New Stylised facts on Firm and Polarization Dynamics

Andrea Caggese* Madalena Gaspar† Carolina Villegas-Sanchez‡

Robert Wojciechowski§

(Preliminary, please do not cite nor distribute)

October 15, 2026

Abstract

We analyse administrative data covering the near-universe of French non-financial firms over 1997-2019. Aggregate productivity dispersion (polarization) rises only modestly, masking substantial industry heterogeneity: most manufacturing industries experience an increase in polarization, whereas in services rising and falling industries are roughly balanced, in contrast with the services-led pattern documented for other European economies. Our main contribution is to decompose the contribution to polarization from different firm dynamic factors. Contrary to traditional narratives that an increase in polarization is driven by technology developments that make it more difficult to challenge the dominance of top firms, we find that industries with larger increases in polarization did not experience greater persistence of top firms or larger declines in entry, and that polarization is driven by compositional turnover at the upper-tail boundary of the productivity distribution rather than by further advances by industry leaders. This is particularly evident in the services sector, where the productivity premium of firms that remain at the top narrows relative to the median, while turnover at the upper-tail boundary pushes polarization in the opposite direction. In particular, top dropouts tend to leave from the lower part of the upper tail, leaving behind a more selected group of top firms. A firm dynamics model with entry, exit, and incremental and radical innovation best matches the observed decomposition when polarization rises because of shocks to entry conditions, namely higher entry costs and higher volatility of entrants' productivity. Overall, the evidence is difficult to square with a "superstar firms" view and suggests that the link between technological change, firm dynamics and polarization is more complex than previously understood.

*UPF, BSE, and CREI, andrea.caggese@upf.edu

†UPF and BSE, madalena.gaspar@upf.edu

‡ESADE, URL, ECB and CEPR, carolina.villegas@esade.edu

§UAB and BSE, robert.wojciechowski@bse.eu

1 Introduction

The overwhelming majority of governments and companies around the world do not engage in significant research and development (R&D) expenditures (see Keller, 2004; Comin and Mestieri, 2014). Hence, technology diffusion, the rate at which companies adopt new frontier technologies, is key for aggregate productivity growth. Several authors argue that the rate of technology diffusion slowed down in recent decades, and that such phenomenon is important in explaining why there has been a decline in competition, an increase in concentration within industries, and an increase in markups and profitability of dominant firms (e.g. Akcigit and Ates, 2021).

One stylised fact that has been mentioned in support of the "declining technology diffusion" hypothesis is the widening productivity gap, emerged in the last 40 years, between the top-productivity firms and the rest (e.g. Andrews et al., 2019). We define this gap as "polarization".

It has been argued that technology diffusion and polarization are negatively related because of technological developments that make it easier for industry leaders to innovate and maintain their advantage over other firms, or that make it harder for laggards to imitate them (e.g. Akcigit and Ates, 2021 and Olmstead-Rumsey, 2022). These technological developments include, for example, network effects that arise with the increasing importance of big data. According to this conjecture, slower technology diffusion has increased the productivity gap between leaders and laggard firms in most industries, and has caused the increase in polarization.

However, the relation between polarization and firm dynamics is in general complex, and an increase in polarization is both consistent with increases and decreases in technology diffusion, depending on the type of firms at the top of the productivity distribution. On the one hand, it might be that polarization increases because already top productive firms become even larger or more productive. On the other hand, it might be that polarization

increases because of innovations introduced by entrants, or because of larger productivity improvements of laggard firms, if such improvements push these new or laggard firms to the top of the productivity distribution and beyond former leaders.

Therefore, to understand the relation between polarization dynamics and technology diffusion, it is essential to carefully analyse not only polarization changes across industries, but also to quantify the contribution from different firms (entering, exiting, laggard and top firms) to these changes. This paper provides such novel analysis.

Our approach is inspired by the productivity literature, which decomposes the contributions of entering, exiting, and incumbent firms to aggregate productivity changes.[1] Our innovation is to apply this type of decomposition to other moments of the productivity distribution; in particular, to the weighted average of the distance from the productivity of the median firm to firms on the upper tail of the productivity distribution. With respect to simply analysing the dynamics of given percentiles of the distribution, our approach has the advantage of being able to identify the mapping from the dynamics of different types of firms, such as entering firms, exiting firms, firms moving up and down the productivity distribution, and "continuing top firms" (defined as those that remain at the top over consecutive years), to the overall changes in polarization.

We consider a dataset that comprises the near-universe of French firms, built by combining the FICUS (1997–2007) and FARE (2008–2019) administrative panels.[2] The combined panel covers all French non-financial corporations subject to corporate tax over 1997–2019, providing very broad coverage of the size distribution – including small and young firms that are essential for our entry, exit and switching components. Our main results are as follows:

i) Aggregate polarization in France increases modestly over the 1997–2019 period,

[1] In particular, we build on the decomposition suggested by Melitz and Polanec (2015).

[2] We use the merging procedure developed by Burstein, Carvalho and Grassi (2025) and Di Giovanni, Levchenko and Mejean (2024). We are grateful to these authors for sharing their code with us.

but this masks substantial industry-level heterogeneity. The majority of manufacturing industries (77% for top 50% polarization, 68% for top 25% polarization) experienced an increase in polarization. The positive trend is less prevalent in services industries (56% and 51% respectively), where rising and falling industries are roughly balanced. This contrasts with the services-led pattern documented for some other European economies.[3]

ii) Industries with greater increases in polarization did not exhibit a corresponding rise in top persistence (the probability that top-productivity firms remain in this position in the following year). Nor did they exhibit a more pronounced decline in entrants' sales share – if anything the cross-industry relationship runs in the opposite direction, with the industries that became more polarized being those in which entry held up best.

iii) When we decompose the change in polarization into its determinants, we find that it is driven by compositional turnover at the upper-tail boundary rather than by leaders pulling ahead: the relative productivity of firms that remain at the top, compared to the median, does not widen over the period and in the services sector actually narrows marginally. Where polarization rose, it did so because the upper-tail group was selectively refreshed – top dropouts, which tend to leave from the lower part of the upper-tail group, left behind a more selected residual upper-tail group – and not because continuing top firms accelerated.

iv) We develop a firm dynamics model with entry, exit, and incremental and radical innovation decisions that allow laggard firms to catch up with the leaders. In the model, polarization is affected by all these margins. For a preliminary version of the model, we calibrate and simulate the transition dynamics after different types of shocks that generate increases in polarization, and compare the model's decomposition to the one observed in the data. This analysis shows that the model best matches the data when polarization increases because of shocks that change entry conditions, namely higher entry costs and higher volatility of productivity of new entrants.

[3]In particular, see Andrews et al., 2019. Our results are however not directly comparable to theirs because they pool together several countries and use a dataset (ORBIS), which does not cover the entire population of firms. Conversely we focus on the entire distribution of firms in one specific country (France).

Overall, this evidence sheds new light on the factors that drive observed polarization. In particular, findings (ii), (iii) and (iv) are difficult to square with a "superstar firms" view in which polarization is the symptom of frontier firms entrenching themselves and pulling further ahead. We consider this as evidence that the relation between technological change, firm dynamics and polarization is complex; that it is not easily explained by frameworks in which a single technological development towards less competition explains productivity trends jointly with concentrations and markup trends; and that it requires further theoretical and empirical analysis.

1.1 Related literature

This paper is related to the literature that analyses long-run trends in entry, exit, markups and productivity, and relates them to changes in the type and rate of diffusion of new technologies. In particular, Akcigit and Ates (2021) show that a model with slower technology diffusion is able to jointly explain the increase in polarization, and other trends related to firm dynamics (the decline in entry rates), to markups, and to industry concentration (the higher and more concentrated markups and market shares, and the decline in labor share).

Other potential explanations for the rise in polarization include limits to competition of a technological nature, such as an increase in entry costs. For example, the rise of digitally intensive sectors and of big data, which allow larger firms to exploit data-network effects, and hence make it more difficult for new firms to enter and compete (e.g. see Calligaris, Criscuolo and Marcolin (2018), De Ridder (forthcoming), among others). Relatedly, Olmstead-Rumsey (2022) finds that in recent years younger firms appear to be less able to do high-impact innovations to catch up with leaders.

The contribution of our paper is to provide a better understanding on how entry-exit and growth dynamics, and polarization trends, interact with each other. Our main motivation is that it is important to better understand the dynamics of firms at the top of the productivity distribution, because while most theoretical models identify as top firms

those that are simultaneously larger, older, more productive and with larger markups, this correlation is far from clear in the data. In this respect, our work relates to Gutiérrez and Philippon (2020), who notice that most of the literature on frontier/superstar firms measure these firms with productivity. However, when superstar firms are identified as the largest firms, there is no longer evidence of "rise of superstars", meaning that the productivity advantage of the largest dominant firms relative to other firms has not increased over time.

2 Methodology

This section develops the analytical framework used in the empirical analysis. A prominent hypothesis in the recent literature holds that rising productivity polarization reflects declining business dynamism: entry rates fall, top firms become more entrenched, and the gap between top and laggard firms widens as competitive pressure weakens. Testing this hypothesis requires three complementary objects. First, a measure of polarization that tracks the upper-tail productivity premium relative to the median. Second, standard measures of business dynamism (top persistence and young-firm activity) that the existing literature uses to evaluate the entrenchment hypothesis. Third, a decomposition that goes beyond these aggregate measures and maps the dynamics of specific firm groups directly into changes in polarization, distinguishing whether rising polarization is driven by continuing top firms pulling ahead, by compositional turnover at the upper-tail boundary, or by changes in the median itself. We develop each in turn.

2.1 Polarization

We define polarization as the productivity gap between upper-tail firms and the median firm within an industry. Let $\varphi_{i,j,t}$ denote the productivity of firm i operating in industry j in year t, and let $\varphi_{j,t}^{50}$ denote the median productivity in industry j and year t. For any percentile threshold $p \geq 50$, our polarization measure is the weighted-average distance from the median

among firms with productivity above the p-th percentile:

$$\Phi^{p}_{j,t} = \sum_{i \in \mathcal{I}^{p}_{j,t}} s_{i,j,t} \hat{\varphi}_{i,j,t}, \tag{1}$$

where $\hat{\varphi}_{i,j,t} \equiv \varphi_{i,j,t} - \varphi^{50}_{j,t}$ is firm i's productivity premium relative to the contemporaneous industry median. The upper-tail group is $\mathcal{I}^{p}_{j,t} = \{i : \varphi_{i,j,t} > \varphi^{p}_{j,t}\}$, where $\varphi^{p}_{j,t}$ is the p-th percentile of the unweighted firm-level productivity distribution in industry j and year t.[4] The weight of firm i within the upper-tail group is

$$s_{i,j,t} = \frac{\omega_{i,j,t}}{\sum_{n \in \mathcal{I}^{p}_{j,t}} \omega_{n,j,t}}, \tag{2}$$

where $\omega_{i,j,t}$ is firm i's sales.

Since productivity is in logs, $\exp(\Phi^{p}_{j,t}) - 1$ gives the proportional increase in the median firm's productivity needed to reach the sales-weighted upper-tail geometric mean; for small values, $\Phi^{p}_{j,t}$ approximates this percentage gap. The existing literature often captures upper-tail productivity dispersion using percentile distances, such as the 90th-to-50th or 95th-to-50th gap (see, e.g., Andrews et al., 2019; Berlingieri et al., 2020). Our measure is closely related but has the key advantage of being decomposable into the contributions of distinct firm groups, as described in Section 2.3. Sales weights make the measure sensitive both to productivity dispersion and to the economic importance of upper-tail firms.

Our baseline analysis focuses on two thresholds: $p = 50$ and $p = 75$, which we call *top 50% polarization* and *top 25% polarization*, respectively.

Productivity measurement. Our baseline productivity measure $\varphi_{i,j,t}$ is log revenue total factor productivity (log TFPR), estimated following the control function approach of Ackerberg, Caves and Frazer (2015) within the production-function framework of De Loecker

[4] Percentile thresholds are computed from the unweighted firm-level productivity distribution within industry-year cells. We then compute weighted averages among firms above the relevant cutoff, so that the upper-tail group is a set of firms rather than a sales-weighted quantile.

and Warzynski (2012).[5] We use TFPR as our baseline because the paper studies productivity differences as they appear in firm revenues and market outcomes. TFPR captures not only technical efficiency but also price, demand, and markup components, all of which may affect the measured economic importance of upper-tail firms. We therefore interpret our baseline measure as revenue productivity rather than as pure physical efficiency.

