Technology-Driven Market Concentration through Idea Allocation*

Yueyuan Ma[†] Shaoshuang Yang[‡]

This Version: December 2025

Abstract

Using a newly-created measure of technology novelty, this paper identifies periods with and without technology breakthroughs from the 1980s to the 2020s in the US. It is found that market concentration decreases at the advent of revolutionary technologies. We establish a theory addressing inventors' decisions to establish new firms or join incumbents of selected sizes, yielding two key predictions: (1) A higher share of inventors opt for new firms during periods of heightened technology novelty. (2). There is positive assortative matching between idea quality and firm size if inventors join incumbents. Both predictions align with empirical findings and collectively contribute to a reduction in market concentration when groundbreaking technologies occur. Quantitative analysis shows that the slowdown in technological breakthroughs predicts faster growth in average firm quality, but this effect is more than offset by slower growth in net firm entry. The slowdown also accounts for a large share of both the upward trend and the fluctuations in market concentration.

Keywords: technological waves, HHI, startups, incumbent firms.

*We would like to thank Ufuk Akcigit, Thomas Chaney, Murat Celik, Emin Dinlersoz, Jesus Fernandez-Villaverde, April Franco, Ted Frech, Jeremy Greenwood, Jeremy Pearce, Gerard Hoberg, Ayse Imrohoroglu, Boyan Jovanovic, Yueran Ma, Peter Rupert, Mehmet Yorukoglu, and seminar participants at UBC, UCR, USC, UF, UPenn, the NBER PIE conference, the Tepper-Laef conference, the Duke Macro Jamboree, and the Cement Workshop for helpful comments. We also gratefully acknowledge the Google patent data provided by Swapnika Rachapalli. Any views expressed are those of the authors and not those of the U.S. Census Bureau. The Census Bureau's Disclosure Review Board and Disclosure Avoidance Officers have reviewed this information product for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to this release. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2125 (CBDRB-FY24-P2125-R11032, CBDRB-FY26-P2125-R12802)

[†]Affiliation: University of California, Santa Barbara. Email: yueyuanma@ucsb.edu.

[‡]Affiliation: The Chinese University of Hong Kong, Shenzhen. Email: yangshaoshuang@cuhk.edu.cn.

1 Introduction

The rise in U.S. market concentration and the slowdown in growth since the late 1990s have raised significant concerns. The period has witnessed declining firm entry and an expansion of productive incumbent firms (Autor et al. (2020)). The existing literature attributes this shift to the increasing market power of incumbents, which strategically deter the entry of new businesses (Cunningham, Ederer and Ma (2021); Akcigit and Goldschlag (2023)), thereby hindering the diffusion of innovative ideas. This paper offers a new perspective, providing empirical evidence and structural analysis to show that technological waves play an important role in shaping growth and concentration by reallocating innovative ideas between incumbents and entrants. The slowdown in technological breakthroughs contributes to slower growth and rising concentration, consistent with the view that good ideas are becoming harder to find (Bloom et al. (2020)).

Using a newly developed measure of technological novelty, this paper identifies distinct periods in the U.S. from the 1980s to the 2020s characterized by breakthrough innovations or incremental advances that build on existing technologies. The analysis reveals a declining long-term trend in technological novelty, punctuated by cyclical waves. During peaks, groundbreaking innovations that depart significantly from existing technologies emerge, whereas during troughs, most new technologies reflect a mature and incremental phase of development.

Surprisingly, we find that market concentration, as measured by the Herfindahl-Hirschman Index (HHI) of firm sales, employment, or payroll, exhibits both a rising trend and a cyclical pattern that is notably negatively correlated with waves of technological novelty. This pattern suggests that the emergence and maturation of novel technologies play a significant role in shaping the dynamics of market concentration.

How are technological waves and market concentration connected? A potential channel is through the allocation of ideas. Since firm size is to a large extent impacted by firm productivity and new ideas are important sources of productivity growth, where new ideas contribute their value will determine the firm size distribution, and therefore, market concentration. Combining the Longitudinal Business Database (LBD) from the Census Bureau and the patent information from the USPTO, this paper tracks the affiliation of patents at their formation. It is shown that at the peaks of the technological waves, a larger share of patents are forming in new businesses, while at the troughs, a larger share of patents come from incumbent firms. Besides, among patents from incumbent firms, there is a positive relationship between patent citations, a quality measure of the ideas behind them, and the size of the firm. These patterns indicate that technological waves affect the number of firm entries and the way new ideas combine with firms of different sizes.

Further patent-level regression analysis reveals that incumbent firm size positively affects the private economic value of patents, given their scientific value, indicating synergy between inventors and incumbent firms. The scientific value of patents has a positive impact on the economic value, while this impact decreases in the aggregate technological novelty, indicating adoption frictions for novel ideas within incumbents.

Based on the empirical findings, this paper proposes a theory about inventors' choice of where to contribute the value of their ideas, and how it connects the technological waves, growth, and market concentration. The technological novelty is assumed to be a random aggregate shock capturing the random arrival of ground-breaking innovations in a period. Each inventor is endowed with an idea of idiosyncratic quality. The inventor needs to choose between forming a new firm of a random size (à la Romer (1990)) with a partner or joining an incumbent firm and climbing the quality ladder. In the case of the latter, she must also decide on the size of the incumbent firm to join. It is frictional for an incumbent firm to adopt new technology as in Greenwood and Yorukoglu (1997), and the friction increases in the aggregate technological novelty. Hence, the arrival of groundbreaking technologies induces more inventors to found startups. These choices shape firm entry, quality upgrading, and, ultimately, the firm size distribution. Simulations of the calibrated model show that a slowdown in breakthrough innovation reduces aggregate growth, even as average firm quality grows faster. Moreover, the concentration dynamics generated by technological waves account for 83.4% of the observed upward trend in market concentration and 43.6% of the detrended volatility.

The model in this paper includes three key elements: realization potential of ideas under adoption frictions, commercialization synergy, and inventor-firm contracts. The realization potential captures the economic value of innovation when an incumbent firm, already utilizing existing technologies, incorporates new ideas into its production processes. The imperfect substitution between the new idea and existing technologies reduces the economic value of the idea, highlighting the adoption frictions. These frictions intensify during peaks of technological waves, when new ideas diverge further from existing technologies. Consequently, startups, free from these adoption frictions, emerge as more attractive platforms for innovation. Commercialization synergy pertains to the added value incumbent firms can provide through their production and commercialization capacities, which startups typically lack. Larger firms offer more synergy, especially to high-quality ideas. Inventor-firm contracts govern the collaboration between inventors and firms, ultimately shaping the allocation of ideas. Given the risky nature of R&D and the unobservable effort of inventors, these contracts are designed to elicit optimal effort through a combination of equity and wages. Larger firms face more severe incentive challenges due to exposure to greater shocks not related to innovation, which diminishes the effectiveness of equity as an incentive mechanism for R&D.

Inventors must consider the realization potential, synergy, and contract terms when choosing between startups and incumbent firms. Startups, while free from adoption friction and offering more aligned incentives, lack the capacity to generate synergy. In contrast, incumbents face adoption friction, with larger firms providing weaker incentives but better synergy. These trade-offs guide inventors in their strategic decision-making regarding firm affiliation.

The model has two major predictions. First, a larger share of inventors choose to start new firms to develop their ideas during periods of high technological novelty since the realization potential at incumbent firms is lower. Second, among inventors that choose to do R&D in incumbent firms, there is positive assortative matching between idea quality and firm size. Therefore, firms already with a larger size attract ideas of higher value. These two predictions are consistent with observations in data and collectively contribute to a reduction in market concentration when the economy is closer to the peak of the technological waves. The upsurge in new startups leads to a proliferation of firms in the market. Besides, given that new startups are less affected by the positive matching between idea quality and firm size, they offer a counterbalance to the tendency of larger firms to further expand.

To quantify how technological novelty waves affect economic growth and market concentration through the allocation of new ideas, we calibrate the model and run counterfactual simulations that vary the economy-wide novelty of new technologies over time. In these simulations, we hold all parameters fixed except the year-specific novelty parameter for each year after 1986, the first peak of the technological waves in our sample. We then compute annual growth rates and decompose them into (i) growth driven by improvements in average firm quality and (ii) growth driven by net expansion in firm mass. We also generate time paths for two key moments: (1) the Herfindahl–Hirschman Index (HHI) of firm size, and (2) the ratio of ideas developed in new firms relative to incumbent firms. The simulated paths are compared to their empirical counterparts.

The evolvement of the technological novelty leads to a lower aggregate growth rate over 1986-2016. Although the average firm quality grows at a higher rate, the growth rate in firm mass decreases more. The simulated paths for the HHI and idea allocation closely track their empirical counterparts. In particular, the model-generated HHI reproduces 83.4% of the observed upward trend. Moreover, the correlation between the detrended simulated and actual HHI is 0.552, and the corresponding correlation for the detrended ratio of ideas in new versus incumbent firms is 0.706. Together, these results indicate that technological waves are a key driver of both market concentration and idea allocation.

To decompose the effect of the two channels, changes in firm numbers (extensive margin) and the positive assortative matching (intensive margin) between idea quality and firm size, on the evolvement of market concentration, we track the HHI change driven by firm numbers in the simulation process. The decomposition shows that both margins contribute to the rising trend in market concentration.

Related Literature

In the endogenous growth literature, innovation is commonly modeled as either variety expansion (as in Romer (1990)) or quality-ladder improvement (as in Aghion and Howitt (1990); Grossman and Helpman (1991); Klette and Kortum (2004); Akcigit and Kerr (2018)). However, their distinct implications for growth and market concentration have not been systematically examined. On the growth side, this paper shows that inventors endogenously choose between these two innovation modes by trading off adoption frictions against synergistic gains. This trade-off shapes the aggregate growth rate and varies over technological waves. On the market concentration side, if each firm produces a single variety, variety expansion predicts an increase in the number of firms and a decrease in concentration. In contrast, quality-ladder growth has ambiguous implications for concentration, depending on how idea quality is matched with firm quality. This paper provides both empirical evidence and theoretical predictions of positive assortative matching, implying that quality-ladder-driven growth increases market concentration.[1]

Our empirical and theoretical analyses highlight the role of technological novelty in shaping where inventors conduct R&D and, in turn, market concentration. This offers a new perspective on the connection between idea allocation and concentration. Prior studies (e.g., Cunningham, Ederer and Ma (2021); Akcigit and Goldschlag (2023)) emphasize that incumbents may strategically acquire and shelve external innovations to deter competition, contributing to a decline in novelty. Our findings do not contradict these conclusions, but instead provide a complementary mechanism. Specifically, this paper implies that technological novelty and market concentration can mutually reinforce each other, strengthening the negative association between technological novelty waves and market concentration.

This paper offers a new explanation for the rise in U.S. market concentration since the late 1990s. Although this trend has coincided with gains in allocative efficiency and productivity (Autor et al. (2020); Ganapati (2021)), it has also been accompanied by declining firm entry and a growing gap between large and small firms (Akcigit and

[1] Literature on the interaction between the firm-size distribution and growth includes Chatterjee and Rossi-Hansberg (2012); Perla, Tonetti and Waugh (2021); Fons-Rosen, Roldan-Blanco and Schmitz (2021); Cassiman and Veugelers (2006); Bena and Li (2014); Akcigit, Celik and Greenwood (2016); Cavenaile, Celik and Tian (2019); Liu and Ma (2021); Ma (2022); Yang (2023).

Ates (2023); Olmstead-Rumsey (2019)), as well as increasing innovation difficulty (Bloom et al. (2020)). We propose that a slowdown in radical technological breakthroughs helps reconcile these observations. Our Novelty Index reveals that major waves of innovation peaked in the mid-1980s and mid-1990s, with a prolonged 17-year lull until a resurgence in the early 2010s. During this stagnation, inventors increasingly turned to incumbent firms, enabling large incumbent firms to further expand and driving up concentration.

Finally, our analysis delves into the implications of the introduction of groundbreaking technologies. Bowen III, Frésard and Hoberg (2023) show empirically that in an era with rapid evolving technologies, more startups remain independent rather than being sold out. Dinlersoz, Dogan and Zolas (2024) discover a surge in AI business applications after 2016. Greenwood and Yorukoglu (1997) and Greenwood and Jovanovic (1999) establish that technological revolutions lead to deterioration in the stock value of existing firms. Jovanovic and Rousseau (2014) shows that at the advent of new technologies, incumbent firms decrease investment due to lack of compatibility while new firms increase investment. This paper extends the existing literature by investigating how a leap in technological progress affects the distribution of firm sizes, primarily due to the frictions when integrating inventors' novel ideas into incumbent firms. It is shown that market concentration is another important outcome of technological revolutions. This paper demonstrates that apart from the high-frequency business cycle influenced by productivity fluctuations (Kydland and Prescott (1982)), the economy may also be susceptible to a low-frequency cycle driven by the waves of technological novelty.

The rest of the paper is organized as follows. Section 2 introduces measures of the technological waves, market concentration, and the allocation of ideas, and subsequently presents their patterns. Section 3 constructs a model where inventors make decisions between initiating new ventures or joining established incumbents at specific sizes. We derive predictions about the mapping between the quality of inventors' ideas and their optimal choices. Section 4 calibrates the model. Section 5 simulates the model to evaluate the degree to which technological waves can account for changes in market concentration through the idea allocation channel. Section 6 concludes.

2 Empirical Patterns

This section exhibits empirical patterns of the technological waves, market concentration, and the choices of the inventors on where to contribute their ideas. A description of the data used in this section is provided in Appendix A.

2.1 Technological Waves

Technology waves capture the extent of new technology breakthroughs over time. At the peak of the technological waves, significantly novel technologies emerge that are often incompatible with existing technologies; at the trough of the waves, most of the technologies in the economy have reached a mature state, and the improvement over existing ones is incremental.

2.1.1 Measurement

To measure the technological waves, we create a Novelty Index of the new technologies in each year using the patent citation data. Specifically,

$$\text{Novelty}_t = \frac{\sum_{i\in I_t}\sum_{s=0}^{5}\text{Forward Citations}_{i,t+s}}{\sum_{i\in I_t}\sum_{s=0}^{5}\text{Forward Citations}_{i,t+s} + \sum_{i\in I_t}\sum_{s=0}^{5}\text{Backward Citations}_{i,t-s}}, \quad (1)$$

where I_t is the set of the new patents granted in year t. The numerator is a summation of the number of forward citations (citations by others) each new patent gets within the next five years. The denominator is a summation of the number of forward citations plus a summation of the number of backward citations (citation on others) each patent makes on other patents granted within the previous five years. The five-year window is to ensure every year in the sample is compared on the common ground, since more recent patents are more likely to receive fewer forward citations due to the right-censoring issue. The rationale for this measure is that groundbreaking innovations typically exhibit lower similarity to current technologies, but pave the way for subsequent patents to emulate them. Since the forward citations capture the overlap of future patents with the focal patent, while the backward citations capture the overlap of the focal patents with previous patents, the relative number of the former provides a measure of patent novelty. The Novelty Index is in the range between zero and one. A higher index indicates that the year witnesses significant breakthroughs in new technologies; a lower index indicates that most of the technologies have evolved into a mature stage in that year.

The data used to generate the Novelty Index comes from the USPTO patent and citation data. The USPTO records all patents granted after 1976 and all the patents they cite. To get a smoother trend, we take a three-year average for each observation,[2]

$$\text{Novelty\_avg}_t = \frac{1}{3}\sum_{h=-1}^{1}\text{Novelty}_{t+h}. \quad (2)$$

[2]The smoother does not change the original pattern, as shown in figures without the smoothing techniques in Appendix B.1

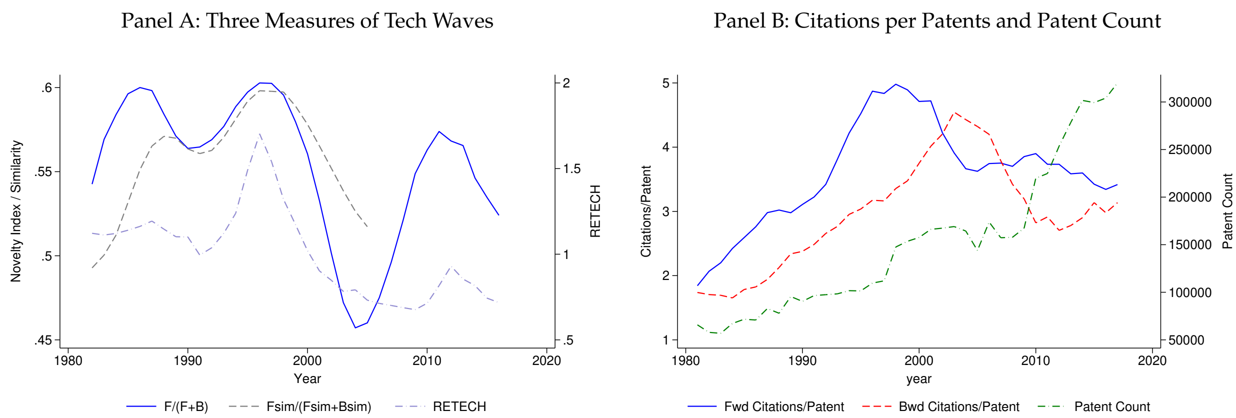

Figure 1: Technological Waves

Notes: Panel A illustrates three measures of the technological waves. The blue solid curve is based on the methodology defined in this paper, while the gray dashed curve is based on the similarity metric in Kelly et al. (2021). They use the left y-axis. The purple dash-dot curve shows the"RETech" index, a measure of patent novelty from Bowen III, Frésard and Hoberg (2021), which assesses patent novelty by the prevalence of vocabularies that are growing in use in the patent description. It uses the right y-axis. Panel B shows the number of forward citations per patent, the number of backward citations per patent, and patent count (without smoothing) separately.
Sources: USPTO patent and citation data.

