Creative Destruction through Innovation Bursts[*]

Giuseppe Berlingieri[‖] ESSEC
Maarten De Ridder[‡] LSE
Danial Lashkari[§] FRBNY
Davide Rigo[¶] LSE

April 2025

Please Click Here for the Latest Version.

Abstract

In theories of creative destruction, product innovation is a key driver of aggregate growth. In this paper, we confront the predictions of these theories about product dynamics with empirical patterns in product-level data on the near-universe of French manufacturing firms. We find that the process of product innovation frequently exhibits *bursts*–episodes in which firms rapidly add multiple products to their portfolio. Such bursts lead to substantial shifts in revenue and explain the majority of the variance in firm-level growth. We introduce a model of firm product innovation compatible with such a process that also nests the canonical models of creative destruction. We show that innovation bursts alter the equilibrium composition of age, size, and innovation efficiency of firms, and further explain the concentration of production among superstar firms. Our model thus enables the joint study of the determinants of industry concentration and growth in a setting consistent with the empirical patterns of product dynamics.

Keywords: Productivity, Endogenous Growth, Firms, Innovation

[*]We thank Philippe Aghion, David Argente, Ben Moll, Nicolas Crouzet, Xavier Gabaix, Pete Klenow, Munseob Lee, Sara Moreira, Michael Peters, and Fabrizio Zilliboti, and numerous seminar participants for helpful feedback and discussions. This work is supported by a public grant overseen by the French National Research Agency (ANR) as part of the 'Investissements d'Avenir' program (reference: ANR-10-EQPX-1, CASD).

[‡]*Address:* ESSEC Business School, Department of Economics, 3 Avenue Bernard Hirsch, Cergy-Pontoise, France *E-mail*: giuseppe.berlingieri@essec.edu. *Web:* https://sites.google.com/site/giuseppeberlingieri/.

[§]*Address:* London School of Economics, Department of Economics, 32 Lincoln's Inn Fields, London, U.K. *E-mail*: m.c.de-ridder@lse.ac.uk. *Web:* http://www.maartenderidder.com/.

[¶]*Address:* Federal Reserve Bank of New York, Department of Research and Statistics, 33 Liberty St, New York, U.S.A. *E-mail*: danial.lashkari@ny.frb.org. *Web:* https://sites.google.com/view/danial-lashkari/home.

[‖]*Address:* London School of Economics, Department of Economic Geography, London, U.K. *E-mail*: d.rigox @lse.ac.uk. *Web:* https://daviderigo.weebly.com/.

1. Introduction

Product innovation lies at the heart of modern theories of firm growth through creative destruction (Aghion and Howitt, 1992; Grossman and Helpman, 1991; Aghion et al., 2014). According to these theories, innovative firms invest in research and development (R&D) to create new products superior to those of competitors, and to expand their market share, output, and profits. The canonical approach, pioneered by Klette and Kortum (2004), portrays product innovation as a gradual process, where firms incrementally add new products one at a time through independent innovation events.[1] This characterization implies steady firm growth, limiting the frequency of rapid expansions and the emergence of superstar firms.

In this paper, we use newly assembled, comprehensive product-level data from French manufacturing to show that such an account of product innovation may be at odds with empirical patterns. We show that firms experience what we term *innovation bursts*, relatively rare but transformative episodes in which a firm rapidly expands its product portfolio, introducing many new products and gaining significant market share.

Such bursts of product creation may stem from innovations that simultaneously spawn multiple commercial applications. Consider, for instance, the case of Laboratoire Science & Nature (LSN), a French manufacturing firm specializing in green and natural cosmetic products. For many years after its inception in the late 1980s, LSN produced only a handful of products. Then in the mid-2010s, the firm developed a new technology to use mineral-rich seawater from the marshlands of Guérande, in western France, to create a new line of anti-aging skincare products such as deodorants, body lotions, and makeup removers. In a short span of time, LSN expanded its number of products by several folds and substantially grew both its employment and revenue.[2] Through the lens of the canonical theories of creative destruction, the story of LSN is exceedingly unlikely. Our data suggest that such bursts of product innovation are far more frequent than these theories predict.

What are the consequences of such burst-like behavior in product innovation? To answer this question, we construct a model of growth and firm dynamics that can match the observed empirical distribution of product creation, as well as a number of other facts in our data on the firm-level growth contributions of product creation, product destruction, and growth on continuing products, and on the evolution of product-level revenue within firms. Our model nests the canonical theories of creative destruction and allows us to examine the interplay between industry concentration and aggregate growth. We directly calibrate this model using the data on product-level dynamics, providing a good fit with the observed facts. Using the calibrated model, we study one example of the implications of innovation bursts for growth, showing that the relationship between firm size and innovation efficiency is much weaker in our model compared to the standard theories of creative destruction. In

[1] Since their seminal contribution, these theories have been extended in many directions to study the consequences of product innovation for many macro-level phenomena such as aggregate productivity growth, business dynamism, competition, and product-market concentration (e.g., Lentz and Mortensen 2008, Akcigit and Kerr 2018, Acemoglu et al. 2018, Cavenaile et al. 2020, Peters 2020, Aghion et al. 2023, Akcigit and Ates 2023).

[2] This information is in part based on an interview conducted in 2024 with Olivier Guilbaud, president of LSN.

contrast to the latter theories, we find that size-dependent innovation policies are far less cost-effective in spurring aggregate growth in our framework.

We begin the paper by using our product-level data from French manufacturing (discussed in detail in Section 2) to present multiple facts on product creation and destruction (presented in Section 3). First, we verify some of the core assumptions of the standard theories of creative destruction on the invariance of the rates of product creation and destruction with respect to firm size and age. Next, we show that the distribution of the number of products is highly concentrated among large firms and has a thick, Pareto-like right tail. More importantly, we also document that the distribution of newly created products also exhibits a substantial degree of concentration, with a similar Pareto-like right tail. In other words, we find bursts of product innovation, such as those exhibited by LSN, to be far from rare: a small share of firms is typically responsible for a large share of all product innovations in every year. We provide various robustness checks to ensure that what we observe in the data as bursts of new products indeed constitute episodes of meaningful product innovation, and are not the byproducts of the specifics of data construction or reporting. Together, these facts point to the dynamics of product innovation as an important determinant of the high degree of concentration of production often observed in firm-level data.

We also use our data to measure the contribution of creative destruction to revenue growth at the firm and aggregate levels, defined as the share of growth explained by changes in revenue due to the introduction of new products or the loss of old ones. This contribution, in contrast to that of revenue growth on continuing products, is more important for firms experiencing rapid growth or decline. This is in line with the important role of innovation bursts in shaping firm dynamics, and further leads to the fact that creative destruction plays a major role in explaining the observed variations in revenue growth across firms.

Finally, we examine the evolution of product-level revenue growth. The initial revenue of new products exhibits a persistent, firm-specific component. Beyond the first year, we find that growth in the revenue of a product exhibits a marked decline over the course of its life cycle. Thus, not only does product innovation grow firm revenue in the first year of the introduction of a new product, it further enables the firm to sustain a faster rate of growth in the ensuing next few years compared to the growth it can generate on a mature product.

In Section 4 we rationalize these facts in a model of endogenous growth in which firms invest in innovation to create new products or to improve the quality of their existing products. In the model, a firm's product innovation may bring about a technology with applications across multiple product lines, leading to a burst of new products. This burst-like stochastic process of product creation is consistent with the high concentration of new products in the data. It naturally nests the canonical Klette and Kortum (2004) model for the special case in which an innovation only expands firms' portfolio by a single product.

We model two additional features to help explain our facts on the evolution of product-level revenue. First, we allow for heterogeneity in innovation efficiency, to capture the observed persistent, firm-specific component of initial product revenue (Lentz and Mortensen, 2008; Acemoglu et al., 2018).

Second, we allow investments in quality upgrading on existing products to exhibit diminishing returns; successive innovations to improve quality of a product using the same firm-level technology become increasingly incremental over its life cycle (Akcigit and Kerr, 2018; Acemoglu et al., 2022). As a product matures, firms find it increasingly *harder to find ideas* to improve the production process (Bloom et al., 2020).

Innovation bursts help explain the high degrees of production concentration typically observed in firm-level data, associated with the Pareto distribution of firm size (Axtell, 2001; Luttmer, 2010). As is well known, canonical theories of firm dynamics through creative destruction can only match such high degrees of concentration if extended to allow for growth in the total number of products (Luttmer, 2011).[3] Here, we propose a fundamentally distinct driver of the concentration of production, based on the lumpy nature of innovative ideas, as reflected in the empirically observed concentrated distribution of product creation.

In Section 5, we quantify the model and assess its normative implications for innovation policy. Given the data's close links to the model, we calibrate most parameters directly from moments in our product-level data. For comparison, we similarly calibrate a conventional model of creative destruction without innovation bursts. We first confirm that, unlike the conventional model, ours can match the empirical distribution of product creation (unconditionally or conditional on the firm's product count), as well as the high degree of production concentration. Our model also fits the life cycle of product revenue growth. While untargeted, the model also reasonably matches the decomposition of revenue growth to the contributions of product creation and destruction and that of continuing products.

Finally, we use the calibrated model to study an example of the normative implications of innovation bursts. Due to heterogeneity in innovation externalities across firms, the social planner would like to reallocate R&D resources in the market equilibrium toward firms who use them more efficiently. Under the conventional model, firms with higher efficiency are far more likely to accumulate many products and grow large. Thus, a government can use observable firm size as a proxy for unobservable firm-level efficiency. Innovation bursts weaken this link: we show that the more burst-like the process, the higher the share of low-efficiency firms among large firms. Relative to the a similarly calibrated standard model, the share of these firms among those with more than 10 products rises more than 3 times in our calibrated model. As a result, we find that a size-dependent 1% R&D subsidy to product innovation for firms with more than 10 products is up to half as cost effective, in terms of growth per subsidized dollar, in our model compared to the standard models.

Related Literature. Our paper contributes to the large body of research in the tradition of the Schumpeterian growth models of creative destruction, confronting them with data on the dynamics of product churn and turnover. Most prior attempts for testing the predictions of these models regarding

[3] In our data, the total number of products is fairly stable over time. Luttmer (2011) shows, even in an environment with sustained growth in the number of products, standard models predict that the largest firms in the data may take centuries to grow to their current scales. He proposes a theoretical solution in which young firms go through a phase of fast product creation before transitioning into their long-run behavior. For other examples of similar approaches to match empirical firm-size concentration, see, e.g., Cao et al. (2017); Acemoglu et al. (2022); Peters and Walsh (2021), and for an alternative approach based on a static assignment model, see Geerolf (2017).

the dynamics of product innovation have instead relied on indirect evidence drawn from patent data (e.g., Akcigit and Kerr 2018).[4]

As a notable exception, using bar-code-level retail scanner data on non/semi-durable consumer products, Argente et al. (2024) show that product creation is a major contributor to firm growth, and that revenue growth falls within each bar code. They provide evidence that a combination of creative destruction by competitors and cannibalization by the firm's own innovation contribute to these patternwhs, and provide a model to study the implied interaction between firms' product creation and own-product innovation. Our paper is distinct from theirs in our focus on the heterogeneity in the *distribution* of product creation and the evidence in favor of innovation bursts. Our data, furthermore, covers the entire manufacturing sector, including durables, and defines products at a level broader than bar codes, ensuring clear differentiation among them. Based on our empirical results, we attribute the falling revenue growth of continuing products in our data to the possibility that ideas for improving quality *get harder to find* as products mature.

Beyond theories of firm innovation and dynamics, Bernard et al. (2010) show that product dynamics, at the level of 5-digit industry codes, are consistent with a multi-product extension of the Melitz (2003) model, while Broda and Weinstein (2010) use bar-code-level scanner data to study the implications of product dynamics for the measurement of aggregate prices.[5] Due to the wider availability of customs records, the dynamics of exported products have also been widely studied (e.g., Fitzgerald et al., 2024; Albornoz et al., 2023). Others have explored the implications of product dynamics at the business cycle frequency as a channel to amplify shocks (e.g., Dekle et al., 2015; Benguria et al., 2022).

Finally, multiple recent studies have documented the fact that industry concentration has been rising over the past few decades, in the US and globally (e.g. Autor et al. 2020; Kwon et al. 2024; Ma et al. 2024). Concurrently with the rise of large firms, there has been a gradual decline in the growth of total factor productivity, despite a sustained rise in R&D investments (Bloom et al. 2020). By constructing a theory of creative destruction that directly matches the firm-level evidence on concentration and product innovation, our paper offers a laboratory for studying the interplay between concentration, innovation, and growth.[6]

2. Data

We use detailed data on the composition of the product portfolio of firms in the French manufacturing sector. The data combines firm-level income statement and balance sheet data from tax records with a detailed survey of the product portfolio of firms, covering revenues and quantities sold at the level of detailed product categories.

[4] In a recent contribution, Argente et al. (2020) build an explicit empirical connection between firms-level patents and products and show their association is weaker among market leaders.

[5] Hottman et al. (2016) use similar data to examine the contribution of the product margin to firm heterogeneity.

[6] Recent examples of Schumpeterian growth models that relate trends in industry concentration and growth include Akcigit and Ates (2021), Olmstead-Rumsey (2022), Akcigit and Ates (2023), Aghion et al. (2023), Cavenaile et al. (2023), Weiss (2023), De Ridder (2024), and Yao (2025).

2.1. Sources

Our main source of information on firm product portfolios is the *Enquête Annuelle de Production* (EAP), which is a survey collected by the French statistical office (INSEE). The EAP surveys the universe of French manufacturing firms with at least 20 employees or 5 million euros in revenue, comprising around 90% of the value of aggregate manufacturing output.[7] The data is available from 2009 to 2019. We start our analysis in 2010, when the survey methodology was finalized. We drop firms in EAP that do not belong to the manufacturing sector, such as those in mining, repairs, and installation industries.

The survey contains revenue, quantity, and average unit values for each product category that a firm produces each year at the level of 10-digit product codes following the PRODFRA classification, which is the official classification of French products published by INSEE. The high level of detail in these codes, distinguishing among 4500 distinct products, enable us to investigate changes in each firm's product portfolio. The first six digits are the harmonized European classification of products by activity codes (CPA), which are sufficiently narrow to identify customs policies. The remaining digits contain a further sub-classification that is produced particularly for France by INSEE and that nests the 8-digit European PRODCOM classification.[8] To ensure consistency of product codes over time, in our baseline analysis, we use a modified 10c-digit product classification obtained through a concordance procedure that affects around 10% of products (see Online Appendix B.1 for further details). All main results are robust to using the original 10-digit classification, where we instead drop products with changes in codes rather than attempt to concord them.

We combine this product data with the *Fichier Approaché des Résultats d'Esane* (FARE). FARE provides firm-level income statement and balance sheet data such as total sales, wage bill, and capital, for the universe of French firms. Our baseline sample is the intersection of FARE and EAP. Additionally, we obtain the number of plants from social security filings (*Déclaration Annuelle de Données Sociales*, DADS) and determine firms' ownership status (whether independent or part of a group) using the *Liaisons Financières entre Sociétés* (LiFi) data. All data is merged using a common firm identifier, the SIREN code.

2.2. Summary Statistics

Table 1 provides the summary statistics. Due to the size cut-off for inclusion, firms in our data are typically larger and older than typical French manufacturing firms. Focusing on the product-level information, we find that the median firm in a typical year has a single product and adds or loses no product every year. The picture is somewhat different for the average firm that produces 1.93 products, adds 0.13 new products, losing 0.2.

The table does not fully reflect the high degree of concentration of the distribution of product counts. Figure 1a plots the distribution and its degree of concentration by plotting the firm's product

[7] The survey additionally includes a random sample of small firms. As this sample is redrawn every year, it is not suitable for our analysis of firm-level product dynamics, which requires panel data.

[8] For example, the 8-digit code 20421945 includes 'pre-shave, shaving and after-shave preparations', while the 10-digit codes 2042194510 and 2042194520 distinguish between lotions and 'foams and gels'.

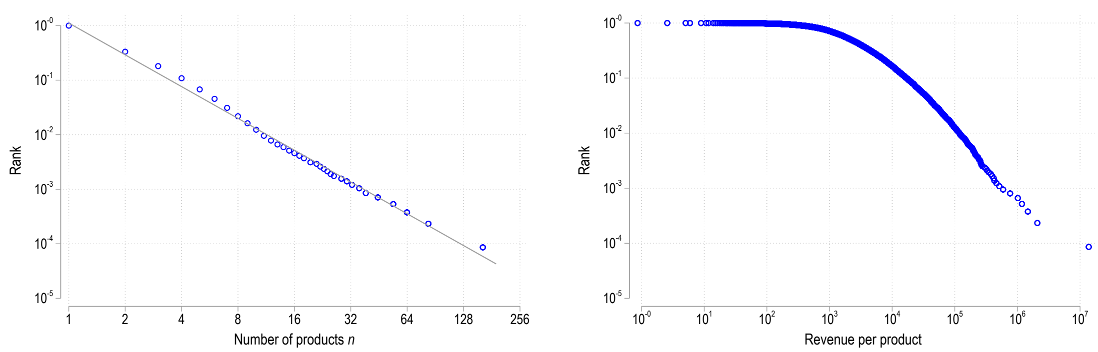
Table 1: Summary Statistics

Variable	Mean	St. Dev.	Median	10th Pct.	90th Pct.	Observations
Sales	22291	283587	3319	584	29533	223883
Age	28	19	24	8	52	223883
Employees	67	307	20	4	120	223883
Revenue per Product	9607	179085	2001	316	15635	223883
Number of Products	1.93	2.64	1	1	4	223883
Lost Products	.2	.81	0	0	1	203156
New Products	.13	.8	0	0	0	193718
Continuing Products	1.81	2.32	1	1	4	193718

Notes: Age and employment are defined in years and full-time-equivalence, respectively. Sales are 1000s of 2015 euros.

count against its ranking in the sample (logarithmic scale). The relationship between product count and its rank is approximately linear. Thus, the number of firm-level products follows an approximate power law with a Pareto-like right tail.The slope of the count-rank relationship implies that the largest 20% of firms produce 46% of all products.[9] This is in line with the widely documented fact that other measures of firm size (e.g., revenue and employment) also follow a power law and is highly concentrated among a small share of superstar firms (see, e.g. Axtell 2001, Luttmer 2007, Gabaix 2009, Garicano et al. 2016).[10] Figure 1b shows that the right tail of the distribution of revenue per product in our data also follows a power law.

Figure 1. Decomposing Firm Concentration into Products and Revenue Per Product



(a) Number of Products　　(b) Revenue per Product

Notes: The figures plot the relationship between a firm's size (horizontal axes) and the firm's rank (vertical axes). Size is either product count or revenue per product (in 1000s of 2015 euros). Rank is measured as the ratio of firms' rank starting from the largest firm, divided by the total number of observations in the data, so that rank equals the share of firms with more than n products. Plots are based on the FARE-EAP for 2019. The distribution for 2010 to 2018 is plotted in Appendix Figure C.2.

[9] A regression of the log-rank on log-firm size relationship yields an tail index of 1.94.

[10] Among firms in our sample, the largest 20% of firms are responsible for 84% of revenue and employ of 78% workers. This is closely in line with the heuristic 80/20% Pareto principle, stating that 80% of the outcome of interest typically belongs to 20% of agents. If we assume a Pareto distribution, the percentage y of the number of products that is produced by the largest x% of firms equals $y = x^{(\theta-1)/\theta}$, where θ is the tail index. For the 80/20% rule, we find a Pareto tail index of around 1.16, fairly close to the tail index of 1 implied by Zipf's law.

Table 2: Product Creation and Destruction

By Size	All	1	2	3	4-5	6-8	8+
Product creation rate	0.066	0.066	0.067	0.068	0.058	0.056	0.083
Product destruction rate	0.082	0.065	0.111	0.117	0.117	0.110	0.129

By Age Bins	All	0-5	5-10	10-15	15-20	20-25	25-50	50+
Product creation rate	0.066	0.075	0.061	0.063	0.059	0.064	0.068	0.074
Product destruction rate	0.082	0.094	0.090	0.080	0.080	0.080	0.079	0.081

Notes: Product creation: number of products that a firm starts producing divided by its original number of products. Product destruction: number of products that a firm stops producing divided by its original number of products. Different columns report the results conditional on the firm's number of products in preceding period n_{t-1} (upper panel) or age bins (lower panel).

3. Stylized Facts on Product Creation and Destruction

In this section, we use our data to document a number of facts on the dynamics of firm-level product portfolios and how they shape the dynamics of firm revenue growth across firms and within each firm's life-cycle.

3.1. How Do the Rates of Product Creation/Destruction Vary with Firm Size/Age?

We begin our investigation by testing one of the key predictions of the standard theories of creative destruction on the relationship between the rates of product creation, destruction, and firm size. Following Klette and Kortum (2004), these theories predict that the firm-level rates of product creation and destruction are independent of a firm's product count. In addition, they posit that product destruction occurs in an undirected and random fashion across all existing products. Together, these predictions yield a strong form of Gibrat's law: the growth rate of a firm's product portfolio is independent of its size.

Empirically, we define the product creation rate as the number of product codes for which a firm earns revenue at time t but not at $t-1$ divided by the number of product codes for which the firm earns revenue at $t-1$. The product destruction rate is the share of the firm's products at $t-1$ for which it does not report revenue at t.

Table 2 presents average product creation and destruction rates. We present unconditional averages and averages conditional on the number of products (top panel) and firms' age (bottom panel). The first row displays the product creation rate, the bottom row gives the destruction rate. We find that firms on average add 0.066 products to their portfolio for every product that they initially produce, while they stop producing 0.082 products for every such product. The gap between incumbents' product creation and destruction rates means that the number of products that an incumbent firm produces on average shrinks over time. That is expected when new firms enter the economy, while the economy's total number of firm-products is approximately constant over time–which is the case in our data. The gap between product creation and destruction rates thus quantifies the contribution of entrants to product creation.

Turning to the results by initial product count, the table shows no particular pattern for product creation rates across firms with different initial sizes: the rate at which small and large firms expand their portfolio is similar, ranging from 0.056 to 0.083 per products initially produced. This means that, in line with the assumptions of Klette and Kortum (2004) and the following literature, the expected number of new products firms add to their portfolios in expectation scales linearly with the sizes of their portfolios. The rate of product destruction is similar for all firms with more than two products, but it is lower for firms that produce a single product. The unusually low rate among single-product firms is likely to stem from selection in the data: the EAP has a size threshold, implying that the single-product firms in the data earn higher revenue on their product compared to other firms.

Across different age groups, no systematic pattern emerges in firm creation or destruction rates. The youngest firms (aged 0-5 years) exhibit slightly higher rates of churn than firms of neighboring ages, but their average product creation rates are only marginally higher than those of the oldest firms (aged 50+ years).

3.2. Innovation Bursts and the Concentration of Product Creation

We next investigate the distribution of product creation across firms.[11] As mentioned in the introduction, all models building on the innovation process proposed by Klette and Kortum (2004) predict that, conditional on the firm's product count, the number of newly created products follows a distribution with a thin right tail. Thus, it is highly unlikely for firms to experience rapid product innovations in these theories.

Figure 2 plots the distribution of a firm's number of new products. We plot the relationship between the firm's number of new products and its rank in that distribution, as we did for firm size in Figure 1. Figure 2(a) plots the distribution across all firms, showing that it is is highly concentrated. The number of new products follows a Pareto-like distribution, with the linear regression having a near perfect fit.[12]

In Figure 2(b,c,d), we plot the distribution for the number of products, but now condition on firm's initial size. Doing so is useful as we have already seen in Figure 1 that the cross-sectional distribution of firms' product count is highly concentrated. That, in combination with the fact that product creation rates in Table 2 were stable in size, could suffice to make the distribution of product creation resemble the distribution in Figure 2(a). Conditioning on size thus answers the question of whether the concentration of product creation is merely a reflection of the cross-sectional concentration of firms' product counts.