2.2 Top Persistence, Young-Firm Activity, and Entry

A natural starting point for evaluating the entrenchment hypothesis is to ask whether the industries with the largest increases in polarization are also those where top firms became harder to displace and where new firms lost market share. The existing literature uses two classes of measures for this purpose: top persistence and entry rates. We construct both.

Top persistence is the fraction of upper-quartile firms in year $t-1$ that remain in the upper-quartile in year t. We use the upper-quartile as our baseline threshold, corresponding to top 25% polarization in equation (1); analogous measures can be constructed for other thresholds.

$$P_{j,t} = \frac{\sum_i \mathbb{I}\big(i \in \mathcal{I}^{75}_{j,t} \cap \mathcal{I}^{75}_{j,t-1}\big)}{\sum_i \mathbb{I}\big(i \in \mathcal{I}^{75}_{j,t-1}\big)}. \tag{3}$$

Firms that were in the upper-quartile in $t-1$ but exit the market before t are counted as non-persistent, so that persistence can decline either because top firms fall below the threshold or because they exit. A rise in $P_{j,t}$ indicates lower turnover among top-productivity firms, consistent with (though not sufficient evidence of) greater entrenchment or weaker competitive pressure from below.

Young-firm activity is captured by two measures. The *numeric young-firm share* $YS_{j,t} = N^Y_{j,t}/N_{j,t}$ is the fraction of active firms that are young (age < 10). Because this measure can over-represent economically small firms, we focus on the *sales share of young*

[5] Estimation details are in Appendix A3.1.

firms:

$$SYS_{j,t} = \frac{\sum_{i \in \mathcal{Y}_{j,t}} \omega_{i,j,t}}{\sum_{n \in \mathcal{M}_{j,t}} \omega_{n,j,t}}, \tag{4}$$

where $\mathcal{Y}_{j,t}$ is the set of young firms and $\mathcal{M}_{j,t}$ the set of all active firms in industry j and year t. Where the data permit identification of true firm births, we also construct analogous entry measures based on firms active in t but not in $t-1$. In the empirical analysis, we also report a stricter market-entry measure based on first sample appearance, age, and firm size, which is designed to identify genuine new firms rather than all young incumbents; details are provided in Section 4.1.

These measures are informative benchmarks but have a fundamental limitation for our purpose: they describe aggregate patterns at the industry level but cannot identify whether it is continuing top firms pulling away, compositional turnover at the upper-tail boundary, or changes in the median that drive changes in the polarization measure. The decomposition in Section 2.3 is designed precisely to make that distinction.

2.3 Decomposing Polarization

The main analytical advantage of equation (1) is that it can be decomposed into contributions from distinct groups of firms. Our decomposition adapts the dynamic productivity decomposition of Melitz and Polanec (2015) from aggregate productivity — a first moment of the distribution — to an upper-tail distributional moment. As in Melitz and Polanec (2015), the decomposition is an exact accounting identity: it attributes changes in polarization to mutually exclusive groups of firms but does not identify causal mechanisms. What it adds, relative to tracking percentiles over time or measuring aggregate persistence and entry rates, is the ability to map the dynamics of specific firm groups (i.e., firms entering or exiting the upper-tail, firms that persist at the top, firms that fall out) directly into changes in the aggregate polarization measure.

To simplify notation, we suppress the percentile superscript p throughout this subsec-

tion.

Backward-looking classification. Firms in the upper-tail group $\mathcal{I}_{j,t}$ are partitioned into three mutually exclusive groups according to their status in year $t-1$:

1. *Continuing top firms* (C): firms in the upper-tail group in both $t-1$ and t.

2. *Young entrants to the top* (YE): young firms (age < 10) that enter the upper-tail group in t, meaning that they belong to the upper-tail group in t but did not belong to it in $t-1$.

3. *Mature entrants to the top* (ME): firms aged 10 or older that enter the upper-tail group in t, meaning that they belong to the upper-tail group in t but did not belong to it in $t-1$.

Entry into the upper-tail group should not be confused with entry into the market. Young and mature entrants to the top are firms newly crossing the upper-tail productivity threshold; they are not necessarily recently established firms. The young/mature distinction is motivated by the paper's focus on technology diffusion and market dynamism. Young firms may be more likely to embody new technologies or business models, while mature firms newly entering the upper-tail capture upward movement among established incumbents, which may reflect catch-up, reallocation, or changes in relative productivity.

Using this classification, polarization at time t is

$$\Phi_{j,t} = \underbrace{S^{C}_{j,t}\Phi^{C}_{j,t}}_{\text{continuing top}} + \underbrace{S^{YE}_{j,t}\Phi^{YE}_{j,t}}_{\text{young entrants to the top}} + \underbrace{S^{ME}_{j,t}\Phi^{ME}_{j,t}}_{\text{mature entrants to the top}}, \tag{5}$$

where $S^{K}_{j,t}$ is the sales share of group K within the upper-tail group and $\Phi^{K}_{j,t}$ is the sales-weighted average productivity premium within group $K \in \{C, YE, ME\}$.

Forward-looking classification. Firms in the upper-tail group at t are partitioned into three mutually exclusive groups according to their status in $t+1$:

1. *Continuing forward top firms* (CF): firms in the upper-tail group in both t and $t+1$.

2. *Exiting top firms* (X): firms in the upper-tail group at t that cease operations before $t+1$.

3. *Top dropouts* (TD): firms in the upper-tail group at t that remain active in $t+1$ but fall below the upper-tail threshold.

Top dropouts do not necessarily move to the bottom of the productivity distribution; they simply fall below the upper-tail threshold. Polarization at time t can be written as

$$\Phi_{j,t} = \underbrace{S_{j,t}^{CF}\Phi_{j,t}^{CF}}_{\text{continuing forward top}} + \underbrace{S_{j,t}^{X}\Phi_{j,t}^{X}}_{\text{exiting top}} + \underbrace{S_{j,t}^{TD}\Phi_{j,t}^{TD}}_{\text{top dropouts}}, \tag{6}$$

where $S_{j,t}^{K}$ and $\Phi_{j,t}^{K}$ are defined analogously for $K \in \{CF, X, TD\}$.

Changes in polarization. Combining both classifications yields a decomposition of yearly changes in polarization. Because group shares sum to one, $S_{j,t}^{C} = 1 - S_{j,t}^{YE} - S_{j,t}^{ME}$ and $S_{j,t-1}^{CF} = 1 - S_{j,t-1}^{X} - S_{j,t-1}^{TD}$. We can therefore rewrite $\Phi_{j,t}$ around the continuing-top group at t and $\Phi_{j,t-1}$ around the continuing-forward-top group at $t-1$. Noting that the set C at time t and the set CF at time $t-1$ contain exactly the same firms — those in the upper-tail group in both $t-1$ and t — subtracting the two expressions gives:

$$\begin{aligned}
\Delta\Phi_{j,t} &\equiv \Phi_{j,t} - \Phi_{j,t-1} \\
&= \underbrace{\left(\Phi_{j,t}^{C} - \Phi_{j,t-1}^{CF}\right)}_{\text{continuing top}} + \underbrace{S_{j,t}^{YE}\left(\Phi_{j,t}^{YE} - \Phi_{j,t}^{C}\right)}_{\text{young entrants to the top}} \\
&\quad + \underbrace{S_{j,t}^{ME}\left(\Phi_{j,t}^{ME} - \Phi_{j,t}^{C}\right)}_{\text{mature entrants to the top}} + \underbrace{S_{j,t-1}^{X}\left(\Phi_{j,t-1}^{CF} - \Phi_{j,t-1}^{X}\right)}_{\text{exiting top}} \\
&\quad + \underbrace{S_{j,t-1}^{TD}\left(\Phi_{j,t-1}^{CF} - \Phi_{j,t-1}^{TD}\right)}_{\text{top dropouts}}.
\end{aligned} \tag{7}$$

The first term captures changes in the upper-tail premium among firms that remain

continuously in the upper-tail group across both years. The remaining four terms capture compositional turnover at the boundary of the upper-tail group. Young and mature entrants raise polarization when their productivity premium exceeds that of continuing top firms. Exiting top firms and top dropouts raise polarization when their productivity premium in $t-1$ falls short of that of firms that remain in the upper-tail; in practice, their departure leaves behind a more selected group.

This structure highlights that rising polarization need not reflect a fixed set of top firms pulling further ahead in absolute terms. It can also arise because the median falls, because market shares within the upper-tail shift toward firms with higher productivity premia, or because of selective compositional turnover at the upper-tail boundary, with no acceleration among continuing top firms whatsoever. The decomposition makes these forces separately observable, whereas aggregate persistence and entry measures cannot.

Multi-period averages. For longer horizons, we report the arithmetic average of each yearly component in equation (7). This avoids compounding year-to-year changes and allows the multi-period contribution of each margin to be interpreted as its average annual contribution over the window.

Together, equation (1), equation (7), and the measures in Section 2.2 provide a framework for studying industry polarization. The polarization index tracks the upper-tail productivity premium relative to the median; the persistence and young-firm measures describe aggregate patterns of business dynamism; and the decomposition maps the dynamics of continuing top firms, young and mature entrants to the top, exiting top firms, and top dropouts directly into changes in polarization. Section 3 describes the data before turning to the main results in Section 4.

3 Data

Our empirical analysis draws on the near-universe of French non-financial corporations over the period 1997–2019. The underlying administrative records are obtained from two complementary data sets maintained by the Institut National de la Statistique et des Études Économiques (INSEE): the FICUS panel, which covers 1994–2007, and its successor the FARE panel, which covers 2008–2019. Both data sets originate from annual tax statements submitted to the French tax administration, and contain firm-level income statements, balance sheets, and demographic information for each corporation that files a corporate tax return. Because the obligation to file is universal, coverage of the size distribution is very broad and includes small and young firms that are essential for computing our entry, exit, and top-switcher components. Following the methodology described in Burstein, Carvalho and Grassi (2025), we harmonize the industry classification across the two panels. Firms are assigned to industries according to the Nomenclature des Activités Françaises (NAF Rev. 2) five-digit classification, which maps closely to the four-digit NACE Rev. 2 classification. For firms observed in both sub-panels, the post-2008 NAF Rev. 2 code is applied to all years; for firms observed in only one sub-panel, the code most frequently associated with the observed code in the crosswalk sample is assigned. This procedure yields a consistent industry classification over the entire 1997–2019 window.

Starting from the harmonized FICUS-FARE panel, we drop firm-year observations with missing core identifiers or missing values of any primary flow or stock variable; we remove, at the firm level, any firm that ever records a negative value for sales, materials, the wage bill, or capital; and we require all four variables to be strictly positive in every observation retained. We deflate nominal monetary values using industry-specific indices from the EU KLEMS database. Output (sales) and intermediate inputs are deflated using gross-output and intermediate-input price indices at the industry level; the wage bill uses the aggregate consumer price deflator; and capital uses the Gross Fixed Capital Formation

(GFCF) deflator.[6] Real value added is then computed as deflated sales minus deflated materials.

The final dataset comprises 25.2 million firm-year observations for 3.4 million unique firms over 1997–2019, spanning 65 two-digit NACE industries – 22 in manufacturing (NACE2 codes 10–33) and 43 in services. The average firm contributes 7.4 years of observations, which reflects the coexistence of long-lived incumbents and a substantial flow of entrants and exiters. Services account for 88.1 percent of firm-year observations, broadly in line with the structural weight of the sector in the French economy.

Table 1 reports summary statistics overall and separately for manufacturing and services. The median French firm employs two workers, has been active for eleven years, and records deflated annual sales of approximately €239,000 and value added of approximately €100,000. The size distribution is strongly right-skewed: mean employment is 9.5 workers and the 90th-to-50th percentile ratio of deflated sales exceeds eight. Manufacturing firms are systematically larger – the median employs three workers versus two in services, with correspondingly higher deflated sales (€315,000 vs. €231,000) and capital (€97,000 vs. €45,000). TFPR dispersion is somewhat higher in services (standard deviation 0.29) than in manufacturing (0.22), consistent with the greater product-market heterogeneity of the service sector. Two data-quality breaks are worth noting: the 2007–2008 FICUS-to-FARE migration introduces a gentle discontinuity in panel coverage and the 2017–2018 introduction of INSEE capital imputations for BNC firms causes a spike in recorded capital and a corresponding artificial increase in entry counts.