Panel A of Figure 1 plots the technological waves defined in this paper and compares them with two text-based measures from the literature. One measure is derived from Kelly et al. (2021)'s patent-similarity metric.[3] Specifically, we compute, for each year, the average ratio of patents' forward similarity to the sum of forward and backward similarity. The second measure is the "RETech" index developed by Bowen III, Frésard and Hoberg (2023).[4] The three series are strongly positively correlated, with largely synchronized peaks and troughs, supporting the robustness of the patterns across alternative measures. The figure indicates major technological breakthroughs in the mid-1980s, the mid-1990s, and the early 2010s, with the third peak notably smaller. In contrast, the period around 1990 and the mid-2000s appear to be phases of relative maturity. The citation-based measure in this paper complements the text-based measure in the literature and offers several advantages. First, it does not rely on the digitization quality of patent abstracts, thereby avoiding issues of inaccuracy. Second, it is unaffected by strategic language use in patent abstracts or changes in language over time. Third, its definition is more transparent and not constrained by computational resources.

A separate examination of forward citations per patent, backward citations per

[3] Kelly et al. (2021) use textual analysis to assess patent importance based on its similarity to patents filed five years before and after the focal patent's filing year.

[4] Bowen III, Frésard and Hoberg (2023) analyze the full text of U.S. patents and define a patent as revolutionary if the vocabulary it employs is rapidly growing in usage across the overall patent corpus ("RETech"). Kalyani (2024) use new phrases in patents to measure creativity.

patent, and the total number of granted patents (reported in Panel B of Figure 1) suggests that fluctuations in the Novelty Index are driven primarily by citation dynamics rather than by patent volume. In particular, forward citations per patent rise sharply and peak earlier, while backward citations per patent respond with a noticeable lag.

Field-level Novelty indices, constructed by the first digit of the International Patent Classification (IPC) defined by the World Intellectual Property Organization (WIPO), are reported in Figure 12 in Appendix B.2. These series exhibit both broad co-movement and substantial heterogeneity across fields.[5]

2.1.2 Contributors to the Tech Waves

Which classes of technology contributed to the three peaks of the technological waves? Who were the major applicants for breakthrough patents—incumbents, startups, or public institutions?

To answer the first question, we decompose the Novelty Index into the contribution of each three-digit IPC code using the following method,

$$\text{Novelty}_t = \sum_{j \in J} \frac{\sum_{i \in I_{jt}} \sum_{s=0}^{5} F_{ij,t+s}}{\sum_{i \in I_{jt}} \sum_{s=0}^{5} F_{ij,t+s} + \sum_{i \in I_{jt}} \sum_{s=0}^{5} B_{ij,t-s}} \frac{\sum_{i \in I_{jt}} \sum_{s=0}^{5} F_{ij,t+s} + \sum_{i \in I_{jt}} \sum_{s=0}^{5} B_{ij,t-s}}{\sum_{i \in I_{t}} \sum_{s=0}^{5} F_{i,t+s} + \sum_{i \in I_{t}} \sum_{s=0}^{5} B_{i,t-s}}, \tag{3}$$

where J is the set of 3-digit IPC code and I_{jt} is the set of patents belonging to the IPC code j granted in year t. Intuitively, the contribution of each technology class in a given year is determined by the IPC-specific Novelty Index multiplied by the share of forward and backward citations of that class. Table 1 lists the top three contributors at the three peaks of the technological novelty waves. Medical or Veterinary Science and Hygiene contribute most to the first peak, while Computing; Calculating or Counting is the leading contributor to the second and third peak.

Table 1: Major Contributors to the Technological Novelty Peaks

	First Peak (1985-1987)	Second Peak (1995-1997)	Third Peak (2010-2012)
1	Medical or Vet. Sci.; Hygiene	Computing; Calculating or Counting	Computing; Calculating or Counting
2	Electric Elements	Medical or Vet. Sci.; Hygiene	Medical or Vet. Sci.; Hygiene
3	Measuring; Testing	Electric Communication Technique	Electric Communication Technique

Notes: This table shows the major technological classes of the top three fields with the highest Novelty index at the technological novelty peaks in the period between 1981 and 2017.

[5] The field-level Novelty Index is defined analogously to the aggregate index, except that the patent set I_t includes only patents in the corresponding technology field. The nine fields are: human necessities; performing operations and transportation; chemistry and metallurgy; textiles and paper; fixed constructions; mechanical engineering; lighting; heating; weapons; blasting; physics; and electricity.

To examine the second question, we draw on the "historically significant patents" compiled by Kelly et al. (2021) from online sources. Our sample period includes 54 such breakthrough patents. To assess whether these patents meaningfully contribute to aggregate novelty, we compute a patent-level Novelty Index using the same methodology applied to the IPC-level index.[6] We then rank each of the 54 patents within the unconditional distribution of Novelty Index scores over the sample period. The mean and median percentile ranks fall in the top 24% and top 7% of all patents, respectively, indicating a strong concordance between Kelly et al. (2021)'s list and our novelty measure. We categorize the applicants of these breakthrough patents by institution type and find that 41% are incumbent firms, 26% are startups, 20% are universities, and the remainder come from other institutions.[7] The wide range of sources demonstrates that highly novel technologies emerge across diverse organizational settings instead of a dominant sector.

While this paper primarily focuses on analysis within the United States, we also calculate the Novelty Index for several European countries with the highest patenting activity. Figure 14 in Appendix B.4 illustrates the technological waves in six European countries based on PATSTAT data. The figure reveals a declining trend in technological novelty among all the six countries from the 1980s through the 2010s.

2.2 Market Concentration

The Herfindahl-Hirschman Index (HHI), a widely adopted measure of market concentration, serves as the primary metric. The analysis relies on two datasets: the Census Bureau's Longitudinal Business Database (LBD) and Compustat Fundamentals Annual. The LBD provides employment and payroll information for all U.S. employer businesses. Compustat complements these data with sales measures, though only for publicly listed U.S. firms. In Compustat, we focus on U.S.-headquartered industrial firms. The HHI is constructed through several steps. First, in the LBD, the squared ratios of each firm's employment or payroll to the total industry employment or payroll are computed within each industry, defined by the 6-digit NAICS code, for each year. In Compustat, the squared ratios of each firm's sales to total industry sales are calculated, defined by the 4-digit SIC code, for each year. These squared ratios are then summed across firms in each industry to derive the annual industry-level HHIs. Each industry is weighted by its total employment (for the LBD) or total sales (for Compustat), and a weighted average

[6] Namely,

$$\text{Novelty}_t = \sum_{i \in I_t} \frac{\sum_{s=0}^{5} F_{i,t+s}}{\sum_{i \in I_t} \sum_{s=0}^{5} F_{i,t+s}} \times \frac{\sum_{s=0}^{5} F_{i,t+s}}{\sum_{i \in I_t} \sum_{s=0}^{5} F_{i,t+s} + \sum_{i \in I_t} \sum_{s=0}^{5} B_{i,t-s}}.$$

[7] A full decomposition is shown in Figure 13 of Appendix B.3.

across industries is computed. To smooth the trend, a three-year average is applied to each observation.[8]

Panel A of Figure 2 displays the annual Herfindahl-Hirschman Index (HHI) for firm employment, payroll, and sales, all of which exhibit similar trends and fluctuations, except for the last ten years. The pairwise correlations among these measures are high: 0.96 between employment and payroll, 0.73 between sales and employment, and 0.66 between sales and payroll. Panel B illustrates the relation between market concentration, measured by the employment-based HHI, and the technological waves. The two series are negatively correlated. The technological waves exhibit a downward linear trend, while the HHI shows an upward trend. The cross-correlation between the detrended HHI (x_t) and the detrended technological waves (y_{t+k}) are -0.736, -0.775, and -0.720 respectively when k equals to 0, -1, -2. This suggests that changes in market concentration closely follow technological waves with a lag.[9]

To assess the robustness of market concentration patterns, the share of sales by top firms is calculated using the cleaned data series from Kwon, Ma and Zimmermann (2023), which is based on IRS data covering the entire population of U.S. corporations. Figure 15 in Appendix B.7 shows that the HHI exhibits similar upward trends and cyclical patterns to the top sales shares.

A more granular analysis is conducted using regression analysis, as specified in Equation (4):

$$\text{HHI}_{st} = \beta_0 \text{Novelty Index}_{st} + \beta_1 \text{Size}_{st} + \theta_s + \mu_t + \epsilon_{st}. \tag{4}$$

Here HHI_{st} is the Herfindahl–Hirschman Index for 6-digit NAICS industry s in year t based on the Census data. The right–hand side includes the industry's Novelty Index, industry fixed effects θ_s, and year fixed effects μ_t. To rule out the effect of technological novelty on HHI through changes in market size (Campbell and Hopenhayn (2005)), we control for industry size, measured by total employment or payroll in each 6-digit NAICS industry. The Novelty Index is constructed at the 4-digit IPC level and mapped to 6-digit NAICS using a concordance that we build by tabulating, in the Census data, the industry composition of patent owners within each IPC category.

Instruments Because the Novelty Index may be correlated with unobserved shocks to concentration and reverse causality may exist, we construct a shift-share instrument

[8] Figure 10 in Appendix B.1 shows the patterns of HHI (based on sales in Compustat) and the Novelty Index without smoothing. Their correlation is similar to the smoothed version.

[9] Cross-correlations for $k \in \{-3, -2, -1, 0, 1, 2, 3\}$ are reported in Table 11 in Appendix B.5. The strongest correlation, in absolute value, occurs at $k = -1$.

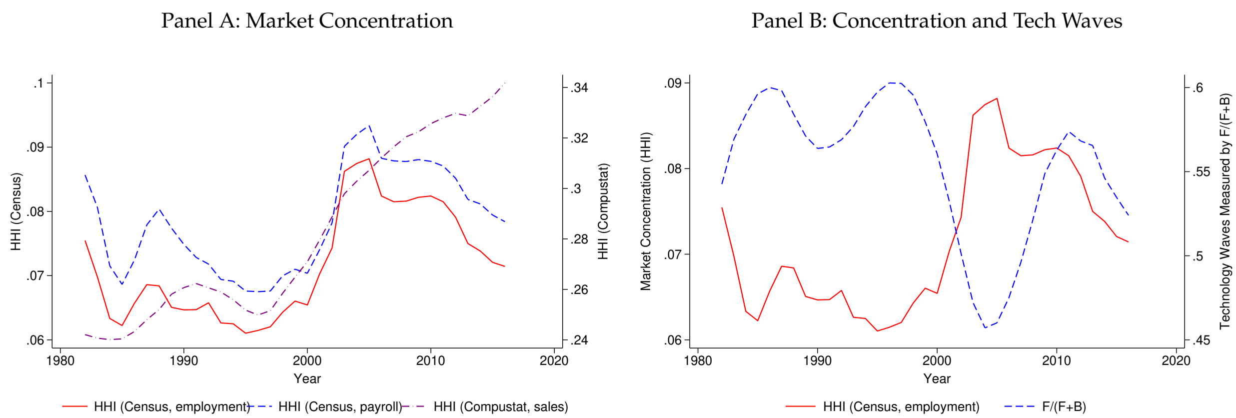

Figure 2: Technological Waves and Market Concentration

Notes: Panel A displays the annual HHI for employment (red solid curve), payroll (blue dash curve), and sales (purple dash-dot curve). The first two are derived from the LBD dataset from the Census, while the last one is based on Compustat. Panel B illustrates the technological waves alongside the trend of market concentration over time. The red solid curve shows the HHI for employment. The blue dashed curve, following the methodology defined in this paper, represents the relative ratio of forward citations to the sum of forward and backward citations. The scale of y-axes are shown on the left and right, respectively.
Sources: Longitudinal Business Database (LBD), Compustat Fundamental Annuals, and USPTO patent data.

that combines the U.S. innovation network with lagged technological waves. We build a time–varying diffusion matrix $\Omega_{c'\to c,t,\tau}$ at the IPC-class level from U.S. patent data, capturing how knowledge in an upstream (cited) class c' diffuses to a downstream (citing) class c in year t with lag $\tau \in \{1,\dots,5\}$. For each (c,t,τ) we expand patent citations to IPC$\times$IPC pairs with fractional weights, drop within–class links, aggregate to (c',c,t,τ), and normalize so that $\sum_{c'\neq c}\Omega_{c'\to c,t,\tau} = 1$. We then combine this diffusion structure with the Novelty Index $I_{c',t-\tau}$ to form a shift–share instrument in the spirit of Acemoglu, Akcigit and Kerr (2016) and Liu et al. (2025):

$$\lambda_{c,t} = \frac{1}{5}\sum_{\tau=1}^{5}\sum_{c'\neq c} \Omega_{c'\to c,t,\tau}\, I_{c',t-\tau}.$$

Intuitively, $\lambda_{c,t}$ loads each downstream class on lagged novelty in technologically connected upstream classes. The instrument is mapped from IPC to NAICS using the same concordance.

The regression results are reported in Table 2. Columns (1)–(4) examine the effect of contemporaneous technological novelty, columns (5)–(8) use a one-year lag for both the Novelty Index and its shift-share instrument, and columns (9)–(12) use a two-year lag. For each set of specifications, we report the OLS estimates first, followed by the 2SLS estimates. The corresponding first-stage results and F-statistic of the instrument are reported in Table 12 in Appendix B.6. Across all columns, higher technological novelty is associated with lower market concentration, and the effect is stronger for lagged novelty.

Table 2: Relationship between HHI and Novelty Index at the 6-digit NAICS code

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	OLS		2SLS		OLS		2SLS		OLS		2SLS	
Novelty	-0.142***	-0.0284	-0.241***	-0.306**								
	(0.0343)	(0.0367)	(0.0850)	(0.139)								
Novelty (1-year lag)					-0.137***	-0.0218	-0.238***	-0.367**				
					(0.0336)	(0.0363)	(0.0831)	(0.148)				
Novelty (2-year lag)									-0.130***	-0.0555	-0.330**	-0.407***
									(0.0327)	(0.0341)	(0.144)	(0.147)
Observations	37500	37500	37500	37500	37500	37500	37500	37500	36500	36500	36500	36500
R-squared	0.011	0.69			0.011	0.69			0.01	0.701		
Industry size	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES
Industry FE	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES

Notes: Standard errors are clustered at the 6-digit NAICS and year level. Columns (1)–(4) examine the effect of contemporaneous technological novelty, columns (5)–(8) use a one-year lag, and columns (9)–(12) use a two-year lag. All specifications control for industry size. The inclusion of year and industry fixed effects is indicated in the last two rows. To comply with Census Bureau disclosure requirements, the number of observations is rounded to the nearest hundred. *** Significant at the 1 percent level; ** significant at the 5 percent level; * significant at the 10 percent level.

The aggregate negative correlation between technological waves and market concentration is also evident in Europe, as shown by the declining trend of the Novelty Index in Appendix B.4 and the increasing market concentration across European countries, as measured by the Herfindahl-Hirschman Index (HHI) and top sales share in recent studies (e.g., Bighelli et al. (2023) and Ma, Zhang and Zimmermann (2024)).

2.3 Allocation of Ideas

One potential link between the technological waves and the market concentration is inventors' choices of where to do innovation. They can work independently and start their own businesses or contribute their innovation efforts to incumbent firms. In the latter case, they also choose the size of incumbent firms to work in. This section describes the flow of the new ideas using the Census and Compustat data.

