

Radical innovation and firm productivity growth

Francesco Venturini

University of Urbino (IT)
NIESR & ESCoE (UK)

ESCoE Conference
London
19th May 2026

Acknowledgement

Coauthors

- Andrea Fronzetti Colladon (University of Roma Tre)
- Barbara Guardabascio (University of Perugia)
- Ludovica Segneri (University of Urbino Carlo Bo)
- Alessandro Sterlacchini (Polytechnic University of Marche)

*This paper has been developed within **INNODATA** project and has been funded by the European Union - Next Generation EU, Mission 4 Component 1, PRIN 2022YT5HAL, CUP I53D23003490008*

Background

- Firms innovate to gain competitive advantage through new products, processes, and business models (Schumpeter 1934, Nelson and Winter 1982)
- Radical innovation vs incremental innovation
- Radical innovations linked to:
 - Technological paradigms and industrial revolutions (Dosi 1982, Mokyr 1990, Lipsey et al. 2005).
 - Market disruption and technological discontinuities (Carpenter et al. 1981)
 - Industry emergence and structural change (Henderson 1993)

Characteristics of radical innovations

- **New knowledge** or **novel recombinations** of existing knowledge across domains and organizational boundaries (Fleming 2001, Schoenmakers and Duysters 2010)
- **Novel** as they depart from the existing technological base, and **impactful** as they shape how subsequent innovators design and produce new technologies (Dahlin and Behrens 2005)
- **Breakthrough** when are both novel and impactful (Kelly et al. 2017)

Challenges

- **Novelty** reflects the distance of innovation from existing technologies (ex ante) $\Rightarrow$ Firm positioning in the current state of knowledge
- **Impact** reflects the closeness of innovation to future technologies (ex post) $\Rightarrow$ Tracking tech transformations and new major trajectories

Aims of the paper

- Measure **novelties** at firm level analyzing the patent texts (28,500 Italian firms between 2012 and 2020)
- Estimate the **causal link** between **novelties** and **productivity growth**
- Examine the **mechanisms** of transmission and **heterogeneity** of effects across firms

Preview of the results

- Short-run decline in the rate of output per worker growth after the launch of novelty
- This effect reverses over time
- Medium-run acceleration in productivity growth ⇒
Net effect +8%
- Effects driven by employment adjustment
- Stronger impact in services and for small firms
- Results robust to reverse causality and selectivity
- Patent texts contain more information on novelties than citations

Baseline model: DiD

$$\Delta \ln p_{it} = \alpha_i + \beta N_{it} + \gamma X_{it} + \mu_{jt} + \epsilon_{it}, \quad (1)$$

- $\Delta \ln p_{it}$ = labour productivity growth (log change)
- N = Novelty (0/1 – Absorbing treatment)
- X = Controls
- α_i = firm FE
- μ_{jt} = sector-by-year FE
- FE-OLS regression with HAC standard errors

$\beta (\times 100)$: % productivity growth premium for firms with novelties (ATT)

Rob checks: Causality

1) Local Projections DiD (Dube et al. 2025)

$$\Delta \ln p_{it+h} - \Delta \ln p_{it} = \sum_{h=-2}^2 \beta_h N_{it} + \gamma X_{it} + \sum_{p=-1}^1 \zeta_p N_{it-p} + \mu_{jt} + \varepsilon_{it}.$$

2) IV-DiD regression (2 stages)

1st
$$N_{it} = \alpha + \text{prd}_{rt} + \text{copub}_{rt} + \vartheta Z_{it} + \mu_j + \mu_{jt} + \mu_{rt} + \varepsilon_{it}$$

2nd
$$\Delta \ln p_{it} = \alpha_i + \sum_{h=-2}^2 \beta_h \hat{N}_{it} + \gamma X_{it} + \mu_{jt} + \epsilon_{it}$$

IVs: (i) share of public R&D expenses on regional value added; (ii) number of scientific publications co-authored by public and business-sector researchers in region r

Rob checks: Selectivity

Matching procedure

- Match companies with novelties (treated) to those without novelties (control) having similar characteristics
- Propensity Score Matching ([PSM](#)) with 10 nearest neighbours (and alternative criteria)
- Replicate the [LP-DiD](#) regression with matched sample

Data sources

- **Moody's ORBIS IP**: balance sheets for companies with at least one patent application from Patstat
- 28,611 companies with 30,000 patents
- 1,510 **treated** firms (vs 27,101 control firms) when novelty is measured with text indicator

Dependent variable

Main

- **Labour productivity growth**: annual rate of change in value added (or sales) per employee

Mechanism

- **Output growth**: annual rate of change in value added
- **Employment growth**: annual rate of change in employees

Explanatory variables

- **Novelty**: Text- or citation-based binary indicator
- **Controls**: dynamic adjustment, structural characteristics, tech capabilities, financial conditions, proximity effects