[6] Where the GFCF deflator is missing or non-positive for a given observation, the consumer price deflator is used as a fallback.

Variable	N	Mean	S.D.	P10	P25	Median	P75	P90
Panel A: All firms								
Employment (workers)	25,219,947	9.53	319.9	0	0.75	2	5	13
Age (years)	25,219,947	14.14	12.97	3	5	11	19	29
Sales (€ 000s)	25,219,947	1,925	55,584	52.8	103.5	239.0	668.4	1,971.8
Value added (€ 000s)	25,219,947	618.0	21,259	20.1	42.0	100.3	267.9	710.3
Capital (€ 000s)	25,219,947	902.3	75,009	7.2	19.7	49.2	135.2	399.0
log TFPR	25,219,947	1.091	0.284	0.765	0.878	1.057	1.281	1.466
Panel B: Manufacturing (NACE2 10–33)								
Employment (workers)	2,990,780	19.66	215.7	0	1	3	9	28.8
Age (years)	2,990,780	16.97	14.70	3	6	13	23	35
Sales (€ 000s)	2,990,780	4,653	101,473	63.3	129.6	315.0	1,069.9	4,181.0
Value added (€ 000s)	2,990,780	1,260	19,410	23.8	53.4	137.8	420.3	1,358.2
Capital (€ 000s)	2,990,780	2,205	60,933	15.5	37.6	97.4	299.8	1,285.2
log TFPR	2,990,780	1.047	0.218	0.804	0.910	1.031	1.163	1.314
Panel C: Services (NACE2 35–96)								
Employment (workers)	22,229,167	8.16	331.4	0	0.5	2	5	11
Age (years)	22,229,167	13.76	12.67	3	5	10	19	29
Sales (€ 000s)	22,229,167	1,558	46,030	51.6	100.7	231.0	631.6	1,784.2
Value added (€ 000s)	22,229,167	531.7	21,494	19.8	40.9	96.5	253.1	647.1
Capital (€ 000s)	22,229,167	727.1	76,704	6.6	18.5	45.0	120.8	337.3
log TFPR	22,229,167	1.097	0.292	0.761	0.873	1.064	1.298	1.480

Table 1: Summary statistics. Monetary variables are deflated using EU KLEMS industry-specific price indices (gross output, intermediate input, and investment deflators) and expressed in thousands of euros at constant prices. Log TFPR is our baseline productivity measure, i.e. the log revenue total factor productivity estimated following the control function approach of Ackerberg, Caves and Frazer (2015) within the production-function framework of De Loecker and Warzynski (2012).

4 Polarization Trends and Firm Dynamics

This section documents the evolution of productivity polarization in France and relates it to standard indicators of business dynamism. The objective is to assess whether the patterns in the French data are consistent with a simple view in which declining entry and greater entrenchment of top firms generate rising productivity polarization.

We proceed in four steps. First, we document the evolution of entry and persistence at the top of the productivity distribution. Second, we describe aggregate and industry-level polarization trends. Third, we examine whether industries with larger increases in polarization are also those with larger increases in top persistence or larger declines in entry.

Finally, we use the decomposition introduced in Section 2.3 to quantify the firm groups that account for changes in polarization.

All polarization measures are computed at the NACE Rev. 2 2-digit industry level and then aggregated to the sector level, distinguishing manufacturing and services.[7] We require each industry to contain at least 100 active firms in every year of the sample, which removes a small number of very thin industries.[8] Sector aggregates use fixed industry weights equal to each industry's average value-added share over 1997–2019, so that sector-level changes reflect within-industry changes rather than changes in the composition of industries.

4.1 Entry and Persistence at the Top

A prominent explanation for rising productivity dispersion is that top firms become increasingly insulated from competitive pressure because new firms enter less often, or because firms below the frontier find it harder to catch up. Such an environment should be associated with weaker entry, greater persistence among top firms, and rising polarization. We begin by documenting the first two moments in the French data.

Entry. We define a market entrant in year t as a firm that makes its first sample appearance in year t, is at most five years old according to its incorporation date, and employs at most 100 workers.[9] This definition is distinct from the "young entrants to the top" in the decomposition below, which are young firms that newly cross the upper-tail productivity threshold.

Figure 1 shows the evolution of entrants' sales share between 1998 and 2019, measured as entrants' total sales divided by total sales in the relevant sector. The left panel reports

[7] Appendix A3.3 lists the industries included in the analysis. We exclude financial services.

[8] This restriction ensures that percentile-based measures are not driven by extremely small industry-year cells. The baseline sample contains 22 manufacturing industries and 43 services industries.

[9] The five-year cutoff is used only for the market-entry measure, where the objective is to identify genuine new firms and avoid counting older firms that enter the administrative panel for other reasons. This differs from the ten-year cutoff used in the decomposition to distinguish young from mature firms entering the upper-tail productivity group. The latter split is intended to describe firm dynamics within the productivity distribution rather than to identify firm birth.

the aggregate series pooling manufacturing and services, while the right panel reports the two sectors separately. The market presence of entrants declines over the sample period. The decline is visible in both sectors, but is stronger in services than in manufacturing. In services, entrants' sales share falls from levels above 2 percent in the late 1990s and early 2000s to roughly 1.3–1.5 percent by the mid-2010s. In manufacturing, the decline is more gradual, from roughly 0.8–1.2 percent to around 0.6 percent.

Several short-run movements coincide with changes in administrative coverage. The increase around 2008 coincides with the transition from FICUS to FARE, while the spike around 2018 reflects the integration of many very small BNC firms into FARE. Because these firms are economically small, the spike is much less pronounced in the sales-weighted series than in entrant counts. Overall, Figure 1 points to a persistent decline in the market importance of new firms.

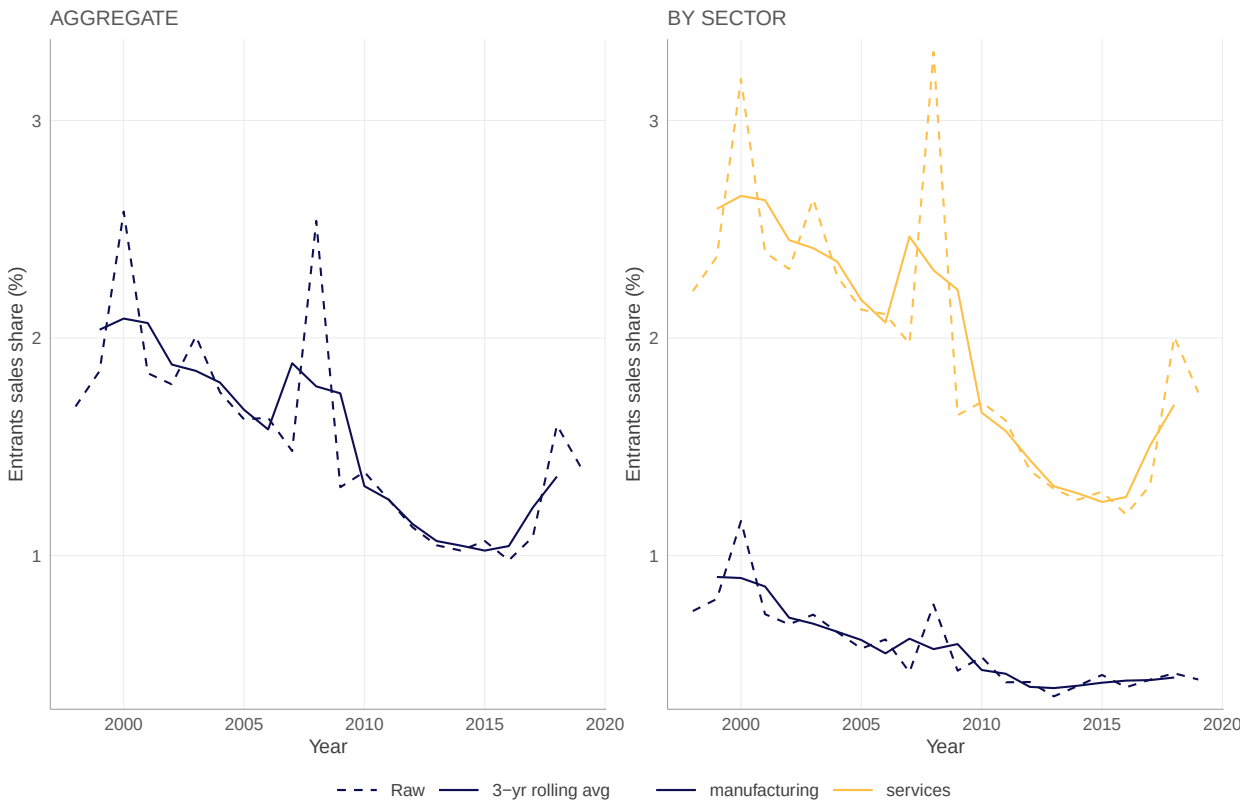


Figure 1: Entrants' sales share in manufacturing and services, 1998–2019

Persistence at the top. We next measure top persistence as defined in Section 2.2: the share of firms in the upper-quartile in year $t-1$ that remain in the upper-quartile in year t. The distribution and the transition are defined within 2-digit industries and then aggregated across industries using time-invariant value-added weights. Firms that were in the upper-quartile in $t-1$ but exit before t are counted as non-persistent.

Figure 2 plots top persistence between 1998 and 2019, pooling manufacturing and services in the left panel and separating the two sectors in the right panel. Top persistence rises mildly over the sample period for both manufacturing and services, with services top persistence remaining below that of manufacturing over the sample period. The raw series exhibits considerable year-to-year variation, but the overall upward trend is clear with top persistence rising by approximately five percentage points in both sectors over the sample period. Increasing top persistence is consistent with a broad decline in business dynamism and greater stability among top firms.



Figure 2: Top persistence (upper quartile) in services and manufacturing, 1998 – 2019

Together, Figures 1 and 2 show that France exhibits two patterns emphasized in the recent literature on declining business dynamism: the market presence of new firms falls, and the identity of top-productivity firms becomes mildly more persistent. We now ask whether these patterns are accompanied by a broad increase in productivity polarization.

4.2 Polarization Trends

We first focus on $\Phi_{j,t}^{50}$, the sales-weighted average gap between log TFPR of firms in the upper-half of the productivity distribution and median productivity within the same industry-year cell. We then repeat the analysis for $\Phi_{j,t}^{75}$, the corresponding top 25% polarization. As discussed in Section 2.1, an increase in these measures indicates a rise in the upper-tail productivity premium relative to the median. Such an increase can reflect movements in the upper-tail, movements in the median, reallocation of sales within the upper-tail, or changes in the composition of firms in the upper-tail group.

Figure 3 reports aggregate top 50% polarization for manufacturing and services. Average polarization is higher in services than in manufacturing throughout the sample period, but the services series is essentially flat. Manufacturing displays a modest increase, rising by slightly more than two percentage points between 1997 and 2019. Figure 4 reports the corresponding top 25% polarization. The qualitative pattern is similar: there is some evidence of an increase in manufacturing, but little evidence of a broad aggregate rise in services.



Figure 3: Top 50% polarization in manufacturing and services, 1998–2019



Figure 4: Top 25% polarization in manufacturing and services, 1998–2019

The modest sector-level movements mask substantial industry-level heterogeneity. For each 2-digit industry, we estimate the linear trend

$$\Phi^{p}_{j,t} = \alpha_j + \beta_j t + \varepsilon_{j,t}, \tag{8}$$

separately for $p = 50$ and $p = 75$.[10] Table 2 reports the number of industries with a positive estimated trend. In manufacturing, 17 out of 22 industries, or 77 percent, display a positive trend in top 50% polarization; for top 25% polarization, the corresponding share is 68 percent. In services, positive and negative trends are more balanced: 56 percent of industries display a positive trend in top 50% polarization and 51 percent in top 25% polarization.

The degree of heterogeneity is economically meaningful. For top 50% polarization,

[10]Before estimating equation (8), we trim the top and bottom 1 percent of yearly industry-level polarization changes to avoid contamination by extreme annual movements. The conclusions are unchanged without this trimming.