2.3.1 Entrants or Incumbent Firms

Information on the affiliations of inventors at the outset of research projects is not directly available. However, affiliation can be inferred by observing the age of the firm to which a patent is granted. Specifically, when a patent is granted to a firm aged zero to five years, it indicates that the initial idea was developed in a startup five years earlier. In contrast, when a patent is granted to a firm older than five years, it implies that the idea was developed internally by an incumbent firm. A five-year window is adopted based on estimates from the USPTO, which suggest that the average time between patent application and issuance is around two to three years. It is also assumed that the completion of a research project takes a similar amount of time. Based on these assumptions, the ratio between the number of ideas in new firms to the number of ideas absorbed in incumbent firms is captured by the following ratio:

$$\text{New-to-Incumbent Ratio}_t = \frac{\sum_{i \in I_{t+5}} \text{Granted in Firm(Age} \leq 5)_{i,t+5}}{\sum_{i \in I_{t+5}} \text{Granted in Firm(Age>5)}_{i,t+5}}, \tag{5}$$

where I_{t+5} denotes the set of patents granted five years after time t, and the variables "Granted in Firm(Age$\leq$5)" and "Granted in Firm(Age$>$5)" indicate whether patent i was granted to a firm aged five years or less, or older than five years, respectively.

The data used to observe patent affiliations is constructed by combining the Longitudinal Business Database (LBD) from the Census and the USPTO patent data. The combined dataset can track the age of firms at patent issuance.

As in previous analyses, a three-year moving average is applied to smooth the time series, and the resulting trend is shown in Figure 3.[10] For comparison, the Novelty Index introduced earlier in this paper is also plotted. The New-to-Incumbent Ratio exhibits pronounced cyclicality, with its peaks and troughs closely aligned with those of the technological novelty waves. In addition to this cyclical pattern, a declining trend is observed in both series over time. The cross-correlation between the detrended New-to-Incumbent Ratio (x_t) and the detrended Novelty Index (y_{t+k}) is computed across different time lags and reported in Table 11 in Appendix B.5. The maximum absolute correlation, 0.612, occurs at $k = 0$, indicating that the two series move in tandem contemporaneously.

2.3.2 Granular Relationship between Tech Waves and the New-to-Incumbent Ratio

The positive correlation between technological waves and the New-to-Incumbent Ratio is evident across most major technological fields, including human necessities, performing operations, chemistry, textiles, fixed constructions, physics, and electricity. Graphs by field illustrating this relationship are presented in Figure 16 in Appendix B.8. Analysis at a more granular level is conducted using a regression, as specified in Equation (6):

$$\text{New firm}_{ict} = \beta_0 \text{Novelty Index}_{ct} + \beta_1 \text{Citation}_{ict} + \theta_c + \mu_t + \epsilon_{ict}. \tag{6}$$

The dependent variable is an indicator equal to one if a patent is granted to a new firm (age $\leq$ 5). We regress it on the Novelty Index measured at the 4-digit IPC (c) by year

[10]Figure 11 in Appendix B.1 presents the unsmoothed series for both the New-to-Incumbent Ratio and the Novelty Index, which display a similar correlation structure to their smoothed counterparts.

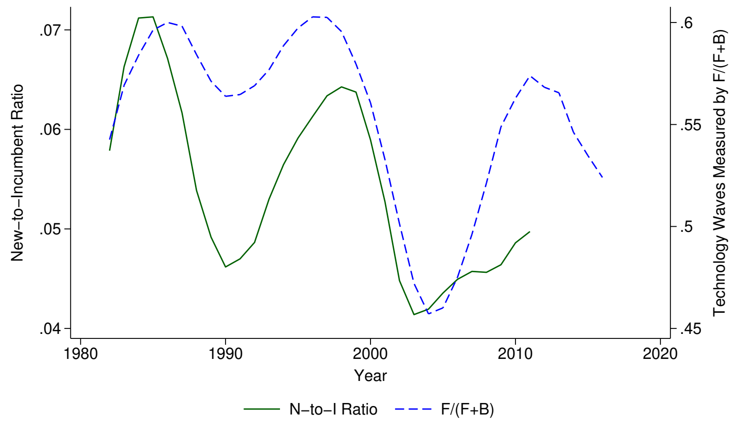

Figure 3: Technological Waves and Idea Allocation

Notes: This figure shows the technological waves and the idea allocation between new and incumbent firms over time. The blue dashed curve, based on the methodology defined in this paper, calculates the relative ratio of forward citations to the sum of forward and backward citations. The green solid curve displays the "New-to-Incumbent Ratio" defined in the paper, capture where new ideas contribute their value. The two curves have different y-axises, which are shown respective on the left and right.
Sources: Longitudinal Business Database (LBD) and USPTO patent and citation data.

(t) level, controlling for the patent's scientific value proxied by the number of forward citations it receives within five years of issuance, and including IPC and year fixed effects. To address the endogeneity issue, we apply the shift-share instrument that combines the U.S. innovation network with lagged technological waves.[11] The regression results are reported in Table 3. Columns (1)–(2) present OLS estimates, while columns (3)–(4) report the 2SLS results. The first-stage regression results are presented in Table 12 in Appendix B.6. We report specifications both with and without IPC and year fixed effects. Across all specifications, a significant positive relationship is found, indicating that idea holders are more likely to develop innovations in new firms when breakthrough technologies occur. This association remains robust after accounting for business-cycle fluctuations at the IPC and year levels.

2.3.3 Size of Incumbent Firms

When inventors opt to contribute their ideas to incumbent firms, they are also making a choice regarding the size of the firm, as it impacts the potential economic value that their innovations can attain. We establish a connection between the quality of inventors' ideas and the size of the incumbent firms they select by examining a subset of patents that have been granted to firms with a history of at least five years in operation. This subset

[11] Since in this regression, the Novelty Index is at the 4-digit IPC level, there is no need to map the Index to the 6-digit NAICS codes.

Table 3: Starting New Businesses and Novelty Index at the Patent Level

	(1)	(2)	(3)	(4)
	OLS		2SLS	
Novelty	0.0560***	0.014	0.127***	0.179***
	(0.00996)	(0.0141)	(0.0477)	(0.0514)
Ln(1+Citations)	0.0104***	0.0113***	0.00966***	0.0108***
	(0.000947)	(0.000926)	(0.000935)	(0.000952)
Observations	4032000	4032000	4032000	4032000
R-squared	0.003	0.013		
IPC FE	NO	YES	NO	YES
Year FE	NO	YES	NO	YES

Notes: Standard errors are clustered at the 4-digit IPC and year level. Columns (1) and (3) include no fixed effects. Columns (2) and (4) incorporate both IPC and year fixed effects. To comply with Census Bureau disclosure requirements, the number of observations is rounded to the nearest thousand. *** Significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

serves as the denominator when calculating the "New-to-Incumbent Ratio," as described in Section 2.3.1. Idea quality is proxied by the number of forward citations each patent receives within five years of issuance. Patents issued to both new and incumbent firms across various years are pooled, and citation quartiles are computed to categorize patents into four groups based on their relative citation counts. For patents granted to incumbent firms, the average firm size—measured by the number of employees in the Longitudinal Business Database (LBD)—is then calculated within each citation quartile. The average firm size for the lowest citation quartile is normalized to one, and the relative firm size across quartiles is plotted in Figure 4.

The figure reveals a pattern of positive assortative matching between idea quality and firm size: higher-quality ideas are more likely to be matched with larger firms. This pattern remains when we measure firm size using employment five years before the patent is granted. To check whether the positive relationship between idea quality and firm size exists for new firms, we calculate the average size of firms the patents in each quartile are granted to if they are granted to new firms—firms with less than five years of operation. The results show that average firm size remains relatively constant across quartiles, suggesting positive assortative matching only holds when ideas are contributed to incumbent firms.

2.4 Economic Value of Patents over Tech Waves

To explore the underlying channels driving the co-movement between technological waves and the allocation of ideas, as well as the positive assortative matching between idea quality and firm size, we perform patent-level regressions. We use an extended

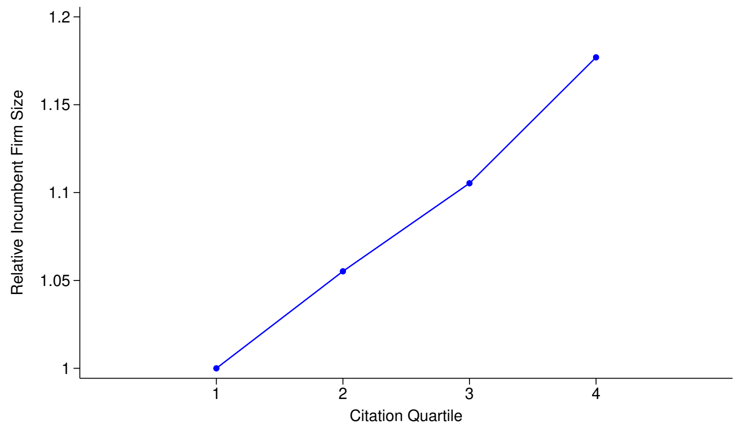

Figure 4: Mapping between Patent Citations and Incumbent Firm Size

Notes: This figure shows the mapping between inventors' idea quality and firm size if inventors opt to develop their ideas in incumbent firms. The idea quality is measured by the number of patent citations and is classified into four quartiles. The firm size is measured by the number of employees. The average employment of firms corresponding to the first citation quartile is normalized to be one.
Sources: Longitudinal Business Database (LBD) and USPTO patent and citation data.

version of the sample constructed by Kogan et al. (2017), which includes more recent years. Kogan et al. (2017) leverages the stock market's response to patent news to estimate the private economic value of patents. Since the sample encompasses all patents granted to publicly listed firms in the US, it provides valuable insights into factors affecting the economic value of patents in incumbent firms over technological waves. The following regression analysis is conducted,

$$\begin{aligned} \ln(\text{economic value}_{ijct}) =& b \ln(\text{Firm size}_{jt}) + \iota \ln(1 + \text{Citations}_{ijct}) + \\ & \phi \text{Novelty}_t \times \ln(1 + \text{Citations}_{ijct}) + \theta_{ct} + \gamma_{jt} + \epsilon_{ijct}. \end{aligned} \tag{7}$$

where i, j, c, and t are indexes for patents, firms, patent technology classes at the 4-digit IPC level, and years. The dependent variable corresponds to the economic value of the patents. Firm size is measured by either the employment or sales of the firm to which the patent belongs. The number of forward citations received within five years of issuance is used to measure the scientific value of patents, serving as a proxy for idea quality. The interaction term between the Novelty Index and the citations captures the impact of technological waves on the relationship between the scientific and economic value of patents. The model controls for IPC-by-year fixed effects in all specifications and firm-by-year fixed effects in some specifications.

Table 4 presents the results using firm employment as the measure of firm size. Columns (1)-(3) exclude the technological wave measure, focusing solely on the properties of patents and firms. Column (4) displays results of Equation (7). In Columns (5), the yearly Novelty Index is replaced by the IPC-by-year Novelty Index.

Firm size has a significantly positive effect on the economic value of patents, given the idea quality. This suggests that the synergy between inventors and firms increases with firm size. Additionally, idea quality positively impacts the economic value of patents, but this impact diminishes with higher aggregate technological novelty, as indicated by the negative coefficients of the interaction terms. This finding implies that, under high aggregate novelty, ideas realize lower value within incumbents, especially high-quality ideas.

Table 4: Factors of Patent's Economic Value for Incumbent Firms

	Ln(Patent's Economic Value)				
	(1)	(2)	(3)	(4)	(5)
Ln(1+Employment)	0.329***				
	(0.064)				
Ln(1+Sales)		0.335***			
		(0.031)			
Ln(1+Citations)	0.085***	0.092***	0.004***	0.032**	0.013**
	(0.017)	(0.017)	(0.001)	(0.012)	(0.006)
Ln(1+Citations)$\times$Novelty$_t$				-0.053**	
				(0.022)	
Ln(1+Citations)$\times$Novelty$_{ct}$					-0.018*
					(0.010)
Observations	1,107,289	1,113,710	1,103,218	1,103,218	1,103,218
R-squared	0.369	0.470	0.884	0.884	0.884
IPC $\times$ Year FE	Yes	Yes	Yes	Yes	Yes
Firm $\times$ Year FE	No	No	Yes	Yes	Yes

Notes: Standard errors are clustered at the 4-digit IPC-by-year level. Columns (1)-(3) exclude the technological wave measure and focus solely on the property of the patents and firms. Columns (4) shows coefficients of the regression equation (7). Columns (5) replaces the yearly Novelty Index by the IPC-by-year Novelty Index. The regressions control for IPC-by-year across all specifications. The firm-by-year fixed effects are controlled in columns (3)-(5). *** Significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

3 Model

To clarify the mechanism through which technological waves influence market concentration, we develop a general equilibrium model with two groups of individuals—households and inventors—and two types of firms: intermediate goods producers and final goods producers. In each period, inventors receive ideas of heterogeneous quality. They decide either to establish new intermediate-goods firms (Romer-type innovation) or to join incumbent firms of chosen size (quality-ladder innovation). The economy is subject to MIT aggregate shocks that change the average quality of new ideas. The shocks shift the relative payoff to inventors between the two innovation types. Inventors choose their innovation type based on both the aggregate shock and the quality of their individual ideas.

3.1 Preferences

There is a long-lived representative household in the economy. She works in the production sector, supplies one unit of labor to firms inelastically, and consumes final goods. She also owns all the firms in the economy. The household's utility function is

$$U_H = \int_0^\infty e^{-\rho t} log(C_H(t))dt, \tag{8}$$

where $\rho > 0$ is the discount rate and $C_H(t)$ is the consumption of the household.

Inventors are the ones who work in the R&D sector. In each period, there is a continuum of inventors of measure one. An inventor, with a short-lived lifespan of dt time periods, dedicates effort $e_I(t)$ to create innovations within either an incumbent firm or a new business. Simultaneously, they engage in consumption. Inventors are risk-averse and have a mean-variance utility:

$$U_I\left(c_I(t), e_I(t)\right) = \mathbb{E}\left(c_I(t)\right) - A\frac{var\left(c_I(t)\right)}{V(t|Q(t))} - R\left(e_I(t)\right) V(t|Q(t)), \tag{9}$$

where $c_I(t)$ is the consumption, $e_I(t)$ is the effort level, and $R\left(e_I\right) V(t|Q(t))$ is the associated cost. $Q(t)$ (defined below) is the aggregate quality in the economy at time t, and $V(t|Q(t))$ is the value of a firm whose size equals to the aggregate quality. The variance and cost are normalized by $V(t|Q(t))$ to keep the problem stable over time. Denote the inventors' aggregate consumption using $C_I(t)$, i.e., $C_I(t) = \int_0^1 c_{Ii}(t)di$.

3.2 Technology

The economy features two types of firms: intermediate goods producers and final goods producers. The setup is similar to Akcigit and Kerr (2018). Both types of firms are owned by the household. The former hires inventors to create innovations, and produce intermediate goods. The latter assembles intermediate goods and produces final goods.

The final good producers produce final goods using a continuum of intermediate goods $j \in [0, N_F(t)]$:

$$Y(t) = \frac{1}{1-\beta}\int_0^{N_F(t)} q_j^\beta(t) y_j^{1-\beta}(t)dj. \tag{10}$$

In this function, $q_j(t)$ is the quality of the intermediate good j, and $y_j(t)$ is its quantity. $N_F(t)$ is the total mass of intermediate goods producers. We normalize the price of the final good to be one in every period. The final good producers are perfectly competitive, taking the input prices as given.

The intermediate goods producers are a continuum of risk-neutral firms. Each firm produces one type of good, with a linear technology using only labor:

$$y_j(t) = Q(t) l_j(t), \tag{11}$$

where $l_j(t)$ is the labor input; $Q(t) = \int_0^{N_F(t)} q_j(t) dj$ is the aggregate quality level of the economy. It implies that improvement in both q_j (intensive margin) and N_F (extensive margin) has positive externalities (Romer, 1986; Aghion and Howitt, 1990). The cost is linear in wage $w(t)$, which intermediate firms take as given. The labor market satisfies the constraint:

$$\int_0^{N_F(t)} l_j(t) dj \leq 1. \tag{12}$$

3.3 Innovation

Innovation in this economy can either create new intermediate-good producers (*Romer-type*) or improve the quality of incumbent intermediate-good producers (*Quality-ladder*). *Romer-type* innovation is led by an inventor who pairs with an outside, risk-neutral partner. *Quality-ladder* innovation can be led either by an inventor or by an incumbent. In the former scenario, the inventor partners with an incumbent firm; in the latter scenario, the firm purchases an idea from an inventor and undertakes the innovation internally. If the inventor leads innovation, the inventor's effort is unobservable to the partner, generating a moral-hazard problem; if the incumbent leads, no such incentive problem arises. In both cases, firms engage in a Betrand competition when hiring inventors. The two types of innovation takes place sequentially. At each instant, the inventor-led R&D happens first, the rest of innovation ideas are then led by firms. Table 5 summarizes these features. The following sections describe each innovation method in detail.