Controls

- **Convergence**: Lagged labour productivity
- **Size and Age**: Employment and years from incorporation
- **Capital intensity**: Capital stock per worker (p.w.)
- **Financial conditions**: cash-flow ratio (cash equivalents / output) or short-term debt ratio (current liabilities / total assets)
- **Tech capabilities**: Patents p.w. Intangibles share on total assets
- **Competition effects**: # of patenting firms in the same region-sector pair
- **Knowledge spillovers**: # of patent applications filed in the same region-sector pair

Semantic novelty measure (Text-based)

- 29,332 applications with priority date from 2012 to 2020 with English-language title and abstract
- N compares each **patent p** with all those filed in the previous 5-year rolling window
- N is maximum distance of **patent p** from prior patents:

$$N = 1 - \max_{q \in \{P\}} \rho_{p,q} \quad (2)$$

$\rho_{p,q} \in (0, 1)$ is **cosine similarity** between patents

- Patents classified as **novel** if N falling above the **90th percentile** of the score distribution

Backward-citations measure

$$N = 1 - \frac{BWC_p}{\max(BWC_t)} \quad (3)$$

- BWC_p is the number of backward cites made by patent p
- t represents its filing year
- $\max(BWC_t)$ is the maximum number of backward citations in the same year (cohort)

Descriptive Statistics: Text-based novelty

	Treated (1,510 firms)			Control (27,101 firms)		
	Mean	Median	SD	Mean	Median	SD
Labour productivity growth (%)	0.28	-0.005	78.98	-0.71	-0.09	88.36
Employment growth (%)	0.18	0	41.35	0.14	0	46.36
Value Added growth (%)	0.46	-0.11	65.85	0.07	-0.14	74.12
Controls						
Capital per worker (mln)	32.9	9.2	278.4	65.4	6.4	1411.2
Patents per worker	0.18	0.03	1.34	0.39	0.06	14.65
# innovators by sector-region	10.6	0	27.63	0.9	0	8.8
# patents by sector-region	213.7	56	326.8	142.9	27	262.8
Cash flow ratio	2.3	0.1	79.7	23.9	0.1	3975.0
Intangible ratio	0.24	0.11	0.29	0.18	0.05	0.27
Debt Ratio	0.68	0.71	0.21	0.71	0.76	0.23

Baseline regression

	(1)	(2)	(3)	(4)	(5)	(6)
	Text					Backward cites
Novelty (post $\times$ treat.)	-0.061*** (0.019)	-0.329*** (0.031)	-0.333*** (0.036)	-0.294*** (0.043)	-0.433*** (0.043)	-0.288*** (0.045)
Novelty t-1		0.350*** (0.030)	0.304*** (0.039)	0.294*** (0.044)	0.486*** (0.047)	0.262*** (0.046)
Novelty t-2			0.081*** (0.023)	0.022 (0.031)	0.032 (0.030)	0.062* (0.036)
Novelty t-3				0.055** (0.024)	0.058** (0.023)	0.007 (0.027)
Novelty (cumulative)				0.076** (0.034)	0.141*** (0.030)	0.044 (0.039)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Labour productivity	Value added p.w.	Value added p.w.	Value added p.w.	Value added p.w.	Sales p.w.	Value added p.w.
Obs.	189,802	189,802	166,431	142,493	179,304	142,493
R-squared	0.56	0.56	0.571	0.585	0.586	0.585

Measurement issues

	(1)	(2)	(3)	(4)	(5)	(6)
Novelty (contemp.)	-0.294*** (0.043)	-0.311** (0.128)	-0.320*** (0.031)	-0.418*** (0.056)	-0.201*** (0.060)	-0.250* (0.138)
Novelty (cumulative)	0.076** (0.033)	0.331*** (0.115)	0.059** (0.025)	0.095** (0.043)	0.049 (0.091)	0.081 (0.163)
Indicator	Max novelty	Max novelty	Max novelty	Mean novelty	Impact	Breakth- rough
Threshold (percentiles)	90th	99th	75th	90th	90th	90th
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	142,493	142,493	142,493	142,493	44,198	44,198
R-squared	0.585	0.585	0.586	0.585	0.755	0.755

Novelty: $N_p = 1 - \max \rho_{p,q}$ q are past patents

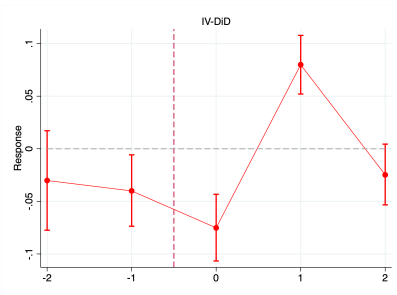
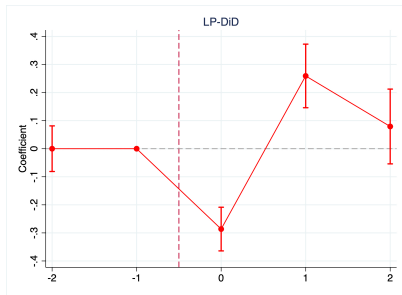
Impact $I_p = 1 - \min \rho_{p,q}$ q are future patents