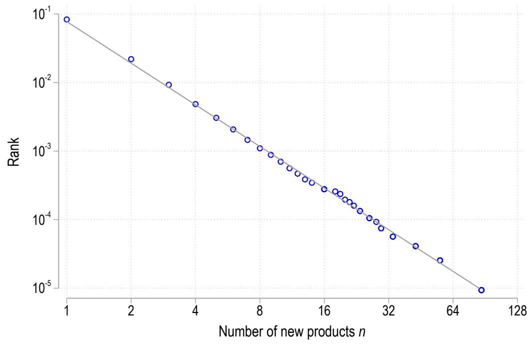
Figure 2(b,c,d) show, to the contrary, that the high degree of concentration in product creation persists when we *condition on the firm's initial product count.*[13] Even among firms that initially produce only a single product, some add tens of new products to their portfolio. In other words, firms

[11] We focus on creation rather than destruction because the number of products that firms lose is bounded by the number they produce.

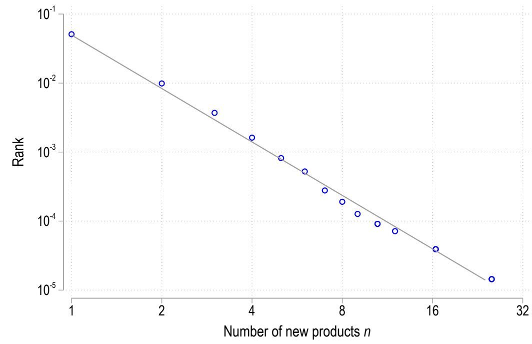
[12] The distribution in Figure 2(a) implies that 76% of product creation comes from just 20% of firms.

[13] Note that the slope of the log-rank, log-new products relationship in Figure 2 only accounts for firms that add at least one product to their portfolio. A large share of firms (66%) report no product creation at all. For single-product firms concentration is even more extreme, with all product creation coming from just 7% of firms.

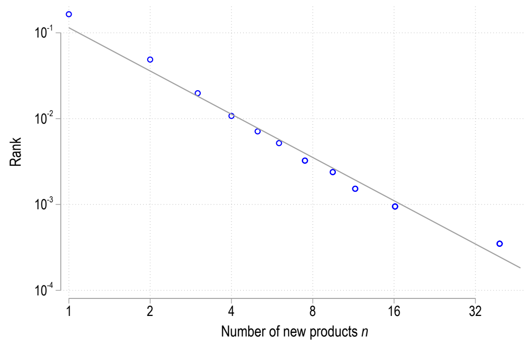
Figure 2. Distribution of Number of New Products



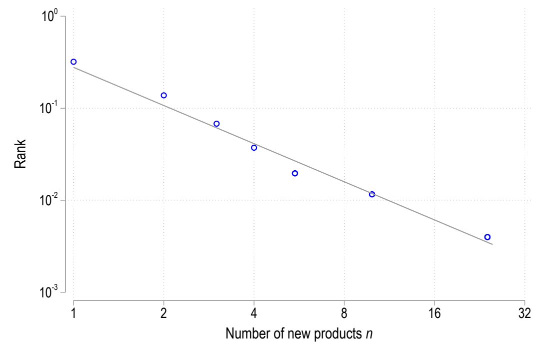
(a) All Firms



(b) Firms with $n_{t-1} = 1$



(c) Firms with $n_{t-1} = 5$



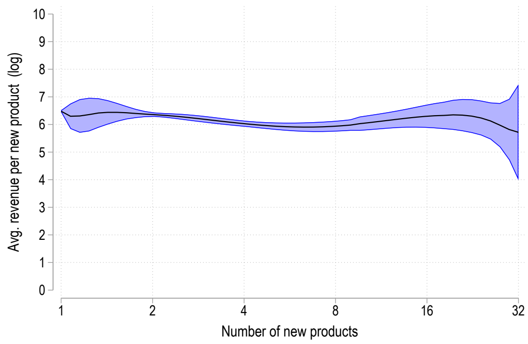
(d) Firms with $n_{t-1} = 10$

Notes: The figures plot the relationship between a firm's number of *new* products (horizontal axes) and the firm's rank (vertical axes), conditional on the firm's product count in the prior period (n_{t-1}). The (rescaled) rank is measured as the ratio of firms' rank starting from the largest firm, divided by the total number of observations in the data. Linear reference lines exclude the observation with the greatest number of added products; certain data points are combined to respect confidentiality rules.

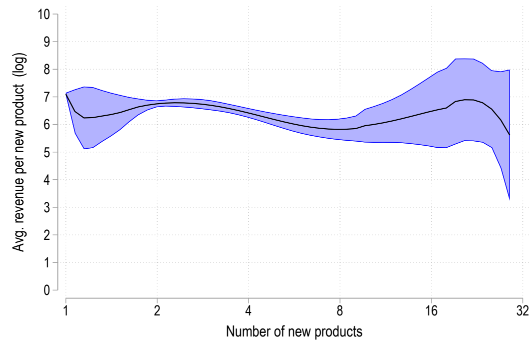
can experience sudden bursts of product creation that, as we will see, are exceedingly unlikely under benchmark theories of creative destruction. This pattern is widespread across firms with any initial number of products in our data. Figures 2(c) and 2(d) show that a sizable share of firms with initially 5 and 10 products also experience bursts of product creation.

Firms that experience an innovation burst also grow rapidly in terms of sales. Figure 3 uses the product-level revenue in the EAP data to plot average revenue earned *on new products* against the firm's number of new products. The left-hand figure plots this for all firms, the right-hand figure conditions on size by only including single-product firms. The flat line implies that a firm that has created, say, 10 products, will on average earn 10 times more revenue from newly created products than a firm that created a single product. Figure C.3 in Appendix C shows the same pattern holds for firms starting with 5 and 10 products. If innovation bursts were an artifact of some firms classifying similar products in many different product classification codes, one would expect average revenue per new product to fall sharply in the number of new products. This would also be the case if the new products were close substitutes. To the contrary, we find that average revenue per product is similar across firms reporting different numbers of new products.

Figure 3. Average Revenue Per New Product



(a) All Firms



(b) Firms with $n_{t-1} = 1$

Notes: Figures plot the logged ratio of total revenue earned on new products divided by the total number of new products on the vertical axis, against the number of new products on the horizontal axis. The line represents a kernel-weighted local polynomial smoothing of degree 5, together with its 95% confidence band.

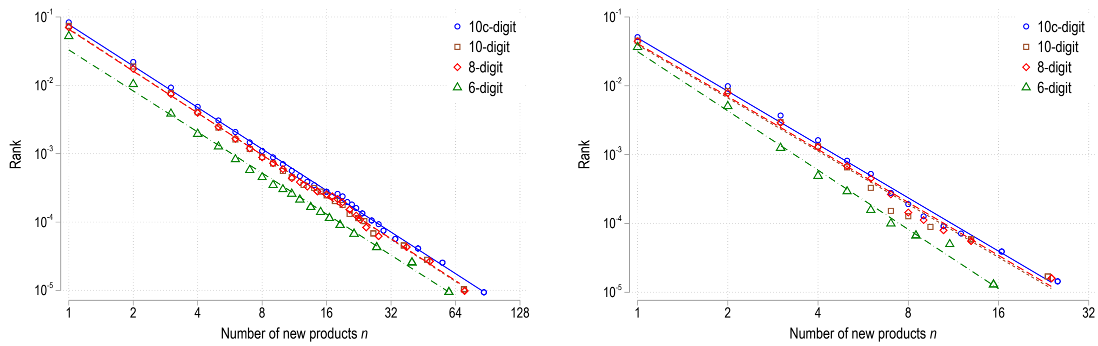
We next present various robustness checks to ensure that innovation bursts are indeed episodes in which firms expand their product portfolio. We show that innovation bursts are visible at different levels of aggregation, reflect organic growth rather than acquisitions, and do not cannibalize a firm's existing products. Moreover, we show that bursts are prevalent among firms of different ages and innovation histories, and across industries.

Level of Aggregation Figure 4 shows that the distribution of product creation is similar if we define a "product" at different levels of aggregation. If innovation bursts are merely driven by the accuracy with which different firms classify and report their revenue across product codes, the distribution should become less concentrated if we define products at higher levels of aggregation. The figure shows that this is not the case: the distribution of the number of new products is similar at the 6-digit, 8-digit and the 10c-digit level. Figure C.4 (Online Appendix C) presents similar results conditional on producing 5 and 10 products. We thus conclude that innovation bursts are episodes of rapid firm expansion, rather than an artifact of the way that some firms report changes to their product portfolio.

Mergers and Acquisitions We also verify that innovation bursts are driven by organic growth rather than mergers and acquisitions (M&A). If a firm's product portfolio changes due to the latter, those changes are unlikely to involve true product creation.

As M&A is difficult to track in the data, we use three different approaches to proxy for firm boundaries, two of which are plotted in Figure 5. As M&As are more prevalent among business groups, Figure 5(a) considers the set of single-product firms that do *not* belong to business groups; since M&As are likely to change the number of establishments a firm owns, Figure 5(b) considers the set of single-product firms that do *not* add any new plants between the two periods under consideration. In both cases, we find that the distribution remains fairly similar to our baseline, suggesting that innovation bursts are not likely to be driven by M&A. Appendix Figure C.5 show that this is also the case for firms

Figure 4. Distribution of Number of New Products by Level of Aggregation



(a) All Firms

(b) Firms with $n_{t-1} = 1$

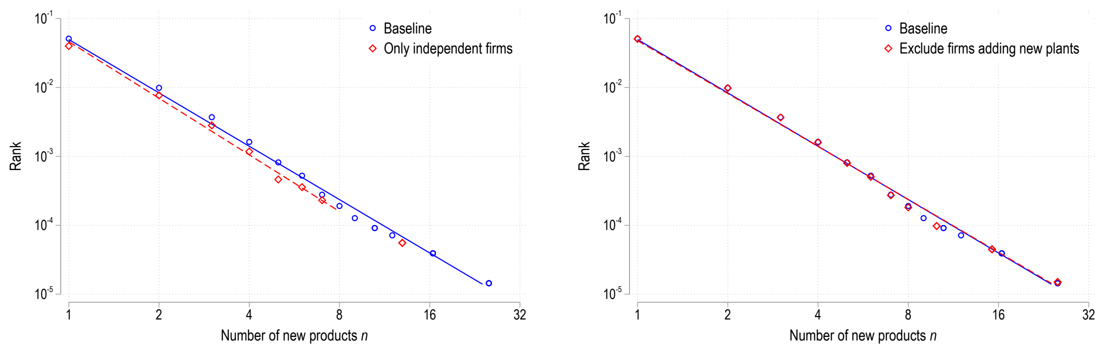
Notes: The figures plot the relationship between a firm's number of new products (horizontal axes) and the firm's rank (vertical axes). The figure provides overlapping plots of the log rank against the log number of products added at the 10c-digit concorded product level (blue circles), 10-digit PRODFRA level (brown squares), 8-digit (red diamonds) and 6-digit level (green triangles).

of different sizes. Finally, we are also able to track whether firms experience a change in intangible capital. Intangible capital's chief component is goodwill, which rises when firms acquire a competitor. Appendix Figure C.6 shows that, when excluding firms with changes in intangibles, the distribution or product creation remains almost unchanged.

Cannibalization We next show that innovation bursts have limited effects on firms' existing products. If bursts cannibalized firms' other products, firms that expand their product portfolio would be more likely to cease producing their older goods, or earn less revenue on them. Argente et al. (2024) find evidence for such cannibalization in the barcode-level, nondurable, convenience consumer goods using scanner data. In our data, we only find modest cannibalization. Appendix Table C.3 shows that the probability that a firm stops producing an existing product increases by only 8% if its product creation rate increases by 100%. Full cannibalization would imply an increase in those odds of 100%. In Section 3.4, we further show that bursts have little effect on per-product revenue of continuing products.

Innovation Bursts by Firm Age, Innovation History, and Sector Online Appendix C presents various graphs to show that product creation is concentrated in bursts for firms of different ages, innovation histories, and sectors. Figures C.7 show that the slope of the log-rank, log-number of new products distribution of firms with below or above-median age is similar. Figure C.8 finds the same for firms that innovated in the previous period and firms that did not. Moreover, we do not find any evidence for autocorrelation in bursts. That is, firms adding many products in their previous episode of product creation are not more likely to be in the tail of product creation the next time they add products. Turning to differences across sectors, although our data limits us to the manufacturing sector, we find that innovation bursts are prevalent across different broad industries within this sector. Figures C.9 and Figure C.10 offer two distinct ways to determine the most prominent broad-industry affiliation corresponding to different positions within the distribution of product creation. The figure shows that

Figure 5. Distribution of Number of New Products by Proxies of Fixed Firm Boundary ($n_{t-1} = 1$)



(a) Independent Firms Only

(b) No Plant Change

Notes: The figures plot the relationship between a firm's number of new products (horizontal axes) and the firm's rank (vertical axes). The (rescaled) rank is measured as the ratio of firms' rank starting from the largest firm, divided by the total number of observations in the data. The two panels plot the baseline Figure 2(b) against the distribution of new products of firms that do not belong to a business group (left panel) and of firms that do not add new plants over the period (right panel).

all industries are represented across the distribution; hence, we conclude that innovation bursts are broad-based and not limited to specific industries.

3.3. How Important Is Product Creation/Destruction for Firm Growth?

So far, we have focused on how firm product portfolios evolve through product creation and destruction. In this section, we use our data to study the contribution of such churn to firm dynamics, by decomposing firm-level revenue growth to the distinct contributions of product creation/destruction and the growth in the revenues of existing products.

We measure firm growth using the symmetric growth rate (Davis et al. 2006), which divides the change in size between period t and $t-h$ by average size across both periods. In contrast to changes in log size, this growth rate can be calculated for continuing firms as well as firms that enter or exit, which is why it has become a standard measure in work on firm dynamics. We decompose the growth in revenue of firm i form time $t-h$ to t as

$$\frac{R_{it} - R_{it-h}}{\frac{1}{2}(R_{it} + R_{it-h})} = \frac{R^N_{it}}{\frac{1}{2}(R_{it} + R_{it-h})} - \frac{R^L_{it-h}}{\frac{1}{2}(R_{it} + R_{it-h})} + \frac{R^-_{it} - R^+_{it-h}}{\frac{1}{2}(R_{it} + R_{it-h})}. \tag{1}$$

The first two terms capture growth in revenue due to product creation and destruction, the third term captures growth on products that firms continue to produce. Here, R^N_{it} denotes revenue on products produced at time t but not at $t-h$; R^L_{it-h} is revenue that the firm earned at $t-h$ on products that it stopped producing between $t-h$ and t; R^-_{it} is the revenue that the firm earns at time t on the products that it was already producing at $t-h$, and R^+_{it-h} is the revenue earned at $t-h$ on this same set of continuing products.

We quantify the contribution of product innovation and destruction in Table 3. In the top panel, we first present the average size of each term over a single and a five-year horizon. Observations are weighted by the denominator on the left-hand side of equation (1) to measure contributions to aggre-

Table 3: Contribution of Product Innovation and Destruction to Revenue Growth

	Overall Growth	Product Innovation	Product Destruction	Continuing Products
Mean value				
1-year	0.004	0.022	-0.024	0.006
5-year	0.027	0.076	-0.075	0.026
Shapley-Owen Contribution				
1-year	100.0	16.0	55.2	28.8
5-year	100.0	37.4	49.7	12.9

Notes: The table decomposes revenue growth into revenue loss from product loss, revenue gain from product gain, and changes in revenue on products the firm continues to produce. Revenue is deflated by the GDP deflator. Observations in the upper panel are weighted by denominator $0.5(R_{it} + R_{it-h})$.

gate revenue growth. The first row shows that, even though each of the two creation/destruction have sizable effects, their net effect is small.

This simple exercise may thus suggest from the table that the net effect of product creation and destruction, henceforth labeled "creative destruction," is small or even negative.[14] However, this average result masks substantial heterogeneity across firms in the growth contribution of creative destruction. To show this, the bottom panel of Table 3 measures how much of the variation in revenue growth is due to creative destruction. It does so using the Shapley value, which quantifies the marginal contribution of each component in decomposition (1) to the R^2 in regressions on all possible combinations of the components (see, e.g., Ozkan et al., 2023). The table shows that creative destruction explains over 70% (87%) of the variation in revenue growth over 1-year (5-year) horizons.[15] A variance decomposition as performed in Hottman et al. (2016) yields quantitatively similar results.

In Figure 6, we show that this large contribution of creative destruction is due to its effect on the tails of firm revenue growth, Following the the decomposition suggested first by Garcia-Macia et al. (2019), we split the sample of firms into 40 equally sized bins based on the firms' revenue growth.[16] The top figure quantifies the importance of each bin for explaining the aggregate revenue growth. Moving from zero to the right side, the figure plots the ratio of the total change in revenue across the firms in each bin by the total change in revenue across all firms with positive growth. Thus, the height of the bars accounts for the contribution of firms in each bin to the overall growth of all growing firms. Moving from zero to the left, we similarly report the ratio of the total fall in the revenue of all firms within each bin to the overall fall in the revenues of all shrinking firms.

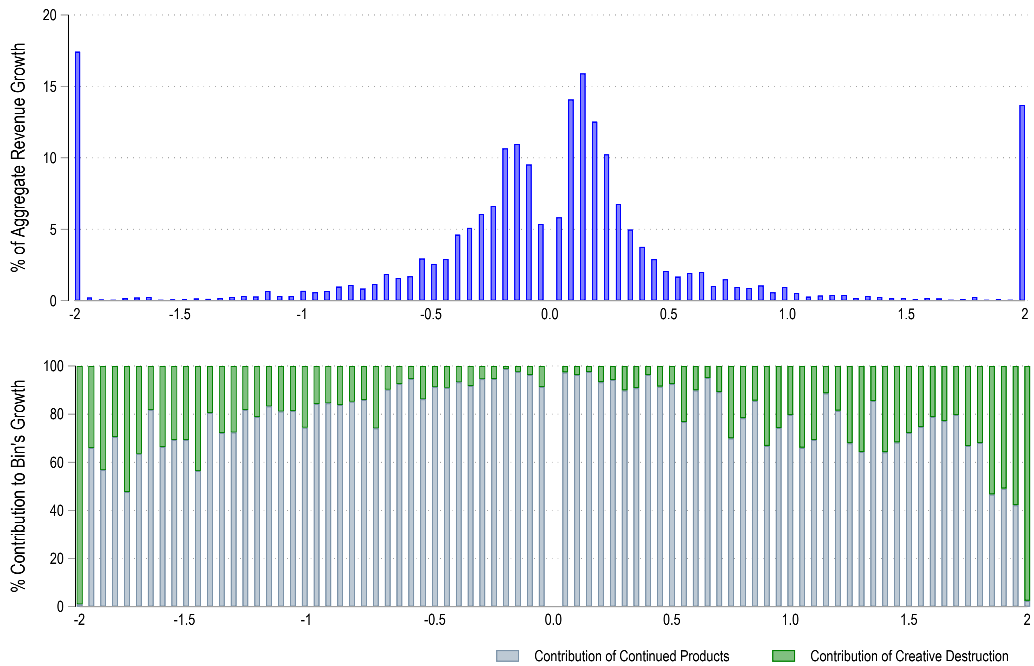
The figure shows that 18% of revenue destruction comes from firms with revenue growth between -1.95 and -2, while 14% of revenue creation comes from firms with revenue growth between 1.95 and 2.

[14] Note that we exclude from the decomposition firms in their first year in the product-level EAP data if they have already appeared in the previous years of the comprehensive FARE dataset. As a result, innovation bursts that push firms over the size threshold to feature in the EAP are therefore excluded from the data. Thus, our results are likely to underestimate the importance of product innovation for revenue growth.

[15] The contribution of product creation is more similar to product destruction over the longer horizon, which is likely because the sales of a new product does not reach its full scale in the first year. For example, if firms start producing new goods uniformly over the year, revenue for new products will on average only capture half a year of sales (Bernard et al., 2010) This lowers the measured contribution of product innovation to revenue growth.

[16] Garcia-Macia et al. (2019) use this approach to account for the contribution of firms with different growth rates to overall job creation and job destruction.

Figure 6. Creative Destruction and Aggregate Revenue Growth



Notes: The horizontal axis measures firm growth through the symmetric growth rate, defined as the change in revenue between t and $t-1$ divided by average revenue in t and $t-1$. Growth rates are separated into 20 negative bins and 20 positive bins. The top figure presents the contribution of changes in revenue across firms in a particular growth bin as a percentage of total revenue creation (the sum of increases in revenue across growing firms) for positive bins or as a percentage of total revenue destruction (the sum of decreases in revenue across shrinking firms) for negative bins. The bottom panel decomposes a bin's overall revenue change into changes coming from continuing products and the net of product innovation and destruction – creative destruction.

While most changes in overall revenue growth originate from the many firms that have modest growth rates, the tails of revenue growth still matter in the aggregate. In the bottom figure, we decompose a bin's total change in revenue into changes in revenue for continuing products (grey) and the net change from creative destruction (green). As we move toward the tails of the distribution, the contribution of creative destruction gradually rises. In the extreme tails, the lion's share of the growth comes from creative destruction.

In Online Appendix D, we show that the contribution of creative destruction to overall revenue growth is similar for firms of different age and size, although the contribution is (by definition) exhaustive for entrants.

Concentration of Revenue Growth We have so far shown that creative destruction plays an important role in the tails of firm growth and that product creation is highly concentrated. A natural question is whether overall revenue growth is also highly concentrated and well described by a power law. Online Appendix D confirms that the distributions of revenue and employment growth are much more concentrated than a log-normal distribution would predict. The log-normal distribution is a natural starting point as, motivated by Gibrat (1931), firm dynamics models with random shocks to productivity typically assume that firm growth is log-normal (see, e.g., Hopenhayn 2014).[17] The appendix also shows that the fat-tailed distribution of revenue growth is visible for firms of different age and size.

3.4. Does Creative Destruction Matter for the Evolution of Product-Level Revenue?

In Section 3.3 we saw that while creative destruction is key for the two tails of firm growth, the major share of the overall revenue growth comes from products that firms continue to produce. In the last part of our empirical investigation, we use our data to study how creative destruction matters for the evolution of product-level revenue. We first show that the initial revenue of new products is driven by a firm-level factor, and then show that, after each instance of product creation, the growth in the product-level revenue subsequently falls over the course of the product life cycle.

3.4.1. Revenue of Newly Created Products Are Auto-correlated at the Firm Level

What determines the initial revenue that a new product brings to an innovating firm? Standard models of creative destruction (e.g., Lentz and Mortensen, 2008; Acemoglu et al., 2018) often assume ex-ante heterogeneity in innovative capacity of firms, which influences how much their new products upgrade the quality of the prior state-of-the-art techniques. Through the lens of these models, the initial revenue brought about by a new product reflects this degree of upgrading. Thus, we begin by studying whether the initial product-level revenue is auto-correlated for firms that experience multiple episodes of product creation.[18]

For firms that experience at least two episodes of product creation, Figure 7 shows the binscatter plot of the average revenue per new product in the current period against the average revenue per new product in their the previous episode of product creation. We take out 10-digit product fixed effects. The left-hand panel shows that on average, revenue on new products is higher for firms whose past instances of product creation had brought in relatively more revenue. The coefficient from a corresponding linear regression is 0.39 and is highly significant. Moreover, the positive relationship disappears when we additionally control for firm fixed effects in the right-hand panel, suggesting that the driver of the persistence is indeed a firm-level factor. This is in line with the existence of ex ante firm-level heterogeneity in the quality improvements embodied in product creation.