Table 5: Features of Different Innovation Methods

Innovation Type	**Leader**	**Partner**	**Incentive problems**
Romer	Inventor	Outside partner	Yes
Quality-ladder	Inventor	Incumbent firm	Yes
	Incumbent firm	None	No

3.3.1 Inventor-led R&D

If inventors lead innovation, they endogenously exert effort e_I. Given the level of effort, the success rate of an innovation follows an instantaneous Poisson flow rate:

$$\lambda(e_I) = \lambda_0 e_I. \tag{13}$$

Exerting effort is costly for inventors. The instantaneous cost of effort e_I is given by $R(e_I)V(t|Q(t))$, where $R(e_I) = \frac{1}{1+\delta}e_I^{\delta+1}$, an increasing and convex function of effort.[12]

Each inventor is endowed with a single innovative idea of random quality z_0, drawn from a random distribution, $\Psi(z_0)$. The inventor may either establish a new intermediate-goods firm with an outside partner or join an incumbent firm. In the case of a startup, the inventor retains full control over the innovation process, and the realized quality of the innovation, $zQ(t)$, is drawn from a uniform distribution,

$$\text{Uniform}\left((1-\phi)z_0Q(t),\ (1+\phi)z_0Q(t)\right). \tag{14}$$

On average, higher-quality ideas lead to higher-quality innovations, but the parameter $\phi \in (0,1)$ introduces randomness into this mapping by capturing variability in how idea quality translates into realized innovation quality.

If the inventor with idea quality z_0 chooses to lead innovation within an incumbent firm of quality q, the resulting quality improvement, $x(z_0, \tilde{q}(t))Q(t)$, is a stochastic variable drawn from a uniform distribution:

$$\text{Uniform}\left((1-\phi)x_0(z_0,\tilde{q}(t))Q(t),\ (1+\phi)x_0(z_0,\tilde{q}(t))Q(t)\right), \tag{15}$$

where $\tilde{q}(t)$ denotes the incumbent's relative quality, defined as $\tilde{q}(t) = \frac{q}{Q(t)}$. For convenience in the subsequent analysis, we adopt the normalization $\tilde{m} = \frac{m}{Q(t)}$ for all variable m. The effective idea quality, $x_0(z_0, \tilde{q}(t))$, depends positively on both the inventor's original idea quality and the relative firm size, i.e.,

$$\frac{\partial x_0(z_0,\tilde{q}(t))}{\partial z_0} > 0 \quad \text{and} \quad \frac{\partial x_0(z_0,\tilde{q}(t))}{\partial \tilde{q}(t)} > 0.$$

[12] The innovation production function and the cost functions are based on the growth theory literature (Romer, 1990; Klette and Kortum, 2004; Akcigit and Kerr, 2018). In the calibration, we choose $\delta = 1$ following the literature. The R&D cost paid by inventors is modeled as a dis-utility. Alternatively, it can be modeled as a cost that can be converted one-to-one from final goods. Switching from one modeling choice to another does not have material impact on our results.

The function $x_0(z_0, \tilde{q}(t))$ takes the form:

$$x_0(z_0, \tilde{q}(t)) = \chi(\tilde{q}(t))\gamma(z_0) = q_0 \left(\frac{\tilde{q}(t)}{q_0}\right)^b \left(B^\eta + z_0^\eta\right)^{\frac{1}{\eta}}, \ \eta < 0.$$

The first term, $\left(\frac{\tilde{q}(t)}{q_0}\right)^b$, captures the synergistic benefits provided by the incumbent firm, which increase with the firm's relative quality. The second term, $\left(B^\eta + z_0^\eta\right)^{\frac{1}{\eta}}$ with $\eta < 0$, is a constant-elasticity-of-substitution (CES) function that models the interaction between the inventor's idea and the existing technological base of the incumbent firm in the co-invention process. Here, B denotes the stock of backward citations and serves as a proxy for the maturity of the underlying technology.

The CES structure with a negative elasticity implies complementarity between the new idea z_0 and the technological stock B, highlighting that the new idea cannot fully substitute for existing product lines within incumbent firms. Consequently, the realized innovation quality is generally lower than if the idea were implemented in a startup, reflecting frictions in adopting and integrating new technologies.

Inventors bear the direct cost of their effort, yet this effort is unobservable to outside partners or incumbent firms. Without performance-based incentives, an inventor receiving only a flat wage would choose $e_I = 0$. To induce effort, outside partners and incumbents must therefore design compensation schemes contingent on innovation outcomes.[13] In this paper, we assume that all partners employ a common contract combining a fixed wage with an equity component. The wage enables risk sharing, while the equity stake aligns the inventor's incentives with those of the partner. The contracting problem in the innovation process is specified separately for the cases where the inventor joins an incumbent firm and where the inventor launches a new business.

Joining Incumbent Firms: Incumbent intermediate firms compete à la Bertrand to attract inventors by offering compensation packages consisting of wage rate $T_I V(t|Q(t))$ and equity a. The firm quality evolves stochastically:

$$dq = -q dJ_{\text{exit}} + x(z_0, \tilde{q}(t))Q(t)dJ_{R\&D}(e_I) \tag{16}$$

where J_{exit} and $dJ_{R\&D}(e_I)$ are two jump processes with Poisson arrival rates at τ and $\lambda(e_I) = \lambda_0 e_I$, respectively. The first jump process represents the exogenous exit, while the second one is related to the in-house innovation. The magnitude of the jump is

[13] It is worth noting that the level of effort e_I is unobservable and unverifiable. Consequently, contracts cannot be contingent on the effort level.

$x(z_0, \tilde{q}(t))Q(t)$, the stochastic value of an innovation produced by inventor z_0 in a firm with a normalized size $\tilde{q}(t)$ when the aggregate technology level is $Q(t)$.

The inventor's compensation $c_I(a, T_I, q, z_0, Q(t))$ is:

$$c_I(a, T_I, q, z_0, Q(t))dt = T_I V(t|Q(t))dt + a dV(j, t|q, z_0), \tag{17}$$

where $dV(j, t|q, z_0)$ is the firm value change from t to $t + dt$ when hiring an inventor of quality z_0. Following a principal-agent framework, the compensation package includes, in addition to a fixed wage, an opportunity for the inventor to purchase a fraction a of equity at the fair price. The inventor then sells this equity back to the firm at the end of each period of length dt to finance consumption. In doing so, the inventor bears part of the risk from exit shocks and innovation shocks. A firm j with quality q solves the following problem when an inventor with idea quality z_0 leads the innovation process, where $\Omega_I(z_0, q, t)$ denotes the expected payoff:

$$\begin{aligned}
\Omega_I(z_0, q, t) = \max_{a, T_I} &(1-a)\,\lambda_0 e_I \left[\int V(j, t|q + x(z_0, \tilde{q}(t))Q(t))dx - V(j, t|q)\right] - T_I V(t|Q(t)) \\
\text{s.t. } & e_I = \arg\max \{U_I\,(c_I(a, T_I, q, Q(t)), e_I)\} \\
& U_I\,(c_I(a, T_I, q, Q(t)), e_I) \geq \bar{u}\,(z_0, t) \\
& (1-a)\,\lambda_0 e_I \left[\int V(j, t|q + x(z_0, \tilde{q}(t))Q(t))dx - V(j, t|q)\right] - T_I V(t|Q(t)) \geq 0
\end{aligned} \tag{18}$$

The instant change rate in a firm's expected payoff from hiring the inventor equals the increase in firm value from the innovation accruing to the original shareholders (all shareholders other than the inventor), net of the wage and equity granted to the inventor. Here, $V(j, t|q)$ denotes the firm's value without R&D.

The first constraint is the inventor's incentive compatibility condition, which ensures that her actions align with utility maximization when employed by the firm. Specifically, given the firm-specific contract a, T_I, the inventor selects an effort level e_I to maximize her expected utility, $U_I\,(c_I(a, T_I, q, Q(t)), e_I)$. The second constraint is the participation constraint of the inventor, requiring that she prefers this firm's offer over her outside options. It implies that the firm must provide a utility level at least as high as $\bar{u}(z_0, t)$. The outside option is endogenously determined in equilibrium through Bertrand competition among firms for inventors. Finally, the third constraint is the firm's own participation condition, which ensures that hiring an inventor does not leave the firm worse off.

Inventor's utility function can be expressed as:

$$U_I\left(c_I(a,T_I,q,Q(t)),e_I\right)=\mathbb{E}\left(c_I(a,T_I,q,Q(t))\right)-A\frac{\mathrm{Var}\left(c_I(a,T_I,q,Q(t))\right)}{V(t|Q(t))}-R\left(e_I\right)V(t|Q(t)). \tag{19}$$

Starting up a New Business: If an inventor starts a new business with an outside partner, the partner solves an optimization problem analogous to that of an incumbent firm (Equation 18). The key differences are that there is neither synergy nor adoption friction, so $q=0$ and $x_0=z_0$, and the partner's outside option is zero. The inventor then chooses her effort level to maximize utility.[14]

3.3.2 Firm-led R&D

If an incumbent firm leads the innovation, it purchases an idea from an inventor and selects its own effort level e_F. The innovation's success is then governed by an instantaneous Poisson arrival rate:

$$\lambda(e_F)=\lambda_0 e_F. \tag{20}$$

Exerting effort is also costly for firms. The instantaneous cost of effort e_F is given by $R(e_F)N_F(t)V(t|q(t))$, where $R(e_F)=\frac{1}{1+\delta}e_F^{\delta+1}$.[15] Both the cost function $R(\cdot)$ and the arrival rate $\lambda(\cdot)$ take the same functional forms as in the inventor-led innovation case.[16] The total R&D cost borne by firms is:

$$R_F(t)=\int_0^{N_F(t)}R(e_F(t,q(j))N_F(t)V(t|q(t))dj. \tag{21}$$

Since firms bear both the costs and the benefits of innovation, they are fully incentivized to choose the effort level that maximizes their R&D profits. If the purchased idea is of quality z_0, a successful innovation raises the firm's quality from $q(t)$ to $q(t)+x_F N_F(t)q(t)$, where $x_F N_F(t)q(t)$ is drawn from a uniform distribution similar to (14):

$$\text{Uniform}\left((1-\phi)\kappa x_{F0}N_F(t)q(t),\ (1+\phi)x_{F0}N_F(t)q(t)\right). \tag{22}$$

where $x_{F0}=\gamma(z_0)=\left(B^\eta+z_0^\eta\right)^{\frac{1}{\eta}}$. $\kappa>0$ captures the scale effects of a firm-led innovation. As compensation, the firm makes a fixed transfer $T_F N_F(t)V(t|q(t))$ to the inventor. The

[14] A detailed exposition of the startup problem is provided in Appendix Section C.1.

[15] The introduction of $N_F(t)V(t|q(t))$ makes the cost and benefit of firm-led R&D comparable with innovator-led R&D.

[16] This assumption can be readily extended to allow for alternative cost functions.

firm-led innovation can be characterized by the following instantaneous maximization problem:

$$\Omega_F(z_0,q,t) = \max_{e_F,T_F} \lambda_0 e_F \left[\int V(j,t|q+\kappa x_{F0}N_F(t)q)dz - V(j,t|q)\right] - \frac{1}{1+\delta}e_F^{\delta+1}N_F V(t|q(t)) - T_F N_F V(t|q(t)) \tag{23}$$

Unlike the inventor-led innovation case, we interpret the firm-led case as arising from frictional matching: when inventors fail to match with their optimal firms, innovation is instead carried out in a random-matched firm. As in the inventor-led case, we assume that the transfer T_F equals the incumbent firm's expected value from the innovation.[17]

$$\lambda_0 e_F[\int V(j,t|q+\kappa x_{F0}N_F(t)q)dz - V(j,t|q)] - \frac{1}{1+\delta}e_F^{\delta+1}N_F V(t|q(t)) = T_F N_F V(t|q(t)) \tag{24}$$

3.4 Timeline

At birth, each inventor observes the quality z_0 of her idea. An outside partner, also observing z_0, offers contracts to jointly establish a new intermediate firm. At the same time, incumbent firms observe their corresponding $x_0(z_0,\tilde{q})$ and extend employment contracts to the inventor. After evaluating all offers, the inventor chooses either to join her preferred incumbent firm of relative quality $\tilde{q}^*(z_0,t)$ or to launch a startup with an outside partner. In both cases, the matching process is subject to frictions. If the inventor chooses to lead innovation within an incumbent firm, she joins the optimally sized firm with probability h; otherwise, she sells her idea to an incumbent. Likewise, if she prefers to start a new business, the venture succeeds with probability h_s, while with probability $1-h_s$ the idea is instead sold to an incumbent firm.

The inventor's decision rule is:

$$u(z_0,t) = \max\{hU_I(c_I(z_0,q^*(z_0,t),t), e_I(z_0,q^*(z_0,t),t) + (1-h)\int_q U_I(T_F(z_0,q),0)f_t(\tilde{q})d\tilde{q}, h_sU_I(c_I(z_0,0,t),e_I(z_0,0,t)) + (1-h_s)\int_q U_I(T_F(z_0,q),0)f_t(\tilde{q})d\tilde{q}\}, \tag{25}$$

[17] An alternative assumption is that the incumbent captures all surplus. The resulting expressions are of a similar functional form, since firm value remains linear in size.

where $f_t(\tilde{q})$ is the firm quality distribution endogenously determined in the equilibrium.

3.5 Entry and Exit

Incumbent intermediate firms face an exogenous exit rate τ, which is independent of their size and is a risk unrelated to innovation. After a firm exits the market, the technology it owns is distributed to new entrants randomly. Entry of new intermediate firms occurs through successful innovations by inventors launching startups with outside partners. The entry rate is denoted by $\lambda_I(t)$:

$$\lambda_I(t) = \int_{z_0 \in \{q^*(t)=0\}} h_s \lambda_0 e_I(z_0, 0, t)\, d\Psi(z_0). \tag{26}$$

Upon entry, an entrant draws a relative quality level $\tilde{q}_I$ from the distribution $f_{I,t}(\tilde{q}_I)$, where it satisfies

$$f_{I,t}(\tilde{q}_I) = \frac{\lambda_I(t)}{N_F(t)} f_t(\tilde{q}), \tilde{q}_I = \frac{\tau N_F(t)}{\lambda_I(t)} \tilde{q}.$$

Therefore,

$$\lambda_I \int \tilde{q}_I f_{I,t}(\tilde{q}_I) d\tilde{q}_I = N_F \int \tau \tilde{q} f_t(\tilde{q}) d\tilde{q} = \tau.$$ [18]

The new firm incurs a cost equal to the value of the technology associated with its drawn quality $q_I = \tilde{q}_I Q(t)$. It then augments a quality update by applying the innovation of the inventor.

The law of motion for the mass of intermediate-goods firms is

$$\dot{N}_F(t) = \lambda_I(t) - \tau N_F(t) \tag{27}$$

3.6 Household's and Firms' Choices

The household maximization problem reveals the relationship between the growth rate, $g(t)$, and the time discount factor,

$$\frac{\dot{C}_H(t)}{C_H(t)} = r(t) - \rho. \tag{28}$$

The final good producer chooses $\{y_j(t)\}_j$ to maximize its profit using the technology described in Section 3.2, which yields the demand function faced by intermediate

[18] In absolute term, upon entry, an entrant draws a quality level q_I from the distribution $f_{I,t}(q_I)$, where it satisfies

$$\lambda_I \int q_I f_{I,t}(q_I) dq_I = N_F \int \tau q f_t(q) dq = \tau Q(t).$$

goods producers: $p_j(t) = q_j(t)^{\beta} y_j(t)^{-\beta}$. The intermediate good producers engage in monopolistic competition.[19] Their FOC yields,

$$y_j(t) = q_j(t) \left(\frac{Q(t)\,(1-\beta)}{w(t)} \right)^{\frac{1}{\beta}}, l_j(t) = \frac{y_j(t)}{Q(t)}, p_j(t) = \frac{w(t)}{Q(t)\,(1-\beta)}. \tag{29}$$

In each period, the labor market clearing condition pins down the wage:

$$w(t) = (1-\beta)\,Q(t). \tag{30}$$

Thus, both the production output $y_j(t)$ and profit $\pi_j(t)$ are linear in quality,

$$y_j(t) = q_j(t), \pi_j(t) = \beta q_j(t), Y(t) = Q(t). \tag{31}$$

The HJB equation of an incumbent firm's value is

$$\begin{aligned} r(t)V(j,t|q) - \dot{V}(j,t|q) &= \pi_j(t) + \Omega_I(z_0,q,t) + \Omega_F(z_0,q,t) - \tau V(j,t|q) \\ &= \pi_j(t) - \tau V(j,t|q) \end{aligned} \tag{32}$$

The second line follows from the scarcity of inventors and the assumption that competition for their ideas ensures that inventors capture the full value of innovations.

