(k)$, the length of k-gram k
- The final score from the matching of patent i to standard j is defined as

$$Score(i,j) = \sum_{k \in \mathcal{A}(j)} \sqrt{\left(\frac{n(k,i)}{|\mathcal{B}(i)|}\right)^{s(k)}} IDF(k)\left(|\mathcal{A}(j) \cap \mathcal{B}(i)|\right)$$