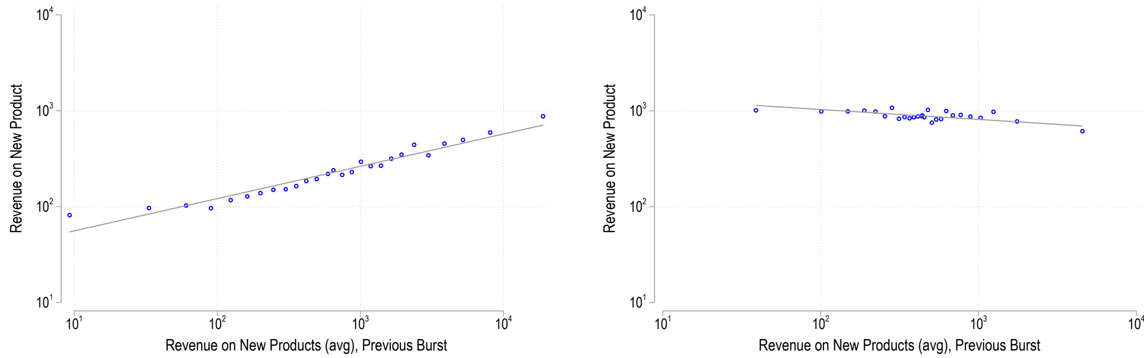
3.4.2. Product-Level Revenue Growth Falls over the Product's Life Cycle

Lastly, we show that the growth of revenue of a continuing product gradually falls over the course of its life cycle. Thus, by introducing new products that have higher potential for within-product quality improvements, creative destruction makes an indirect contribution to firm-level revenue growth on continuing products.

[17] Similar patterns have recently been documented using ORBIS (Jaimovich et al., 2023) and Compustat (Melcangi and Sarpietro, 2024) firm-level data. Our results echo recent findings in the literature on income dynamics which shows that individual income growth is also better described by a fat-tailed distributions (see, e.g., Guvenen et al. (2021). Gabaix et al. (2016) discuss the importance of the income growth process to explain stationary distributions of income and wealth.

[18] Since the rate of product creation is fairly small, it is difficult to use the data to uncover systematic firm-level differences in the rates of product creation given the limited time span of our sample.

Figure 7. Firm-Level Persistence in Average Revenue from New Products



(a) Without Firm Effects

(b) With Firm Effects

Notes: The figure presents a binned scatter plot. Horizontal axes give average revenue per new product that an innovating firm earned the previous year that it added products to its portfolio. Vertical axes gives revenue on a product that the respective firm is producing for the first time in that year. Both axes are on the log scale. We demean the data using 10-digit product fixed effects, and additionally by firm fixed effects on the right-hand figure. Appendix Table C.2 gives the regression table for the linear fits.

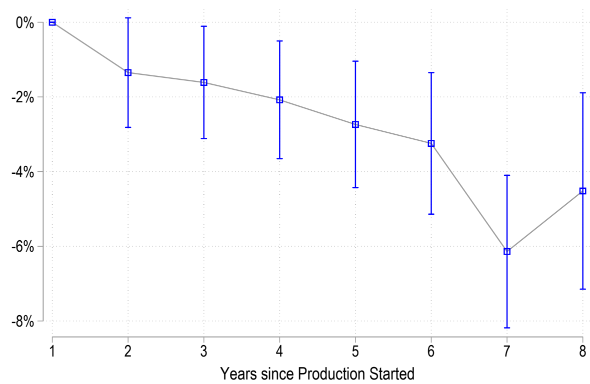
Figure 8 plots the relationship between revenue growth and product tenure—the time since a firm began producing a product. Before calculating average growth by tenure, we subtract each firm-product's average revenue growth over its observed life cycle to avoid selection bias: revenue growth may influence how long a firm keeps producing a product. Thus, the figure shows how a firm-product's revenue growth deviates from its life cycle average as tenure increases. Each point represents the average revenue growth at a given year of tenure, relative to growth at tenure 1, and after absorbing firm-product fixed effects.

The figure shows a significant decline in revenue growth over a product's life cycle, averaging a decrease of about 0.5 percentage points per year. The decline of revenue growth over the life cycle of a product means that creative destruction and revenue growth on continuing products interact. For firms to sustain revenue growth on their products, they must improve the quality or cut production costs of the products that they sell. When firms begin to produce a new product, their revenue growth is rapid and exceeds the average growth over the entire spell of their product. As time passes, revenue growth falls over the product's life cycle. Thus, it appears that the firm-level opportunities to sustain growth through the upgrading of own products decline as the product matures. This means that creative destruction offers two benefits: besides directly contributing to revenue growth of the firm engaging in product creation, it also enables a sequence of follow-up innovations.

We should note that the falling growth over the life cycle in Figure 8 appears not to be driven by cannibalization as a side effect of product creation in our data.[19] The figure is similar for products owned by firms that engage in product creation over the horizon, for example, and firms that do not

[19] As mentioned in the introduction, Argente et al. (2024) document a similar pattern to Figure 8 at the bar-code level using scanner data from the US. They find that around two-fifths of the life cycle pattern is driven by cannibalization. Our finding that cannibalization plays a minor role might be due, in part, to the fact that we look at 10-digit product codes rather than bar codes.

Figure 8. Life Cycle of Revenue Growth



Notes: Vertical axis: avg. product revenue growth relative to growth at tenure 1 and after absorbing firm-product fixed effects. Horizontal axis: years since firm started producing the product. Confidence bounds (90% level) are based on clustered std. errors. To avoid mismeasuring tenure as a result of the concordance procedure, the figure is constructed using the original 10-digit PRODFRA codes. Figure C.11 shows similar results using the 10c-digit concorded product codes.

(Appendix Figure C.12). A more formal analysis in Appendix Table C.4 shows that product creation has limited effects on a firm's existing products.

Conditioning on Size, Age or Sector The decline in revenue growth in a product's tenure is present across firms of different initial size, age, and industry. In Online Appendix Figure C.13 and C.14, we respectively show that revenue growth declines in product tenure for firms across these different classes of firms. In Appendix Figure C.15, we further show that the revenue decline in tenure also similarly appears across diffrent Broad Economic Classification (BEC) classes of industries in our data.

4. The Model

In this section we describe our parsimonious model of creative destruction through innovation bursts that rationalizes the facts documented in the previous section. We start by describing the static and dynamic optimization of the firm, and then characterize the model's stationary balanced growth path equilibrium.

4.1. Households, Firms, and the Aggregate Econonmy

An infinitely-lived representative household has log utility over consumption C_t, which is discounted at a rate ρ over time. Time is continuous and indexed by t, which is omitted when convenient. The household supplies a single unit of labor inelastically. Consumption is a CES aggregate of a continuum of differentiated goods with a constant measure:

$$C = \left[\int_0^1 \left(\sum_{i \in I_j} q_{ij} y_{ij} \right)^{\frac{\epsilon-1}{\epsilon}} dj \right]^{\frac{\epsilon}{\epsilon-1}} \tag{2}$$

where $\epsilon > 1$ is the elasticity of substitution, goods are indexed by j, firms are indexed by i, and I_j denotes the set of firms that own the technology to produce good j at quality q_{ij}.

Firm i is defined by the combination of the products for which it owns the leading level of quality as well as its innovation efficiency. Firm i can produce the quantity of output y_{ij} in product j that it produces with a production function $y_{ij} = l_{ij}$, where l_{ij} denotes production labor, hence marginal costs are equal across products and firms. All firms that can produce each j compete à la Bertrand, but subject to an infinitesimally small cost for operating in that market. Given the identical marginal costs across firms, this means that only the firm with the highest-quality technology to produce j will enter and produce in equilibrium, charging the monopoly markups $\frac{\epsilon}{\epsilon-1}$, as in Acemoglu et al. (2018).

Letting the aggregate price index of the households be the numeraire, the demand for product j from the firm i with the highest quality is $y_{ij} = C\, p_{ij}^{-\epsilon}\, q_{ij}^{\epsilon-1}$ where C denotes aggregate consumption and p_{ij} the product price. As we saw this product price is given by a constant markup $\epsilon/(\epsilon-1)$ over marginal cost, which equals the wage rate w. Thus, product-level demand is an isoelastic function of quality:

$$y_{ij} = Y\left(\tfrac{\epsilon}{\epsilon-1}\, w\right)^{-\epsilon} q_{ij}^{\epsilon-1}, \tag{3}$$

where we have also used that aggregate consumption C in equilibrium equals output Y.

Defining aggregate productivity Q as the ratio of output to production labor, we have

$$Q^{\epsilon-1} = \int_0^1 \int q_{ij}^{\epsilon-1}\, di\, dj. \tag{4}$$

The equilibrium wage is a constant markdown over aggregate productivity, $w = \left(\frac{\epsilon-1}{\epsilon}\right) Q$. Substituting this result in Equation (3) allows us to write the profit of firm i on its sales of product j as a function of product quality relative to aggregate productivity $\widehat{q}_{ij} \equiv q_{ij}/Q$ of each product relative to the economy's productivity index $\pi(\widehat{q}_{ij}) = \frac{1}{\epsilon}\, \widehat{q}_{ij}^{\epsilon-1}\, Y$. Thus, a firm with a product portfolio J_i has total profits $\frac{Y}{\epsilon} \sum_{j\in J_i} \widehat{q}_{ij}^{\epsilon-1}$.

As we will see below, incumbent firms and potential entrants also hire workers to conduct research and development (R&D). Defining total production labor L as all labor used to produce goods, $L = \int_0^1 \int l_{ij}\, di\, dj$, labor market equilibrium requires $L = 1 - L^e - L^{rd}$ where L^{rd} and L^e denote labor involved in R&D by incumbents and entrants, respectively.

4.2. Innovation Bursts

Firms expand their portfolio of products through innovation bursts that occur according to a stochastic Poisson process with endogenous arrival rates. Each burst generates an "idea" that enables the production of one or more products the firm has not previously offered, each drawn randomly from the continuum of product lines in the economy. Each of these new products surpasses the quality of the prior incumbent in the corresponding product line. Large innovation bursts are breakthroughs with applications across a broader range of technologies, allowing a firm to advance the frontier quality of multiple product lines using a single idea. However, such bursts are uncommon; more often, firms develop innovations that enable them to add very few products to their portfolio.

Formally, we model each innovation burst as a set with n_i^c elements, where each element corresponds to a randomly chosen, newly created product $j \in [0,1]$ with a corresponding quality improvement $\lambda_{b,ij} > 1$ over the product line's incumbents' quality q_{-ij}, such that the quality q_{ij} at which the innovator produces its new products is given by $q_{ij} = q_{-ij}\,\lambda_{b,ij}$.

The number n_i^c of such new products in the burst is a random variable that follows a Zeta distribution, with the probability mass function given by

$$P(n_i^c = n) = \frac{n^{-\theta}}{\zeta(\theta)}, \quad \text{where} \quad \zeta(\theta) = \sum_{i=1}^{\infty} \frac{1}{i^{\theta}}, \tag{5}$$

where $\zeta(\cdot)$ is the zeta function. The Zeta distribution is the discrete counterpart of the Pareto distribution. Tail parameter $\theta > 1$ determines the thickness of the tail of innovation bursts. As θ approaches one, innovation bursts have a Zipf's distribution. As θ increases, innovation bursts become less dispersed. The average number of product innovations in a burst is given by the ratio of $\zeta(\theta-1)$ and $\zeta(\theta)$, which is finite as long as θ exceeds two. In the limiting case where θ approaches infinity, every burst has a single product, $n_i^c = 1$, and the model reduces to the standard process of Klette and Kortum (2004).

To achieve a particular Poisson arrival rate $x_{b,i} \geq 0$ of innovation bursts, firm i must hire R&D researchers $z_{b,i}$. The relationship between the arrival rate of bursts and the number of researchers is given by

$$z_{b,i} = \eta_b\, x_{b,i}^{\psi}\, n_i^{-(\psi-1)\sigma}, \tag{6}$$

where n_i denotes the firm's current product count. The number of researchers that the firm employs is convex in the rate of innovation ($\psi > 1$) and declines in the number of goods that the firm produces ($\sigma > 0$). The former implies diminishing returns to R&D and the latter implies that current access to more leading technologies makes firms more productive innovators. Following Klette and Kortum (2004), this assumption reflects the idea that a firm's product count also proxies for its knowledge or organizational capital, making its innovation activity more productive. As such, the arrival rate of innovation bursts will increase with product count. As in Akcigit and Kerr (2018), we allow the extent to which innovation scales in size to be governed by σ. The standard case in which the rate of product creation scales linearly with product count corresponds to the case of $\sigma = 1$.

4.3. Firm Heterogeneity, Entry, and Exit

In addition to the heterogeneity in the number and qualities of their products, firms are also heterogeneous in terms of the efficiency with which they use R&D researchers to generate innovation bursts. We characterize this heterogeneity through variations in types h of the distribution $H_{b,h}(\lambda)$ from which firm i draws the quality improvement $\lambda_{b,ij}$ embodied in its new product j (Lentz and Mortensen, 2008). Firms that belong to a higher innovation type h on average achieve higher quality improvements in their innovation bursts, in the sense that the average moment $\overline{\lambda}_{b,h}$ defined as $\overline{\lambda}_{b,h} \equiv \mathbb{E}_{b,h}[\lambda^{\epsilon-1}]^{1/(\epsilon-1)}$, is increasing in h. Since subsequent innovators improve the quality of the product in proportion to its current quality, these firms also generate a greater positive externality to other firms. We assume that the firm's innovation efficiency is a fixed, firm-level characteristic.[20]

Firms endogenously enter and exit the economy. There is a unit measure of potential entrants that hire researchers to create a new product. To achieve an innovation arrival rate x_e, the potential entrant must hire $z_e = \eta_e x_e^{\psi}$ researchers. These entrepreneurs are ex-ante identical: they know the probability $G(h)$ that they will have innovation efficiency type h but only learn about their type after creating their new product.

Firms stop producing a product when a different firm comes up with a higher-quality version of that product. The rate at which such displacement occurs is the rate of creative destruction, which is endogenously determined by the rate of entry and the rate of innovation bursts by incumbents. As we will see later when characterizing the behavior of firms along a stationary equilibrium path, firm i chooses its rate of innovation bursts as a function of the two sources of heterogeneity: its innovation efficiency h_i and its product count n_i, so that $x_{b,i} = x_{b,h_i}(n_i)$. Accordingly, the rate of creative destruction is given by

$$x_d = x_e + \frac{\zeta(\theta-1)}{\zeta(\theta)} \sum_{h,n} M_h(n)\, x_{b,h}(n), \tag{7}$$

where $M_h(n)$ is the measure of firms with efficiency h and product count n. Creative destruction due to incumbents, the second term on the right, equals their rate of innovation bursts times the expected number of new products $\zeta(\theta-1)/\zeta(\theta)$ per burst. As usual, firms exit when innovation by other firms causes them to cease producing their sole good.

4.4. Firm Innovation on Own Products

In addition to expanding their product portfolio through innovation bursts, firms also invest in innovation to upgrade the quality of the products currently in their portfolios. Quality raises demand, so higher quality raises revenue on firms' existing products.

To engage in own-product innovation, firms make targeted R&D investments. If a firm is successful, it raises the quality by which own-product j is produced by a factor $\lambda_{o,ij} \equiv 1 + \lambda \beta^{s_{ij}}$. Here λ is a random component drawn from a common distribution $H_o(\lambda)$, s_{ij} denotes the firm's number of past own-product innovations on product j, and $0 < \beta \leq 1$. When $\beta < 1$, quality improvements on own products diminish as innovations accumulate, allowing us to match the life cycle of revenue growth within the firm. This is similar to how diminishing returns over product life cycles are modeled by Acemoglu et al. (2022).

To achieve an arrival rate of $x_{o,ij}$ of own-product innovations, firm i must hire $z_{o,ij}$ researchers along

$$z_{o,ij} = \eta_o\, x_{o,ij}^{\psi}\, c_o(\widehat{q}_{ij}, s_{ij}),$$

where $\eta_o > 0$ is the efficiency of the quality upgrading innovation on own products. R&D on own-product innovation rises in the relative quality of the product and falls in the number of previous process innovations. For ease of analysis, and because own-product innovation is not a key theoretical

[20] In principle, innovation efficiency h can evolve stochastically over time, or be affected by age (e.g., Acemoglu et al., 2018). As we explain in Section 5, our evidence suggests that efficiency is a persistent firm characteristic.

contribution of the paper, we choose a functional form for $c_o(\widehat{q}_{ij}, s_{ij})$ such that the equilibrium rate of own-product innovation does not depend on $\widehat{q}_{ij}$ and s_{ij}. As a result, firms choose equal innovation rates on all products within their portfolio, irrespective of product quality or the number of past innovations.

In addition to quality changes arising from innovation, we assume product quality may also change over time due to idiosyncratic shocks to consumer taste. These shocks are independently and identically distributed over time and across products. Importantly, we assume such shocks apply to all potential producers within the same product line, so that they affect the producer's revenue and profits, while not altering the identity of the producer.

4.5. Innovation Bursts and the Equilibrium Distribution of Innovation Efficiency

In an economy where innovation bursts can be large (that is, θ is finite) firms can grow rapidly. Experiencing a burst of product creation allows a firm to substantially grow its product portfolio, even if the quality improvements associated with these new products are smaller due potentially to the firm's low innovation efficiency.

Let us derive an expression for the innovation-efficiency composition of firms along a stationary equilibrium path. Letting $x_{b,h}(n)$ and $M_h(n)$ stand, as in Equation (7), for innovation burst rate and the measure of firms with n products and with innovation efficiency h, the change in the measure of single-product firms over time is given by

$$\dot{M}_h(1) = G(h)\, x_e + 2M_h(2)\, x_d - M_h(1)\left(x_{b,h}(1) + x_d\right).$$

The first term on the right-hand side captures new h-type firms from entry, the second term captures the inflow from firms of type h that used to produce two products but that have lost one due to creative destruction, and the final term captures outflow either through creatively destroyed firms exiting or through expansion by firms that draw an innovation burst.

For multi-product firms, the equation is

$$\dot{M}_h(n) = \sum_{k=1}^{n-1} M_h(k)\,\frac{(n-k)^{-\theta}}{\zeta(\theta)}\, x_{b,h}(k) + (n+1)\, M_h(n+1)\, x_d - M_h(n)\left(x_{b,h}(n) + n\, x_d\right).$$

The first term on the right-hand side shows how bursts alter the composition of firms. In a world without innovation bursts, i.e. when product creation is a Poisson process ($\theta \to \infty$), only firms that produce $n-1$ products can become producers of n products. Because of larger bursts, however, some firms are able to jump from being small producers to being large firms, and the likelihood that this happens increases as θ declines.

The stationary composition of firms is found by setting $\dot{M}_h(n)$ to zero for all sizes and innovation efficiencies. For single-product firms, the resulting measure of firms by type is $M_h(1) = (x^e / x^d)\, G(h)$. For multi-product firms, the measure is given by the sequence

$$M_h(n+1) = \frac{1}{(n+1)\, x_d}\left(M_h(n)\left(x_{h,b}(n) + n\, x_d\right) - \sum_{k=1}^{n-1} M_h(k)\,\frac{(n-k)^{-\theta}}{\zeta(\theta)}\, x_{h,b}(k)\right). \tag{8}$$

4.6. Firm Value and Innovation Decisions

As labor supply is constant, the growth g in aggregate productivity is the sole source of output growth, which grows at the same rate g along the balanced growth path. Given the standard structure of household consumption-saving decisions, we find the standard Euler equation, $r - g = \rho$, where r and g are the interest rate and growth rate of output along the path. This allows us to next set up the dynamic optimization problem of a firm i, which involves choosing the optimal arrival rate of innovation bursts, $x_{b,i}$, as well as the optimal rate of quality improvements for each of its existing products, $x_{o,ij}$.

Let $V_{h,t}(\widehat{\boldsymbol{q}}_i)$ denote the value function for a firm with innovation efficiency h and with a product portfolio $\widehat{\boldsymbol{q}}_i$. The portfolio is a matrix with n_i rows, where each row has information on a product for which the firm has the highest quality product: the relative quality $\widehat{q}_{ij}$, which determines the profit, and the number of prior own-product innovations s_{ij}, which determines the efficacy of its own-product innovation activity. Along the balanced growth path, aggregate quality Q grows at rate g and firm value satisfies:

$$
\begin{aligned}
& rV_{h,t}(\widehat{\boldsymbol{q}}_i) - \dot{V}_{h,t}(\widehat{\boldsymbol{q}}_i) = \\
& \max_{x_{b,i},\{x_{o,ij}\}} \left\{ \begin{array}{c}
\frac{Y_t}{\epsilon} \sum_{j\in J_i} \widehat{q}_{ij}^{\epsilon-1} - g \sum_{j \in J_i} \widehat{q}_{ij} \frac{\partial V_{h,t}(\widehat{\boldsymbol{q}}_i)}{\partial \widehat{q}_{ij}} \\
+ \sum_{j\in J_i} x_d \left[V_{h,t}\left(\widehat{\boldsymbol{q}}_i \setminus \{\widehat{q}_{ij}, s_{ij}\}\right) - V_{h,t}(\widehat{\boldsymbol{q}}_i) \right] \\
+ \sum_{j\in J_i} x_{o,ij} \mathbb{E}\left[V_{h,t}\left(\widehat{\boldsymbol{q}}_i \setminus \{\widehat{q}_{ij}, s_{ij}\} \cup_+ \left\{\lambda_{o,j}(s_{ij})\, \widehat{q}_{ij},\, s_{ij}+1\right\}\right) - V_{h,t}(\widehat{\boldsymbol{q}}_i)\right] \\
+ x_{b,i} \sum_{k=1}^{\infty} \frac{k^{-\theta}}{\zeta(\theta)} \mathbb{E}_h \left[V_{h,t}\left(\widehat{\boldsymbol{q}}_i \cup_+ \{\lambda_{b,i\ell}\, \widehat{q}_{i\ell}, 0\}_{\ell=1}^{k}\right) - V_{h,t}(\widehat{\boldsymbol{q}}_i) \right] \\
- w_t \eta_b x_{b,i}^{\psi} n_i^{-\sigma(\psi-1)} - w_t \eta_o \sum_{j\in J_i} x_{o,ij}^{\psi} c_o(\widehat{q}_{ij}, s_{ij}) - F(\widehat{\boldsymbol{q}}_i)
\end{array} \right\},
\end{aligned} \tag{9}
$$

where $\dot{V}_{h,t}$ denotes the change in firm value $V_{h,t}$ with time.

The first line on the right-hand side contains the sum of the flow of profits, $\frac{Y_t}{\epsilon} \widehat{q}_{ij}^{\epsilon-1}$, and the decline in profits over time from the gradual increase in average quality. The second line contains the expected change in value if the firm stops producing j because of creative destruction. The term $V_{h,t}\left(\widehat{\boldsymbol{q}}_i \setminus \{\widehat{q}_{ij}, s_{ij}\}\right)$ denotes the value of producing the portfolio of products $\widehat{\boldsymbol{q}}_i$ except product j with relative quality $\widehat{q}_{ij}$, on which the firm has implemented s_{ij} own-product improvements. The third line contains change in value from own-product innovation. The fourth row contains the expected increase in value due to innovation bursts. This is equal to the arrival rate of innovation bursts, $x_{b,i}$, multiplied by the expected increase in the firm's value if it acquires an innovation burst. The term $V_{h,t}\left(\widehat{\boldsymbol{q}}_i \cup_+ \{\lambda_{b,i\ell}\, \widehat{q}_{i\ell}, 0\}_{\ell=1}^{k}\right)$ denotes the rise in value if the burst contains k new products, with the ℓ-th product having an initial relative quality $\widehat{q}_{i\ell}$. This value is weighted by the probability density function of k, which is $k^{-\theta}/\zeta(\theta)$. The first and second term in the last row contain the firm's total R&D expenditure. The final term $F(\cdot)$ is a fixed cost that firms must pay to operate, which we assume exactly equals the option value of own-product innovation, to simplify exposition.

Next, we characterize the firm value function along a balanced growth path.