Solving the firm-led innovation problem derives the following condition:

$$\lambda_0 [\int V(j,t|q + \kappa x_{F0} N_F(t) q) dz - V(j,t|q)] = e_F^{\delta} N_F(t) q.$$

Since the condition holds for any $N_F(t)q$, it implies that the value of an intermediate firm is linear in its quality, i.e.,

$$V(j,t|q) = \nu(t) q. \tag{33}$$

Therefore, incumbent firm's effort, $e_F = (\lambda_0 \kappa x_{F0})^{\frac{1}{\delta}}$, is invariant to time and firm size[20]. The arrival rate, $\lambda(e_F) = \lambda_0^{\frac{1+\delta}{\delta}} (\kappa x_{F0})^{\frac{1}{\delta}}$, is constant across firms. The transfer to the inventor, $T_F(t) N_F V(t|q(t)) = \frac{\delta}{1+\delta} (\lambda_0 \kappa x_{F0})^{\frac{1+\delta}{\delta}} \nu(t) N_F q(t)$, is linear in $\nu(t)$. We assume that their ideas are allocated to firms randomly based on the incumbent firm size distribution, $f(\tilde{q}(t))$.

In the inventor-led innovation problem, the inventor's effort e_I is independent of time when (33) holds, the consumption $c_I(a, T_I, q, z_0, Q(t))$ and utility $U_I(c_I, e_I)$ are both linear in $\nu(t)$, The inventor's choice of optimal firm only depends on parameters, which

[19]The profit maximization problem and the solution process of the final good and intermediate good producers are shown in Section C.2 in the Appendix.

[20]See Appendix C.3 for details.

is invariant to time.[21]

3.7 Aggregate Variables

The final goods are consumed by the household and inventors, or invested into R&D by firms. The resource constraint of the economy is:

$$Y(t) = C_H(t) + C_I(t) + R_F(t), \tag{34}$$

Inventors' total consumption, $C_I(t)$, can be expressed as the following equation by applying the law of large numbers :

$$\begin{aligned} C_I(t) = \nu(t)Q(t) \int_{z_0 \in \{z_0 | \tilde{q}^* > 0\}} \left(h\lambda_0 e_I\left(z_0, \tilde{q}^*\right) x_0\left(z_0, \tilde{q}^*\right) + (1-h)T_F\right) d\Psi(z_0) \\ + \nu(t)Q(t) \int_{z_0 \in \{z_0 | \tilde{q}^* = 0\}} \left(h_s\lambda_0 e_I\left(z_0, \tilde{q}^* = 0\right) z_0 + (1-h_s)T_F\right) d\Psi(z_0). \end{aligned} \tag{35}$$

Firms' R&D costs, $R_F(t)$, can be expressed by

$$R_F(t) = \frac{1}{1+\delta} e_F(t)^{\delta+1} \nu(t) Q(t) \left[(1-h) \int_{z_0 \in \{z_0 | \tilde{q}^* > 0\}} d\Psi(z_0) + (1-h_s) \int_{z_0 \in \{z_0 | \tilde{q}^* = 0\}} d\Psi(z_0)\right] \tag{36}$$

The growth is from a single source—innovation. Although the innovation value is uncertain, we can apply the law of large numbers, since there is a continuum of firms. The aggregate growth can be written as,

$$\begin{aligned} g(t) = \frac{\dot{Q}(t)}{Q(t)} = \int_{z_0 \in \{z_0 | \tilde{q}^* > 0\}} \left(h\lambda_0 e_I\left(z_0, \tilde{q}^*\right) x_0\left(z_0, \tilde{q}^*\right) + (1-h)\,\lambda_0 e_F \kappa x_{F0}\right) d\Psi\left(z_0\right) \\ + \int_{z_0 \in \{z_0 | \tilde{q}^* = 0\}} \left(h_s\lambda_0 e_I\left(z_0, \tilde{q}^* = 0\right) z_0 + (1-h_s)\,\lambda_0 e_F \kappa x_{F0}\right) d\Psi\left(z_0\right) \end{aligned} \tag{37}$$

The aggregate growth rate only depends on the allocation of inventor-led innovations across firms, which is invariant to time, implying that:

$$g(t) = g.$$

[21] See Appendix C.4 for the formal proof.

Note that the growth rate of the aggregate quality can be decomposed into two parts:

$$\frac{\dot{Q}(t)}{Q(t)} = \frac{\dot{N}_F(t)}{N_F(t)} + \frac{Q(t)\dot{/}N_F(t)}{Q(t)/N_F(t)} \tag{38}$$

The first term captures growth from the extensive margin (firms' mass), and the second term captures growth from the intensive margin (average quality).

3.8 Equilibrium

This section characterizes an equilibrium of the economy in which aggregate variables (Y, C, R, w, Q) grow at a rate g.

Definition An equilibrium of this economy is the mapping between $\tilde{q}^*(t)$ and z_0, the allocation $\left(\left\{y_j^*(t)\right\}_j, Y^*(t), C_I^*(t), C_H^*(t)\right)$, the prices $\left(w^*(t), r^*(t), \left\{p_j^*(t)\right\}_j\right)$, the growth rate g^*, the entry rate $\lambda_I^*(t)$, and the measure of firms $N_F^*(t)$, such that (1) for any $j \in [0,1]$, $y_j^*(t)$ and $p_j^*(t)$ satisfy Equation (29); (2) the wage $w^*(t)$ satisfies Equation (30); (3) the interest rate $r^*(t)$ satisfies Equation (28); (4) the measure of the intermediate producers $N_F^*(t)$ satisfies Equation (27); (5) the mapping between $\tilde{q}^*(t)$ and z_0 comes from inventors' optimal choices; (6) the entry rates $\lambda_I^*(t)$ satisfy Equation (26); (7) R&D spending $C_I^*(t)$ and $R_F^*(t)$ satisfy Equation (35) and Equation (36), respectively; (8) the aggregate output Y^* satisfies Equation (31); (9) the aggregate consumption C_H^* satisfies Equation (34); and (10) the growth rate g satisfies Equation (37).

Proposition 1. *Consider the above-described economy starting with an initial condition $Q(0) > 0$. Then there exists a unique equilibrium. In this equilibrium, growth is always balanced, i.e., $g(t) = g^*, \forall t$. The aggregate quality, $Q(t)$, aggregate output, $Y(t)$, and aggregate consumptions, $C_H(t)$ and $C_I(t)$, all grow at the rate g^*.*

Lemma 1. *There is no transition dynamics.*

The proofs of Proposition 1 and Lemma 1 are provided in Appendix C.5. In the remainder of the paper, we work with variables normalized by aggregate quality $Q(t)$,

$$\tilde{q}(t) = \frac{q(t)}{Q(t)}.$$

We omit the time subscript t whenever doing so does not cause confusion.

3.9 Inventor's Choice in a Close Form

To illustrate the intuition and main channels in the model, this section presents a simplified setup that characterizes the inventor's trade-off in closed form. Two adjustments are introduced in this setup: (1) the innovation value x is drawn from a distribution with mean $x_0(z_0,\tilde{q})\,\nu$ and second moment $e_I^{-1}x_0(z_0,\tilde{q})^2\,\nu^2$, instead of from a uniform distribution $U((1-\phi)\,x_0(z_0,\tilde{q})\,\nu,(1+\phi)\,x_0(z_0,\tilde{q})\,\nu)$; and (2) there are no matching frictions, i.e., $h=h_s=1$. Under the first assumption, the variance of the inventor's consumption becomes

$$\operatorname{Var}\left(c_I\left(a,T_I,\tilde{q},z_0\right)\right)=\nu^2 Q(t)^2\left(\underbrace{\tau\tilde{q}^2}_{\operatorname{Var}\tilde{V}(\tilde{q})}+\underbrace{\lambda_0\left(x_0\left(z_0,\tilde{q}\right)^2\right)}_{\operatorname{Var}(\text{innovation})}\right).$$

Innovation-related uncertainty no longer depends on the effort level.

The parameter δ in the effort cost function, $R(e_I)$, is assumed to take the value of 1, consistent with empirical estimates in the literature (e.g., Akcigit and Kerr (2018)). Using backward induction, firms anticipate that the inventor will chooses the following effort:

$$e_I=\lambda_0 a x_0\left(z_0,\tilde{q}\right).$$

When an inventor holds a larger equity share, a, or when the potential value of her innovation, $x_0(z_0,\tilde{q})$, is higher, she is incentivized to exert greater effort. In both cases, the marginal return to effort increases, making additional investment more rewarding.[22]

The firm's problem in Equation 18 yields,

$$a^*=\frac{1}{1+2\frac{A}{\lambda_0}\left(\frac{\tau\tilde{q}^2}{\lambda_0 x_0(z_0,\tilde{q})^2}+1\right)} \tag{39}$$

The optimal equity level, a^*, decreases in the firm size $\tilde{q}$ when $b<1$. This is because a^* is determined jointly by two forces: the commercialization value $x_0(z_0,\tilde{q})$, and the non-innovation-related shock—the exit shock, $\tau\tilde{q}^2$. The firm size $\tilde{q}$ affects both factors but in opposite directions. Larger firms can provide the inventor with more synergy, leading to a greater commercialization value. This raises the equity share of the inventor, since it is more worthwhile to incentivize her effort. Meanwhile, larger firms face a higher exit risk. A lower equity share will expose inventors less to risks unrelated to innovation. The relationship between the equity a and the firm size $\tilde{q}$ depends on the relative strength of the two channels. When $b<1$, as found in the model calibration, the second channel

[22]The derivation of the contract solution is provided in Section C.6 of the Appendix.

dominates. Therefore, larger firms optimally offer less equity to an inventor.

The optimal compensation scheme is (a^*, T_I^*), where the wage T_I^* is determined by the zero profit, due to Betrand competition.

$$T_I^* = (1 - a^*)\, \lambda_0 e_I x_0 \left(z_0, \tilde{q}\right).$$

Upon reviewing all contracts, an inventor with idea quality z_0 chooses the firm $\tilde{q}$ she would like to work for. The first-order-condition yields

$$\frac{\partial x_0 \left(z_0, \tilde{q}\right)}{\partial \tilde{q}} = \frac{2A\tau}{4A\tau \frac{\tilde{q}}{x_0(z_0,\tilde{q})} + \frac{2A\lambda_0 + \lambda_0^2}{\tilde{q}/x_0(z_0,\tilde{q})}} \tag{40}$$

The left-hand-side element is the benefit of joining a larger firm—higher synergy and hence better commercialization. The right-hand-side element is the cost—the inventor gets a lower equity share when combining her idea with a firm with a higher risk unrelated to innovation. The optimal firm size is

$$\tilde{q}^* = \left(\frac{\left(2A\lambda_0 + \lambda_0^2\right) \left(\gamma\left(z_0\right)\right)^2 b}{2A\tau q_0^{2b} \left(1 - 2b\right)} \right)^{\frac{1}{2-2b}}.$$

Proposition 2. *When $b < 0.5$, $\frac{\partial \tilde{q}^*}{\partial z_0} > 0$.*[23]

When b is low, the synergy does not grow unboundedly with firm size and the balancing role of a lower equity share ensures the existence and uniqueness of a solution. The model predicts that among incumbent firms, better-quality innovations are more likely to be created in larger ones. The reasons are twofold. First, better ideas benefit more from synergy. Second, they generate relatively greater innovation-related uncertainty, making inventors less vulnerable to incentive problems.

Proposition 3. *When z_0 is drawn from an alternative distribution, $\Psi_{new}(z_0)$, which first-order dominates the original distribution $\Psi(z_0)$. i.e., idea becomes better, a larger share of inventors opt for joining startups firms.*[24]

The quality of new idea increases both the technological output in incumbent firms and in startups. However, since the stock the technology is immune to the change in idea qualtiy, the effect is weaker in incumbent firms. As a result, when there is a new tech wave, the economic value of innovations increases less in incumbent firms than in new firms, making incumbents relatively less attractive.

[23]See Appendix C.7 for the proof.

[24]See Appendix C.8 for the proof.

Proposition 4. *When $b < \frac{\min\left(z_0^{-\eta}\right)}{\min\left(z_0^{-\eta}\right)+\max(B^{-\eta})}$, there exists a cutoff $\bar{z}_0\left(B\right)$, such that all inventors with $z_0 < \bar{z}_0\left(B\right)$ opt for incumbent firms.*[25]

The effects of both incumbents' adoption efficiency, $\gamma(z_0)$, and the synergy-adjusted innovation value, $x(z_0, \tilde{q})$, are driven by the underlying idea quality, z_0. High-quality ideas experience greater losses due to adoption frictions, but also benefit more from synergies when implemented by incumbent firms. However, if the synergy gain does not increase too sharply with firm size ($b < \frac{\min\left(z_0^{-\eta}\right)}{\min\left(z_0^{-\eta}\right)+\max(B^{-\eta})}$),then the adoption friction dominates, and inventors with high-quality ideas are more likely to start new businesses.

4 Calibration

We calibrate the full model to target the average US economy from 1982 to 2016. Patents are used as a surrogate for innovations. An innovation's idea quality, denoted by z_0, and the realized value, x (z in the context of a startup), correspond to the patent's citation (scientific importance) and the patent's economic value, respectively. Additionally, we assume that the idea quality z_0 follows the Pareto distribution characterized by a scale factor z_m and a shape factor α.

4.1 Identification

Parameters in the model are categorized into two groups. The first group is calibrated by a prior information from the aggregate statistics or the literature. The second group is calibrated by estimation from the micro-level data or through the model. Table 6 reports the parameters in the first group, $(\rho, \beta, \tau, A, \delta)$. The discount rate, ρ, is set to 0.02 to match the average interest rate in the sample period. The production function quality share, β, is 0.109, following Akcigit and Kerr (2018). The firm exit rate, τ, is 0.06, targeting the average exit rate of firms above 5 years old during our sample period based on the Business Dynamics Statistics (BDS).[26] The risk aversion parameter, A, and the effort cost elasticity, δ, are set to be 0.5 and 1, respectively, which are commonly used in the literature (Hall and Van Reenen, 2000).

We calibrate the ten remaining parameters in the second group, (λ_0, α, z_m, ϕ,η, B, b, q_0, h, h_s, κ), using the minimum distance method, inspired by Lentz and Mortensen (2008). Further, we assume that the degree of friction is the same whether the inventor's

[25]See Appendix C.9 for the proof.

[26]The BDS data is compiled from the Longitudinal Business Database (LBD) by the Census Bureau.

Table 6: Parameter Values from a Priori Information

Parameter	Description	Value	Identification
ρ	Discount rate	0.02	Interest Rate
β	Production function quality share	0.109	Firm profitability
τ	Exo. exit rate	0.06	BDS
A	Risk aversion	0.5	Risk aversion
δ	Effort cost elasticity	1	Effort cost elasticity

Notes: This table shows parameter values from the literature or direct estimation.

optimal choice is an incumbent firm or a startup; that is, $h = h_s$. The parameters, along with their corresponding moments are in Table 7.

Table 7: Parameters from the Minimum Distance Estimation

Para.	Description	Identification
λ_0	Innovation arrival rate	Growth rate
α	Shape of idea quality distribution	SD/Mean of patent citations
z_m	Scale of idea quality distribution	Aggregate economic value of innovations
ϕ	Innovation value dispersion	SD/Mean of economic value cond. on cites
B	Maturity of technology	Technology Novelty index
η	Elasticity of substitution	Estimation of the economic-value function
b	Exponent of the synergy function	Estimation of the economic-value function
q_0	Scale of the synergy function	New-to-incumbent ratio
h, h_s	Matching friction (incumbent & startup)	Firm size ratio by cite quartiles (Q4/Q1)
κ	Scale of the firm-led innovation	Entrant/incumbent citation ratio

Notes: Parameters in this table are jointly calibrated to minimize the distance between the model and data moments.