Table 2: Industry polarization trends, France 1997–2019

		Industries	$\hat{\beta}_j > 0$	Share with $\hat{\beta}_j > 0$ (%)
Top 50% polarization	Manufacturing	22	17	77%
	Services	43	24	56%
Top 25% polarization	Manufacturing	22	15	68%
	Services	43	22	51%

Notes: The table reports the number of 2-digit industries with a positive linear trend in polarization. An industry has a positive trend if $\hat{\beta}_j > 0$ in equation (8). Trend regressions are estimated after trimming the top and bottom 1 percent of yearly industry-level polarization changes.

the interquartile range of cumulative changes spans from -2.4 to $+4.8$ percentage points across services industries and from $+1.0$ to $+4.6$ percentage points across manufacturing industries. Thus, the aggregate averages conceal industries with substantially different polarization paths. Appendix Figures 10 and 11 show the underlying industry-level time series.

Thus, although France exhibits declining entry and mildly rising persistence at the top, it does not exhibit a broad aggregate increase in productivity polarization. The increase is concentrated in manufacturing; in services, where the decline in entrants' sales share is particularly visible, aggregate polarization is essentially flat.

Stylised Fact 1. *In France, productivity polarization does not rise broadly across all sectors. It increases modestly in manufacturing, where most industries display positive polarization trends, but is essentially flat in services, where rising and falling industry trends are roughly balanced.*

4.3 Cross-Industry Evidence

The previous subsections show that France experienced lower entry and mildly higher top persistence, but only limited aggregate growth in polarization. We now ask whether the industries with larger increases in polarization are precisely those with larger increases in persistence or larger declines in entry.

For this cross-industry analysis, we report both polarization thresholds, $\Phi^{50}_{j,t}$ and $\Phi^{75}_{j,t}$. Top 50% polarization, $\Phi^{50}_{j,t}$, captures a broader notion of upper-tail polarization and is less sensitive to movements among a small number of firms at the very top. Top 25% polarization, $\Phi^{75}_{j,t}$, is the closest counterpart to our top persistence statistic, since both are based on the upper-quartile of the productivity distribution. Showing both measures allows us to assess whether the conclusions depend on the chosen upper-tail threshold.

Figure 5 plots the cross-industry relation between changes in top persistence and changes in polarization. The left panel uses top 25% polarization, while the right panel uses top 50% polarization. If rising polarization were driven by increasingly entrenched top firms, one would expect industries with larger increases in top persistence to also display larger increases in polarization. The data do not show such a pattern. In neither panel is there evidence of a positive relation between changes in top persistence and changes in polarization; if anything, the fitted relationship is mildly negative in both cases.

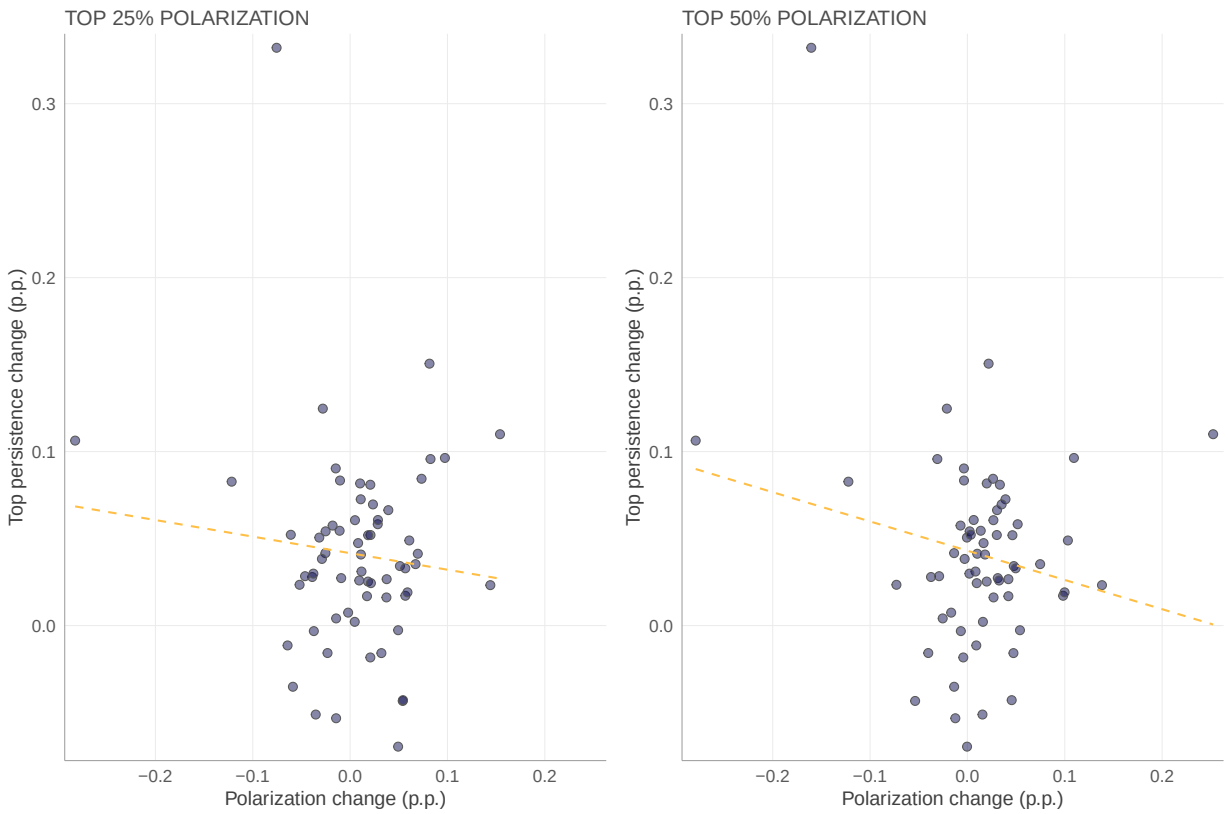


Figure 5: The figure plots, across 2-digit industries, changes in top persistence against changes in polarization. The left panel uses top 25% polarization, $\Phi_{j,t}^{75}$; the right panel uses top 50% polarization, $\Phi_{j,t}^{50}$. Top persistence is defined using the upper-quartile of the productivity distribution.

Figure 6 performs the analogous exercise for entry. It plots changes in polarization against changes in entrants' sales share, again showing both top 50% and top 25% polarization. A simple entry-barrier interpretation would predict a negative relation: industries where entry declined more should be those where polarization rose more. The data do not show this pattern. For top 25% polarization, the relationship is clearly positive, while for top 50% polarization it is flatter but still positive. In neither case is there evidence for the negative relation predicted by a simple entry-barrier story.

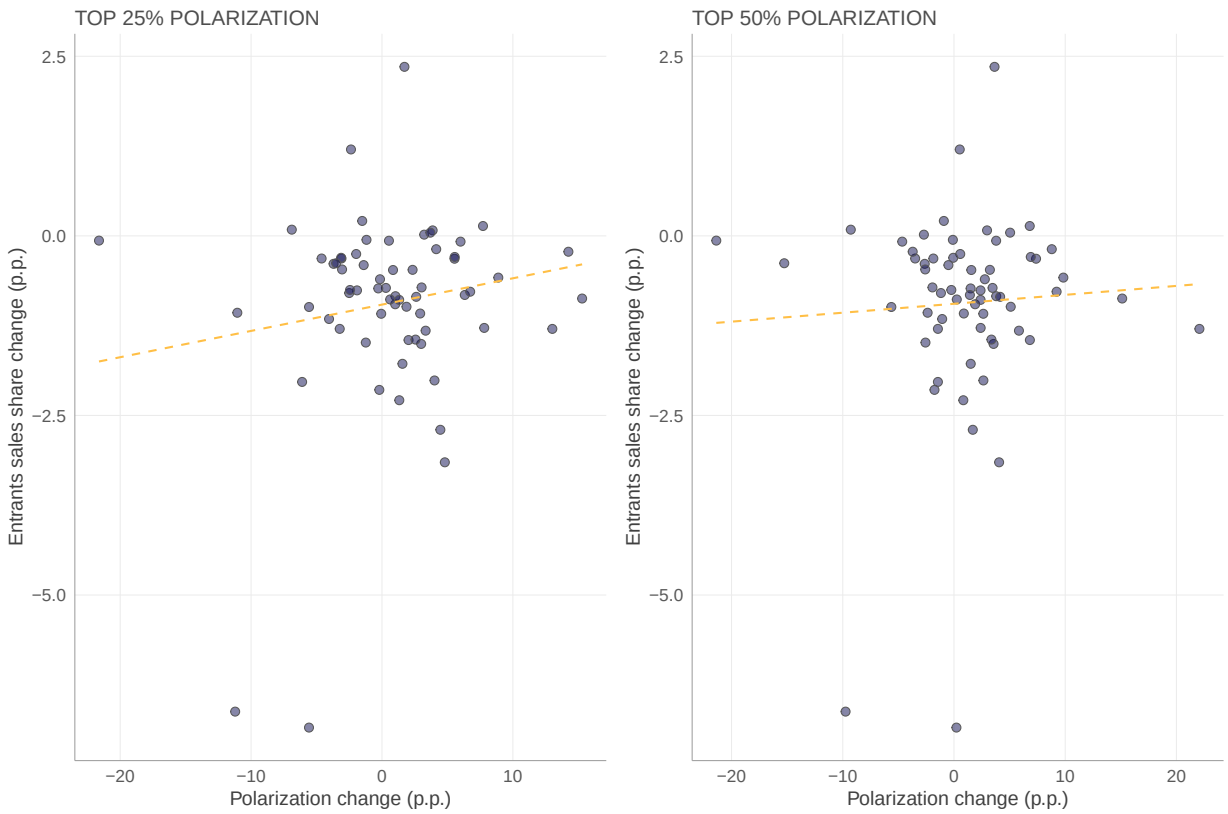


Figure 6: The figure plots, across 2-digit industries, changes in entrants' sales share against changes in polarization. The left panel uses top 25% polarization, $\Phi^{75}_{j,t}$; the right panel uses top 50% polarization, $\Phi^{50}_{j,t}$.

These cross-industry correlations are descriptive and should be interpreted cautiously: the slopes may be affected by a small number of industries, and the figures do not identify causal effects. Nevertheless, they reinforce the aggregate message: rising polarization is not concentrated in industries with the strongest decline in entry or the strongest increase in top persistence.

Stylised Fact 2. *Industries with larger increases in polarization do not display larger increases in top persistence or larger declines in entry. This holds for both top 50% and top 25% polarization. In neither case do the data support the view that polarization rose most where entry declined most.*

We now apply the accounting decomposition from Section 2.3 to identify which firm

groups account for changes in polarization. The decomposition separates changes in $\Phi^p_{j,t}$ into contributions from continuing top firms, young entrants to the top, mature entrants to the top, exiting top firms, and top dropouts. In the figures below we report four margins. Continuing top firms, young entrants to the top, and exiting top firms appear separately. Mature entrants to the top and top dropouts are combined into a single switching margin, which captures compositional turnover among firms that cross the upper-tail threshold without necessarily entering or leaving the market.

Figure 7 reports the decomposition of changes in top 50% polarization over 1997–2019. Bars correspond to average annual percentage-point contributions (p.p./year) over the period, aggregated across 2-digit industries within each sector using time-invariant value-added weights.

In manufacturing, top 50% polarization increases by approximately +2.2 p.p. in total between 1997 and 2019, as shown in Figure 3. Continuing top firms contribute about +0.1 p.p./year. Young entrants to the top and exiting top firms contribute about −0.2 and −0.1 p.p./year, respectively, while the switching margin contributes about +0.2 p.p./year. Scaled over the sample period, continuing top firms account for roughly +1.4 p.p. of the total increase, while the remaining +0.8 p.p. reflects net compositional turnover at the upper-tail boundary. Continuing top firms therefore matter, but the increase in manufacturing polarization is not solely a continuing-top-firm phenomenon.

In services, the total change in top 50% polarization is essentially zero. However, this masks strongly offsetting forces. Continuing top firms contribute about −0.2 p.p./year, implying a cumulative contribution of roughly −4.1 p.p.: the productivity premium of firms that remain continuously in the upper-tail narrows substantially relative to the median. Young entrants to the top and exiting top firms contribute about −0.2 and −0.1 p.p./year, while the switching margin contributes about +0.5 p.p./year. Together, the turnover margins imply a positive compositional contribution of approximately +4.0 p.p. over the sample period. The flat aggregate trend in services is therefore not the result of all components

being small; it conceals large offsetting movements, with continuing top firms narrowing their premium relative to the median while compositional turnover pushes strongly in the opposite direction.