Proposition 1. The value of a firm with innovation efficiency h that produces a portfolio $\widehat{\boldsymbol{q}}$ with n products grows at rate g along the balanced growth path and is given by

$$
V_{h,t}(\widehat{\boldsymbol{q}}) = w_t \left(v^* \sum_{j\in J} \widehat{q}_j^{\epsilon-1} + n\, O_h(n) \right), \tag{10}
$$

where $v^* \equiv \frac{L/(\epsilon-1)}{\rho+(\epsilon-1)g+x_d}$ and where $O_h(n)$ denotes per-product option value of product creation, satisfying

$$(\rho + n x_d) O_h(n) = (n-1) x_d O_h(n-1) + \frac{(\psi-1)/\psi}{(\psi\eta_b)^{\frac{1}{\psi-1}}} \left(\frac{\zeta(\theta-1)}{\zeta(\theta)} \overline{\lambda}_{b,h}^{\epsilon-1} v^* + \sum_{k=1}^{\infty} \frac{k^{-\theta}}{\zeta(\theta)} (n+k) O_h(n+k) - n O_h(n) \right)^{\frac{\psi}{\psi-1}} n^{\sigma-1}. \tag{11}$$

Moreover, optimal innovation policies for product creation and own-product quality upgrading are characterized by

$$x_{b,h}(n) = \left(\frac{1}{\psi\eta_b} \left(\frac{\zeta(\theta-1)}{\zeta(\theta)} \overline{\lambda}_{b,h}^{\epsilon-1} v^* + \sum_{k=1}^{\infty} \frac{k^{-\theta}}{\zeta(\theta)} (n+k) O_h(n+k) - n O_h(n) \right) \right)^{\frac{1}{\psi-1}} n^{\sigma}, \tag{12}$$

$$x_o = \left(\frac{v^*}{\psi\eta_o} \right)^{\frac{1}{\psi-1}}. \tag{13}$$

Proof. See Appendix A.2.1.

As Equation (12) shows, the rate of innovation bursts increases in the expected number of products in an innovation burst (that is, falls in θ), falls in the cost of innovation η_b, rises in the present value of expected profits v^*, in expected product improvement $\overline{\lambda}_{b,h}$, and in the firm's product count. As expected, Equation (13) shows that the own-product innovation falls in the cost of innovation η_o and rises in the present value of expected profits v^*. The constancy of own-product innovation in the number of past innovations s is a consequence of our simplifying assumption about the fixed costs of each product. This leaves variation in quality step sizes over the product life cycle as the sole driver of the declining relationship between product-level revenue growth and tenure.

4.7. Innovation Bursts and Growth

Along the balanced growth path, productivity growth is the sum of contributions from creative destruction and from own-product improvements:

$$g = \frac{\zeta(\theta-1)}{\zeta(\theta)} \sum_h \left(\frac{\overline{\lambda}_{b,h}^{\epsilon-1} - 1}{\epsilon-1} \right) \left(\sum_n M_h(n)\, x_{b,h}(n) \right) + \left(\frac{\sum_h G(h) \overline{\lambda}_{b,h}^{\epsilon-1} - 1}{\epsilon-1} \right) x_e + \left(\frac{\mathbb{E}_s \left[\overline{\lambda}_o(s)^{\epsilon-1} \right] - 1}{\epsilon-1} \right) x_o, \tag{14}$$

where $\overline{\lambda}_o(s)^{\epsilon-1} \equiv \mathbb{E}_{H_o} \left[\left(1 + \lambda \beta^s\right)^{\epsilon-1} \right]$ gives an average moment of expected quality improvements for products with s past innovations and $\mathbb{E}_s[\cdot]$ denotes the expectation over the numbers of past innovations across incumbent-owned products.

We can use Equation (14) to examine the impact of innovation bursts on aggregate productivity growth. First, innovation bursts appear directly in the expression through the term $\zeta(\theta-1)/\zeta(\theta)$ that determines the average size of bursts. However, as in Equation (7), a higher average size of bursts matches the same rate of growth and creative destruction by a correspondingly lower average rate of the arrival of bursts $x_{b,h}(n)$.

The more important effect of bursts on growth stems from their impact on the distribution of innovation efficiency and size across incumbent firms, captured in Equation (14) through the term $\sum_n M_h(n)\, x_{b,h}(n)$. The presence of innovation bursts changes the size composition of firms with innovation efficiencies h, as given by the measure $M_h(n)$ and characterized by Equation (8). Since firms of different size and innovation efficiency generically make different contributions to aggregate growth (reflected in the variations in their innovation rates $x_{b,h}(n)$), this implies that *innovation bursts matter for the overall rate of aggregate growth*. The only exception is perhaps the baseline setting in Klette and Kortum (2004): without heterogeneity in innovation efficiency and with a linear relationship between size and innovation ($\sigma = 1$). Any deviations from this special case would lead to the prediction that the per-product rate of innovation $x_{b,h}(n)/n$ varies across firms with different efficiency h or size n.[21]

To better understand this influence, let us focus on the empirically relevant case of $\sigma = 1$, where the main driver of variations in per-product innovation rates $x_{b,h}(n)/n$ is the heterogeneity in efficiency types h. This setting is in line with the facts we documented in Section 3: the rate of product creation does not systematically vary with firm-level product count and that the revenue from newly created products is auto-correlated at the firm level. In this environment, larger firms are more likely to have higher efficiency types since this allows them to accumulate more products over time (Lentz and Mortensen, 2008). However, as the bursts of product creation become more thick-tailed, large firms become increasingly likely to have been those with a large single burst of product creation rather than to have had multiple successful smaller instances of product creation. As such, the connection between firm product count and innovation efficiency weakens with burst-like innovation. As Equation (14) shows, the resulting changes in the composition of size and innovation efficiency lead to different predictions regarding the contributions of large and small firms to aggregate growth. In Section 5.3 below, we examine these differences in the context of a specific size-dependent innovation policy.

5. Quantification

We next quantify the framework and shows that it replicates the empirical facts from Section 3. We use the quantified model to illustrate its normative implications, by comparing the design of innovation policy with and without accounting for innovation bursts.

5.1. Calibration

To calibrate the model, we infer most model parameters directly from the data on product dynamics. This contrasts with the common approach to calibrating models of creative destruction, which typically relies only on indirect inference based only on firm-level information. Table 4 summarizes the parameters.

Product Innovation The key novel parameter of our model is the tail of the zeta distribution, θ, which determines the thickness of the tail of the distribution of product creation. We calibrate θ to match the relationship between the number of newly created products among single-product firms and their

[21] While here we focus on the analysis of the stationary equilibrium, innovation bursts also matter for the speed of transitions in response to shocks, e.g., by accelerating the transitional dynamics after the entry of firms with different characteristics from incumbents.

Table 4: Summary of Parameter Values

Parameter	Description	Value
θ	Tail parameter of the zeta distribution	3.10
σ	Degree of returns to product counts in product innovation	1.00
η_e	Product creation R&D cost scalar for entrants	14.9
η_b	Product creation R&D cost scalar	10.4
ψ	Convexity of R&D costs in innovation rate	2.00
ρ	Discount rate	0.02
β	Decline rate of follow-up process innovation size	.830
x_o	Poisson rate of process innovations	1.00
λ	Process innovation step size	.042
Ξ	Standard deviation of idiosyncratic quality shocks	.272
ϵ	Elasticity of substitution	4.00
ρ	Discount rate	0.02
G	Entrant share of type L and H	[0.91,0.09]
$\lambda_{b,h}$	Average quality improvement size of product innovation L and H	[1.0001,1.0870]

corresponding rank in the distribution, as in Figure 2b. At $\theta = 3.1$, the model matches the slope of -2.57 in the figure.[22]

To account for the ex-ante heterogeneity in innovation efficiency, we introduce two types of firms, $h \in \{L, H\}$. Both draw quality improvements from Pareto distributions $H_{b,h}(\lambda)$ with a minimum value of 1 but with different tail parameters, so that average draws equal $\lambda_{b,L}$ and $\lambda_{b,H}$, where $\lambda_{b,L} < \lambda_{b,H}$. Since innovation by high-efficiency types improves product quality more than innovation by low-efficiency types, this type of heterogeneity leads to an autocorrelation in the size of revenue from new products. Therefore, to calibrate the relative efficiency $\lambda_{b,H}/\lambda_{b,L}$, we target the slope of the line in Figure 7, which is 0.39.

To pin down the remaining degree of freedom between $\lambda_{b,L}$ and $\lambda_{b,H}$, we choose the average degree of quality improvement that delivers the aggregate productivity growth rate of 1.6%, matching the French average in the Penn World Table (Feenstra et al. 2015). Combined with the own-product innovation parameters, this is achieved by setting $\lambda_{b,L}$ to about 1 and $\lambda_{b,H}$ to 1.087. This resembles the average quality improvements by efficient and inefficient innovators in Lentz and Mortensen (2008).

To calibrate the type distribution of entrants, we target the ratio of the variances of revenue on products that are new to a firm and the variance of revenue on continuing products, which is higher if the shares of firm types are close. This ratio is 1.07 in the data.

Turning to the frequency of innovation bursts among incumbents, we have two key parameters: (1) the innovation cost scalar η_b which governs the average rate of arrival of bursts, and (2) σ, which controls how product creation scales with size. We set σ to 1 to match the lack of a clear relationship between firm product count and product creation rates in Table 2. We calibrate η_b to match observed incumbent product creation rates, which average 0.066 (Table 2), while the average number of products per burst, $\zeta(\theta-1)/\zeta(\theta)$, is 1.29. Thus, η_b is set to 10.4 so that bursts arrive at an average rate of

[22] As firms may experience multiple innovation bursts within a year, the distribution of the number of newly created products differs from the distribution of the size of innovation bursts. Hence, we cannot infer θ directly from the slope of the rank-size relationship in Figure 2. Instead we perform a grid search to find the θ that delivers the log-rank, log-new products relationship that we observe in the figure.

0.051, the ratio of these two values. Finally, we set the curvature parameter ψ to 2, so the cost elasticity of R&D equals 1, a standard value (e.g., Bloom et al. 2002).

As for the rate of product creation by entrants, we do not directly observe these rates in the data due to the fact that the EAP data covers only firms beyond a size threshold. Instead, we rely on the fact that the total number of products is constant over time in our model, a fact that also provides a good approximation for the observed pattern in our data. When weighted by size, firms lose an average of 0.1 products for every product that they initially produce, while they only gain 0.066 products (see Table 2). Therefore, the contribution of entrants x_e is implied to be 0.034. This is achieved by setting η_e to 14.9.

Own-Product Innovation The quality improvements due to own-product innovation $\overline{\lambda}_o$ and their rate of arrival x_o play an interchangeable role in the model. We set the Poisson arrival rate x_o to 1, such that firms on average improve the quality of their existing products once per year. We then calibrate $\overline{\lambda}_o$ to 0.042 to match the average rate 0.6% of growth in real revenue on firms' continuing products, as found in the upper panel of Table 3. The degree of decreasing returns to follow-up innovation, β, determines the rate at which quality growth declines in a product's tenure. The expected annual decline in the growth of revenue with tenure is given by $\mathbb{E}\left[\frac{\partial^2 \ln p_{ijt} y_{ijt}}{\partial t \partial s}\right] = \lambda \beta^s (\epsilon - 1) \ln \beta$ where s is the product tenure. We calibrate β to 0.83, which delivers the decline in revenue growth in Figure 8.

Other Parameters The discount rate ρ is set to 0.02. The idiosyncratic product-level quality shocks cause changes in sales unrelated to innovation. Greater volatility in these shocks raises the contribution of continuing products to the overall variance of firm revenue growth. We assume a mean-zero, log-normal distribution for productivity shocks with standard deviation Ξ, set at 0.272 to match the 28.8% Shapley value of continuing products in revenue growth (Table 3). Finally, the elasticity of substitution ϵ, which controls the markup, is set to 4 to reflect a markup of 1.33, consistent with evidence for French manufacturing firms (De Ridder et al. 2022).

Parametrization of the Benchmark Model We compare the performance of the model with innovation bursts to a benchmark model where product creation is a Poisson process, which is the assumption underlying standard models such as Klette and Kortum (2004), Akcigit and Kerr (2018), Acemoglu et al. (2018) and Garcia-Macia et al. (2019). To implement this, we set θ to a large number and adjust the innovation costs such that firms still add an average of 0.066 products per product that they initially produce. This is achieved when we set η_b to 7.3. The remainder of the parameters are unchanged.

5.2. Model Performance

Next, we assess the model's ability to match the targeted and untargeted moments, including those we presented in our discussion of the stylized facts in Section 3. The model's performance on targeted moments is summarized in Table 5, which shows that all moments are matched with precision. The

Table 5: Performance on Targeted Moments

Parameter	Moment	Target	Model
Product Innov.			
θ	Regression of new products (log) on rank (log) if N_{t-1} =1	-2.57	-2.57
σ	Scaling of number of innovation bursts in firm size	1.00	1.00
η_b	Arrival rate of innovation bursts (per existing product)	.051	.051
η_e	Entry rate	.034	.034
λ_b	Growth rate of total factor productivity	.016	.016
Own Products			
x_o, λ	Revenue growth on continuing products	.006	.006
β	Change in revenue growth in product tenure	-.005	-.005
Ξ	Shapley-Owen contrib. of own products to revenue growth	28.8%	28.8%
Firm Types			
$\lambda_{b,L}/\lambda_{b,H}$	Persistence of revenue per new product within firms	0.39	0.39
$G(h)$	Relative standard deviation of existing vs new product revenue	1.07	1.07
Other			
ϵ	Markup	1.33	1.33

results in the remainder of the section come from a simulation of the quantified economy for 10,000 firms over 40 years.[23]

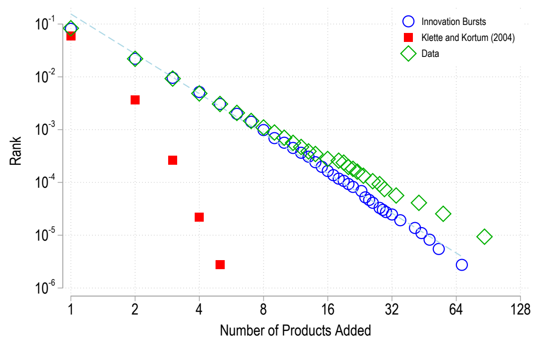
Figure 9 compares the distribution of product creation in the data and the model. The empirical distribution is plotted with green diamonds, while simulations from the model with innovation bursts are plotted with blue circles. To understand the importance of the fact that product innovations come in bursts, the figures also plot the distribution of product creation from simulations of a model with an equal amount of overall innovation, but for the benchmark model where product creation is a Poisson process.

The results in Figure 9 shows that the model accurately matches the high concentration of product innovation in the data. Pooling firms of all size in the top-left panel, the linear relationship between the level and the rank of the number of new products fits almost perfectly with an R^2 of 0.99. The similarity of the new-product distribution in the simulation and the data suggests that product innovation has a similar degree of concentration in the model and the data. Conditioning on size, the model with innovation bursts again closely matches the empirical distributions. The benchmark model in which the instances of product creation do not come in bursts, in contrast, fits the empirical pattern poorly. In this case, the resulting distribution of product creation is close to a Poisson distribution, which leads to substantially lower degrees of concentration. Even though the simulation tracks 10,000 for 40 years, no firm adds more than 5 products in any year, in sharp contrast to the data.

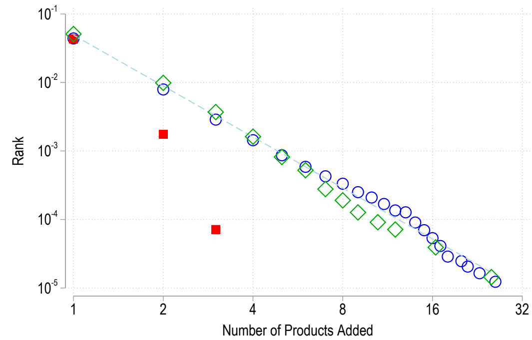
In addition to matching the empirical pattern of product creation, innovation bursts also help explain the overall concentration of production observed in the data. Figure 10 plots the distribution of firm product count where, as before, the x- and y-axes show size and rank, respectively, in logarithmic

[23] We simulate the economy for 40 years rather than the 10 years of the data that we have for the EAP survey. This is because the fat-tailed distribution of product innovation implies that in small samples, the convergence of sample moments to true moments may be slow.

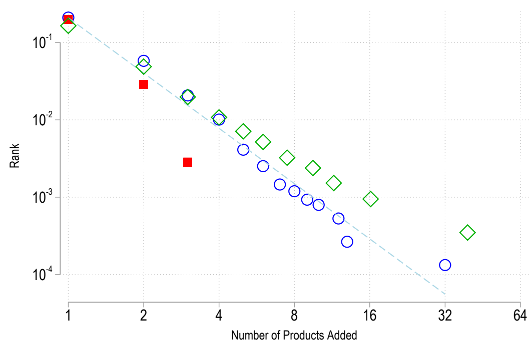
Figure 9. Distribution of Number of New Products: Data and Model



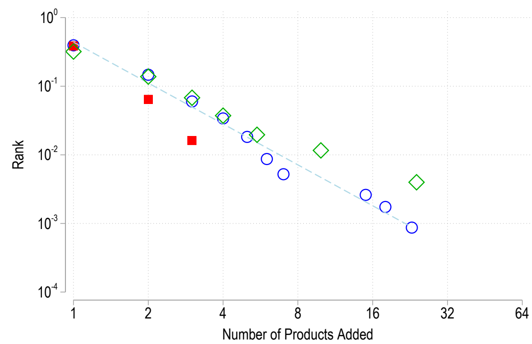
(a) All Firms



(b) Firms with $n_{i,t-1} = 1$ (target)



(c) Firms with $n_{i,t-1} = 5$



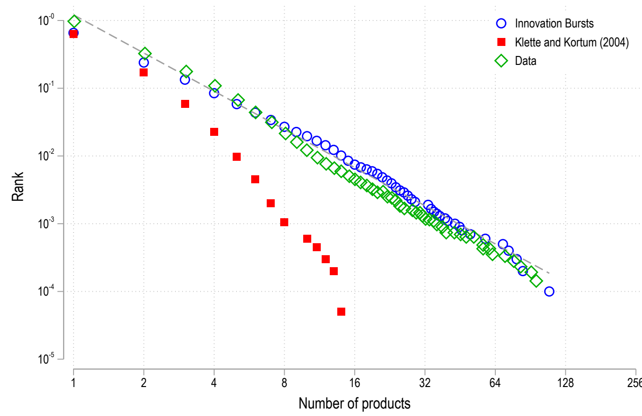
(d) Firms with $n_{i,t-1} = 10$

Notes: Figures plot the distribution of the number of products firms add to their portfolio. Green diamonds are data. Blue circles are from a model with innovation bursts. Red squares are based on a model with Poisson product innovation.

scale. The figure shows that innovation bursts are vital for explaining the observed concentration of production across large firms. When product innovation comes in bursts, the distribution of the number of products in the model is similar to the distribution of the model. Instead, under the benchmark model of one product per innovation, the number of large firms is negligible. This is a familiar shortcoming of standard models of creative destruction (see, e.g., Akcigit and Kerr 2018). This means that innovation bursts offer an alternative explanation for the emergence of large firms through product innovations. Luttmer (2011) shows that models of creative destruction can feature Pareto-tailed firm size distribution in the presence of growth in the total number of products (see also, e.g., Cao et al. 2017, Peters and Walsh 2021). However, as we mentioned earlier, we do not observe such a growth in our data.

Table 6 summarizes the model's ability to explain the fraction of firm growth that is driven by product creation, product destruction and revenue growth on continuing products. As a reminder, we have only targeted the contribution of continuing products to revenue growth over 1-year horizons, while the remainder of the moments are untargeted. The model matches the fact that within the contribution of creative destruction, product creation explains a smaller part of the variation in revenue growth than product destruction. This is due to the fact that new products can be introduced at any

Figure 10. Untargeted Moment: Firm Size Distribution



Notes: The figure plots the firm-size distribution in the log-log space. The vertical axis measures firms' rank divided by the number of observations. The horizontal axis measures the number of products that firms produce. Data is from the EAP. Simulations are based on 10,000 firms with identical type distributions and overall product arrival rates. Product arrival is a Poisson process in Klette and Kortum (2004) simulations and a Poisson-burst process along (5) in the Innovation Burst simulation.

point within the preceding year, leaving less than a full year of time for the firm to earn revenues on them (Bernard et al., 2017). The gap between the contribution of product creation (by incumbents) and product destruction is greater in the data than the model, which may be due to the fact that we underestimate entry in the EAP data.[24] Over 5-year horizons, all moments in the table are untargeted, and the model is able to quantitatively match the fact that product churn is a much larger contributor to overall revenue growth over longer horizons, and to qualitatively match the relative contribution of product destruction and product creation.

Figure 11 presents the life cycle of revenue growth in the model and in the data. Recall that we have targeted the decline in revenue growth in the data between the first and the second years, while the remaining data points are untargeted. The figure shows that the model's pattern of predicted decline in innovation efficiency over the product life cycle beyond the first two years is also in close agreement with the data.

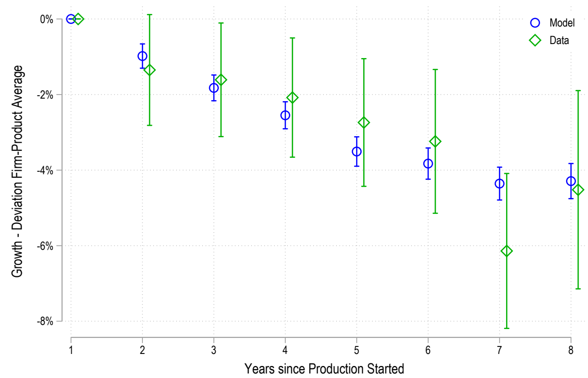
Table 6: Untargeted Moments: Shapley-Owen Decomposition in the Model vs the Data

	Overall Growth	Product Creation	Product Destruction	Continuing Products
Data (1-year)	100.0	16.0	55.2	28.8 (target)
Model (1-year)	100.0	23.8	47.4	28.8 (target)
Data (5-year)	100.0	37.3	49.8	12.9
Model (5-year)	100.0	25.2	66.5	8.26

Notes: The table decomposes total revenue growth into revenue loss from product loss, revenue gain from product gain, and changes in revenue on products that the firm is continuing to produce.

[24] Firms only enter EAP if they earn at least 5 million euros of revenue or employ 20 employees. We only consider a product "new" if we observe the firm prior to introduction of a product, which means that we under-count entry of new products when firms add a product to their portfolio in the year that they first appear in EAP.

Figure 11. Life Cycle of Revenue Growth



Notes: The vertical axis plots average growth of product revenue among all firm-products in the sample, where the average growth rate of a firm-product's revenue is subtracted from growth at each horizon. The horizontal axis plots the number of years that have passed since the firm first started producing the product. Confidence bounds are at the 90% level.

As a final untargeted moment, Appendix Figure C.16 shows that model is able to replicate the broad empirical pattern that the bulk of overall revenue destruction and creation occurs at firms with modest changes in sales, and that continuing products are responsible for the bulk of that growth. The model is also able to match that the tails of growth, in particular firm entry and exit, make up a significant part of overall revenue creation and destruction. The model also qualitatively matches the fact that the larger a firm's change in revenue, the greater the contribution of creative destruction.

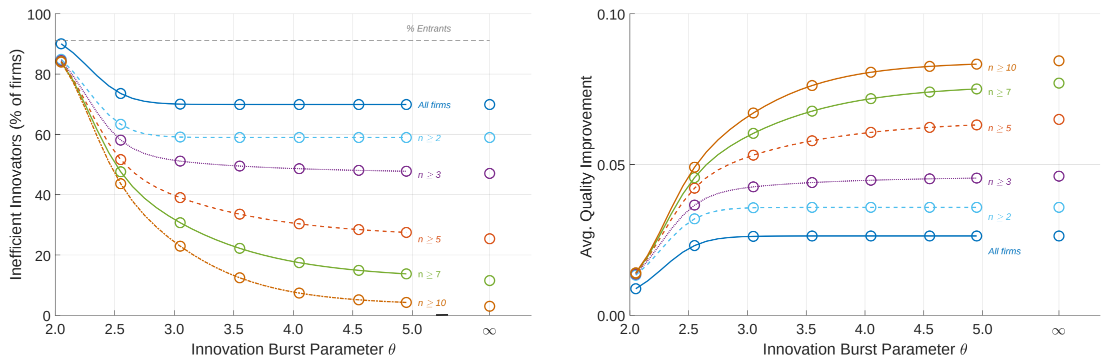
5.3. Application: Size-Dependent Innovation Subsidies

Having established that the model with innovation bursts can match the empirical stylized facts on firm growth and creative destruction, we next study the normative implications of introducing innovation bursts to Schumpeterian growth models. We consider a size-dependent innovation policy in the form of a government subsidy covering a share $\tau(n)$ of R&D costs associated with product creation, which can vary as a function of the firm's product count n. We assume that the government runs a balanced budget and imposes a lump-sum tax on the household to finance the subsidy.[25] To focus attention on the returns to such a subsidy scheme, we study the degree of cost-effectiveness of such policies, defined as the rise in aggregate productivity growth generates per dollar of subsidy.