Growth Rate—Innovation is the sole driver of growth in this model. Consequently, the scale parameter of the innovation arrival rate, λ_0, plays a critical role in determining the aggregate growth rate. A higher arrival rate shortens the average time between innovations, thereby raising the overall growth rate. We calibrate λ_0 so that the model's implied aggregate growth rate matches 2%, consistent with the long-run growth rate in the US.

The S.D.-to-Mean Ratio of Patent Citations—This ratio captures the dispersion in patent citations observed in the data, which reflects the underlying dispersion in inventors' idea quality. The parameter α is the primary driver of this dispersion. Specifically, the standard deviation-to-mean ratio of the idea distribution is given by $\frac{1}{\sqrt{\alpha(\alpha-2)}}$. Although the patents recorded in the USPTO data represent only successful innovations—a selected subset of all ideas—the dispersion in patent citations remains heavily influenced by α. We construct the citation distribution by pooling all patents granted since 1976 and their corresponding citations recorded by the USPTO, and compute the standard deviation-to-mean ratio, which is approximately 2.784.

Aggregate Innovation Value—The economic value of innovations is modeled as directly contributing to firm value. In the model, the economic value, denoted by x (or z for startups), is assumed to follow a uniform distribution with its mean determined by the underlying idea quality, z_0. Given α, the average scientific value of ideas is governed by the scale parameter of the Pareto distribution, z_m. Accordingly, z_m is calibrated to match the aggregate economic value of patents. The estimation method proposed by Kogan et al. (2017) is adopted, in which the stock market's reaction to patent-related news is used to infer patent value. An extended version of their dataset, provided by the authors, is employed. It links patents issued to U.S. firms from 1926 to 2022 with stock returns from CRSP and firm-level data from Compustat. It is found that, at the aggregate level, the total patent value in each year amounts for approximately 0.17 times total firm value. Because the distribution of public firms differs from that of the broader firm population, we apply the statistical model developed by Yang (2023) to estimate the average patent value across all firms using public firm data. After this adjustment, the ratio is 0.15. The parameter z_m is then calibrated to lie in the range $0.15 - 0.17$.

S.D.-to-Mean Ratio of Innovations' economic Value Conditional on Citations—The economic value of innovations is based on the scientific value of ideas but is also subject to additional randomness. In the model, the degree of randomness is governed by the parameter ϕ. Specifically, conditional on idea quality, the standard deviation-to-mean ratio of the uniform distribution of innovation economic value is given by $\frac{\phi}{\sqrt{3}}$. Using the same sample employed to calibrate z_m, the standard deviation-to-mean ratio of patent economic value is estimated while controlling for the number of patent citations. In the data, this ratio is found to be approximately 0.416.

Estimation of the Function of Patents' Economic Values—Both the elasticity of substitution between past knowledge stock and new ideas (η) and the regression coefficient of innovations' economic value on incumbent firm size (b) are estimated directly from the data using nonlinear least squares. We observe, for each patent i granted in firm j and year t, an economic value $x_{ijt}\nu Q(t)$, firm quality q_{jt}, average backward citations of the economy B_t, and patent forward citations z_{ijt}. The structural model for the underlying patent value is

$$x_{ijt}\nu Q(t) = \alpha\left(\frac{q_{jt}}{Q(t)}\right)^b \left(B_t^\eta + z_{ijt}^\eta\right)^{\frac{1}{\eta}} \nu Q(t)\varepsilon_{ijt},$$

where b and η are parameters of interest and ε_{ijt} is a multiplicative disturbance with $E[\ln \varepsilon_{it} \mid q_{jt}, B_t, z_{ijt}] = 0$.

Taking logarithms and absorbing constants into an intercept and year fixed effects we obtain the estimating equation

$$\ln(x_{ijt}\nu Q(t)) = c + b\ln q_{jt} + \frac{1}{\eta}\ln\left(B_t^\eta + z_{ijt}^\eta\right) + \delta_t + u_{ijt},$$

where c collects α and ν, δ_t is a year fixed effect capturing $Q(t)$, and u_{ijt} is an error term satisfying $E[u_{ijt} \mid q_{ijt}, B_t, z_{ijt}, t] = 0$. Let the full parameter vector collect all unknowns,

$$\theta = \left(c,\ b,\ \eta,\ \{\delta_t\}_{t\in\mathcal{T}}\right),$$

and define the residual

$$u_{ijt}(\theta) = \ln(x_{ijt}\nu Q(t)) - c - b\ln q_{jt} - \frac{1}{\eta}\ln\left(B_t^\eta + z_{ijt}^\eta\right) - \delta_t.$$

We estimate θ by nonlinear least squares, i.e.

$$\hat{\theta} = \arg\min_\theta S(\theta), \qquad S(\theta) = \sum_{ijt} u_{ijt}(\theta)^2.$$

At the optimum $\hat{\theta}$, the first-order conditions with respect to every component of θ are satisfied:

$$\frac{\partial S(\theta)}{\partial\theta_k} = 2\sum_i u_i(\theta)\,\frac{\partial u_i(\theta)}{\partial\theta_k} = 0, \qquad k = 1,\ldots,\dim(\theta).$$ [27]

The elasticity η and the scale factor b, included in θ, are identified in the estimation.

Technological Novelty Index—Technological novelty is defined as the ratio of total forward citations to the sum of forward and backward citations for all patents granted in a given year. In the model, the past technological stock, B, which enters the realization potential of new ideas, $\gamma(z_0) = (z_0^\eta + B^\eta)^{1/\eta}$, corresponds to the stock of backward citations, representing the maturity of the existing technological base. The value of B is calibrated so that the model-generated average forward-to-backward citation ratio,

[27] These conditions can be interpreted as a system of moment restrictions. Defining the moment vector

$$g(W_{ijt},\theta) = u_{ijt}(\theta)\,\frac{\partial u_{ijt}(\theta)}{\partial\theta},$$

where $\frac{\partial u_{ijt}(\theta)}{\partial\theta}$ is the gradient of the residual with respect to all components of θ, the population moment conditions are

$$E\left[g(W_{ijt},\theta_0)\right] = 0.$$

Thus, the nonlinear least squares estimator $\hat{\theta}$ can be viewed as a just-identified GMM estimator that jointly estimates $(c, b, \eta, \{\delta_t\}_{t\in\mathcal{T}})$ by enforcing the orthogonality conditions implied by the model.

$B/(B + \int z_0 \, d\Psi(z_0))$, matches $1 -$ the average Novelty Index between 1982 and 2016, where $\int z_0 \, d\Psi(z_0)$ corresponds to the total forward citations of available ideas.[28]

New-to-Incumbent Ratio—The scale parameter in the synergy function, $\tilde{q}_0$, affects the benefit of contributing an idea to an incumbent firm compared to initiating a new venture. Therefore, it is related to inventors' choice between incumbent firms and startups. We calibrate $\tilde{q}0$ using the "New-to-Incumbent Ratio" defined in Section 2.3.1, targeting its average value over 1982–2016.

Firm Size Ratio by Fourth-to-First Quartile of Patent Citations—The model predicts that when inventors choose to join incumbent firms, the size of the firm they select should increase with the quality of their ideas. However, this positive sorting is hindered by matching frictions. Greater frictions weaken the relationship between idea quality and firm size. To calibrate the degree of frictions, h and h_s (assumed equal), the model generates the average firm size by patent citation quartiles, conditional on the patent being developed within an incumbent firm. The ratio of firm size between the fourth and first quartiles is then computed and matched to its empirical counterpart. [29]

Citation Ratio Between New and Incumbent Firms—The model predicts a threshold in idea quality above which inventors prefer founding startups to joining incumbent firms. Consequently, patents from startups should, on average, exhibit higher scientific value than those from incumbents. However, frictions in the startup decision attenuate this effect. The scale effect of frictional matching, κ, affects the realized ratio between frictional matching and the optimal matching. The patent citation ratio between new and incumbent firms captures this relationship and is used to calibrate κ.

4.2 Calibration Results

Table 8 reports the model-generated moments and their counterparts in the data. Overall, the model matches the targeted moments closely. The resulting parameter values are

[28]The Novelty Index defined in Section 2.1 can be expressed as $\frac{F}{F+B}$, where F denotes all forward citations in a given period. Thus, $(B/(B + \int z_0 \, d\Psi(z_0)))$ corresponds to $\frac{B}{B+F} = 1 - \frac{F}{B+F} = 1 -$ the Novelty Index.

[29]Within incumbent firms, innovations originated from an idea of quality z_0 (with $z_0 \in \mathcal{Z}$) fall into three categories: (i) inventor-led innovation; (ii) firm-led innovation that was not matched with an incumbent of the optimal size; and (iii) firm-led innovation that was not matched with a startup. The average size of firms that owns patents z_0 within a range $\mathcal{Z}$ is calculated as following:

$$\begin{aligned}\mathbb{E}_{z_0 \in \mathcal{Z}}(q) = &\int_{z_0 \in \mathcal{Z}} hq^*(z_0)\lambda_I^*(z_0) d\Psi(z_0) + \int_{z_0 \in \mathcal{Z}, q^*(z_0)>0} (1-h)\, \lambda_F^*(z_0) \int qf(q,t) dq d\Psi(z_0) \\ &+ \int_{z_0 \in \mathcal{Z}, q^*(z_0)=0} (1-h_s)\, \lambda_F^*(z_0) \int qf(q,t) dq d\Psi(z_0)\end{aligned}$$

where $\int qf(q,t)dq$ is the average firm size in both firm-led cases, since ideas are allocated to firms randomly based on the incumbent firm size distribution.

reported in Table 9.

Table 8: Moments

Identification Moment	Data	Model
Growth rate	0.020	0.020
S.d.-to-mean ratio of patent citations	2.784	2.823
Aggregate innovation value	0.15-0.17	0.166
S.d.-to-mean ratio of innovation value cond. on citations	0.416	0.416
Elasticity of Substitution between past Knowledge Stock and New Ideas	-0.31	-0.31
Technology Novelty index	0.554	0.554
Regression coefficient of innovation value on firm size	0.32	0.32
New-to-incumbent ratio	0.054	0.054
Firm size ratio by fourth-to-first-quartile of citations	1.18	1.02
Citation ratio between new and incumbent firms	1.361	1.332

Notes: This table compares the moments generated from the calibrated model and the data. In general, the model generated moments match the data well.

Our estimates suggest that, relative to startups, incumbent firms have substantially lower realized returns from a given innovation, largely because of frictions in adopting and integrating new technologies. In particular, the ratio $\frac{\int \gamma(z_0)\, d\Psi(z_0)}{\int z_0\, d\Psi(z_0)} = 0.090$ is significantly below one. At the same time, synergy is a key force in commercialization. The scale parameter in the synergy function is $\tilde{q}_0 = 1.5 \times 10^{-4}$, and the elasticity is -0.31, implying that an incumbent of average size can generate roughly 18 times the commercialization value of a startup through synergy alone. Once adoption frictions are incorporated, the net advantage declines to about 1.64 times.

Technological novelty waves shape idea allocation along both the extensive and intensive margins. The extensive margin determines how many inventors found new businesses, while the intensive margin governs the firm size an inventor selects when joining an incumbent. Figure 5 illustrates both margins by plotting the optimal firm-size choice across idea qualities in a frictionless matching environment ($h = h_s = 1$). A positive firm size corresponds to joining an incumbent, whereas zero indicates founding a new business. The figure indicates positive assortative matching among inventors who join incumbents: higher-quality ideas are matched with larger firms. When an inventor's idea quality, z_0, exceeds a cutoff, she instead chooses to launch a startup.

Although inventors' optimal choices given idea quality, z_0, is time-invariant, the economy's position along technology waves changes the distribution of idea quality: both its support and density shift over time. For instance, in a mature stage, the lower bound of the idea-quality distribution is marked by the dashed vertical line. So, inventors with idea quality between the lower bound and the cutoff join incumbents. As aggregate technological novelty rises, this lower bound moves to the right (as indicated by the

Table 9: Estimated Parameter Values

Parameter	Description	Value
λ_0	Innovation arrival rate	0.91
α	Shape of idea quality distribution	2.12
z_m	Scale of idea quality distribution	0.052
ϕ	Innovation quality draw	0.72
η	CES elasticity of substitution	-0.31
B	Maturity of technology	0.079
b	Exponent of the synergy function	0.32
q_0	Denominator of the synergy function	1.5E-4
h, h_s	Matching friction (incumbent & startup)	0.11
κ	Scale of the firm-led innovation	17.62

Notes: Parameters in this table are jointly calibrated to minimize the distance between the model and data moments.

arrow), implying that a larger share of inventors found startups and aggregate positive assortative matching weakens. Both forces reduce market concentration.

Moreover, as more ideas are developed in startups, growth shifts toward expansion in firm mass rather than increases in average firm quality. Overall, because startups face lower adoption frictions and achieve higher realized value from new ideas, aggregate growth increases.

5 Quantitative Analysis

Using the calibrated model, we simulate the economy starting in 1986, the first peak of the technology wave in our sample. In each year, we adjust the average idea quality of inventors, z_0, so that the model-implied Novelty Index, $\frac{\int z_0\, d\Psi(z_0)}{B+\int z_0\, d\Psi(z_0)}$, matches its empirical counterpart. Equivalently, we hold all parameters fixed except the one governing aggregate technological-novelty shocks as "MIT shocks". The simulation produces annual series for model-implied market concentration and the allocation of innovation. We then compare these series to the data and compute their correlations.

In the initial year, 1986, we match the model to the empirical firm size distribution using the Census Business Dynamics Statistics (BDS). Appendix D.1 provides a detailed description of this procedure. Because the equilibrium features no transition dynamics, the evolution of firm mass, N_F, and the firm size distribution, $f(q)$, can be computed directly. The resulting paths of N_F and $f(q)$ jointly determine the time series of growth rates and market concentration.

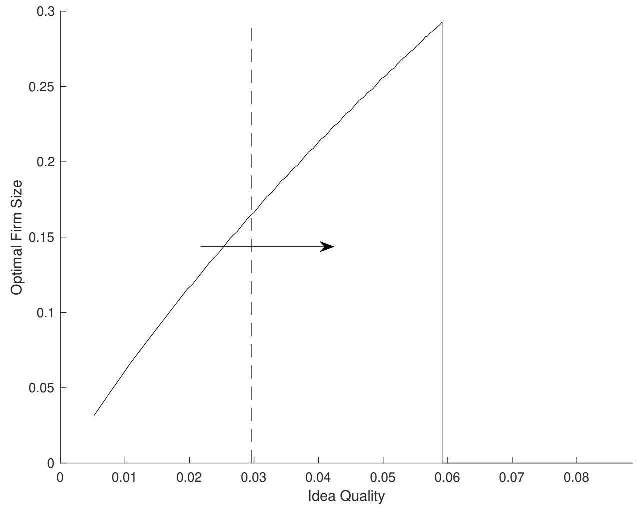

Figure 5: Firm-size Choices of Inventors

Notes: The figure shows the model-implied mapping from an inventor's idea quality to the optimal firm size. A positive firm size corresponds to joining an incumbent of that size, whereas zero indicates founding a new business.

5.1 Growth Rate

The aggregate growth rate, g, can be decomposed into two components: growth in average firm quality and growth in firm mass (Equation (38)). Using the simulated economy, we compute 10-year averages of g and each component, and report them in Figure 6. Two patterns stand out. First, aggregate growth declines over time, with the bulk of the slowdown coming from the decline in firm mass. Second, firms' average quality behaves differently. Growth in average firm quality is relatively strong in 1996–2005, consistent with faster improvements among surviving firms, but it is more than offset by a pronounced contraction in firm mass. In the final decade, the improvement in average firm quality contributes less to growth, reinforcing the downward trend in g. These qualitative patterns align with the evidence in Autor et al. (2020) and Akcigit and Ates (2019): since the late 1990s, productive incumbents have expanded while net firm entry and overall dynamism have declined.