Figure 7: The figure reports the decomposition of changes in top 50% polarization ($\Phi_{j,t}^{50}$) over 1997–2019, using the decomposition in equation (7). Bars correspond to average annual percentage-point contributions (p.p./year). Continuing firms: continuing top firms. Entering firms: young entrants to the top. Exiting firms: exiting top firms. Switching firms: mature entrants to the top plus top dropouts.

Figure 8 repeats the exercise for top 25% polarization. The qualitative conclusions are similar. In manufacturing, continuing top firms contribute about +0.1 p.p./year and do not account for the entire increase on their own; young entrants, exiting top firms, and switchers also contribute materially. In services, continuing top firms contribute about −0.1 p.p./year, young entrants and exiting top firms contribute about −0.2 and −0.1 p.p./year, and the switching margin contributes about +0.4 p.p./year, again offsetting the negative continuing-top contribution.



Figure 8: As Figure 7, for top 25% polarization ($\Phi^{75}_{j,t}$).

Overall, the decomposition evidence provides little support for the view that polarization dynamics are primarily driven by these continuing top firms pulling away. Instead, changes in polarization are strongly shaped by turnover at the upper-tail boundary. Young entrants to the top tend to dilute the average upper-tail premium when they enter below the average premium of the continuing top firms. Exiting top firms contribute through the departure of firms that were still in the upper-tail when they left the market. The switching margin captures firms that cross the upper-tail threshold while remaining active: mature entrants to the top and top dropouts. Top dropouts in particular raise measured polarization when they leave from the lower part of the upper-tail group, thereby leaving behind a more selected set of top firms. These forces can be large and offsetting, as the services decomposition illustrates strikingly.

Stylised Fact 3. *In manufacturing, continuing top firms contributed* $+1.4$ *p.p. to the total increase of* $+2.2$ *p.p., with the remainder accounted for by compositional turnover at the upper-tail boundary. In services, the flat aggregate trend conceals large offsetting forces: continuing top firms contributed* -4.1 *p.p. while compositional turnover contributed* $+4.0$ *p.p. The evidence is not consistent with rising polarization being driven primarily by continuing top firms pulling further ahead.*

Figure 9 plots industry level polarization against the polarization contribution of continuing top firms, switchers, and entry and exit. In manufacturing, industries with larger increases in polarization tend to have larger positive contributions from continuing top firms, but entry and exit also account for a sizeable part of the cross-industry dispersion. In services, the pattern is even less consistent with a simple continuing-top-firm story. Some industries with little aggregate change combine sizeable positive and negative component contributions, while others display polarization changes driven mainly by turnover at the upper-tail boundary. The industry-level evidence therefore reinforces the message from the sector aggregates: polarization dynamics are not explained solely by continuing top firms pulling away, but by the interaction between continuing top firms and compositional turnover within the upper-tail.

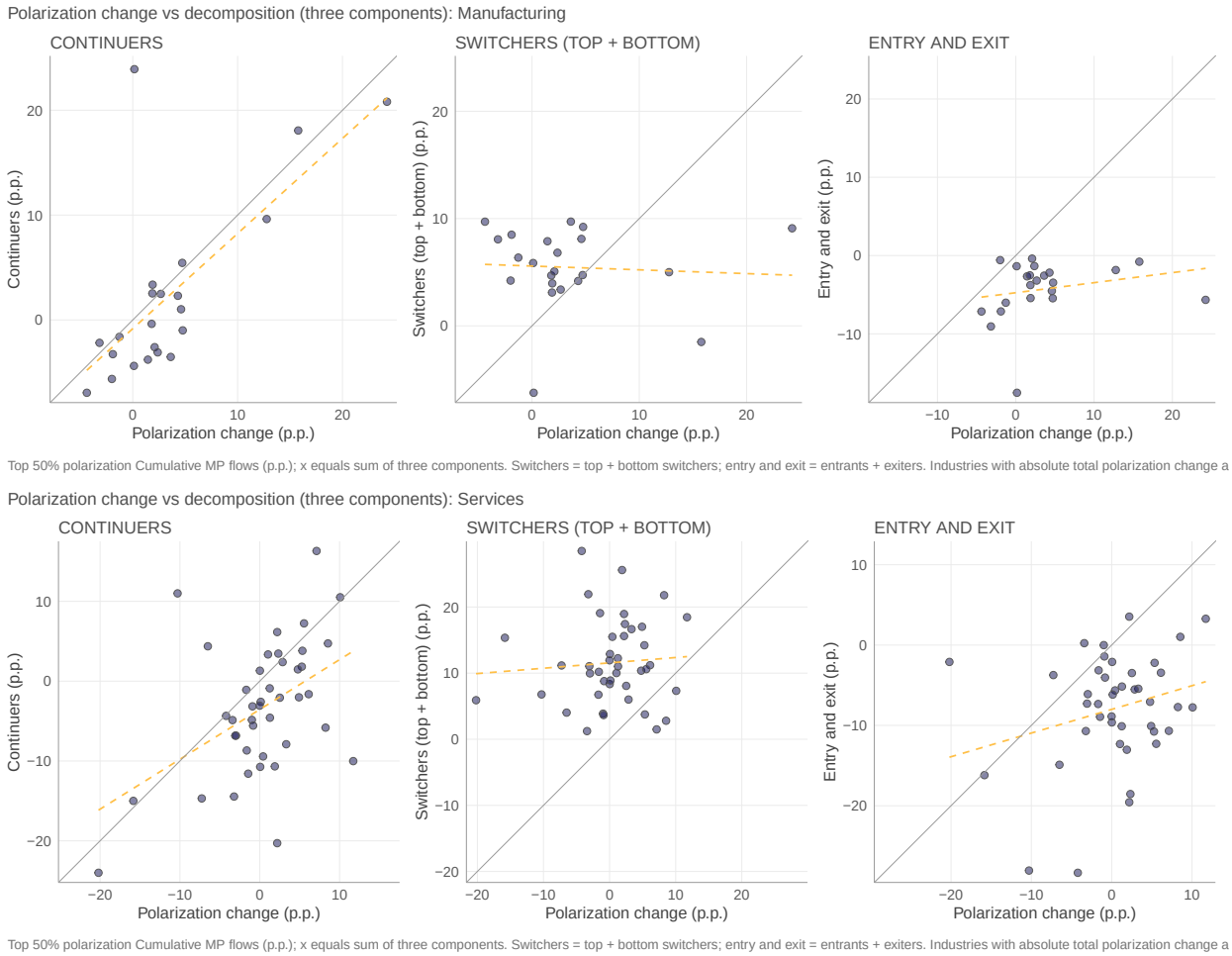


Figure 9: Cross-industry decomposition of changes in top 50% polarization ($\Phi^{50}_{j,t}$). The top panel reports manufacturing industries and the bottom panel reports services industries. Each point is a 2-digit industry. The horizontal axis is the total change in polarization, and the vertical axis reports the contribution of continuing top firms, switchers, and entry and exit. The dashed line is the cross-industry fitted regression line and the grey line is the 45-degree line. Industries with an absolute polarization change above 30 percentage points are excluded.

5 Model

Motivated by the empirical evidence, we propose a firm dynamics model in the style of Melitz (2003) with endogenous innovations decisions. We consider a discrete-time, indexed by t, general equilibrium industry model. Within an industry, there is a continuum of monopolistically competitive firms indexed by $\omega \in \Omega$ that face Constant Elasticity of Substitution

(CES) demands for their differentiated products. They differ in their productivity z (corresponding to φ in the empirical sections) and can invest in R&D to improve it. R&D investments can be applied to two types of innovation: incremental and radical innovation. Firms discount the future at a constant rate $\beta < 1$.

Demand: A representative consumer has CES preferences over a continuum of goods indexed by ω:

$$U = \left[\int_{\omega \in \Omega} q(\omega)^{\rho} \mathrm{d}\omega \right]^{\frac{1}{\rho}}, \ \rho \in (0,1) \tag{9}$$

The elasticity of substitution between the goods is governed by ρ and given by $1/(1-\rho) > 1$. The CES aggregate price index P_t is equal to:

$$P_t = \left[\int_{\omega \in \Omega} p_t(\omega)^{\frac{\rho}{\rho-1}} \mathrm{d}\omega \right]^{\frac{\rho-1}{\rho}} \tag{10}$$

where $p_t(\omega)$ is the price of variety ω at time t.

Consumer optimization gives the following demand function for variety ω:

$$q_t(\omega) = Q_t \left[\frac{P_t}{p_t(\omega)} \right]^{\frac{1}{1-\rho}} \tag{11}$$

where $q_t(\omega)$ is the quantity consumed of variety ω, and aggregate Q_t is given by:

$$Q_t = \left[\int_{\omega \in \Omega} q_t(\omega)^{\rho} \mathrm{d}\omega \right]^{\frac{1}{\rho}} \tag{12}$$

The overall demand is set exogenously to some level $\bar{Q}$.

Production: Each firm in the continuum produces a differentiated variety ω. There is a single input, labor, with elastic supply. We normalize the wage to 1. At any point in time

t, the producer of variety ω produces with a constant marginal product of labor equal to $z_t(\omega)$. The *ex-ante* productivity distribution is bounded between $\underline{z}$ and $\bar{z}$, where $\bar{z}$ defines the technological frontier. Production requires the payment of an overhead cost F_t. As the problem is identical for all varieties, henceforth we omit the ω. Static profits are given by:

$$\pi_t(z_t, \varepsilon_t) = p_t q_t - \frac{q_t}{z_t} - F_t \tag{13}$$

where F_t is given by

$$F_t = (1 + \varepsilon_t) F(z_t), \;\; F'(z_t) > 0 \tag{14}$$

ε_t is a shock drawn from an i.i.d. zero-mean distribution across time and varieties. In particular, we assume ε_t is drawn from a uniform distribution with support $[-\overline{\varepsilon}, \overline{\varepsilon}]$. Following Caggese (2019), $F(z_t)$ is:

$$F(z_t) = \left(\frac{z_t}{\hat{z}_0}\right)^{\kappa}, \;\; \kappa \in \left(0, \;\; \frac{\rho}{1-\rho}\right) \tag{15}$$

Firms choose p_t to maximize static profits in (16):

$$\max_{\{p_t\}} \pi_t(z_t, \varepsilon_t) = p_t \cdot q_t(p_t) - \frac{q_t(p_t)}{z_t} - F_t \tag{16}$$

subject to the demand function in (11). The first-order condition yields the standard pricing function:

$$p_t = \frac{1}{\rho} \cdot \frac{1}{z_t} \tag{17}$$

and quantity produced:

$$q_t = Q_t P_t^{\frac{1}{1-\rho}} (\rho z_t)^{\frac{1}{1-\rho}} \tag{18}$$

It follows that static profits are given by:

$$\pi(z_t, \varepsilon_t) = Q_t P_t^{\frac{1}{1-\rho}} z_t^{\frac{\rho}{1-\rho}} \rho^{\frac{\rho}{1-\rho}} (1-\rho) - F_t \tag{19}$$

Incumbents: An incumbent firm begins period t with productivity z_t and observes the overhead shock ε_t. It produces quantity q_t and receives profits $\pi(z_t, \varepsilon_t)$ according to (18) and (19). After production, the firm decides whether to exit or continue. If it continues, it chooses innovation, the details of which are described below, the outcome of which determines its potential productivity for next period $\tilde{z}_{t+1}$. At the end of the period, the firm exits exogenously with probability d. Conditional on survival, the firm faces a productivity depreciation shock: with probability d_z, the realized productivity z_{t+1} is scaled such that $z_{t+1} = \tilde{z}_{t+1}(1 - g_z)$; otherwise $z_{t+1} = \tilde{z}_{t+1}$. If realized productivity z_{t+1} falls below $\underline{z}$, the firm is forced to exit.