As shown by Lentz and Mortensen (2016) in an environment with heterogeneity in innovation efficiencies similar to ours, optimal innovation policy should aim to shift R&D toward firms with higher innovation efficiency. Such firms generate higher externalities when they create new higher quality products, since part of this higher value will be captured by firms that will replace these products in the future. We assume that the government can observe and verify firm reports on product dynamics but not the underlying innovation efficiency type (h). As such, since high-efficiency firms on average also create more products, product count n might appear like a reasonable proxy for the firm's innovation efficiency.

[25] In our model, the government could equivalently tax worker's income because labor is supplied inelastically.

Figure 12. Effect of Innovation Bursts on Firm-Type Distribution and Average Step Size



(a) Composition of Firms (b) Average Quality Improvement

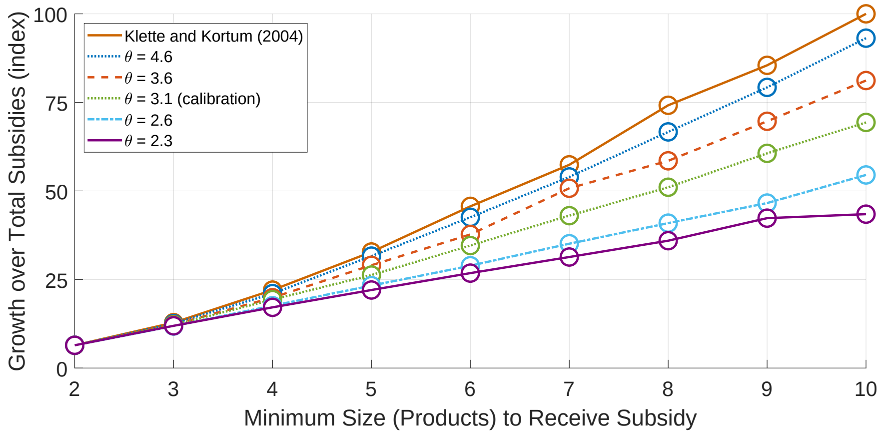
Notes: Left-hand figure: equilibrium % of firms with at least n products that belongs to the lowest innovation efficiency when θ is set to the level on the horizontal axis. Right-hand figure: average % quality improvement of firms with at least n products.

Figure 12 illustrates that in the presence of innovation bursts, firm size may not offer a reasonable proxy for its innovation efficiency. The left-hand figure plots the fraction of firms in our calibration with the low type of innovation efficiency in the stationary equilibrium. The horizontal axis plots the tail parameter θ of the Zeta distribution of innovation bursts, which approaches infinity in the benchmark model with one product per innovation (as in Klette and Kortum 2004), and which is 3.1 in our quantified model (as in Table 4). Each line plots the fraction of low-efficiency innovators among firms of a particular minimum size, which varies from the case where this minimum is 1 and we consider all firms (the top line) to the case of firms with at least 10 products (the bottom line).

As the figure shows, under the benchmark model without bursts, about 70% of firms are of the low-innovation-efficiency type in the stationary equilibrium, while they make up 91% of entrants. In this equilibrium, the share of this type among larger firms rapidly falls as we impose higher minimums on product count. Among firms producing at least 5 products, 25% have the low-efficiency type, while for firms with 10 products, the share falls to 6%. The right-hand plot in Figure 12 translates these percentages to average quality improvements. The figure shows that the average quality improvement of innovations across firms with any number of products equals 2.6% on average, while the average quality improvements of innovations by firms with at least 10 products is 8.4% per innovation. The latter is close to the innovation step size of the high-productivity innovators.

These patterns change in the presence of innovation bursts. As we consider models with lower values of θ corresponding to more burst-like patterns of innovation, moving toward the left on the x-axis in Figure 12, the share of low-efficiency innovators among incumbents rises, regardless of the size minimum imposed. The intuition is simple: if each innovation brings about a single product, a firm that is large must have had a long history of successful past innovations. If products arrive in innovation bursts, a small number of innovations suffice to make a firm large. This weakens the link between firm size and innovation rate, making size-dependent subsidies a poor alternative for targeted innovation policies.

Figure 13. Cost Effectiveness of Size-Dependent Innovation Policy by θ



Notes: The figure measures the cost efficiency of size-dependent subsidies by θ. Cost effectiveness is measured as the ratio of the additional growth induced by the subsidy (compared to a growth path where $\tau = 0 \forall n$), divided by aggregate subsidy spending. Horizontal axis: minimum product count from which firms receive subsidy τ. Vertical axis: the cost effectiveness as a ratio as a percentage of the highest cost effectiveness in the plot (firms with ≥ 10 products, Klette and Kortum calibration $\theta \to \infty$).

We quantify the impact of innovation bursts on the cost effectiveness of size-dependent innovation policies more formally as follows. We analyze the effect of an innovation subsidy that only covers 1% of the costs of product creation for all firms with a current product count beyond a certain size cut-off $\underline{n}$ ($\tau(n) = 0.01$ for $n \geq \underline{n}$ and $\tau(n) = 0$ otherwise). We compare the additional growth generated per dollar in subsidy paid for such policies at different levels of $\underline{n}$, and at different degrees of bursts in innovation. In response to the policy, firms endogenously adjust their innovation investments accounting for the additional incentives to create new products in order to qualify for the subsidy.

When comparing size-dependent innovation policies we solve the model in general equilibrium, taking into account that higher rates of innovation bursts affect the rate of innovation on firms' own products and their average quality improvements, as well as the entry rate. Appendix A.2 details the numerical solution to the model with size-dependent taxes.[26]

Figure 13 presents our measure of the cost-effectiveness of such size-dependent innovation policies. The figure compares cost-effectiveness as a function of different minimum numbers of products $\underline{n}$ at which the subsidy kicks in. Each line represents the predictions of the model under a different value for the tail parameter of innovation bursts, θ. The top (orange-solid) line is the Klette and Kortum (2004) calibration where $\theta \to \infty$, in which case the size-dependent innovation policy appears most cost-effective. The green-dotted line is our main calibration where $\theta = 3.1$. The vertical axis plots the cost effectiveness of the policy, defined as the additional growth that is generated by setting the 1% subsidy for firms with $n \geq \underline{n}$, divided by total spending on the subsidy. The most cost-effective policy in the plot serves as an index to ease interpretation.

[26]As the appendix shows, the value function remains a linear sum of a term capturing the expected net present value of profits until displacement and the innovation option value that captures the effect of this product in the productivity of the firm's investments in product innovation.

There are two clear results. First, the cost effectiveness of the subsidy increases in $\underline{n}$; the greater $\overline{n}$, the larger the fraction of firms subject to the subsidy is made up of high-efficiency innovators. Second, the size-dependent innovation policy is much less effective when product innovations come in bursts. The cost-effectiveness under our calibration of $\theta = 3.1$ is around 30% lower than the cost effectiveness of the same policy in the Klette and Kortum (2004) when firms with at least 10 products are targeted.

6. Conclusion

This paper contributes to the growing literature on firm-level innovation dynamics and their effects on aggregate outcomes such as productivity growth, business dynamism, and the firm-size distribution. Using comprehensive data from the French manufacturing sector, we document a novel fact about firm-level product innovation: the distribution of new products introduced by firms each year exhibits a thick, Pareto-like tail. This empirical pattern is difficult to reconcile with standard theories of creative destruction.

To address this gap, we propose a new model of endogenous growth driven by product innovation that incorporates the concept of innovation bursts. In our framework, innovations can arrive in clusters, enabling firms to rapidly expand their product portfolios through immediate applications. This idea of innovation bursts aligns with the novel empirical pattern we document among French manufacturing firms.

Our proposed model offers a new explanation for the Pareto distribution of firm size commonly observed in the data, which is a well-known limitation of theories of creative destruction. By showing that the equilibrium firm-size distribution can be Pareto for a sufficiently thick-tailed distribution of innovation bursts, we propose the lumpy nature of innovative ideas as the main driver of the concentration of firm-level production.

To illustrate the normative implications of innovation bursts, we show that size-dependent innovation policy is less effective when innovations come in bursts. More broadly, innovation bursts have the potential to affect aggregate productivity growth by affecting the transitional dynamics of the economy in response to policy shocks, and by changing the composition of firm age, firm size, and firm types along the balanced growth path equilibrium.

References

Acemoglu, D., Akcigit, U., Alp, H., Bloom, N., and Kerr, W. (2018). Innovation, reallocation, and growth. *American Economic Review*, 108(11):3450–3491.

Acemoglu, D., Akcigit, U., and Celik, M. A. (2022). Radical and incremental innovation: The roles of firms, managers, and innovators. *American Economic Journal: Macroeconomics*, 14(3):199–249.

Aghion, P., Akcigit, U., and Howitt, P. (2014). What do we learn from schumpeterian growth theory? In *Handbook of economic growth*, volume 2, pages 515–563. Elsevier.

Aghion, P., Bergeaud, A., Boppart, T., Klenow, P. J., and Li, H. (2023). A theory of falling growth and rising rents. *Review of Economic Studies*.

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

Akcigit, U. and Ates, S. T. (2021). Ten facts on declining business dynamism and lessons from endogenous growth theory. *American Economic Journal: Macroeconomics*, 13(1):257–98.

Akcigit, U. and Ates, S. T. (2023). What happened to us business dynamism? *Journal of Political Economy*, 131(8):2059–2124.

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

Albornoz, F., Pardo, H. F. C., Corcos, G., and Ornelas, E. (2023). Sequentially exporting products across countries. *Journal of International Economics*, 142:103735.

Argente, D., Baslandze, S., Hanley, D., and Moreira, S. (2020). Patents to products: Product innovation and firm dynamics.

Argente, D., Lee, M., and Moreira, S. (2024). The life cycle of products: Evidence and implications. *Journal of Political Economy*, 132(2):000–000.

Autor, D., Dorn, D., Katz, L. F., Patterson, C., and Van Reenen, J. (2020). The fall of the labor share and the rise of superstar firms. *The Quarterly Journal of Economics*, 135(2):645–709.

Axtell, R. L. (2001). Zipf distribution of us firm sizes. *science*, 293(5536):1818–1820.

Behrens, K. and Martin, J. (2015). Concording large datasets over time: The C3 method. *Unpublished paper*.

Benguria, F., Matsumoto, H., and Saffie, F. (2022). Productivity and trade dynamics in sudden stops. *Journal of International Economics*, 139:103631.

Bernard, A. B., Bøler, E. A., Massari, R., Reyes, J.-D., and Taglioni, D. (2017). Exporter dynamics and partial-year effects. *American Economic Review*, 107(10):3211–3228.

Bernard, A. B., Redding, S. J., and Schott, P. K. (2010). Multiple-product firms and product switching. *American economic review*, 100(1):70–97.

Bloom, N., Griffith, R., and Van Reenen, J. (2002). Do r&d tax credits work? evidence from a panel of countries 1979–1997. *Journal of Public Economics*, 85(1):1–31.

Bloom, N., Jones, C. I., Van Reenen, J., and Webb, M. (2020). Are ideas getting harder to find? *American Economic Review*, 110(4):1104–44.

Broda, C. and Weinstein, D. E. (2010). Product creation and destruction: Evidence and price implications. *American Economic Review*, 100(3):691–723.

Cao, D., Hyatt, H. R., Mukoyama, T., and Sager, E. (2017). Firm growth through new establishments. *SSRN 3361451*.

Cavenaile, L., Celik, M. A., Perla, J., and Roldan-Blanco, P. (2023). A theory of dynamic product awareness and targeted advertising. Technical report, Working Paper.

Cavenaile, L., Celik, M. A., and Tian, X. (2020). Are markups too high? competition, strategic innovation, and industry dynamics. *Working Paper*.

Davis, S. J., Haltiwanger, J., Jarmin, R., Miranda, J., Foote, C., and Nagypal, E. (2006). Volatility and dispersion in business growth rates. *NBER macroeconomics annual*, 21:107–179.

De Ridder, M. (2024). Market power and innovation in the intangible economy. *American Economic Review*, 114(1):199–251.

De Ridder, M., Grassi, B., Morzenti, G., et al. (2022). The hitchhiker's guide to markup estimation.

Dekle, R., Kawakami, A., Kiyotaki, N., and Miyagawa, T. (2015). Product dynamics and aggregate shocks: Evidence from japanese product and firm level data. *USC-INET Research Paper*, (16-01).

Feenstra, R. C., Inklaar, R., and Timmer, M. P. (2015). The next generation of the penn world table. *American economic review*, 105(10):3150–3182.

Fitzgerald, D., Haller, S., and Yedid-Levi, Y. (2024). How exporters grow. *Review of Economic Studies*, 91(4):2276–2306.

Gabaix, X. (2009). Power laws in economics and finance. *Annu. Rev. Econ.*, 1(1):255–294.

Gabaix, X., Lasry, J.-M., Lions, P.-L., and Moll, B. (2016). The dynamics of inequality. *Econometrica*, 84(6):2071–2111.

Garcia-Macia, D., Hsieh, C.-T., and Klenow, P. J. (2019). How destructive is innovation? *Econometrica*, 87(5):1507–1541.

Garicano, L., Lelarge, C., and Van Reenen, J. (2016). Firm size distortions and the productivity distribution: Evidence from france. *American Economic Review*, 106(11):3439–3479.

Geerolf, F. (2017). A theory of pareto distributions. *UCLA manuscript*.

Gibrat, R. (1931). Les inégalits économiques. *Sirey*.

Grossman, G. M. and Helpman, E. (1991). Quality ladders in the theory of growth. *The review of economic studies*, 58(1):43–61.

Guvenen, F., Karahan, F., Ozkan, S., and Song, J. (2021). What do data on millions of us workers reveal about lifecycle earnings dynamics? *Econometrica*, 89(5):2303–2339.

Hopenhayn, H. A. (2014). Firms, misallocation, and aggregate productivity: A review. *Annu. Rev. Econ.*, 6(1):735–770.

Hottman, C. J., Redding, S. J., and Weinstein, D. E. (2016). Quantifying the sources of firm heterogeneity. *The Quarterly Journal of Economics*, 131(3):1291–1364.

Jaimovich, N., Terry, S. J., and Vincent, N. (2023). The empirical distribution of firm dynamics and its macro implications. Technical report, National Bureau of Economic Research.

Klette, T. J. and Kortum, S. (2004). Innovating firms and aggregate innovation. *Journal of political economy*, 112(5):986–1018.

Kwon, S. Y., Ma, Y., and Zimmermann, K. (2024). 100 years of rising corporate concentration. *American Economic Review*, 114(7):2111–2140.

Lentz, R. and Mortensen, D. T. (2008). An empirical model of growth through product innovation. *Econometrica*, 76(6):1317–1373.

Lentz, R. and Mortensen, D. T. (2016). Optimal growth through product innovation. *Review of Economic Dynamics*, 19:4–19.

Luttmer, E. G. (2007). Selection, growth, and the size distribution of firms. *The Quarterly Journal of Economics*, 122(3):1103–1144.

Luttmer, E. G. (2010). Models of growth and firm heterogeneity. *Annual Review of Economics*, 2(1):547–576.

Luttmer, E. G. (2011). On the mechanics of firm growth. *The Review of Economic Studies*, 78(3):1042–1068.

Ma, Y., Zhang, M., and Zimmermann, K. (2024). Business concentration around the world: 1900–2020.

Melcangi, D. and Sarpietro, S. (2024). Nonlinear firm dynamics. Technical report, Staff Report.

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

Olmstead-Rumsey, J. (2022). Market concentration and the productivity slowdown.

Ozkan, S., Hubmer, J., Salgado, S., and Halvorsen, E. (2023). Why are the wealthiest so wealthy? a longitudinal empirical investigation.

Peters, M. (2020). Heterogeneous markups, growth, and endogenous misallocation. *Econometrica*, 88(5):2037–2073.

Peters, M. and Walsh, C. (2021). Population growth and firm dynamics. *Working Paper*.

Weiss, J. (2023). Market concentration, growth, and acquisitions.

Yao, J. (2025). Knowledge is (market) power. *Available at SSRN 4273942*.

'Growth Through Innovation Bursts'
Appendix - For Online Publication Only

Appendix A. Theoretical Appendix

A.1. Benchmark Theories: the Distribution of Product Creation

In this appendix we discuss the firm-size distribution and the distribution of product creation in benchmark theories of creative destruction. As Klette and Kortum (2004) show (p. 994), for a firm that at time $t = 0$ begins with a single product that creates new products at rate x and loses products at rate τ, the probability $\mathbb{P}_{n,1}(t)$ to have n products that after time t is given by

$$\mathbb{P}_{n,1}(t)(t) = \begin{cases} \frac{\tau}{x}\chi(t), & n = 0, \\ \left(1 - \frac{\tau}{x}\chi(t)\right)\left(1 - \chi(t)\right)\chi(t)^{n-1}, & n \geq 1, \end{cases} \tag{A.1}$$

where the function $\chi(t)$ is given by

$$\chi(t) = \frac{1 - e^{-(\tau - x)t}}{\tau / x - e^{-(\tau - x)t}}. \tag{A.2}$$

Let us compute the tail distribution function (complementary cumulative distribution function) for the net product creation of such a firm to be greater than n_c:

$$\begin{aligned} \mathbb{P}\{n - 1 \geq n_{nc}; t\} &= \sum_{k=1+n_c}^{\infty} \mathbb{P}_{k,1}(t) = \left(1 - \frac{\tau}{x}\chi(t)\right)\left(1 - \chi(t)\right)\sum_{k=n_c}^{\infty} \chi(t)^k, \\ &= \left(1 - \frac{\tau}{x}\chi(t)\right)\chi(t)^{n_c}. \end{aligned}$$

This means that the tail index of the distribution grows indefinitely as n_c rises

$$-\frac{\partial \log \mathbb{P}\{n - 1 \geq n_{nc}; t\}}{\partial \log n_c} = -n_c \log \chi(t), \tag{A.3}$$

since $\chi(t) \in (0,1)$. The tail of this distribution is very thin as a result.

Luttmer (2011) further generalizes this result to characterize the probability $\mathbb{P}_{n,k}(t)$ for a firm starting with k products at time 0 to have n products at time t, which is given by

$$\mathbb{P}_{n,k}(t) = \sum_{m=1}^{\min\{n,k\}} \binom{k}{m} \binom{n-1}{m-1} \left(1 - \frac{\tau}{x}\chi(t)\right)^m \left(\frac{\tau}{x}\chi(t)\right)^{k-m} \left(1 - \chi(t)\right)^m \chi(t)^{n-m}. \tag{A.4}$$

To better understand the distribution in Equation (A.4), note that we can find it as follows. Let $n(t)$ denote the random variable characterized by the distribution in Equation (A.1) for a given firm i with a single initial product. Let $d(t)$ denote a Bernoulli random variable that takes value 0 with probability $\frac{\tau}{x}\chi(t)$ and let $n_c(t)$ denote a random variable characterizing the number of failures until the first success of another Bernoulli distribution with a probability of failure of $\chi(t)$, which has a geometric distribution. Then, it is straightforward to see that the number of products of the firm after time t satisfies $n(t) = d(t)\,(n_c(t) + 1)$. Now, let $N_k(t)$ denote the random variable characterized by the distribution in Equation (A.1) for a given firm i with an initial k product, such that $N_1(t) = n(t)$. Given the assumptions of the benchmark models, we can write $N_k(t) = \sum_{j=1}^{k} n_j(t)$ where $n_j(t)$'s are k independent draws of the random variable $n(t)$. We can then write this random variable as $N_k(t) = K(t) + S(t)$ where $K(t) = \sum_{j=1}^{k} d_j(t)$ is a binomial distribution with k draws and with probability $1 - \frac{\tau}{x}\chi(t)$. Conditional on $K(t) = K$, random variable $S(t) = \sum_{j=1}^{k} d_j(t)\, n_{c,j}(t)$ is distributed according to a negative binomial distribution characterized with the probability mass function

$$\mathbb{P}\{S(t) = S \,|\, K(t) = K\} = \binom{S+K-1}{K-1} \left(1-\chi(t)\right)^K \chi(t)^S.$$

Then, the distribution of the total sum is given by

$$\mathbb{P}\left\{N_k(t) = N\right\} = \sum_{K=0}^{k} \mathbb{P}\{K(t) = K\}\, \mathbb{P}\{S(t) = S \,|\, K(t) = K\},$$

which leads to Equation (A.4) after some algebra.

The main observation from the above discussion is that, as the number of initial product count k grows, the central limit theorem implies that the distribution of the number of products after time t converges to a normal distribution $N_k(t) \sim \text{Normal}\left(k\mathbb{E}[n(t)],\, k\mathbb{V}[n(t)]\right)$. The tail index of the normal distribution for large $N_k(t)$ also grows indefinitely as a quadratic $N_k(t)$ as $N_k(t)$ rises. Thus, the distribution of new products inherits this quality and has a thin tail.