5.2 Technology Waves and the Market Concentration

Figure 7 presents the simulated evolution of market concentration, measured by the Herfindahl-Hirschman Index (HHI), alongside its empirical counterpart and the technological novelty waves. The solid red curve depicts the HHI from the data, normalized to its 1986 level, while the dashed red curve shows the simulated HHI from the model, also normalized by the 1986 value. To illustrate the connection with

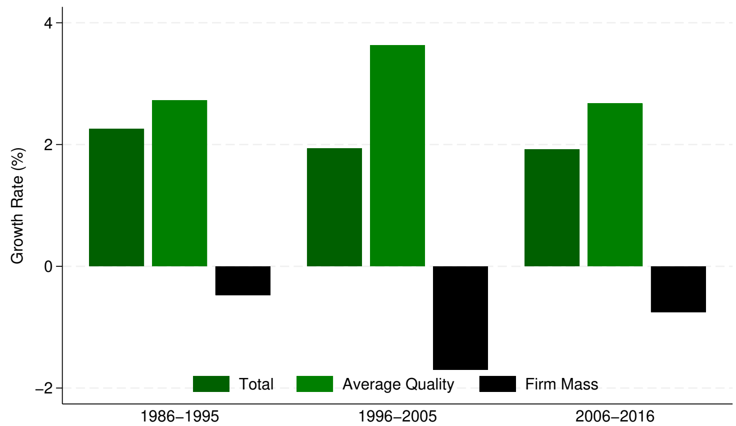

Figure 6: Decomposition of Growth Rates by Period

Notes: This figure reports decade averages of the aggregate economic growth rate and its two components. In each decade, the left bar shows total growth. The middle bar shows the contribution from improvements in average firm quality, and the right bar shows the contribution from changes in firm mass.

technological waves, the figure also includes the relative ratio of forward citations to the sum of forward and backward citations—following the methodology defined in this paper, and consistent with Figure 1. The model-generated HHI closely mirrors its empirical counterpart and exhibits a negative relationship with technological waves.

Although the calibration does not explicitly target any measure of market concentration, the model successfully replicates both the upward trend and cyclical fluctuations observed in the data. To disentangle these components, we fit separate linear trends to the empirical and simulated HHI series and subtract them to obtain the detrended time variations, as shown in Figure 8. Summary statistics are reported in the first two rows of Panel A in Table 10.

The fitted linear trend has a slope of 7.89×10^{-2}, which explains about 83.4% of the observed rise in market concentration over the sample period. The detrended HHI series has a correlation of 0.552 with the data. The standard deviation of the model's detrended series is 0.044, accounting for 43.6% of the empirical volatility. The first-order autocorrelation is 0.847 in the model, compared with 0.920 in the data. Overall, these results indicate that technological waves are an important driver of both the trend and fluctuations in market concentration.

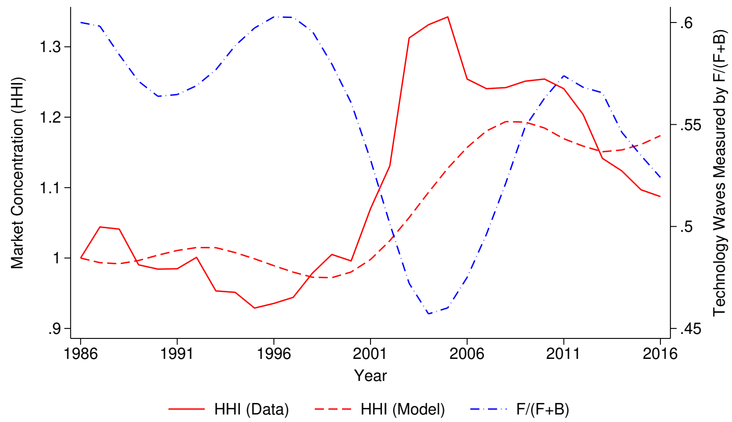

Figure 7: Technology Waves and Model Generated HHI

Notes: This figure shows the technological waves and the trend of model-generated market concentration over time. The blue curve, based on the methodology defined in this paper, calculates the relative ratio of forward citations to the sum of forward and backward citations (same as Figure 1). The red solid curve displays the simulated HHI in each year, which is normalized by the HHI in 1986. The red dashed curve is the empirical HHI moved forward for three years, also normalized by 1986.

Table 10: Comparison between Model and Data

	Original		Detrended		
	Mean	Slope	S.D.	Autocorr	Corr. Data
	Panel A. HHI				
Data	1.099	9.46E-2	0.101	0.847	1
Model	1.068	7.89E-2	0.044	0.920	0.552
Model (Ext. Margin)	0.984	4.47E-2	0.054	0.909	0.487
	Panel B. New-to-Incumbent Ratio				
Data	0.052	-5.05E-4	0.007	0.830	1
Model	0.054	-7.46E-4	0.011	0.914	0.706

Notes: This table shows the trend and detrended time variations of the HHI and New-to-Incumbent ratio in the data and the model.

5.3 Decomposition of the Intensive and Extensive Margins

Market concentration is shaped by both the extensive margin (the mass of firms) and the intensive margin (the firm-size distribution). To disentangle these forces, we simulate HHI using the time path of firm mass while holding the firm-size distribution fixed. Under this counterfactual, changes in HHI are driven solely by entry and exit, effectively

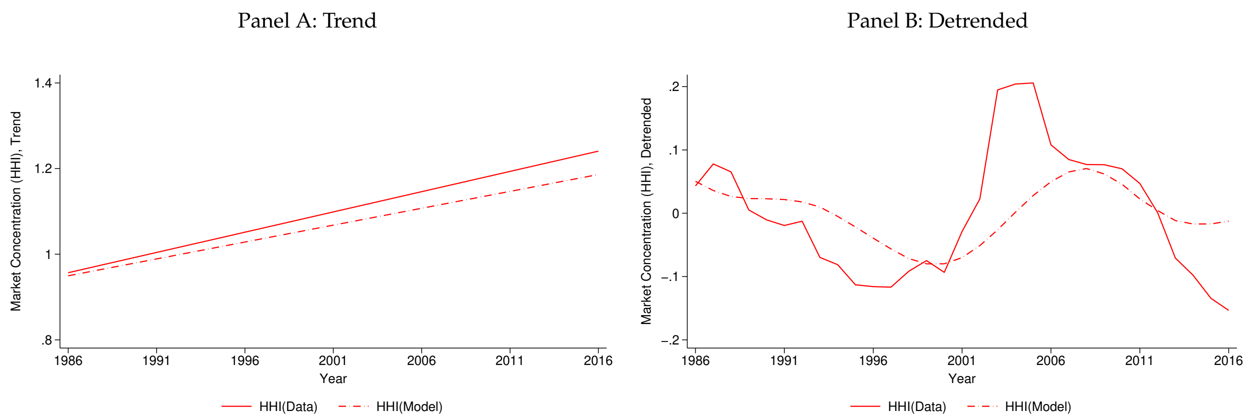

Figure 8: The Trend and Detrended Time Variations of the HHI
Notes: The left panel plots the linear trend in market concentration (HHI) over time in the data (solid line) and in the model simulation (dashed line). The right panel shows the corresponding detrended HHI series, highlighting fluctuations around the trend.

shutting down the intensive margin by removing changes in idea allocation among incumbents. The resulting "extensive-margin-only" HHI is plotted as the dash-dotted line in Panel A of Figure 9, alongside the full-model HHI that incorporates both margins (dashed line). While the extensive-margin-only series exhibits fluctuations similar to the full model, it displays a notably weaker upward trend. Panel A of Table 10 summarizes the corresponding moments: the extensive margin alone accounts for 56.7% of the upward trend in the model-implied HHI and 47.3% of the empirical trend. These results indicate that both the extensive and intensive margins contribute meaningfully to the rise in market concentration over the sample period.

5.4 Allocation of ideas

Empirically, this paper shows that inventors are more likely to form startups when revolutionary technologies appear and join incumbent firms when technologies mature. This is repeatedly shown by the solid curve in Panel B of Figure 9. The New-to-Incumbent ratio generated by the model is shown by the dashed curve in the same figure. They have nearly simultaneous waves. To further evaluate their relationship, we use linear trends to fit the two curves respectively, and then subtract them to get the detrended time variations. Both curves have a negative trend, and the correlation between the detrended model-generated and the detrended actual new-to-incumbent ratio is 0.703. Further summary statistics are displayed in Panel B of Table 10.

The average New-to-Incumbent ratio in the model is 0.054, close to 0.052 in the data. The slope of the linear trend in the model is $-7.46E-3$, slightly more negative than the data counterpart. The detrended series also shows slightly larger variability, with a

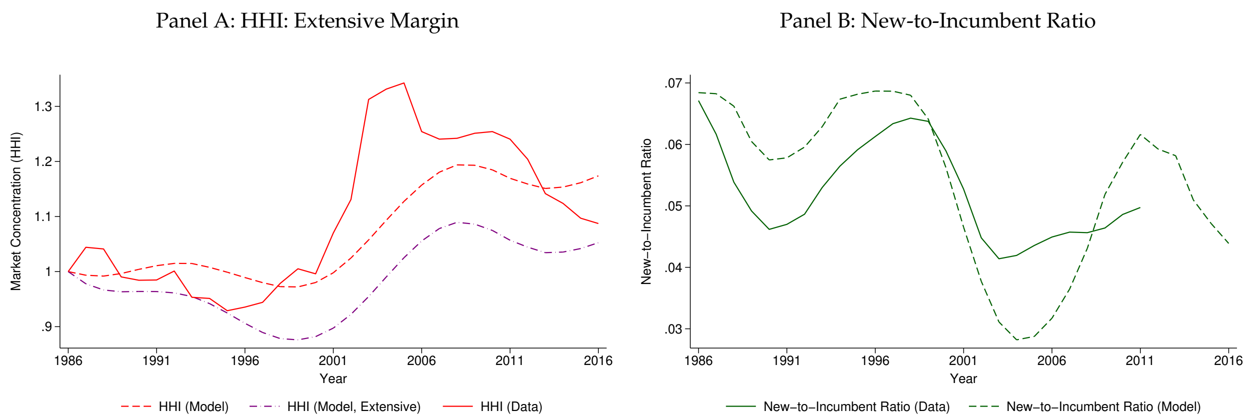

Figure 9: The Extensive Margin of HHI and Idea Allocation

Notes: Panel A shows the extensive margin of the model-generated market concentration over time. The dashed curve and short-dashed curve displays the simulated HHI and the HHI only considering extensive margin, respectively. Both are normalized by the model generated HHI level in 1986. Panel B shows the technological waves and the trend of model-generated share of innovations in startups over time. The dashed curve displays the simulated New-to-Incumbent ratio in each year whereas the solid curve shows the New-to-Incumbent ratio in the data (same as Figure 3).

standard deviation of 0.011 compared to 0.007 in the data, and a slightly higher first-order autocorrelation (0.914 vs. 0.830). The model's larger volatility may be due to the absence of adjustment costs. In our model, inventors are short-lived and choose between joining startups or incumbent firms without regard to their previous affiliations. As a result, the New-to-Incumbent ratio reacts immediately to aggregate shocks.

6 Conclusion

This paper studies how technological waves shape the market concentration, through the reallocation of inventors. It provides empirical evidence and structural analysis showing that market concentration is inversely related to and lagged behind the technological waves. This discovery suggests the presence of a low-frequency business cycle in the economy. We explore one potential channel behind this connection: the allocation of ideas. Using the data from the Longitudinal Business Database (LBD) from the Census Bureau and the patent information from the USPTO, this paper shows that the share of patents formed in new businesses co-move closely with the technological waves. At the peaks of the technological waves, a larger share of patents are forming in new businesses, while at the troughs, a larger share of patents come from existing businesses.

This paper proposes a theoretical framework that elucidates the decision-making process of inventors regarding their choice of innovation pathways, thus providing an explanation for the observed empirical patterns. Inventors are faced with a choice between forming a new business of a random size with a partner or joining an incumbent business of a selected size. This decision hinges on a trade-off: new businesses offer better incentives and adaptability in embracing novel technologies, while incumbents possess synergies and experience in commercialization. Our model effectively captures the relationship between technological waves and market concentration, primarily through the redistribution of innovative ideas. It implies that the deceleration in the emergence of groundbreaking technologies could be a significant contributing factor to the decline in economic growth and the rise in market concentration after the late 1990s.

References

Acemoglu, Daron, Ufuk Akcigit, and William R Kerr. 2016. "Innovation network." *Proceedings of the National Academy of Sciences*, 113(41): 11483–11488.

Aghion, Philippe, and Peter Howitt. 1990. "A model of growth through creative destruction."

Akcigit, Ufuk, and Nathan Goldschlag. 2023. "Where have all the" creative talents" gone? Employment dynamics of us inventors." National Bureau of Economic Research.

Akcigit, Ufuk, and Sina T Ates. 2019. "What Happened to Us Business Dynamism?"

Akcigit, Ufuk, and Sina T Ates. 2023. "What happened to US business dynamism?" *Journal of Political Economy*, 131(8): 2059–2124.

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

Akcigit, Ufuk, Murat Alp Celik, and Jeremy Greenwood. 2016. "Buy, Keep, or Sell: Economic Growth and the Market for Ideas." *Econometrica*, 84(3): 943–984.

Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen. 2020. "The fall of the labor share and the rise of superstar firms." *The Quarterly Journal of Economics*, 135(2): 645–709.

Ayerst, Stephen, Faisal Ibrahim, Gaelan MacKenzie, and Swapnika Rachapalli. 2023. "Trade and diffusion of embodied technology: an empirical analysis." *Journal of Monetary Economics*, 137: 128–145.

Bena, Jan, and Kai Li. 2014. "Corporate Innovations and Mergers and Acquisitions." *The Journal of Finance*, 69(5): 1923–1960.

Bighelli, Tommaso, Filippo Di Mauro, Marc J Melitz, and Matthias Mertens. 2023. "European firm concentration and aggregate productivity." *Journal of the European Economic Association*, 21(2): 455–483.

Bloom, Nicholas, Charles I Jones, John Van Reenen, and Michael Webb. 2020. "Are ideas getting harder to find?" *American Economic Review*, 110(4): 1104–1144.

Bowen III, Donald E, Laurent Frésard, and Gerard Hoberg. 2021. "Rapidly Evolving Technologies and Startup Exits." *Swiss Finance Institute Research Paper*, , (19-22).

Bowen III, Donald E, Laurent Frésard, and Gerard Hoberg. 2023. "Rapidly evolving technologies and startup exits." *Management Science*, 69(2): 940–967.

Campbell, Jeffrey R, and Hugo A Hopenhayn. 2005. "Market size matters." *The Journal of Industrial Economics*, 53(1): 1–25.

Cassiman, Bruno, and Reinhilde Veugelers. 2006. "In Search of Complementarity in Innovation Strategy: Internal R&d and External Knowledge Acquisition." *Management Science*, 52(1): 68–82.

Cavenaile, Laurent, Murat Alp Celik, and Xu Tian. 2019. "Are markups too high? competition, strategic innovation, and industry dynamics." *Competition, Strategic Innovation, and Industry Dynamics (August 1, 2019)*.

Chatterjee, Satyajit, and Esteban Rossi-Hansberg. 2012. "Spinoffs and the Market for Ideas." *International Economic Review*, 53(1): 53–93.

Cunningham, Colleen, Florian Ederer, and Song Ma. 2021. "Killer acquisitions." *Journal of Political Economy*, 129(3): 649–702.

Dinlersoz, Emin, Can Dogan, and Nikolas Zolas. 2024. "Starting Up AI."

Fons-Rosen, Christian, Pau Roldan-Blanco, and Tom Schmitz. 2021. "The Aggregate Effects of Acquisitions on Innovation and Economic Growth."

Ganapati, Sharat. 2021. "Growing oligopolies, prices, output, and productivity." *American Economic Journal: Microeconomics*, 13(3): 309–327.

Greenwood, Jeremy, and Boyan Jovanovic. 1999. "The information-technology revolution and the stock market." *American Economic Review*, 89(2): 116–122.

Greenwood, Jeremy, and Mehmet Yorukoglu. 1997. "1974." Vol. 46, 49–95, Elsevier.

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

Grullon, Gustavo, Yelena Larkin, and Roni Michaely. 2019. "Are US industries becoming more concentrated?" *Review of Finance*, 23(4): 697–743.

Hall, Bronwyn, and John Van Reenen. 2000. "How effective are fiscal incentives for R&D? A review of the evidence." *Research policy*, 29(4-5): 449–469.

Jovanovic, Boyan, and Peter L Rousseau. 2014. "Extensive and intensive investment over the business cycle." *Journal of Political economy*, 122(4): 863–908.

Kalyani, Aakash. 2024. "The Creativity Decline: Evidence from US Patents." *Available at SSRN 4318158*.

Kelly, Bryan, Dimitris Papanikolaou, Amit Seru, and Matt Taddy. 2021. "Measuring technological innovation over the long run." *American Economic Review: Insights*, 3(3): 303–320.

Klette, Tor Jakob, and Samuel Kortum. 2004. "Innovating firms and aggregate innovation." *Journal of political economy*, 112(5): 986–1018.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman. 2017. "Technological Innovation, Resource Allocation, and Growth." *The Quarterly Journal of Economics*, 132(2): 665–712.

Kwon, Spencer Yongwook, Yueran Ma, and Kaspar Zimmermann. 2023. "100 years of rising corporate concentration." *University of Chicago, Becker Friedman Institute for Economics Working Paper*, , (2023-20).