Incremental innovation: Successful incremental innovations determine a potential productivity $\tilde{z}_{t+1}$ according to (20). The rate of growth depends on the distance to the frontier $z_t/\bar{z}$ according to:

$$\tilde{z}_{t+1}(z_t) = \bar{z} \left(\frac{z_t}{\bar{z}} \right)^{\gamma}, \ \gamma \in (0, 1) \tag{20}$$

The function in equation (20) has the following nice properties. First, if a firm has technology at the frontier, i.e. $z_t = \bar{z}$, even if the incremental innovation is successful $\tilde{z}_{t+1} = z_t = \bar{z}$. Second, $\tilde{z}_{t+1}$ is strictly increasing in z_t. Third, the productivity growth rate is decreasing in z_t such that the productivity growth is increasing in the distance to the frontier:

$$\frac{\tilde{z}_{t+1}}{z_t} - 1 = \bar{z}^{1-\gamma} z_t^{\gamma-1} - 1 \tag{21}$$

Firms can influence the probability of successful incremental innovation by choosing how much to invest $M_t \geq 0$ in this type of innovation. The probability of a successful incremental innovation $\mathbb{P}$ relates to the amount invested M_t in the following way:

$$\mathbb{P}(M_t) = \left(\frac{M_t}{M_t + 1}\right)^{\frac{1}{\sigma}}, \ \sigma > 0, \ M_t \geq 0 \tag{22}$$

The probability functional form in (22) implies that if the firm does not invest $M_t = 0$, the probability of success is zero $\mathbb{P}(0) = 0$. Moreover, the probability of success is increasing in the amount M_t with asymptote at 1.

Radical innovation: Alternatively, the firm can invest in radical innovation. This innovation has a fixed cost $\bar{E}$ and a constant success probability $\mathbb{S}$. If successful, the firm draws its potential productivity $\tilde{z}_{t+1}$ from a distribution concentrated near the technological frontier with support $[z_h, \bar{z}]$ where $z_h = \nu_r \underline{z} + (1 - \nu_r)\bar{z}$. However, if unsuccessful, productivity is reduced by a fraction $g > 0$, i.e.:

$$\tilde{z}_{t+1} = z_t(1 - g), \ g > 0 \tag{23}$$

Value functions: The previously described incumbent problem can be summarized by value-functions. First, an incumbent decides whether to continue and receive $V^c(z_t, \varepsilon_t)$ or exit and receive $V^x(z_t, \varepsilon_t)$, such that $V(z_t, \varepsilon_t)$ is given by:

$$V(z_t, \varepsilon_t) = \max\{V^c(z_t, \varepsilon_t), V^x(z_t, \varepsilon_t)\} \tag{24}$$

and

$$V^x(z_t, \varepsilon_t) = \pi_t(z_t, \varepsilon_t) \tag{25}$$

The continuation value $V^c(z_t, \varepsilon_t)$ includes the current static profits and the continua-

tion value. The latter depends on the R&D decision. In particular,

$$V^c(z_t, \varepsilon_t) = \pi(z_t, \varepsilon_t) + \max\{V^i(z_t, \varepsilon_t), V^r(z_t, \varepsilon_t)\} \tag{26}$$

Where $V^i(z_t, \varepsilon_t)$ is the value of incremental innovation and $V^r(z_t, \varepsilon_t)$ is the value of radical innovation. The value of incremental innovation is given by:

$$V^i(z_t, \varepsilon_t) = \max_{\{M_t\}} -M_t + \beta \cdot (1-d)\left[\mathbb{P}(M_t)\mathcal{V}(\tilde{z}_{t+1}) + (1-\mathbb{P}(M_t))[\mathcal{V}(z_t)]\right] \tag{27}$$

and the value of radical innovation is given by:

$$V^r(z_t, \varepsilon_t) = -\bar{E} + \beta(1-d)\left[\mathbb{S}\int_{z_h}^{\bar{z}} \mathcal{V}(z')h(z')dz' + (1-\mathbb{S})\mathcal{V}(z_t(1-g))\right] \tag{28}$$

where z_h is the lower support of the productivity distribution conditional on a successful radical innovation and $h(\cdot)$ is the probability density function. In equations (27) and (28), $\mathcal{V}(\cdot)$ is given by:

$$\mathcal{V}(\tilde{z}) = \mathbb{E}_\varepsilon\left[(1-d_z)V(\tilde{z}, \varepsilon')\mathbb{I}\{\tilde{z} \geq \underline{z}\} + d_z V(\tilde{z}(1-g_z), \varepsilon')\mathbb{I}\{(1-g_z)\tilde{z} \geq \underline{z}\}\right] \tag{29}$$

Entrants: In each period t there is a large number of potential entrants. These potential entrants pay the entry cost F^e to draw their productivity from a right-skewed distribution f with support $[\underline{z}^e, \bar{z}^e]$. We assume that the productivity of entrants is drawn from a uniform distribution where:

$$\underline{z}^e = \tilde{\nu}\underline{z} + (1-\tilde{\nu})\bar{z};\ \ \bar{z}^e = (1-\nu)\underline{z} + \nu\bar{z} \tag{30}$$

where $\tilde{\nu} > 1-\nu$. After paying the entry cost and learning their type, they decide whether they want to enter the market in the next period. The free entry condition requires *ex-ante*,

the expected value of paying F^e conditional on the expectation of the initial values.

$$V^e = \beta \int_{\underline{z}}^{\bar{z}} \max\{\mathbb{E}_{\varepsilon_0}\left[V(z_0, \varepsilon_0)\right], 0\} f(z_0) \mathrm{d}z_0 - F^e = 0 \tag{31}$$

Equilibrium Conditions: We focus on a stationary industry equilibrium where aggregate variables Q_t, P_t, and the distribution of firms over productivity and shocks remains constant over time. Consequently, we drop time subscripts t for aggregate variables. The stationary equilibrium is defined by a set of prices $\{p(z, \varepsilon)\}$, quantities $\{q(z, \varepsilon)\}$, investment decisions $\{M(z, \varepsilon), \mathbb{I}\{\text{incremental}(z, \varepsilon)\}, \mathbb{I}\{\text{radical}(z, \varepsilon)\}\}$, a mass of entrants M_e, and a stationary distribution of firms $\mu(z, \varepsilon)$ such that:

1. Given aggregate demand Q and price index P, the functions $p(z, \varepsilon)$, $q(z, \varepsilon)$, and the innovation decisions maximize the firm value as described by the Bellman equations above.

2. The aggregate price P is such that the net present value of entry is zero, as defined in the free entry condition $V^e = 0$.

3. Aggregate output produced by the continuum of firms clears the exogenously given demand.

4. The mass of firms entering the industry and the transitions between productivity levels (driven by innovation outcomes and depreciation) exactly offset the mass of firms exiting due to exogenous shock d, the depreciation shock d_z driving productivity below $\underline{z}$, and endogenous exit.

5.1 Benchmark Calibration and Counterfactual Simulations

In the steady state equilibrium of our model, polarization Φ_t is constant by construction, and $\Delta\Phi_t = 0$. So changes in polarization, to be analysed using the methodology developed in section 2.3, can be generated by characterizing the transition dynamics of the distribution

after a given exogenous shock. Therefore our strategy is to first calibrate the model to match a set of relevant empirical moments, then to analyse to what extent shocks to entry and exit and to technology diffusion can explain the empirical evidence.

More specifically, we calibrate the model by jointly targeting a set of empirical moments of the services sector in the initial period (1997-2001). The targeted moments are specified in Table 3 and the calibrated parameters are shown in Table 4.

Table 3: Benchmark Calibration Moments

Moment	**Data**	**Model**
Profits sales ratio for the 50th percentile	0.12	0.09
Exit share	0.10	0.09
Average productivity of new firms relative to the frontier	0.67	0.48
Polarization	0.21	0.23
Contribution of continuing polarization changes	-0.02	0.00
Contribution of continuing firms to polarization	0.19	0.23
Contribution of mature entrants to the top to polarization	0.01	0.01
Contribution of top dropouts to polarization	0.01	0.01
Contribution of exiting top to polarization	0.01	0.01

Notes: The table shows targeted empirical moments and their model counterparts. Empirical moments are averages across 2-digit industries within the services sector for the 1997-2001 period. Model counterparts are computed in the stationary equilibrium.

We rely on the calibrated model to study the effects of changes in the structural parameters on polarization and its components during the transition. To compute the transition path between the initial and final steady-states, we assume a perfect foresight transition over a sufficiently long time horizon T. Starting from the initial stationary distribution of firms, we guess a path for the aggregate endogenous variables, solve the firms' dynamic programming problems backward from period T to period 0, and simulate the evolution of the firm distribution forward. We iterate on the aggregate paths until the path of prices clears the markets along the entire transition.

In particular, we construct a series of independent structural shocks, each scaled precisely to generate a uniform $\Delta\Phi = 0.05$ increase in aggregate polarization in the new stationary equilibrium compared to the original equilibrium. The structural decompositions of

Table 4: Benchmark Calibration Parameters

Description	**Parameter**	**Value**
Preferences and Technology		
Gross interest rate ($\star$)	$1/\beta$	1.02
Substitution parameter	ρ	0.68
Productivity cutoff ($\star$)	$\underline{z}$	1
Productivity frontier ($\star$)	$\overline{z}$	10
Upper bound overhead shock	$\overline{\varepsilon}$	0.37
Overhead cost curvature	κ	0.09
Normalizing factor overhead costs ($\star$)	$\hat{z}_0$	1
Exogenous exit probability	d	0.04
Depreciation shock probability	d_z	0.24
Productivity loss from depreciation	g_z	0.20
Innovation		
Innovation step curvature	γ	0.90
Innovation success concavity	$1/\sigma$	0.11
Cost of radical innovation	$\bar{E}$	1.00
Prob. success radical innovation	$\mathbb{S}$	0.07
Regulator for distribution for succ. radical	ν_r	0.11
Productivity loss if failed radical innovation	g	0.01
Entry		
Entrant draw spread	$\tilde{\nu}$	0.79
Entrant location	ν	0.68
Entry cost	F^e	2.60

Notes: The table shows the baseline calibration parameters. Parameters marked with a $\star$ are externally calibrated.

these transitions are reported in Table 5.

Starting with the incremental innovation margin, we find that both a higher innovation step curvature (γ) and a lower innovation success concavity (σ) successfully drive aggregate polarization up. While these changes mirror a slowdown in technology diffusion, making it costlier or less effective for lagging firms to catch up, they generate a negative contribution from entry and exit dynamics (approximately -0.04 in both cases). This implies that a pure decline in technology diffusion suppresses the entry/exit polarization channel, a pattern that directly contradicts the empirical cross-sectional evidence documented in Section 4. Consequently, we conclude that the empirical rise in polarization cannot be primarily driven

by a slowdown in technology diffusion.

Similarly, shocks that facilitate radical innovation, a higher success probability ($\mathbb{S}$) and a lower radical innovation cost ($\bar{E}$), fail to generate meaningful positive variations along the entry and exit margins. In these innovation counterfactuals, the aggregate increase in polarization is almost entirely absorbed by switching firms.

In contrast, adjustments to the firm selection and operating environment match our empirical sign expectations. Elevating overhead costs—either via a wider idiosyncratic shock distribution ($\bar{\varepsilon}$) or steeper cost curvature (κ)—robustly amplify polarization through the entry and exit channel, yielding positive contributions of +0.0331 and +0.0570, respectively.

On the one hand, these results are supportive of empirical evidence provided above, that increases in polarization are related to entry and exit dynamics and not primarily to slower technology diffusion. On the other hand, this evidence is still preliminary because the current version of the model fails along one specific dimension: every single-parameter experiment yields a negative contribution from continuing firms to the change in polarization. This systematically runs counter to our empirical baseline (Figure 7), where continuing firms register a strong positive contribution. A new version of the model, which includes a richer characterization of innovation at the frontier, and which we believe will be able to address this shortcoming, is currently work in progress.