A.2. Solution to the Value Function

A.2.1. Proof of Proposition 1:

The dynamic optimization problem is given by the HJB equation (10). Guess that the solution takes the form $V_{h,t}(\hat{\mathbf{q}}) = w_t \left[\sum_{j\in J_i} v_h(\hat{q}_{ij}, s_{ij}) + nO_h(n)\right]$. Substituting for this expression in the HJB equation, we find

$$(r-g)\left[\sum_{j\in J_i} v_h(\hat{q}_{ij}, s_{ij}) + n\, O_h(n)\right] = \max_{x_{b,i},\{x_{o,ij}\}} \left\{ \begin{array}{c} \frac{L}{\epsilon-1} \sum_{j\in J_i} \hat{q}_{ij}^{\epsilon-1} - g \sum_{j\in J_i} \hat{q}_{ij} \frac{\partial v_h(\hat{q}_{ij}, s_{ij})}{\partial \hat{q}_{ij}} \\ -\sum_{j\in J_i} x_d \left[v_h(\hat{q}_{ij}, s_{ij}) + nO_h(n) - (n-1)O_h(n-1)\right] \\ +\sum_{j\in J_i} x_{o,ij}\, \mathbb{E}\left[v_h\left((1+\lambda\beta^{s_{ij}})\hat{q}_{ij}, s_{ij}+1\right) - v_h(\hat{q}_{ij}, s_{ij})\right] \\ +x_{b,i} \left(\sum_{k=1}^{\infty} \frac{k^{-\theta}}{\zeta(\theta)} \left(k\mathbb{E}_h\left[v_h\left(\lambda_{b,i}\hat{q}, 0\right)\right] + (n+k)O_h(n+k) - nO_h(n)\right)\right. \\ -\eta_b\, x_{b,i}^{\psi}\, n_i^{-\sigma(\psi-1)} - \eta_o \sum_{j\in J_i} x_{o,ij}^{\psi}\, c_o(\hat{q}_{ij}, s_{ij}) - F(\hat{\mathbf{q}}_i) \end{array} \right\}, \tag{A.5}$$

Now substitute $v_h\left(\hat{q}, s\right) \equiv \tilde{v}(s)\, \hat{q}^{\epsilon-1}$ and let $c_o\left(\hat{q}, s\right) \equiv \tilde{c}_o(s)\, \hat{q}^{\epsilon-1}$ and $F(\hat{\mathbf{q}}_i) \equiv \sum_{j\in J_i} \tilde{F}(s)\, \hat{q}^{\epsilon-1}$ in the above equation and collect the terms involving $\tilde{v}(\cdot)$ to find

$$(r-g)\,\tilde{v}(s) = \max_{x_o} \left\{ \begin{array}{c} \frac{L}{\epsilon-1} - g(\epsilon-1)\,\tilde{v}(s) - x_d\,\tilde{v}(s) \\ x_o\left(\mathbb{E}_{H_o}\left[(1+\lambda_o\beta^s)^{\epsilon-1}\right]\tilde{v}(s+1) - \tilde{v}(s)\right) - \eta_o x_o^{\psi}\,\tilde{c}_o(s) - \tilde{F}(s) \end{array} \right\}. \tag{A.6}$$

Collecting the terms involving $O_h(\cdot)$ then yields

$$(r-g+x_d)\, n\, O_h(n) = \max_{x_b} \left\{ \begin{array}{c} x_b \sum_{k=1}^{\infty} \frac{k^{-\theta}}{\zeta(\theta)} \left((n+k)\, O_h(n+k) - n\, O_h(n) + k\,\tilde{v}(0)\,\mathbb{E}_h\left[\lambda^{\epsilon-1}\right]\right) \\ -\eta_b\, x_b^{\psi}\, n^{-\sigma(\psi-1)} - x_d(n-1)\, O_h(n-1) \end{array} \right\}. \tag{A.7}$$

The first order condition for x_o yields

$$x_o(s) = \left(\frac{1}{\psi\eta_o}\frac{\mathbb{E}_{H_o}\left[(1+\lambda\beta^s)^{\epsilon-1}\right]\widetilde{v}(s+1)-\widetilde{v}(s)}{\widetilde{c}_o(s)}\right)^{\frac{1}{\psi-1}}, \tag{A.8}$$

while the first order condition for x_b gives

$$x_{b,h}(n) = \left(\frac{1}{\psi\eta_b}\left(\frac{\zeta(\theta-1)}{\zeta(\theta)}\overline{\lambda}_{b,h}^{\epsilon-1}\widetilde{v}(0)+\sum_{k=1}^{\infty}\frac{k^{-\theta}}{\zeta(\theta)}(n+k)\,O_h(n+k)-n\,O_h(n)\right)\right)^{\frac{1}{\psi-1}} n^{\sigma}. \tag{A.9}$$

Noting $\rho = r - g$, from Equation (A.6), we find

$$\left(\rho+(\epsilon-1)\,g+x_d\right)\widetilde{v}(s) = \frac{L}{\epsilon-1}+\frac{\psi-1}{\psi}\left(\psi\eta_o\widetilde{c}_o(s)\right)^{-\frac{1}{\psi-1}}\left(\mathbb{E}_{H_o}\left[(1+\lambda\beta^s)^{\epsilon-1}\right]\widetilde{v}(s+1)-\widetilde{v}(s)\right)^{\frac{\psi}{\psi-1}}-\widetilde{F}(s).$$

Next, we make the following simplifying assumptions

$$\widetilde{c}_o(s) \equiv \mathbb{E}_{H_o}\left[(1+\lambda\beta^s)^{\epsilon-1}\right]-1,$$
$$\widetilde{F}(s) \equiv \frac{\psi-1}{\psi}\left(\psi\eta_o\right)^{-\frac{1}{\psi-1}}\widetilde{c}_o(s)\left(\frac{L/(\epsilon-1)}{\rho+(\epsilon-1)g+x_d}\right)^{\frac{\psi}{\psi-1}}.$$

Given these assumptions, Equations (A.6) and (A.8) now together imply

$$\widetilde{v}(s) = v^* \equiv \frac{L/(\epsilon-1)}{\rho+(\epsilon-1)g+x_d},$$

and Equations (A.9) and (13). Substituting these results in Equation (A.7), we find Equation (11).

A.2.2. Equilibrium Definition

The following provides the definition of the equilibrium in our model.

theoremdef:eq The economy is in a balanced growth path equilibrium if for every t the variables $\{r, x_d, x_e, x_o, L, L^e, L^{rd}, g\}$ and functions $\{x_{b,h}, M_h\}$ are constant, $\{Y, C, Q, w\}$ grow at a constant rate g that satisfies (14), interest rates follow from $r = g+\rho$, Q is given by Equation (4), Y is given by $Y = QL = \frac{\epsilon}{\epsilon-1}wL$, innovation rates $x_{b,h}$ and x_o satisfy (12) and (13), the entry rate is given by

$$x_e = \left(\frac{v^*+\sum_h G(h)O_h(1)}{\psi\eta_e}\right)^{1/(\psi-1)},$$

which maximizes the profits of entrants, and together they satisfy Equation (14). The stationary distributions of product types M_h satisfy Equation (8), the rate of creative destruction x_d satisfies (7), and both goods and labor markets are in equilibrium so that $Y = C$ and $L = 1-L^{rd}-L^e$, with the labor hired in incumbent R&D and entry sectors given by

$$L^{rd} = \eta_b\sum_n M_h(n)\,x_{b,h}(n)^{\psi}\,n^{-\sigma(\psi-1)}+\eta_o\,x_o^{\psi}\,\mathbb{E}\left[c_o(\widehat{q},s)\right],$$

and $L^e = \eta_e x_e^{\psi}$.

A.2.3. Size-Dependent Taxes

The dynamic problem with size-dependent taxes is similar to the dynamic problem in Section 5. The only changes are that the cost of product innovation in the final line of the HJB equation (10) now includes the size-dependent subsidy $\tau(n)$ and that the innovation-size scalar σ is calibrated to unity. Thus, that term, and the corresponding term in Equation (A.5), are now given by $(1-\tau(n_i))\,\eta_b\, x_{b,i}^{\psi}\, n_i^{-(\psi-1)}$. Accordingly, the optimal arrival rate of innovation bursts is now given by:

$$x_{b,h}(n) = \left(\frac{1}{\psi(1-\tau(n))\eta_b}\left(\frac{\zeta(\theta-1)}{\zeta(\theta)}\overline{\lambda}_{b,h}^{\epsilon-1}\,\widetilde{v}(0) + \sum_{k=1}^{\infty}\frac{k^{-\theta}}{\zeta(\theta)}(n+k)\,O_h(n+k) - n\,O_h(n)\right)\right)^{\frac{1}{\psi-1}} n^{\sigma} \tag{A.10}$$

A.2.4. Computation

To solve the model with size-dependent taxes we deploy the following algorithm. In an outer loop, we guess a growth rate of aggregate productivity g and a rate of creative destruction x^d. With these guesses, we then calculate the expected net present value of future profits for each of innovation efficiency type h.

We then calculate the sequence of innovation option values for $O_h(n)$ in an inner loop. First, we use the fact that $\lim_{n\to\infty} O_h(n) = \overline{O}_h$, where $\overline{O}_h$ is the root of

$$(\rho + x_d)\overline{O}_h = \frac{(\psi-1)/\psi}{\left((1-\overline{\tau})\psi\eta_b\right)^{\frac{1}{\psi-1}}}\left(\frac{\zeta(\theta-1)}{\zeta(\theta)}\left(\overline{\lambda}_{b,h}^{\epsilon-1}\, v^* + \overline{O}_h\right)\right)^{\frac{\psi}{\psi-1}}.$$

where $\overline{\tau}$ is the innovation subsidy for firms with a size $n \geq \overline{n}$.[27] To find the remainder of the option values, we then guess the sequence $O_h(n)$ for all n from 0 to some large N (which we set to 100), using $O_h(1) = 0$ and $O_h(\overline{n}) = \overline{O}_h$, iterating the following update rule until the results are consistent with model-implied sequence for $O_h(n)$ along:

$$O_h(n) = \frac{x_d(n-1)O_h(n-1) + \frac{x_{b,h}(n)}{n}\left(\frac{\zeta(\theta-1)}{\zeta(\theta)}\overline{\lambda}_{b,h}^{\epsilon-1}\, v^* + \sum_{k=1}^{\infty}\frac{k^{-\theta}}{\zeta(\theta)}(n+k)\,O_h(n+k)\right) - (1-\tau(n))\eta_b\left(\frac{x_{b,h}(n)}{n}\right)^{\psi}}{\rho + n x_d + x_{b,h}(n)},$$

where innovation rates $x_{b,h}(n)$ are given by (A.10).

We then use these innovation rates to calculate the equilibrium firm-size and firm-type distributions, and calculate the optimal entry and process innovation rates in response to the guessed growth and creative destruction rates. From the endogenous objects, we then calculate new guesses for the growth and creative destruction rates, and iterate until the guesses converge.

Appendix B. Data Appendix

In this section, we describe the data sources used in our analysis and the procedure we use to merge them and clean the resulting dataset of outliers.

[27] For the standard calibration where $\psi = 2$, this equation has two positive roots. The root that is closest to 0 is well-behaved, in the sense that the option value of R&D increases in taxes and in the optimal rate of R&D. This is the root we use in this step of the algorithm.

B.1. EAP

The data are based on an annual survey of firms' production activities, called the *Enquête Annuelle de Production* (EAP), which is administered by the Institut National de la Statistique et des Études Économiques (INSEE). In accordance with EU regulation, the survey must encompass at least 90 per cent of the annual production of each 4-digit industry. The data contain comprehensive information on sales and the volume of goods. The volume is recorded in units of measurement (number of items, kilograms, litres) that are product-specific, while the value is recorded in current euros. The survey provides information at the 10-digit product level, classified according to the PRODFRA system—-the official French product classification published by INSEE, which includes approximately 4,400 product codes (see Section B.1.2). The survey covers the entire manufacturing sector (NACE rev. 2 section C), except for the agri-food industries (section 10, 11 and 12) and the manufacture of wood (16). While also including the extractive industry; electricity, gas, steam and air-conditioning supply; water supply, sewerage, waste management and remediation; we exclude these specific industries.

B.1.1. Sampling Framework

The survey contains an exhaustive sample of firms with at least 20 employees or revenues higher than 5 million Euros. The sample size varies over time but it is usually around 25,000 firms. To ensure a good level of coverage the survey must cover at least 90 per cent of the total production value of each 4-digit industry (NACE rev. 2). If this threshold is not reached more enterprises are surveyed. Additionally, the survey contains a random sample of firms with less than 20 employees. Its size varies year by year and it is usually around 8/9000 firms. Because this sample does not allow us to observe product creation or loss, we drop this second set of smaller firms from our analysis.

B.1.2. Product Classification

As for the level of product aggregation, the survey classifies products following the PRODFRA system—the official French product classification published by INSEE, which includes roughly 4,500 10-digit product codes. The first eight positions of PRODFRA represent the European PRODCOM classification, where the last two positions are used to refine the nomenclature.[28] Table B.1 presents examples of products in the PRODFRA classification. A few special products, which account for around 5% of the total number of product categories, are identified by a letter (H, N, S, Y) instead of a number in the 9th position of PRODFRA. These product categories are dropped when working with the PRODCOM classification. The PRODCOM classification can be directly linked to the classification of industrial activities in the EU (NACE Rev. 2), as the first four digits of the PC code identify a 4-digit NACE industry.

B.1.3. Product Concordance

A feature of PRODFRA is that it changes over time (from 3 to 5% of the product categories change each year). The use of these product codes in longitudinal studies requires harmonizing the product clas-

[28] For additional information see https://www.insee.fr/en/metadonnees/definition/c1097.

Table B.1: Examples of products in the PRODFRA classification

1812125000	Advertising and similar printed matter (excluding commercial catalogues)
1812199010	Administrative or commercial printed matter, at or continuous, customised or not, and directories
2042194510	Lotions for pre-shaving, shaving, or after-shaving
2042194520	Pre-shaving, shaving, and after-shaving foams and gels
2511235040	Industrial boiler products: not including tanks, boilers, nuclear equipment
3102100010	Wooden kitchen furniture: by mounted elements, including custom
310912502B	Dining and living room furniture other than tables: buffets, credenzas and livings, bookcases, cabinets by element.

sification system over time. To do so, we use the algorithm developed by Behrens and Martin (2015), called "connected components concordance", or *C3* for short. *C3* uses the graph theory to identify stable and comparable groups of products over time while minimizing the size of each group. The identified groups of products are then assigned to a single, time-consistent code. The vast majority of products (almost 90%) are not affected by this concordance procedure, and a marginal fraction of the new product groups include more than three PRODFRA10 codes.

B.2. FARE

The *Fichier approché des résultats d'Esane* (FARE) contains a coherent set of statistics on the universe of French private companies. It combines administrative data obtained from annual profit declarations made by companies to the tax authorities and from annual social data which provide information on employees and data obtained from a sample of companies surveyed by a specific questionnaire to produce structural business statistics (ESA).

B.3. DADS

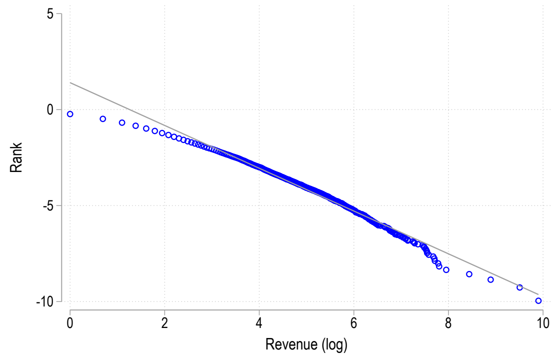
We obtain the number of establishments from *Déclaration Annuelle de Données Sociales* (DADS), which is a matched employer-employee dataset that covers the whole population of private sector workers in France. From the various versions provided by INSEE, we utilize DADS Poste (Fichiers Régionaux des Postes), which offers data at the individual job spell level. Each worker in the dataset is associated with an establishment identifier, and if the same employee works in two different establishments during the same year, only the main job is included in our analysis.

B.4. LiFi

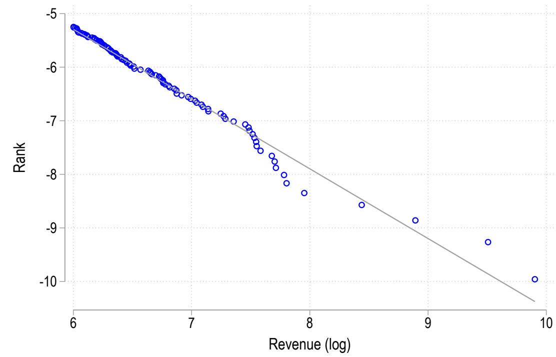
Finally, we determine the ownership status of firms (whether independent or part of a group) using the *Liaisons Financières entre Sociétés* (LiFi) data. This dataset collects information on the financial links between enterprises incorporated in France as well as their foreign owners and affiliates. LiFi serves as the French directory of corporate groups and is based on the most comprehensive knowledge of capital ownership links between companies (or financial connections). It is constructed using multiple sources: data from the *Banque de France* collected as part of corporate credit ratings, tax data (DGFIP), commercial data (ORBIS-BvD), and information available in activity reports published by corporate groups. We define independent firms as those not listed in the LiFi dataset, either as head of the group or as affiliates.

Appendix C. Additional Figures

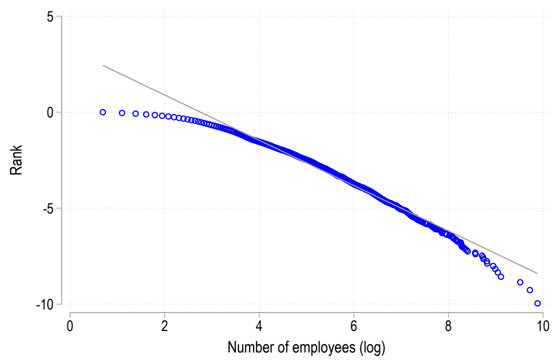
Figure C.1. Size Distribution - Revenue and Employment



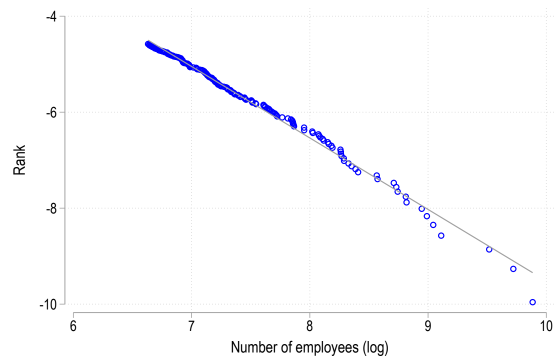
(a) Distribution: Revenue



(b) Distribution: Revenue - Tail



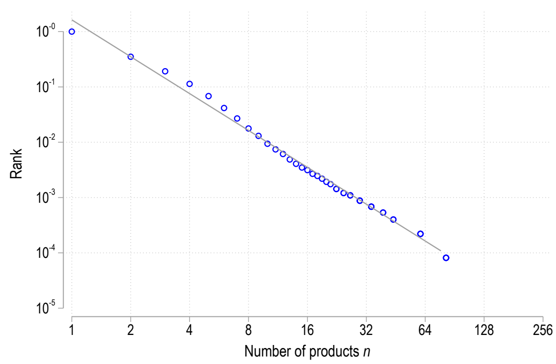
(c) Distribution: Employment



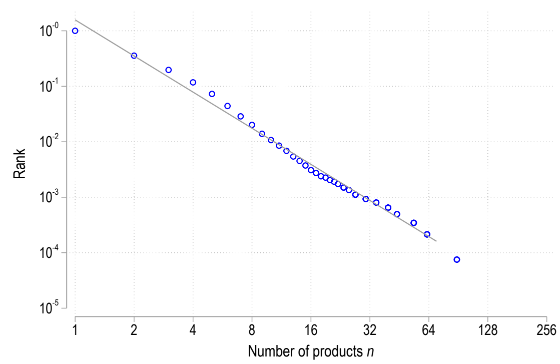
(d) Distribution: Employment - Tail

Notes: The figures plot the relationship between a firm's size (horizontal axes) and the firm's rank (vertical axes). Size is measured through either revenue (in 2015 euros) or employment (in full-time equivalence). The (rescaled) rank is measured as the ratio of firms' rank starting from the largest firm, divided by the total number of observations in the data. Plots are based on the FARE-EAP sample for 2019. Slopes by figure: (a) Slope: -1.1. Standard error: .01. R^2: 0.98; (b) Slope: -1.30. Standard error: .02. R^2: 0.98; (c) Slope: -1.2. Standard error: .01. R^2: 0.97; (d) Slope: -1.5. Standard error: .01. R^2: 0.99.

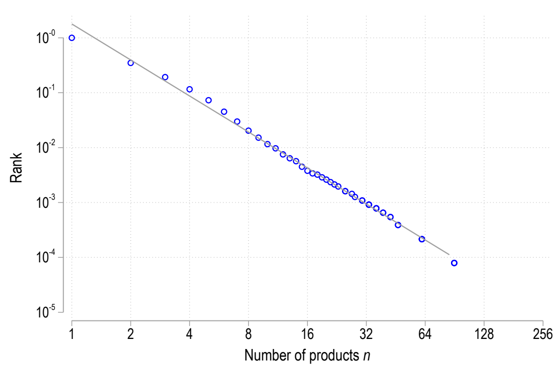
Figure C.2. Distribution of Number of Products: Different Years



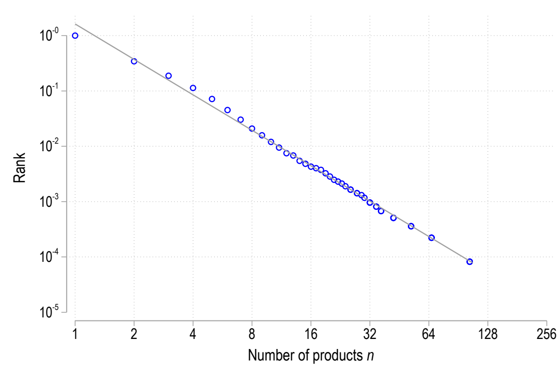
(a) 2010



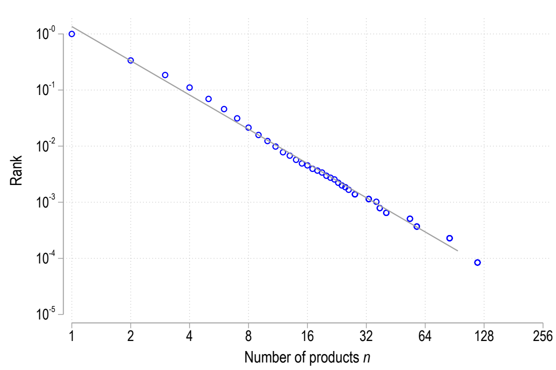
(b) 2012



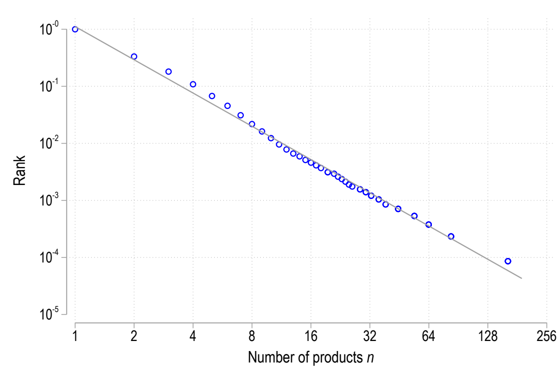
(c) 2014



(d) 2016



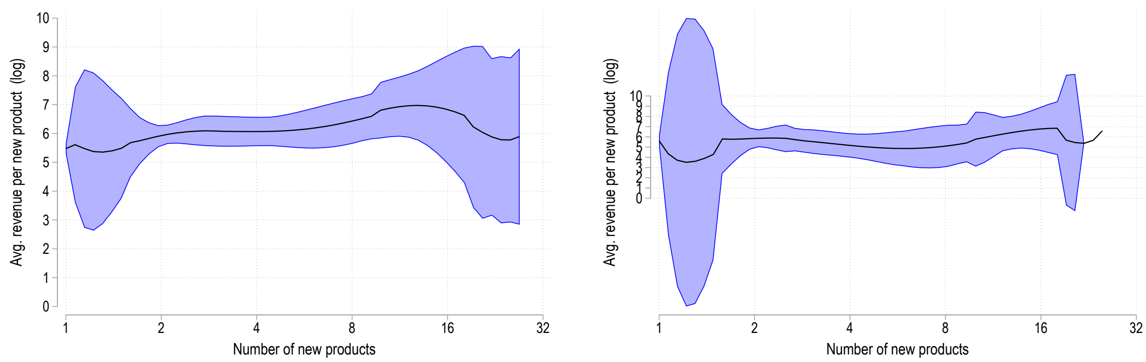
(e) 2018



(f) 2019

Notes: The figures plot the relationship between a firm's number of products (horizontal axes) and the firm's rank (vertical axes). The (rescaled) rank is measured as the ratio of firms' rank starting from the largest firm, divided by the total number of observations in the data. The figures replicate Figure 1(a) for different years. Slopes of figures (a)-(f) are respectively -2.2, -2.2, -2.2, -2.1, -2.0, -1.9, all significant at 1% level.

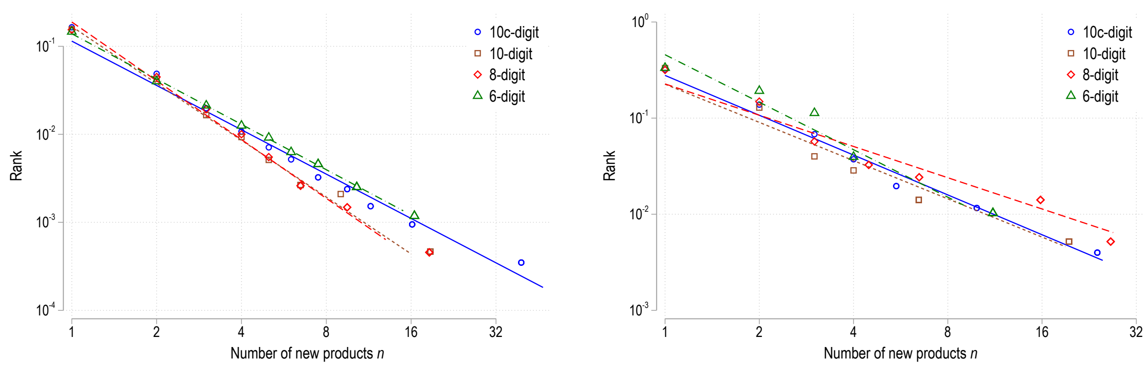
Figure C.3. Average Revenue Per New Product: Different Sizes



(a) Firms with $N_{t-1} = 5$ (b) Firms with $N_{t-1} = 10$

Notes: The figures plot the ratio of total revenue earned on new products divided by the total number of new products on the vertical axis, against the number of new products on the horizontal axis. The line represents a kernel-weighted local polynomial smoothing of degree 5, together with its 95% confidence band.