Breakthrough $B_p = I_p / (1 - N_p)$

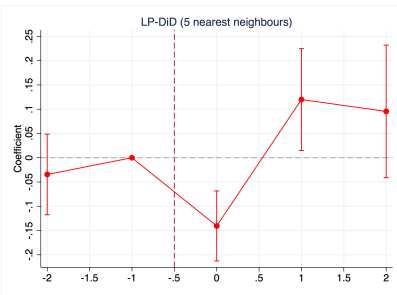
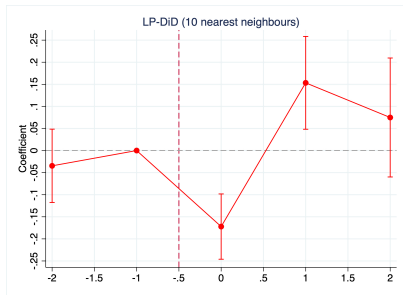
Drivers of novelties: Probit regression

	(1)	(2)	(3)
Regional public R&D	1.120* (0.08)		-0.109 (0.105)
Regional scientific co-authorships		0.471*** (0.033)	0.0879** (0.044)
# of patents by region-sector (log)	-0.0001 (0.0001)	-0.0001 (0.0001)	-0.0002 (0.0001)
# of innovators by region-sector (log)	0.002*** (0.0001)	0.002*** (0.0001)	0.002*** (0.0001)
Controls	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Year-by-Industry FE	Yes	Yes	Yes
Year-by-Region FE	No	Yes	Yes
Observations	102,529	102,529	105,362
Pseudo R-squared	0.153	0.153	0.148
Log-likelihood	-3854	-3854	-3893

Event analysis (Full sample)



Event analysis (Matched sample)



Impact on output and employment growth

	(1)	(2)	(3)	(4)	(5)	(6)
	Text			Backward cites		
	Output growth		Employment growth	Output growth		Employment growth
	Value added	Sales		Value added	Sales	
Novelty (contemp.)	0.029 (0.036)	-0.005 (0.025)	0.436*** (0.031)	-0.009 (0.034)	-0.019 (0.028)	0.383*** (0.033)
Novelty (cumulative)	-0.0213 (0.040)	0.042* (0.025)	-0.093*** (0.021)	-0.058* (0.041)	0.031 (0.028)	-0.092*** (0.027)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	146,901	179,708	196,896	146,901	179,708	196,896
R-squared	0.509	0.511	0.722	0.509	0.511	0.721

Channels by sectors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Manufacturing				Services			
	Labour productivity growth	Output growth		Employment growth	Labour productivity growth	Output growth		Employment growth
		Value added	Sales			Value added	Sales	
Novelty (contemp.)	-0.232*** (0.043)	0.032 (0.037)	0.035 (0.025)	0.334*** (0.031)	-0.519*** (0.106)	-0.017 (0.090)	-0.114* (0.059)	0.746*** (0.066)
Novelty (cumulative)	0.037 (0.034)	-0.008 (0.032)	0.027 (0.024)	-0.068*** (0.016)	0.167** (0.082)	-0.06 (0.075)	0.079 (0.048)	-0.148*** (0.039)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	82,956	85,400	102,582	109,889	59,537	61,501	77,126	87,007
R-squared	0.588	0.530	0.506	0.708	0.586	0.496	0.515	0.736

Channels by firm size

	(1)	(2)	(4)	(5)	(6)	(8)	(9)	(10)	(12)
	Small firms			Medium firms			Large firms		
	Labour productivity growth	Output growth Value added	Employment growth	Labour productivity growth	Output growth Value added	Employment growth	Labour productivity growth	Output growth Value added	Employment growth
Novelty (contemp.)	-0.435*** (0.066)	0.059 (0.058)	0.652*** (0.044)	-0.073* (0.044)	0.001 (0.040)	0.068*** (0.017)	-0.001 (0.077)	0.002 (0.073)	0.009 (0.023)
Novelty (cumulative)	0.170*** (0.051)	0.040 (0.049)	-0.124*** (0.024)	-0.040 (0.040)	-0.103** (0.038)	-0.022 (0.015)	-0.027 (0.071)	-0.023 (0.071)	0.013 (0.023)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	105,185	108,539	148,959	26,578	27,024	33,744	9,165	9,675	12,387
R-squared	0.587	0.507	0.762	0.585	0.531	0.806	0.643	0.492	0.890

Conclusions

- The productivity effects of **novelties** are **non-linear** (in Italy): short-term slowdown is followed by sustained medium-term gains.
- The net effect is roughly **8%** higher rate in labour productivity growth.
- Adjustment dynamics mainly by **employment reconfiguration**
- Non-linear effect confined to **services** and **small firms**

■ Policy and managerial implications:

- Support complementary assets: finance, skills, and R&D networks.
- Invest in training, digital capabilities, and organizational redesign.
- Strengthen regional innovation ecosystems and university–industry links.

■ Methodological contribution:

- NLP-based novelty measures provide **timely, predictive signals** of future productivity gains.

- **Takeaway:** Integrating text mining with economic analysis improves our understanding of how radical innovation translates into long-term productivity growth.