6 Conclusions

This paper examines the drivers of polarization between 1997 and 2019 in France, using the combined FICUS–FARE administrative panel. Aggregate polarization in France is essentially flat over the period, but the majority of manufacturing industries (77% for top 50% polarization) experienced a positive trend, while services industries are roughly evenly split between rising and falling polarization. The market presence of new firms declined persistently over the period, with entrants' sales share falling from approximately 2.5 percent in

Table 5: Counterfactual Decompositions of Polarization Changes

Experiment	$\Delta\Phi$	**Continuing**	**Entry & exit at the top**	**Switching**
Entry and Exit				
Lower exogenous exit probability (d)	0.05	-0.2676	0.0568	0.2608
Higher entry cost (F^e)	0.05	-0.2389	0.0185	0.2703
Lower entrant location (ν)	0.05	-0.0151	-0.1355	0.2006
Higher entrant draw spread ($\tilde{\nu}$)	0.05	-0.2406	0.0166	0.2740
Higher overhead cost shock ($\bar{\varepsilon}$)	0.05	-0.0994	0.0331	0.1163
Higher overhead cost curvature (κ)	0.05	-0.2109	0.0570	0.2040
Incremental Innovation				
Higher innovation step curvature (γ)	0.05	-0.5369	-0.0368	0.6237
Lower innovation success concavity (σ)	0.05	-0.2692	-0.0372	0.3564
Radical Innovation				
Higher radical success probability ($\mathbb{S}$)	0.05	-0.2323	0.0021	0.2803
Lower radical innovation cost ($\bar{E}$)	0.05	-0.2685	-0.0277	0.3461
Productivity Depreciation				
Higher depreciation probability (d_z)	0.05	-1.1959	-0.1976	1.4435
Higher depreciation shock size (g_z)	0.05	-0.4065	-0.1208	0.5574
Macro and Preferences				
Lower gross interest rate ($1/\beta$)	0.05	-0.2419	-0.0503	0.3422

Notes: The table reports the structural decomposition of changes in industry polarization across alternative parameter counterfactuals. Each experiment scales the underlying structural shock to generate a uniform target increase of $\Delta\Phi = 0.05$ in the final stationary equilibrium.

the late 1990s to roughly 1.5 percent by the mid-2010s in services, and from around 1 percent to approximately 0.6 percent in manufacturing. Over the same period, top persistence has risen by approximately 5 percentage points from 60% to 65% in both manufacturing and services. Taken together, these aggregate trends are consistent with the broader decline in business dynamism documented for France and other advanced economies.

We develop a novel application of the Melitz–Polanec decomposition to analyse the contributions of different firm dynamics to both the level and the change of polarization. The analysis reveals two main results. First, in any given year continuing top firms (those that were already in the upper-tail group in the previous year) account for the bulk of the level of polarization (around 85–90%): polarization is largely a property of these continuing

firms. Second, the cumulative changes in polarization over 1997–2019 are not driven by these continuing firms pulling further ahead: their productivity gap to the median is roughly stable in manufacturing and narrows marginally in services. Where polarization rose, it did so because of compositional turnover at the boundary of the upper-tail group: top dropouts (active firms that fall below the upper-tail productivity threshold) left behind a more elite residual group, partially offset by young and mature entrants to the top entering below the average of the continuing group.

These findings are difficult to reconcile with simple versions of theories in which productivity polarization is driven primarily by slower technology diffusion, increased entry costs, or the dominance of incumbent "superstar" firms. Relatedly, we find no systematic relation between polarization increases and indicators of diminished competition such as increased persistence of top firms or reduced entrants' sales share — and the cross-industry correlation between rising polarization and falling entry, far from being negative as the entry-barriers hypothesis would predict, is in fact mildly positive for top 50% polarization. Overall, the evidence presented suggests that the relationship between technological change, firm dynamics, and polarization is complex and multifaceted, and that French data are best described as a regime of declining business dynamism with a stable elite at the top of each industry distribution rather than as one of frontier divergence. Further theoretical and empirical investigation is necessary to fully understand the interplay of these factors in driving polarization.

7 References

References

Ackerberg, Daniel A, Kevin Caves, and Garth Frazer. 2015. "Identification properties of recent production function estimators." Econometrica, 83(6): 2411–2451.

Akcigit, Ufuk, and Sina T Ates. 2021. "Ten facts on declining business dynamism and lessons from endogenous growth theory." American Economic Journal: Macroeconomics, 13(1): 257–298.

Andrews, Dan, Chiara Criscuolo, Peter N Gal, et al. 2019. "The best versus the rest: divergence across firms during the global productivity slowdown." Centre for Economic Performance, LSE.

Berlingieri, Giuseppe, Sara Calligaris, Chiara Criscuolo, and Rudy Verlhac. 2020. "Laggard firms, technology diffusion and its structural and policy determinants." OECD Publishing.

Burstein, Ariel T, Vasco M Carvalho, and Basile Grassi. 2025. "Bottom-up markup fluctuations." The Quarterly Journal of Economics, 140(4): 2619–2684.

Caggese, Andrea. 2019. "Financing constraints, radical versus incremental innovation, and aggregate productivity." American Economic Journal: Macroeconomics, 11(2): 275–309.

Calligaris, Sara, Chiara Criscuolo, and Luca Marcolin. 2018. "Mark-ups in the digital era." OECD Publishing.

Comin, Diego, and Marti Mestieri. 2014. "Technology diffusion: Measurement, causes, and consequences." In Handbook of economic growth. Vol. 2, 565–622. Elsevier.

De Loecker, Jan, and Frederic Warzynski. 2012. "Markups and firm-level export status." American Economic Review, 102(6): 2437–2471.

De Ridder, Maarten. forthcoming. "Market Power and Innovation in the Intangible Economy." American Economic Review.

Di Giovanni, Julian, Andrei A Levchenko, and Isabelle Mejean. 2024. "Foreign shocks as granular fluctuations." Journal of Political Economy, 132(2): 391–433.

Gutiérrez, Germán, and Thomas Philippon. 2020. "Some Facts about Dominant Firms." National Bureau of Economic Research Working Paper 27985.

Keller, Wolfgang. 2004. "International technology diffusion." Journal of economic literature, 42(3): 752–782.

Levinsohn, James, and Amil Petrin. 2003. "Estimating Production Functions Using Inputs to Control for Unobservables." The Review of Economic Studies, 70(2): 317–341.

Melitz, Marc J. 2003. "The impact of trade on intra-industry reallocations and aggregate industry productivity." econometrica, 71(6): 1695–1725.

Melitz, Marc J, and Sašo Polanec. 2015. "Dynamic Olley-Pakes productivity decomposition with entry and exit." The Rand journal of economics, 46(2): 362–375.

Olley, G. Steven, and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." Econometrica, 64(6): 1263–1297.

Olmstead-Rumsey, Jane. 2022. "Market Concentration and the Productivity Slowdown."

Appendix 3

A3.1 Production Function Estimation

Our analysis of the determinants of *polarization* requires the estimation of firm-level productivity. In our preferred specification, we define productivity as the log of TFPR by estimating the following industry-specific Cobb-Douglas production function:

$$q_{it} = \alpha + \beta_\ell \ell_{it} + \beta_m m_{it} + \beta_k k_{it} + \varphi_{it} + \epsilon_{it} \tag{32}$$

where lower cases denote log variables: q_{it} is log output/revenues, ℓ_{it} is log labor/employment, m_{it} is log materials, k_{it} is log capital, φ_{it} stands for firm productivity, and ϵ_{it} is measurement error/unanticipated shocks. The parameters α, β_l, β_m, and β_k are common across all firms within an industry.

The traditional estimation challenge to obtain consistent estimates of the output elasticities is simultaneity bias, due to the possibility that firm productivity (unobserved to the econometrician but known to the firm) is correlated with input choices. As in De Loecker and Warzynski (2012), we follow the control function approach pioneered by Olley and Pakes (1996) and Levinsohn and Petrin (2003) and updated by Ackerberg, Caves and Frazer (2015). The methodology assumes that productivity follows a first-order Markov process and can be expressed as a function of the firm's flexible inputs and capital:

$$\varphi_{it} = h(\ell_{it}, m_{it}, k_{it})$$

Then the method proceeds in two steps. In the first step, one obtains estimates of the

expected output that removes measurement error and unanticipated shocks:[11]

$$q_{it} = \phi_t(\ell_{it}, m_{it}, k_{it}) + \epsilon_{it}$$

In the second step, the method relies on the law of motion for productivity:

$$\varphi_{it} = g_t(\varphi_{i,t-1}) + \xi_{it}$$

where ξ_{it} are the innovation shocks to productivity. Based on these steps, the following moment conditions can be formed to obtain the output elasticity estimates:

$$E\left[\xi_{it}(\beta)\begin{pmatrix}\ell_{i,t-1}\\ m_{i,t-1}\\ k_{i,t}\end{pmatrix}\right] = 0$$

Note that the moment condition for the flexible input uses $t-1$ as an instrument and addresses the Ackerberg, Caves and Frazer (2015) critique. Production functions are estimated separately by industry (2-digit NACE Rev.2 industry classification).

A3.2 Additional Figures for Polarization and Firm Dynamics

This appendix reports additional figures underlying the main analysis in Section 4. The main text focuses on aggregate time-series evidence, cross-industry correlations using both top 50% and top 25% polarization, and the decomposition of polarization changes. Here we report the industry-level polarization series, convergence plots for persistence and entry, robustness checks around administrative data breaks, and sector-specific versions of the cross-industry relationships.

[11] In particular, estimates of expected output are obtained from a second-order polynomial approximation of the production function.

A3.2.1 Industry-Level Polarization Series

Figures 10 and 11 report the industry-level time series for top 50% polarization, separately for manufacturing and services.

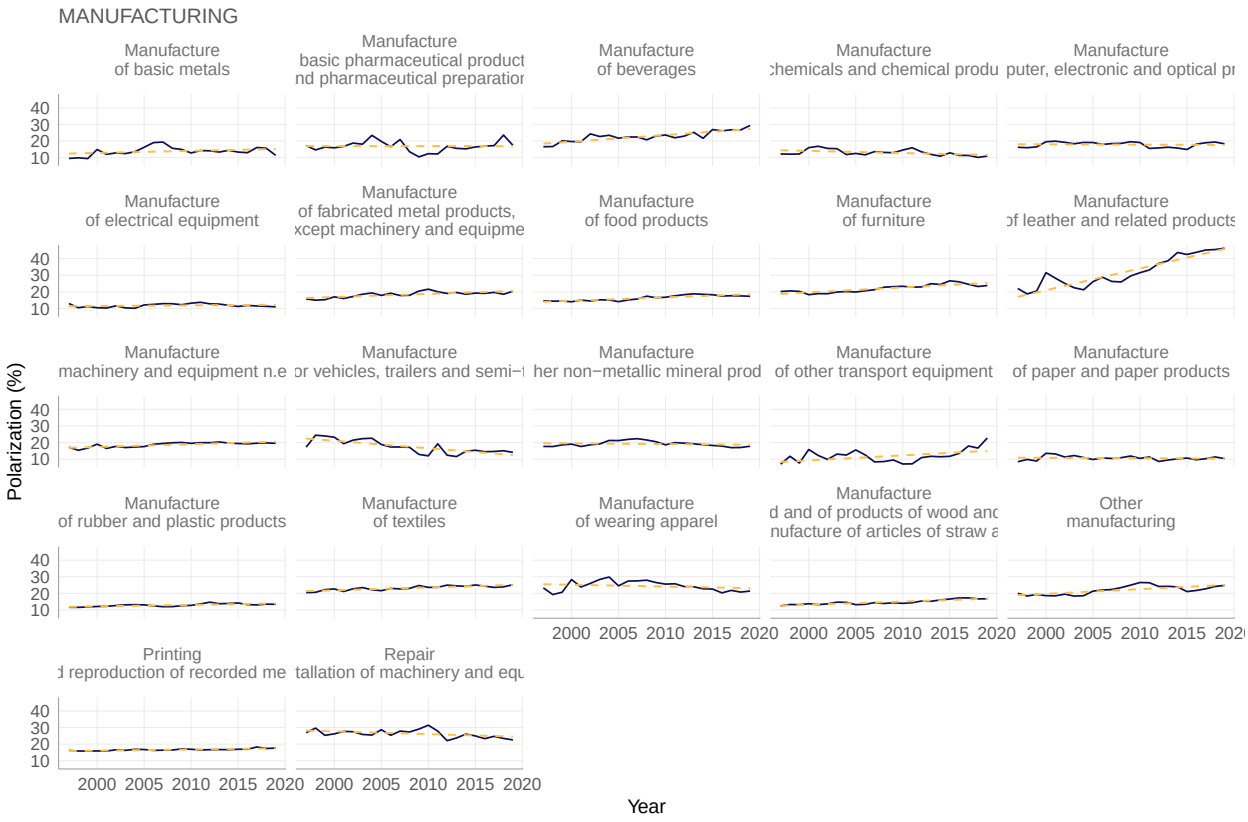


Figure 10: Industry-level top 50% polarization in manufacturing



Figure 11: Industry-level top 50% polarization in services

A3.2.2 Initial and Final Polarization Across Industries

Figure 12 plots end-of-sample polarization against initial polarization across 2-digit industries, separately for manufacturing and services. We pool top 50% and top 25% polarization, shown with different markers. The solid black line is the 45-degree line.