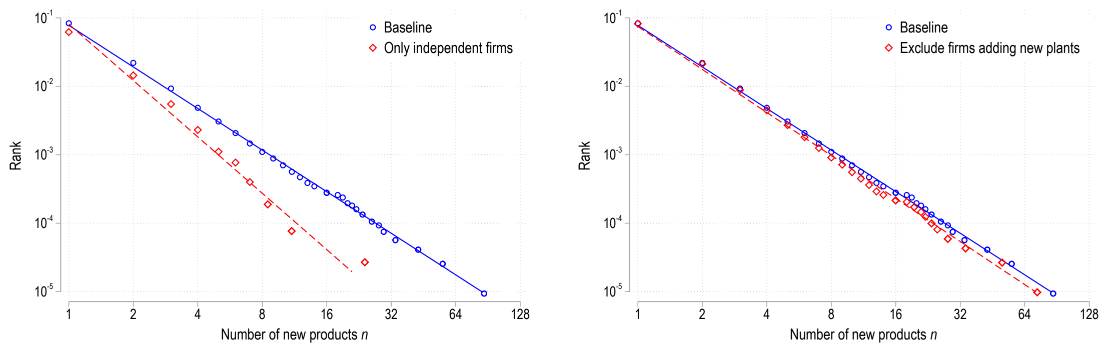
Figure C.4. Distribution of Number of New Products by Level of Aggregation: Different Sizes



(a) Firms with $n_{it-1} = 5$ (b) Firms with $n_{it-1} = 10$

Notes: The figures plot the relationship between a firm's number of new products (horizontal axes) and the firm's rank (vertical axes). The figure provides overlapping plots of the log rank against the log number of products added at the 10c-digit concorded product level (blue circles), 10-digit PRODFRA level (brown squares), 8-digit (red diamonds) and 6-digit level (green triangles). The (rescaled) rank is measured as the ratio of firms' rank starting from the largest firm, divided by the total number of observations in the data.

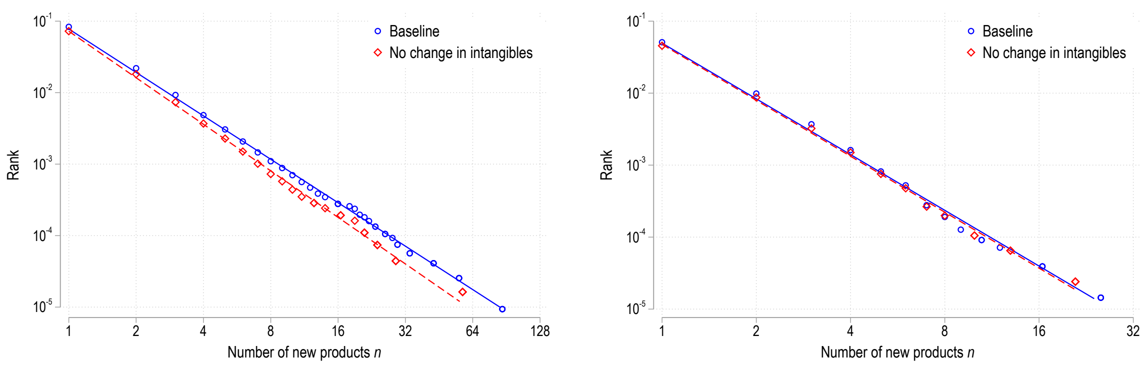
Figure C.5. Distribution of Number of New Products by Proxies of Fixed Firm Boundary (All Firms)



(a) Independent Firms Only (b) No Plant Change

Notes: The figures plot the relationship between a firm's number of new products (horizontal axes) and the firm's rank (vertical axes). The (rescaled) rank is measured as the ratio of firms' rank starting from the largest firm, divided by the total number of observations in the data. The two panels plot the baseline Figure 2(a) against the distribution of new products of firms that do not belong to a business group (left panel) and of firms that do not add new plants over the period (right panel). Note that independent firms are, on average, smaller than firms that are part of a business group.

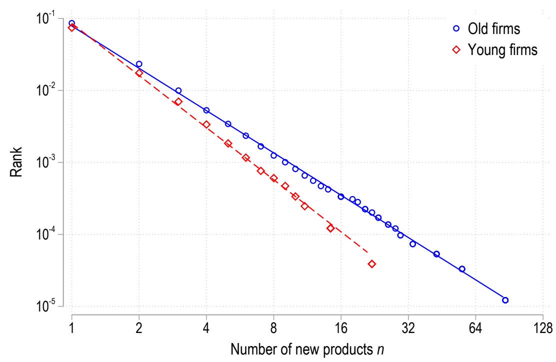
Figure C.6. Distribution of Number of New Products with Fixed Intangible Capital



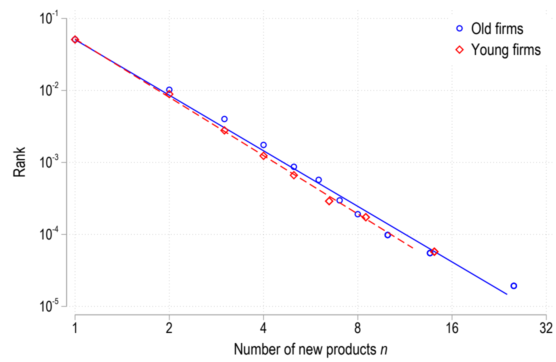
(a) All Firms (b) Firms with $N_{t-1} = 1$

Notes: The figures plot the relationship between a firm's number of new products and its rank. The (rescaled) rank is measured as the ratio of firms' rank starting from the largest firm, divided by the total number of observations in the data. The two panels plot the baseline Figure 2(a) (left panel) and Figure 2(b) (right panel) against the distribution of new products of firms that do not change intangible capital from one period to the next.

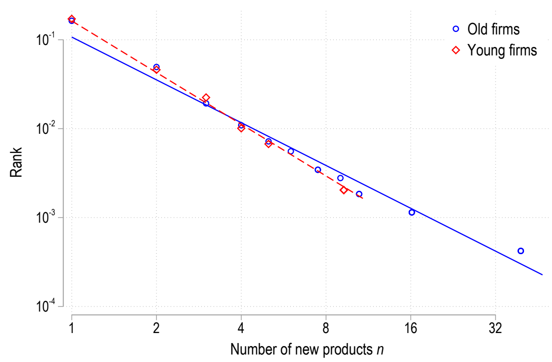
Figure C.7. Distribution of Number of New Products by Age



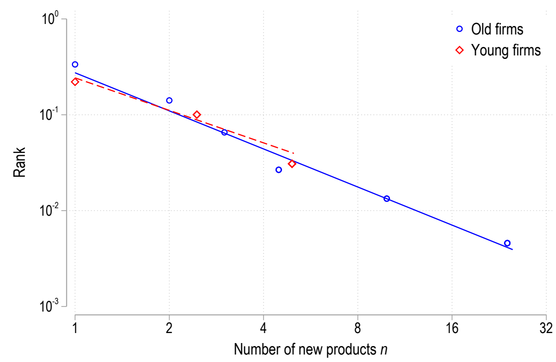
(a) All Firms



(b) Firms with $N_{t-1} = 1$



(c) Firms with $N_{t-1} = 5$



(d) Firms with $N_{t-1} = 10$

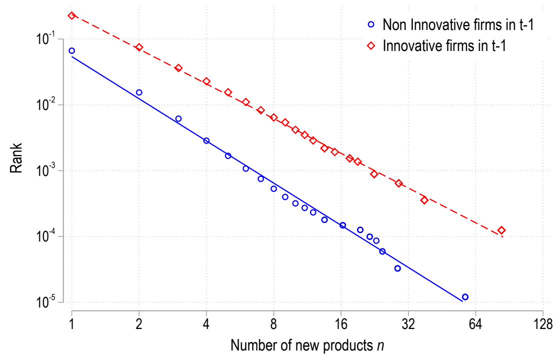
Notes: The figures plot the relationship between a firm's number of new products (horizontal axes) and the firm's rank (vertical axes). The (rescaled) rank is measured as the ratio of firms' rank starting from the largest firm, divided by the total number of observations in the data. The four panels replicate Figure 2 by splitting firms into two groups by age (young if age ≤ 10 years).

Table C.2: Persistence in Revenue per Product of New products

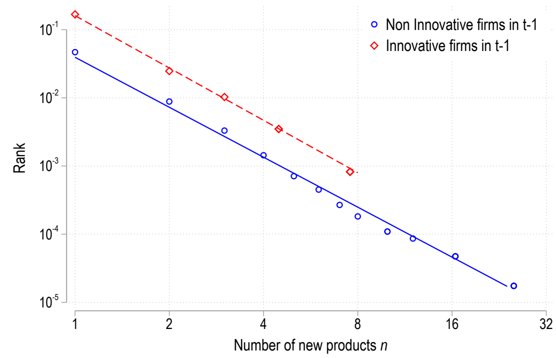
	(1)	(2)
Persistence of product-level revenue on new products	0.39	-0.13
	(0.03)	(0.05)
Product (10-digit) fixed effects	Yes	Yes
Firm fixed effects	No	Yes
R-squared	0.58	0.75

Notes: The table plots the linear regression coefficient from a regression where the dependent variable is the revenue that a firm earns on a newly added product, while the explanatory variable is the average revenue that the firm earned on its new product in its previous episode of product creation. Standard errors in parentheses. Both regressions include 10-digit fixed effects.

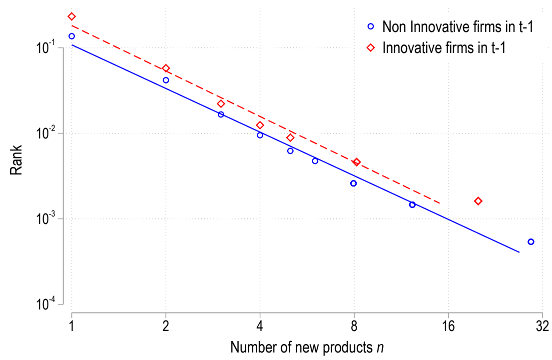
Figure C.8. Distribution of Number of New Products by Innovation History



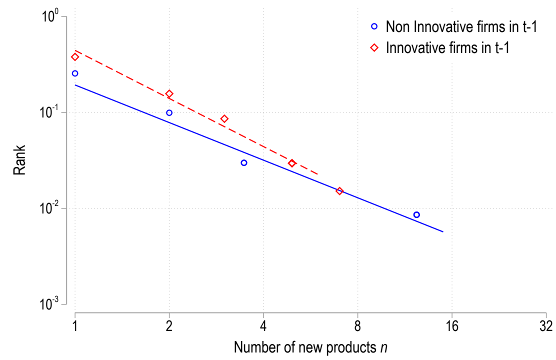
(a) All Firms



(b) Firms with $N_{t-1} = 1$



(c) Firms with $N_{t-1} = 5$



(d) Firms with $N_{t-1} = 10$

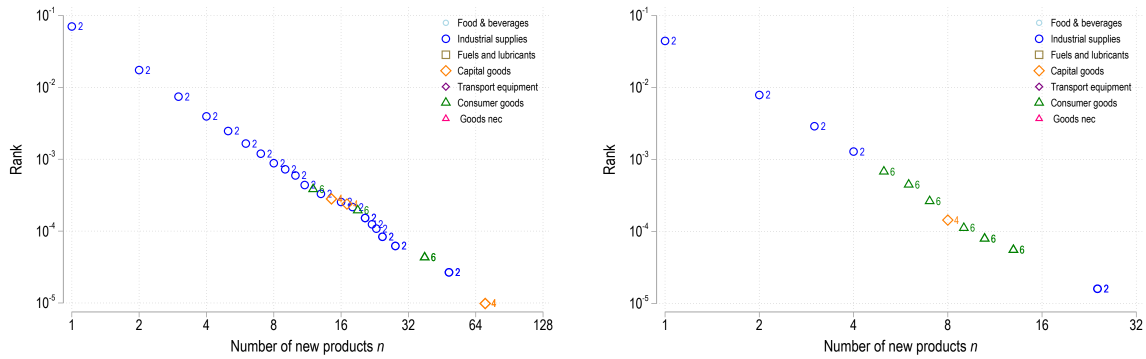
Notes: The figures plot the relationship between a firm's number of new products (horizontal axes) and the firm's rank (vertical axes). The (rescaled) rank is measured as the ratio of firms' rank starting from the largest firm, divided by the total number of observations in the data. The four panels replicate Figure 2 by splitting firms into two groups by innovative status (innovative firms if they introduced new products at $t-1$).

Table C.3: Cannibalization: Effect of New Products on Probability of Product Exit

	(1)	(2)	(3)	(4)	(5)	(6)
β	0.084	0.084	0.079	0.107	0.109	0.071
	(0.044)	(0.044)	(0.041)	(0.045)	(0.045)	(0.032)
R-squared	0.019	0.045	0.085	0.218	0.221	0.346
Controls						
Product count FE	No	Yes	Yes	Yes	Yes	No
Product FE	No	No	Yes	Yes	Yes	No
Firm FE	No	No	No	Yes	Yes	Yes
Additional controls	No	No	No	No	Yes	Yes
Firm-product FE	No	No	No	No	No	Yes

Notes: The table presents regression coefficients for the linear regression along $\mathbb{1}^{exit}_{ijt} = \beta\left(\frac{n^c_{it}}{N_{it-1}}\right) + \gamma' X_{ijt} + \nu_{ijt}$, where $\mathbb{1}^{exit}_{ijt-1}$ is the indicator function that equals 1 for a product that has excited; i is the firm index; t denotes time; N_{it} is the number of products that the firm produces; n^c_{it} is the number of products the firm adds. A coefficient $\beta = 1$ implies full cannibalization, $\beta = 0$ implies no cannibalization. 347,000 observations.

Figure C.9. Distribution of Number of New Products and Product Category (Mode)

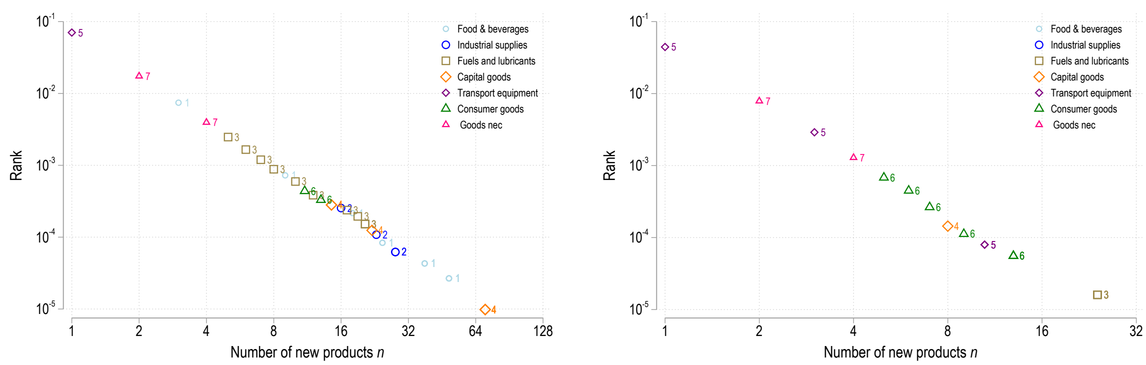


(a) All Firms

(b) Firms with $N_{t-1} = 1$

Notes: The figures plot the relationship between a firm's number of new products (horizontal axes) and the firm's rank (vertical axes). The (rescaled) rank is measured as the ratio of firms' rank starting from the largest firm, divided by the total number of observations in the data. The two panels replicate Figure 2(a) (left panel) and Figure 2(b) (right panel), assigning to each bin of new products n the most frequent 1-digit code from the Broad Economic Categories (BEC) classification. The mode is computed across all new products produced by firms of a given size ('All Sizes', left panel; $N_{t-1} = 1$, right panel) experiencing an innovation burst of n new products.

Figure C.10. Distribution of Number of New Products and Product Category (Relative Frequency)



(a) All Firms

(b) Firms with $N_{t-1} = 1$

Notes: The figures plot the relationship between a firm's number of new products (horizontal axes) and the firm's rank (vertical axes). The (rescaled) rank is measured as the ratio of firms' rank starting from the largest firm, divided by the total number of observations in the data. The two panels replicate Figure 2(a) (left panel) and Figure 2(b) (right panel), assigning to each bin of new products n the most frequent 1-digit code from the Broad Economic Categories (BEC) classification, measured relative to the full population of innovators. The relative frequency is determined by comparing the distribution of BEC codes among all products produced by firms of a given size ('All Sizes', left panel; $N_{t-1} = 1$, right panel) experiencing an innovation burst of n new products, against the distribution of BEC codes for any new product introduced by any firm.

Table C.4: Cannibalization: Effect of New Products on Continuing Products

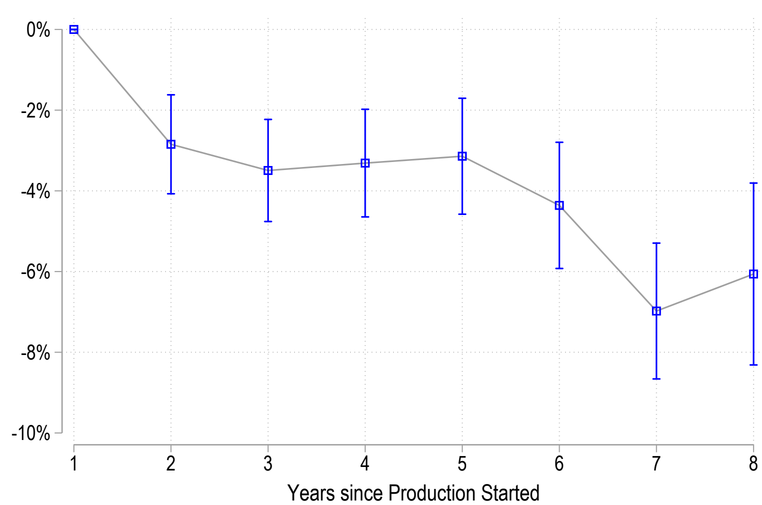
	(1)	(2)	(3)	(4)	(5)	(6)
β	-0.093	-0.095	-0.095	-0.083	-0.085	-0.083
	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
R-squared	0.001	0.007	0.020	0.086	0.088	0.171
Controls						
Product count FE	No	Yes	Yes	Yes	Yes	No
Product FE	No	No	Yes	Yes	Yes	No
Firm FE	No	No	No	Yes	Yes	Yes
Additional controls	No	No	No	No	Yes	Yes
Firm-product FE	No	No	No	No	No	Yes

Notes: The table presents regression coefficients for the linear regression along

$$\frac{R^{-}_{ijt} - R^{+}_{ijt-1}}{0.5(R^{-}_{ijt} + R^{+}_{ijt-1})} = \beta\left(\frac{R^{N}_{it}/n^{-}_{it-1}}{0.5(R^{+}_{ijt} + R^{-}_{ijt-1})}\right) + \gamma' X_{ijt} + \nu_{ijt},$$

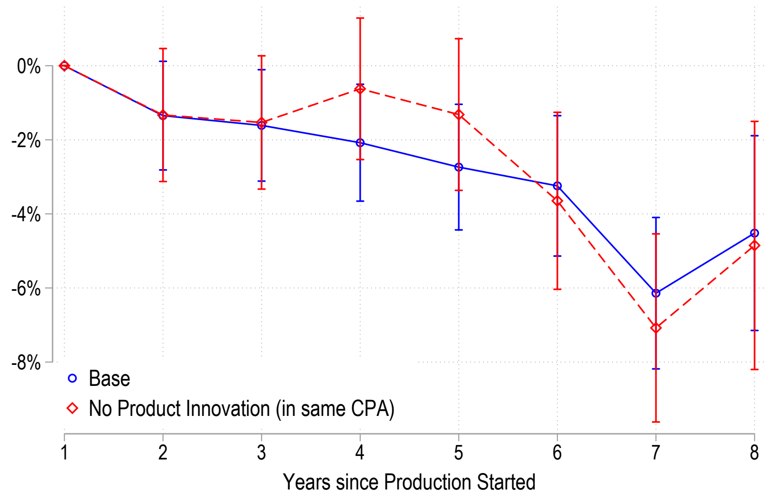
where R^{N}_{it} denotes total revenue from new products and R^{-}_{ijt} denotes revenue for kept product j while R^{+}_{ijt-1} denotes the revenue that the firm earned on that product at $t-1$; i is the firm index; t denotes time; n^{-}_{it} denotes the number of products that firm i produces and keeps producing. A coefficient $\beta = 1$ implies full cannibalization, $\beta = 0$ implies no cannibalization. 347,000 observations.

Figure C.11. Life Cycle of Revenue Growth with 10c-digit Concorded Products



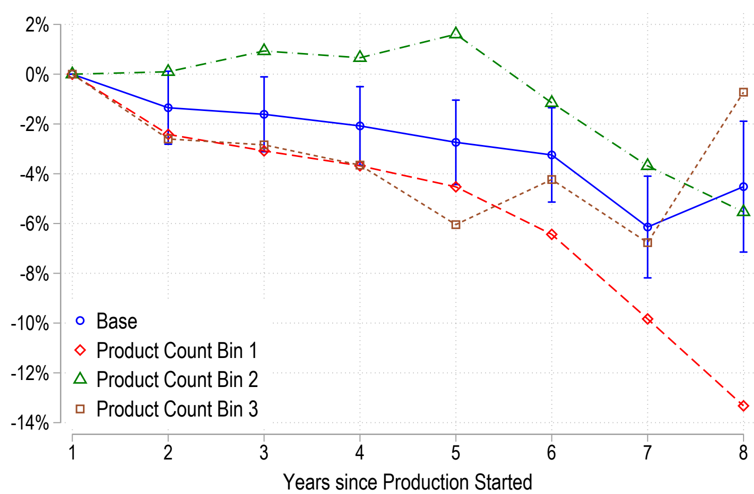
Notes: The vertical axis plots the average growth of product revenue among all firm-products in the sample relative to growth at tenure 1 and after absorbing firm fixed effects. The horizontal axis plots the number of years that have passed since the firm first started producing the product. Confidence bounds are at the 90% level and based on clustered standard errors. Products are defined at the 10c-digit concorded level.

Figure C.12. Life Cycle of Revenue Growth with No Product Innovation



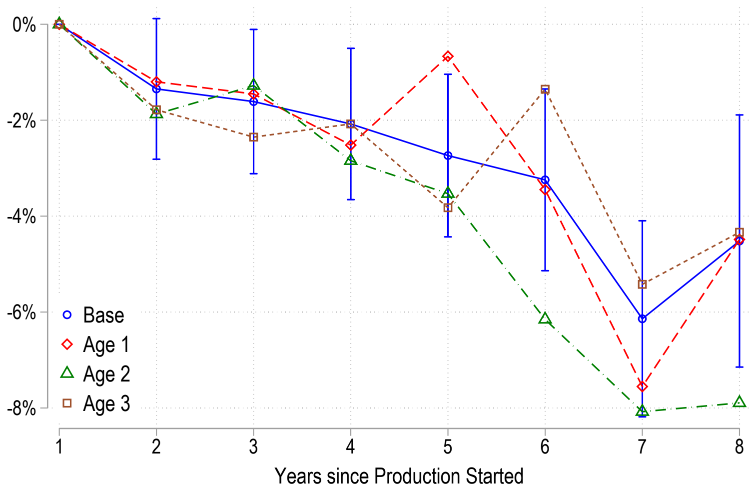
Notes: The vertical axis plots the average growth of product revenue among all firm-products in the sample relative to growth at tenure 1 and after absorbing firm fixed effects. The horizontal axis plots the number of years that have passed since the firm first started producing the product. Confidence bounds are at the 90% level and based on clustered standard errors. The line for 'No Product Innovation' excludes firms that introduce any new product within the same CPA product category (6-digit) over the life cycle of any product.