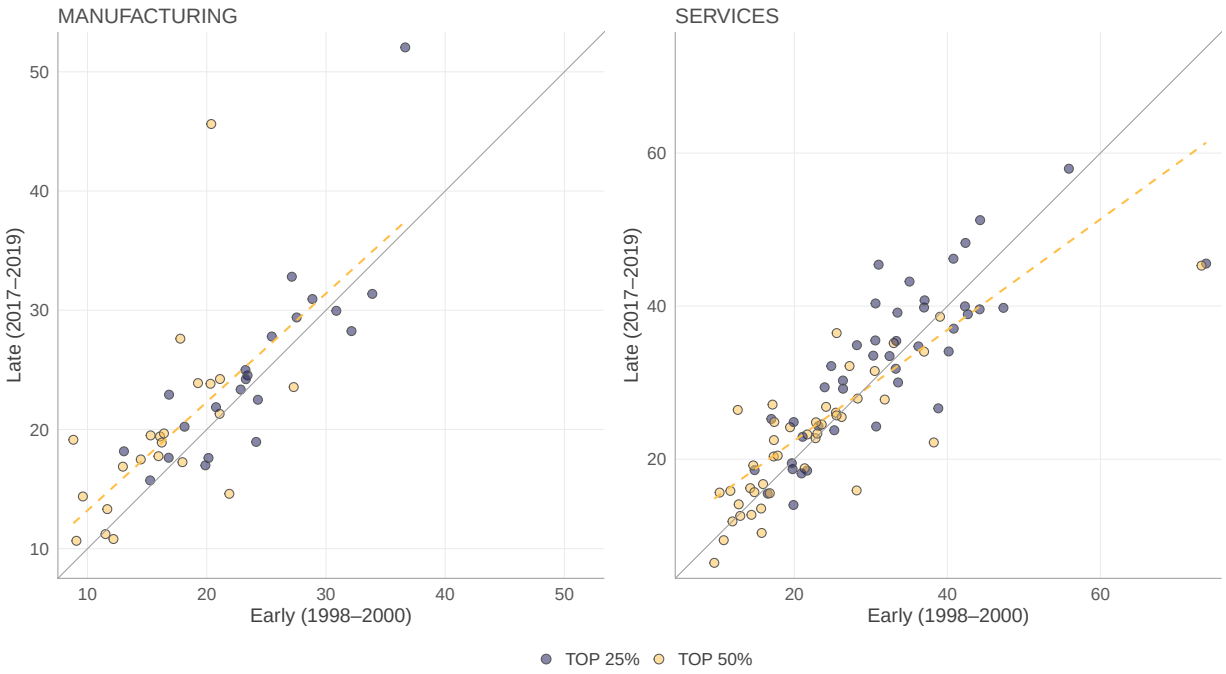


Figure 12: Initial and final polarization across industries

A3.2.3 Top Persistence Across Industries

Figure 13 plots end-of-sample top persistence against initial top persistence across 2-digit industries. The fitted line is flatter than the 45-degree line, indicating convergence in top persistence across industries.

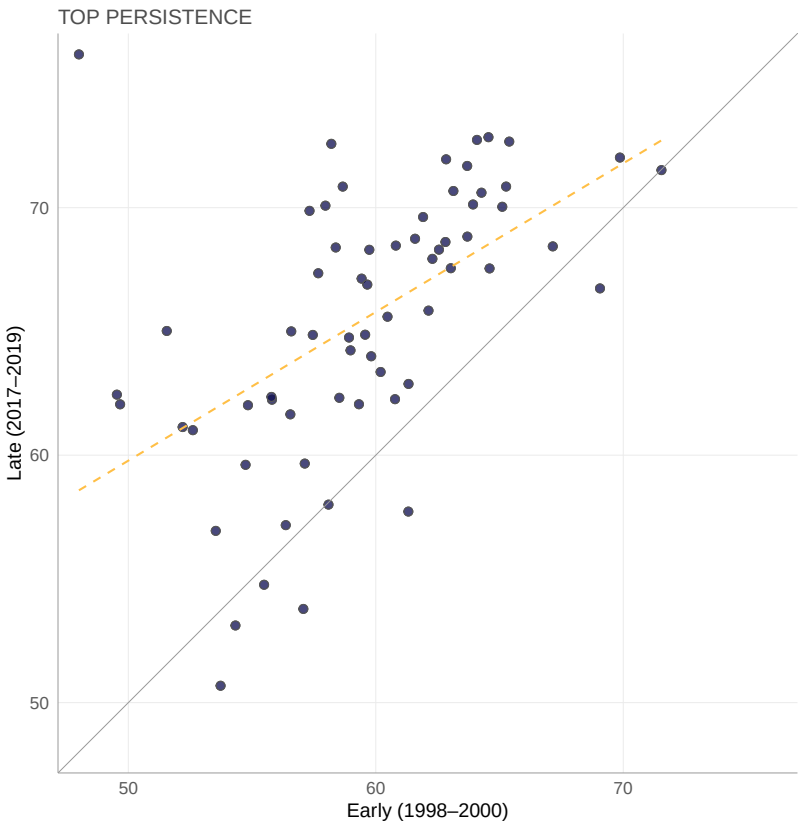


Figure 13: Top persistence converges across industries

A3.2.4 Entry Across Industries

Figure 14 plots initial and final entry measures across 2-digit industries. In both panels the fitted line is flatter than the 45-degree line, indicating convergence: industries with initially high entrants' sales share experienced the largest declines.

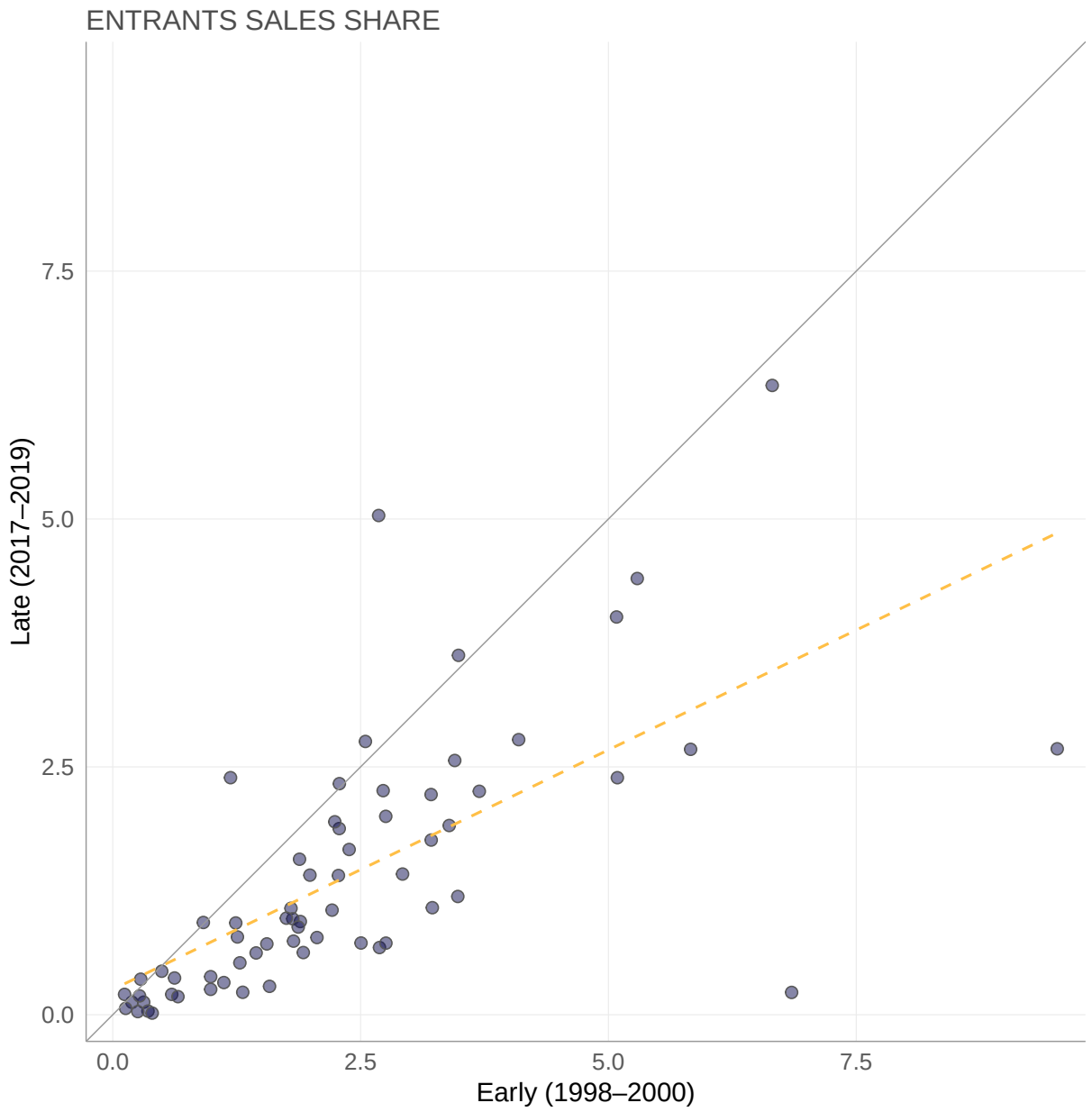


Figure 14: Entrants' sales share converges across industries as high-entry industries decline most

A3.3 Industry Classification

Financial and insurance activities (NACE divisions 64–66, Section K) are excluded from the main analysis due to the distinct accounting and regulatory structure of financial firms.

Manufacturing Industries

NACE2	Description	Section
10	Manufacture of food products	C
11	Manufacture of beverages	C
13	Manufacture of textiles	C
14	Manufacture of wearing apparel	C
15	Manufacture of leather and related products	C
16	Manufacture of wood and of products of wood and cork, except furniture	C
17	Manufacture of paper and paper products	C
18	Printing and reproduction of recorded media	C
20	Manufacture of chemicals and chemical products	C
21	Manufacture of basic pharmaceutical products and pharmaceutical preparations	C
22	Manufacture of rubber and plastic products	C
23	Manufacture of other non-metallic mineral products	C
24	Manufacture of basic metals	C
25	Manufacture of fabricated metal products, except machinery and equipment	C
26	Manufacture of computer, electronic and optical products	C
27	Manufacture of electrical equipment	C
28	Manufacture of machinery and equipment n.e.c.	C
29	Manufacture of motor vehicles, trailers and semi-trailers	C
30	Manufacture of other transport equipment	C
31	Manufacture of furniture	C
32	Other manufacturing	C
33	Repair and installation of machinery and equipment	C

Service Industries

NACE2	Description	Section
41	Construction of buildings	F
42	Civil engineering	F
43	Specialised construction activities	F
45	Wholesale and retail trade and repair of motor vehicles and motorcycles	G
46	Wholesale trade, except of motor vehicles and motorcycles	G
47	Retail trade, except of motor vehicles and motorcycles	G
49	Land transport and transport via pipelines	H
50	Water transport	H
51	Air transport	H
52	Warehousing and support activities for transportation	H
53	Postal and courier activities	H
55	Accommodation	I
56	Food and beverage service activities	I
58	Publishing activities	J

NACE2	Description	Section
59	Motion picture, video and television programme production, sound recording and music publishing activities	J
60	Programming and broadcasting activities	J
61	Telecommunications	J
62	Computer programming, consultancy and related activities	J
63	Information service activities	J
68	Real estate activities	L
69	Legal and accounting activities	M
70	Activities of head offices; management consultancy activities	M
71	Architectural and engineering activities; technical testing and analysis	M
72	Scientific research and development	M
73	Advertising and market research	M
74	Other professional, scientific and technical activities	M
77	Rental and leasing activities	N
78	Employment activities	N
79	Travel agency, tour operator reservation service and related activities	N
80	Security and investigation activities	N
81	Services to buildings and landscape activities	N
82	Office administrative, office support and other business support activities	N
85	Education	P
86	Human health activities	Q
87	Residential care activities	Q
88	Social work activities without accommodation	Q
90	Creative, arts and entertainment activities	R
91	Libraries, archives, museums and other cultural activities	R
92	Gambling and betting activities	R
93	Sports activities and amusement and recreation activities	R
94	Activities of membership organisations	S
95	Repair of computers and personal and household goods	S
96	Other personal service activities	S