Figure C.13. Life Cycle of Revenue Growth by Product Count



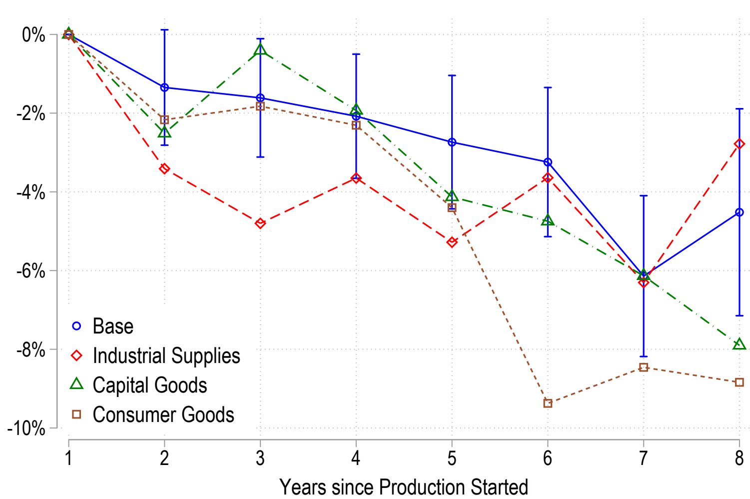
Notes: The vertical axis plots the average growth of product revenue among all firm-products in the sample relative to growth at tenure 1 and after absorbing firm fixed effects. The horizontal axis plots the number of years that have passed since the firm first started producing the product. Confidence bounds are at the 90% level and based on clustered standard errors. Firms are binned into three equisized bins based on the number of products that they are producing.

Figure C.14. Life Cycle of Revenue Growth by Age



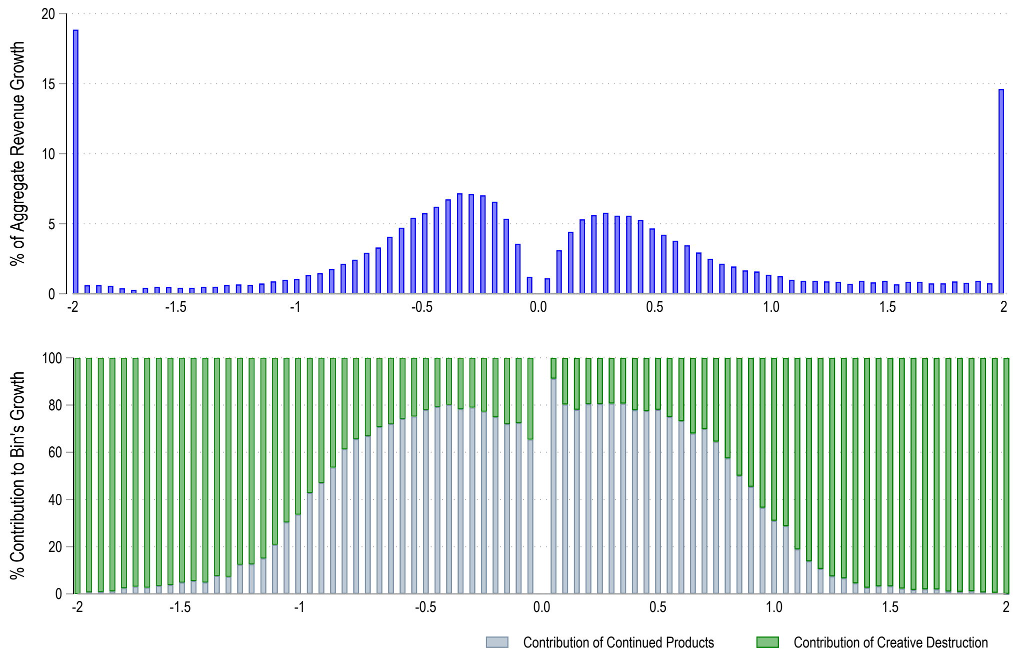
Notes: The vertical axis plots the average growth of product revenue among all firm-products in the sample relative to growth at tenure 1 and after absorbing firm fixed effects. The horizontal axis plots the number of years that have passed since the firm first started producing the product. Confidence bounds are at the 90% level and based on clustered standard errors. Firms are binned into three equisized bins based on their age.

Figure C.15. Life Cycle of Revenue Growth by Sector



Notes: The vertical axis plots the average growth of product revenue among all firm-products in the sample relative to growth at tenure 1 and after absorbing firm fixed effects. The horizontal axis plots the number of years that have passed since the firm first started producing the product. Confidence bounds are at the 90% level and based on clustered standard errors.

Figure C.16. Untargeted Moments: Creative Destruction and Aggregate Revenue Growth



Notes: The horizontal axis measures firm growth through the symmetric growth rate, defined as the change in revenue between t and $t-1$ divided by average revenue in t and $t-1$. Growth rates are separated into 20 negative bins and 20 positive bins. The top figure presents the contribution of changes in revenue across firms in a particular growth bin as a percentage of total revenue creation (the sum of increases in revenue across growing firms) for positive bins or as a percentage of total revenue destruction (the sum of decreases in revenue across shrinking firms) for negative bins. The bottom panel decomposes a bin's overall revenue change into changes coming from continuing products and the net of product innovation and destruction – creative destruction.

Table D.5: Concentration of Firm Growth

	1 Year				5 Year			
	-3 σ	-2 σ	+2 σ	+3 σ	-3 σ	-2 σ	+2 σ	+3 σ
Product Count								
Data	0.98	1.08	1.09	1.00	0.90	0.93	0.99	0.96
Normal Dist.	0.13	2.28	2.28	0.13	0.13	2.28	2.28	0.13
Ratio	7.3	0.5	0.5	7.4	6.7	0.4	0.4	7.1
Revenue								
Data	0.47	0.97	0.84	0.42	0.13	0.34	0.55	0.34
Normal Dist.	0.13	2.28	2.28	0.13	0.13	2.28	2.28	0.13
Ratio	3.5	0.4	0.4	3.1	1.0	0.2	0.2	2.5
Employment								
Data	0.77	1.51	0.95	0.38	0.27	0.75	0.65	0.19
Normal Dist.	0.13	2.28	2.28	0.13	0.13	2.28	2.28	0.13
Ratio	5.7	0.7	0.4	2.8	2.0	0.3	0.3	1.4

Notes: The table presents the percentage of observations in the tail of the data and the percentage of observations in the tail under the normal distribution. Colomn headers indicate tail in terms of standard deviations from the mean.

Appendix D. Distribution of firm growth

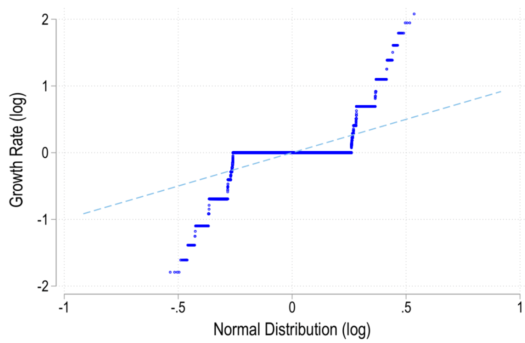
In this section, we show that the distribution of firm *growth* in our data is fat tailed. As with size, we measure firm growth through changes in the number of 10-digit products that it sells, changes in revenue and changes in employment. We analyze either year-on-year or five-year changes in the natural logarithm of each of these variables: $\ln y_{it} - \ln y_{it-h}$, where $h = 1,5$. By focusing on log changes we are able to study the entire distribution of firm growth. The main alternative measure of firm growth is the symmetric growth proposed by Davis et al. (2006), which is bounded by $[-2,2]$ regardless of the underlying distribution of firm size. This makes it impractical for studying the tail of the firm-growth distribution.

To show that the distribution of firm growth is fat tailed, we compare the distribution of log-change to the log-normal distribution. Log-normal is a natural starting point as, motivated by Gibrat (1931), firm dynamics models with random shocks to productivity typically assume shocks (and thus firm growth) are log normal (see, e.g., Hopenhayn 2014).[29]

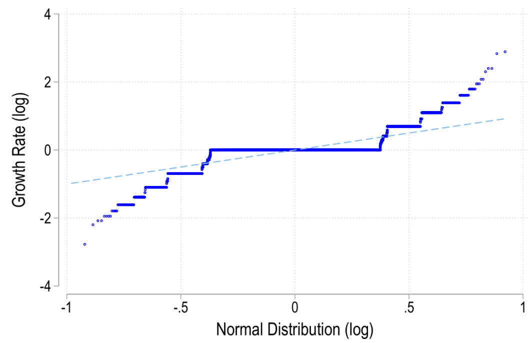
Quantile-Quantile plots We plot the distribution of firm growth in Figure D.17. Each sub-figure contains a Quantile-Quantile (QQ) plot comparing the distribution of a measure of firm growth to a log-normal distribution with the same mean and standard deviation. To produce these figures, we sort firm-years by growth and calculate ranks. The plot than compares a particular quantile's growth rate on the y-axis to what that quantile's growth rate would be under the log-normal distribution. If growth rates are log-normally distributed, the scatters should be positioned on the diagonal reference lines. If

[29] Our results echo recent findings in the literature on income dynamics which shows that individual income growth is also better described by a fat-tailed distributions such as the Pareto distribution (see, e.g., Guvenen et al. (2021). Gabaix et al. (2016) discuss the importance of the income growth process to explain stationary distributions of income and wealth.

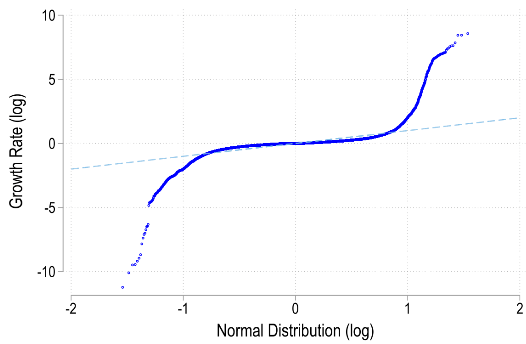
Figure D.17. Distribution of Growth Rates



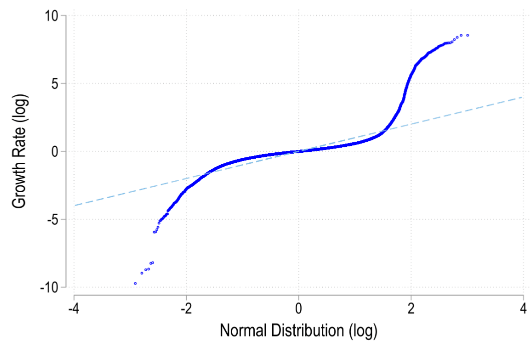
(a) Product Count Growth



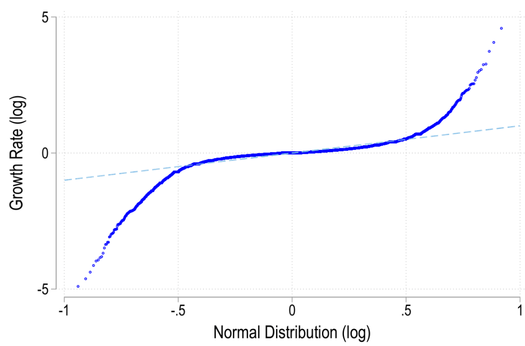
(b) Product Count Growth (5 year)



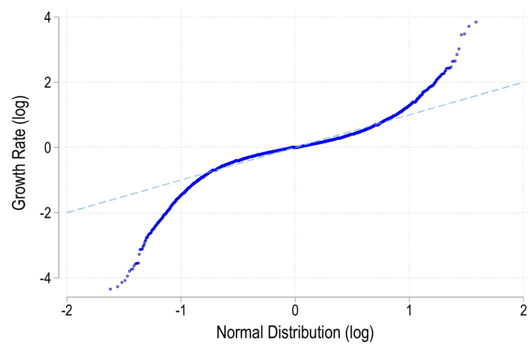
(c) Revenue Growth



(d) Revenue Growth (5 year)



(e) Employment Growth



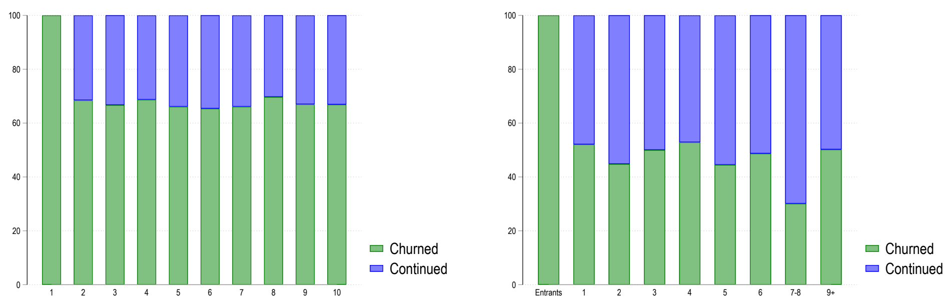
(f) Employment Growth (5 year)

Notes: The figures plot Quantile-Quantile plots comparing the empirical distribution of log change in product count, revenue and employment against the log-normal distribution. Dashed light-blue lines present the reference log-normal distribution.

growth is fat tailed, scatters of the left are below the reference line; scatters on the right are above the reference line.

Figure D.17 shows that firm growth is fat tailed in the data. The upper, middle, and lower figures respectively plot the distribution product count, revenue, and employment. For each measure, both the 1-year (left-hand) and the 5-year growth rate show that large positive and negative growth occurs more often than a log-normal distribution predicts. The deviation is particularly clear for the

Figure D.18. Symmetric Shapley-Owen Decomposition of Revenue Growth by Age and Size



(a) Age (b) Product Count

Notes: The figures decompose total revenue growth into revenue loss from product loss, revenue gain from product gain, and changes in revenue on products that the firm is continuing to produce.

largest changes. For employment, for example, the deviation becomes graphically clear around ± 0.5 log change, which translate to -45% and +65%.

Tail mass Table D.5 quantifies how much the tails of firm growth deviates from the log-normal distribution. It compares the percentage of observations to the left (right) of minus (plus) 2 and 3 standard deviations from the mean for each variable and horizon. For product count, the data shows a close alignment with the normal distribution at the +2 standard deviations and +3 standard deviations levels in the 1-year period, with ratios of 0.5 and 7.4 respectively, indicating that extreme changes in a firm's product count are relatively common. Revenue data again displays a marked deviation from the normal distribution, with changes in annual revenue growth outside the ±3 standard deviations occurring more than three times as often. Results for employment growth are similar.

Log-rank versus log-size Having established that extreme firm growth is much more common than under the log-normal distribution, we next explore what distribution fits better. As growth can be both positive and negative, the standard Pareto or Zeta distribution are poor fits. Instead, we explore whether a two-sided version of these distribution can explain both the tails of firm growth. To see if this is the case, we separately calculate the rank of firm growth firms among the firms with negative growth ($y_{it} < y_{it-h}$) and positive growth ($y_{it} > y_{it-h}$) and plot the log of the inverse rank against the log of firm growth (y_{it}/y_{it-h}). This relationship is linear at either side of zero if firm growth is indeed two-sided Pareto.

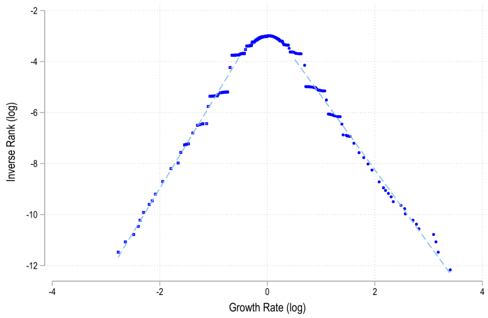
Figure D.19 plots the results. The subfigures have the same order as in Figure D.17. The horizontal axis range is fixed at -4 to 4 in all plots. The figure shows that the two-sided Pareto distribution fits the empirical firm growth well. For product count, the linear reference lines are a near perfect fit for positive growth in both the 1-year and 5-year growth plots, although negative product growth seems somewhat less thick-tailed than the Pareto distribution at the 5-year horizon. For revenue growth and employment growth the tails appear to be *even fatter* than the two-sided Pareto distribution. This is particularly visible for the right tail of the distribution: as the scatter plot is mostly above linear reference lines, extreme growth is more common in the data than in a Pareto distribution.

Conditioning on age or size We next assess how the tails of the firm growth distribution change with age and size. Figure D.20 presents the QQ plots for all measures of 1-year growth, dividing the sample into three equally large age groups on the left-hand figures. Blue-round scatters plot the distributions for the youngest third of firms, while red-square and green-diamond scatters respectively belong to the middle and oldest age group. (describe result: tails decline with size, stronger result than by age).

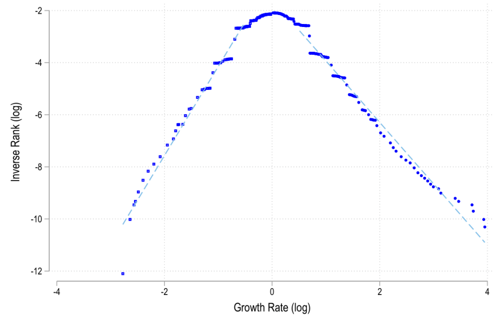
Figure D.18 presents the symmetric Shapley-Owen decomposition separately for firms of different age deciles (left-hand) and initial sizes (right-hand), grouping entry and exit into churn. Size is measured through the number of products that firms produced in the previous year.[30] The figure shows that the contribution of churn is similar for firms of different ages and sizes. The only outlier is entrants, which inherently see their entire growth attributed to churn as they have no continuing products.

[30] Results are similar when measuring size through revenue and employment.

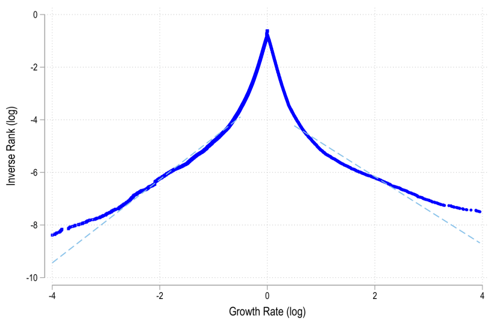
Figure D.19. Two-Sided Pareto Distribution of Firm Growth



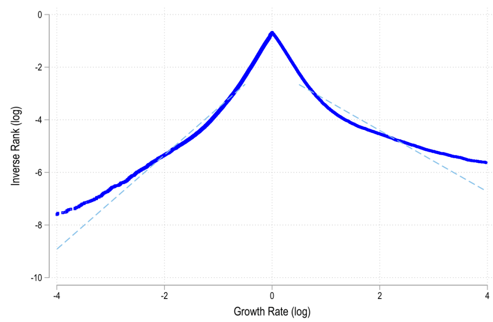
(a) Product Count Growth



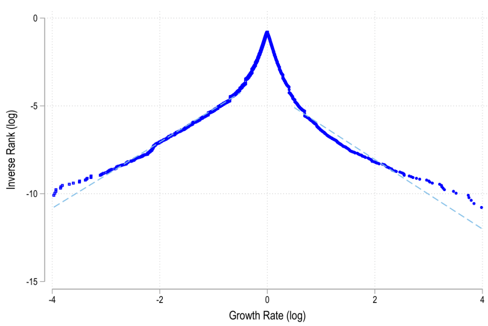
(b) Product Count Growth (5 year)



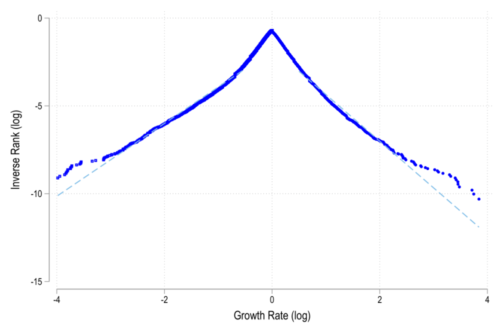
(c) Revenue Growth



(d) Revenue Growth (5 year)



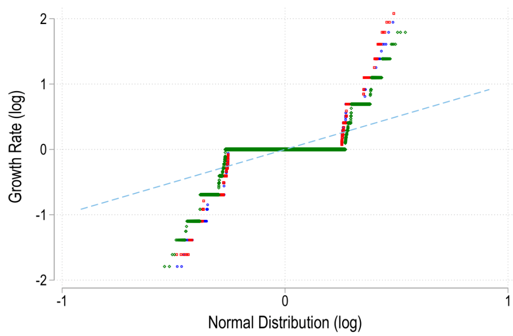
(e) Employment Growth



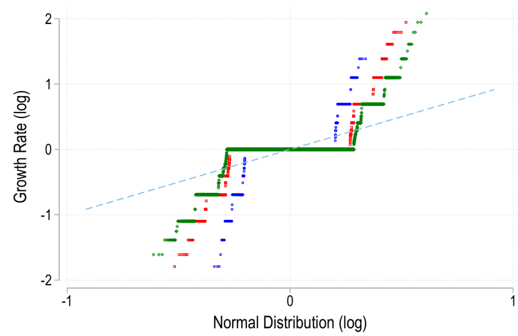
(f) Employment Growth (5 year)

Notes: The figures plot the relationship between the log of a firm's inverse rank for growth and the log of growth. Growth is defined as y_{it}/y_{it-h}. Rank is calculated separately for firms with negative growth and firms with positive growth. Blue-dashed lines are illustrative linear reference lines to asses whether

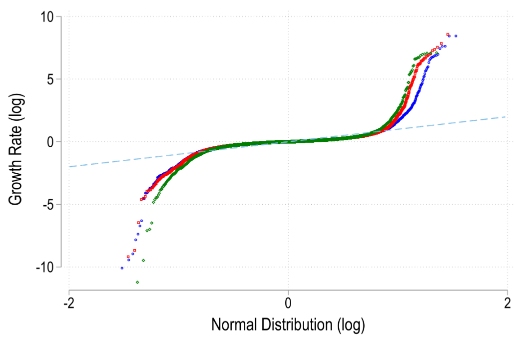
Figure D.20. Distribution of Growth Rates by Age and Initial Size



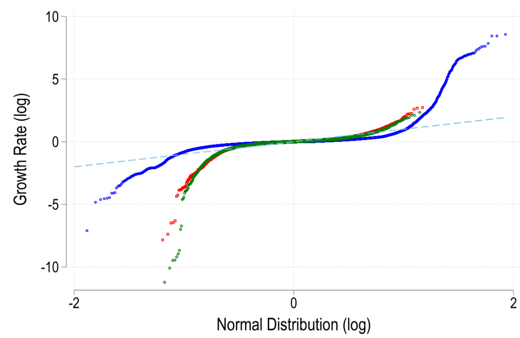
(a) Product Count Growth - by Age



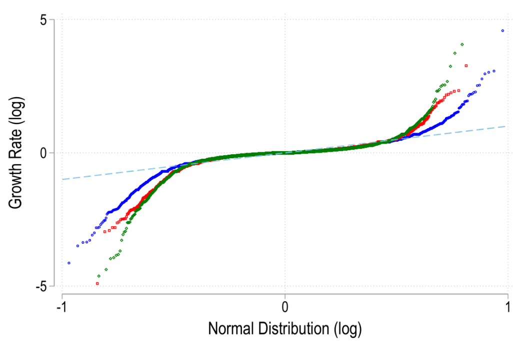
(b) Product Count Growth - by Size



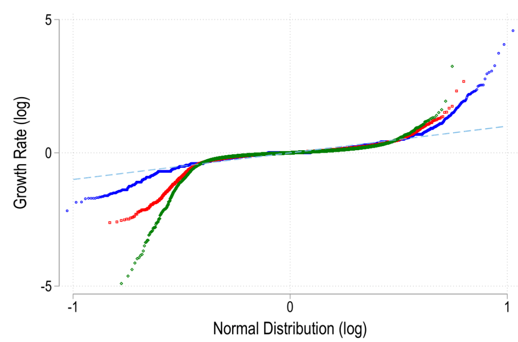
(c) Revenue Growth - by Age



(d) Revenue Growth - by Size



(e) Employment Growth - by Age



(f) Employment Growth - by Size

Notes: The figures plot Quantile-Quantile plots comparing the empirical distribution of log change in product count, revenue and employment against the log-normal distribution. Dashed light-blue lines present the reference log-normal distribution.