Kydland, Finn E, and Edward C Prescott. 1982. "Time to build and aggregate fluctuations." *Econometrica: Journal of the Econometric Society*, 1345–1370.

Lentz, Rasmus, and Dale T Mortensen. 2008. "An Empirical Model of Growth through Product Innovation." *Econometrica*, 76(6): 1317–1373.

Liu, Ernest, and Song Ma. 2021. "Innovation Networks and Innovation Policy."

Liu, Huben, Dimitris Papanikolaou, Lawrence DW Schmidt, and Bryan Seegmiller. 2025. "Technology and Labor Markets: Past, Present, and Future; Evidence from Two Centuries of Innovation." National Bureau of Economic Research.

Ma, Yueran, Mengdi Zhang, and Kaspar Zimmermann. 2024. "Business Concentration around the World: 1900—2020."

Ma, Yueyuan. 2022. "Specialization in a knowledge economy." *Available at SSRN 4052990*.

Olmstead-Rumsey, Jane. 2019. "Market concentration and the productivity slowdown."

Perla, Jesse, Christopher Tonetti, and Michael E. Waugh. 2021. "Equilibrium Technology Diffusion, Trade, and Growth." *American Economic Review*, 111(1): 73–128.

Romer, Paul M. 1986. "Increasing Returns and Long-Run Growth." *Journal of Political Economy*, 94(5): 1002–1037.

Romer, Paul M. 1990. "Endogenous Technological Change." *Journal of Political Economy*, 98(5, Part 2): S71–S102.

Yang, Shaoshuang. 2023. "The Distribution of Innovations across Firms." *Working Paper*.

Online Appendix

A Data Description

The data used in this paper includes the Longitudinal Business Database (LBD), the USPTO patent data, and the Compustat Fundamentals Annual. This section provides details about the information of the datasets and the construction of key variables.

A.1 The USPTO Patent Data

The USPTO patent data contains information of all patents issued between 1976 and 2022. It can be downloaded from the PatentsView website. For each patent, the data documents the patent type (utility, design, plant, etc.), the IPC code indicating its technological class, the grant year, and the patents it cites and it is cited. We keep all the utility patents to focus our attention to the introduction of new products and processes.

Forward Citations Forward citations are citations a focal patent receives from others. It indicates how many patents follow the focal one. This paper calculates the number of forward citations each patent gets within five years after issuance.

Backward Citations Backward citations are citations that other patents receive from the focal patent. It indicates to what extent the focal patent follows the existing technology. This paper calculates the number of backward citations by counting the number of patents cited by the focal patent that were granted within the previous five years.

The Novelty Index According to the definition in the paper, we calculate this index by dividing the number of forward citations received by all the utility patents granted in a year by the summation of the forward and backward citations of those patents. The Novelty Index by IPC is derived in a similar way for each IPC class and each year.

A.2 The Compustat Fundamentals Annual

The Compustat Fundamentals Annual contains information of all the publicly listed firms in the US. It records the firms' net sales, the number of employees, the primary industry (4-digit SIC code), and the headquarter locations of each firm. We keep all the firms that are headquartered in the US.

Primary Industry The primary industry of each firm in Compustat is based on the 4-digit SIC code assigned to each firm in the Fundamentals Annual. The code can be aggregated to different levels. Manufacturing is corresponding to SIC codes 2000-3999; utility and transportation is corresponding to SIC codes 4000-4999; wholesale trade is corresponding to SIC codes 5000-5199; retail trade is corresponding to SIC codes 5200-5999; finance is corresponding to SIC codes 6000-6999; service is corresponding to SIC codes 7000-8999.

The Herfindahl-Hirschman Index (HHI) Following the methods in Grullon, Larkin and Michaely (2019), we first calculate the HHI of each 3-digit SIC code by the squared ratios of firm net sales to the total net sales in that 3-digit industry. To get the aggregate HHI, we sum up the HHIs of all the 3-digit SIC codes and weight them by their total net sales.

A.3 The Longitudinal Business Database (LBD)

The LBD is collected by the US Census Bureau and is an establishment-level data that covers the universe of US businesses with paid employees from 1976 to 2020. The dataset assigns a firm ID to all establishments belonging to the same firm. Using the Business Dynamics Statistics of Patenting Firms (BDS-PF) patent assignee-FIRMID crosswalk from the Census, this paper links the USPTO patent data with firms in the LBD, therefore, derives all utility patents in the US that were granted to employer businesses between 1976 and 2020.

New-to-Incumbent Ratio After merging the patent data with the LBD, this paper can identify the firm each patent was granted to. If the firm is less than or equal to five years old in the patent's grant year, we indicate that the idea behind the patent was absorbed by a new firm 5 years ago. Otherwise, we indicate that the idea was absorbed by an incumbent firm 5 year ago. Then we divide the number of ideas combined with new firms by the number of ideas combined with incumbent firms to get the New-to-Incumbent Ratio.

Firm Size The LBD documents the number of employees each firm hires in each year. We deriving the mapping between patent forward citations and incumbent firm size, we use the number of employees as a proxy for size.

B More Empirical Evidence

B.1 Empirical Patterns without Smoothing

Figure 10 and Figure 11 present the patterns of the technological waves, HHI, and the New-to-Incumbent ratio without smoothing techniques. The negative correlation between the HHI and the Novelty Index, as well as the positive correlation between the New-to-Incumbent ratio and the Novelty Index, remain prominent.

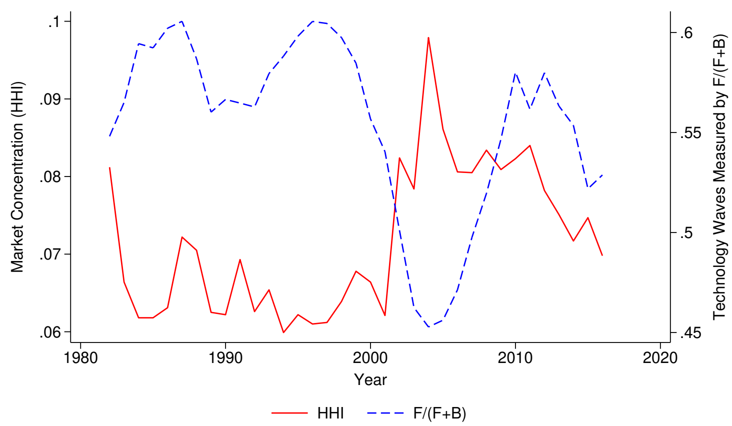

Figure 10: Technological Waves and Market Concentration without Smoothing
Notes: This figure shows the technological waves and the trend of market concentration over time. The blue dashed curve, based on the methodology defined in this paper, calculates the relative ratio of forward citations to the sum of forward and backward citations. The red solid curve displays the HHI in each year, which is the weighted average of the industry-level HHI in each year. The weight is the total sales of firms in each industry. The two curves have different y-axises, which are shown respective on the left and right.
Sources: Compustat Fundamental Annuals and USPTO patent and citation data.

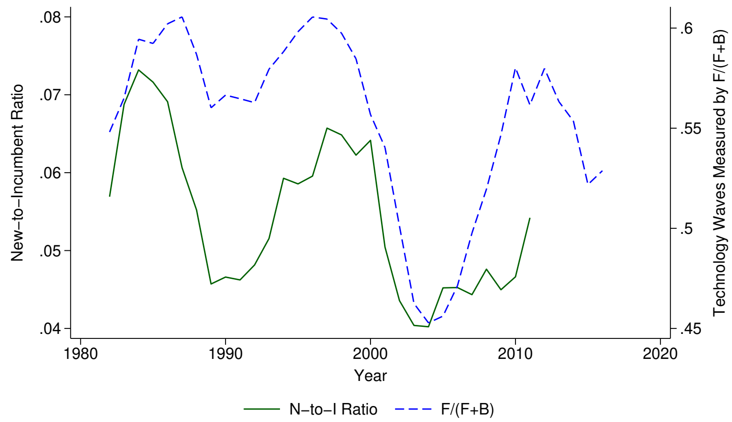

Figure 11: Technological Waves and Idea Allocation without Smoothing

Notes: This figure shows the technological waves and the idea allocation between new and incumbent firms over time. The blue dashed curve, based on the methodology defined in this paper, calculates the relative ratio of forward citations to the sum of forward and backward citations. The green solid curve displays the "New-to-Incumbent Ratio" defined in the paper, capture where new ideas contribute their value. The two curves have different y-axises, which are shown respective on the left and right.
Sources: Longitudinal Business Database (LBD) and USPTO patent and citation data.

B.2 Technological Waves by Technological Field

The Novelty index across the nine technological fields is shown in Figure 12. The index is based on the same algorithm as in Equation 1 except that the forward and backward citations are aggregated across each of the 1-digit IPC code. The top three fields with the highest Novelty index are Human Necessities, Physics, and Electricity at the first peak; Electricity, Physics, and Human Necessities at the second peak; Human Necessities, Chemistry and Metallurgy, and Mechanical Engineering etc. at the third peak.

B.3 Applicants of Breakthrough Patents

To characterize the institutions behind breakthrough patents, we use the set of "historically significant patents" compiled by Kelly et al. (2021) from online sources. Our sample contains 54 such patents. For each patent, we identify the applicant institution using information from official firm websites (when the patent led to a firm) and, when needed, Wikipedia. The resulting composition is summarized below.

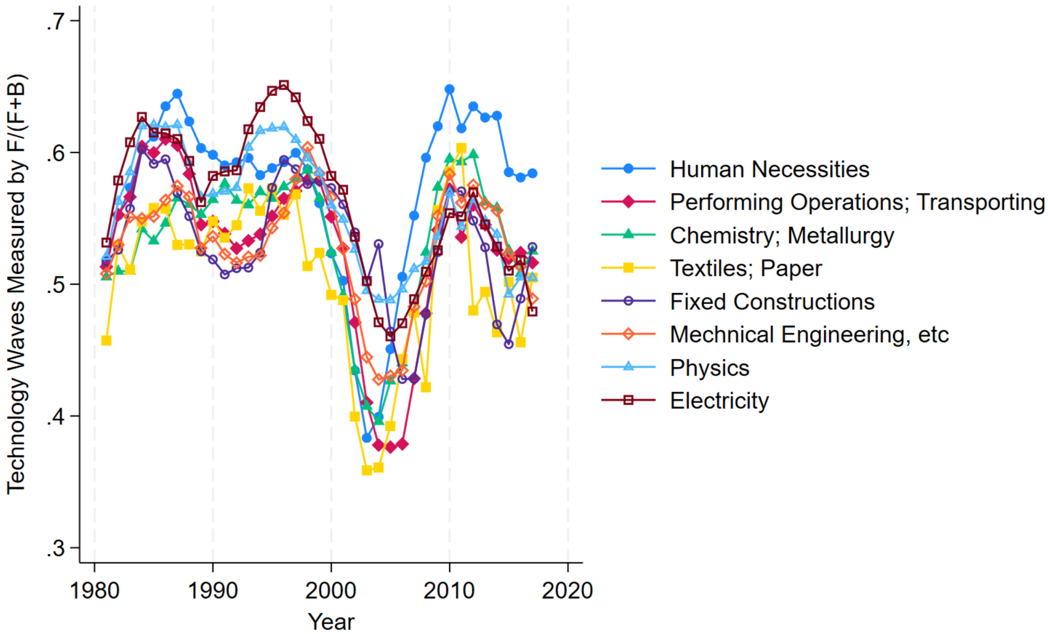

Figure 12: Technological Waves by Technological Fields

Notes: This figure shows the technological waves by the nine technological fields between 1981 and 2017. The nine fields are defined by the 1-digit IPC code. The technological waves are measured by the Novelty index as defined by Equation 1 in the paper.
Sources: USPTO patent and citation data.

B.4 Novelty Index in Europe

To calculate the Novelty Index for European countries with intensive patenting activities, we use data from PATSTAT (Patent Statistical Database), a comprehensive global dataset maintained by the European Patent Office (EPO). PATSTAT provides detailed bibliographic data on patents from various patent offices worldwide, with a particular focus on those filed through the EPO. We restrict our sample to patents with inventors based in European countries. The six countries with the highest number of patent issuances between 1982 and 2016 are Germany, France, the United Kingdom, Italy, Switzerland, and the Netherlands. Using the definition of the Novelty Index as outlined in Equation 1, we calculate the technological waves in these six countries and present them in Figure 14. Across all six countries, we observe an overall declining trend in technological novelty, with common peaks in the mid-1980s and early 2010s. Italy also experienced a distinct peak in the mid-1990s. In general, the technological trends in these European countries with the highest patenting activity mirror those observed in the U.S.[30]

[30]Similar results can be obtained by calculating the Novelty Index using Google Patents data, as cleaned by Ayerst et al. (2023). These results are available upon request.

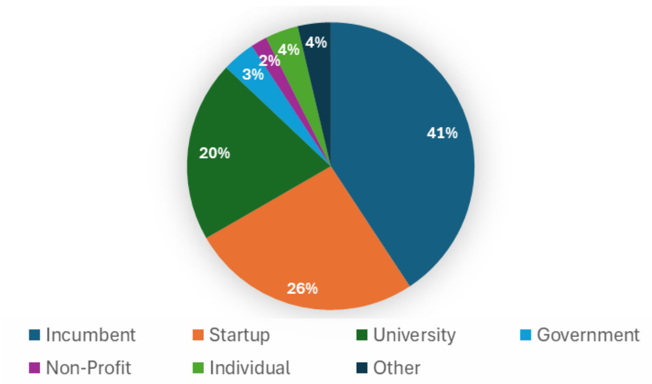

Figure 13: Composition of Applicants of Breakthrough Patents

Notes: The pie chart illustrates the share of applicants from each institutional group for the 54 breakthrough patents. Incumbents are defined as private firms that applied for a patent at least three years after their founding. Startups are private firms that applied within their first three years or before. Others include cases where the applicants are multiple institutions of different types.
Sources: The "historically significant patents" complied by Kelly et al. (2021).

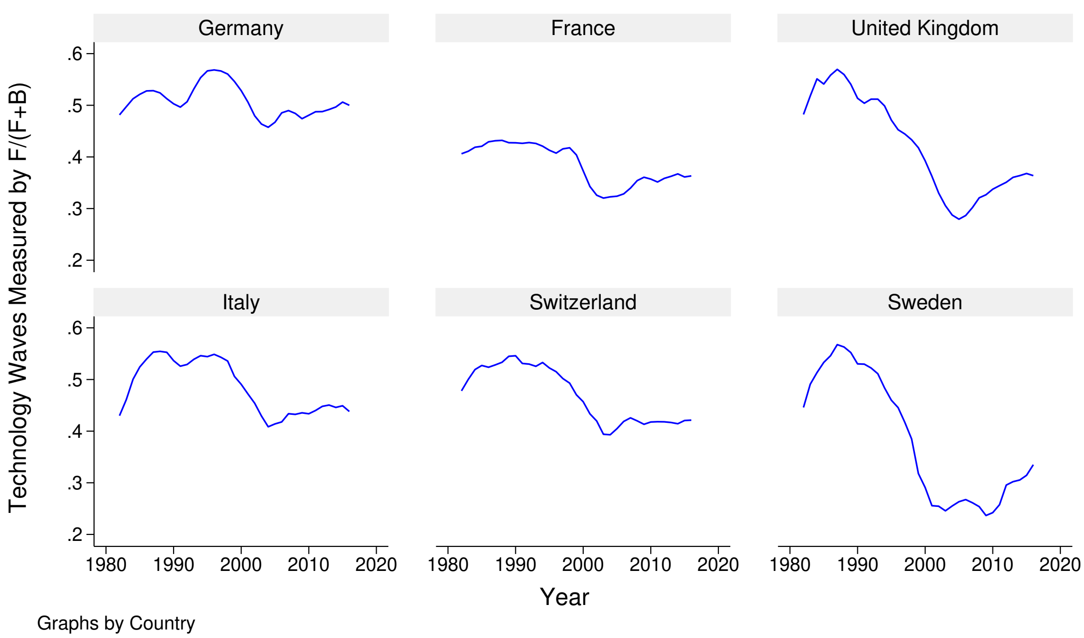

Figure 14: Technological Waves in European Countries

Notes: This figure shows the technological waves in six European countries with the highest number of patent issuances between 1982 and 2016. The technological waves are measured by the Novelty index as defined by Equation 1 in the paper.
Sources: PATSTAT (Patent Statistical Database).

B.5 Relationship with the Technological Waves

Table 11 exhibits the time trend of the technological novelty waves, the market concentration measured by the HHI, and the New-to-Incumbent Ratio of idea allocations (Panel A). It also displays the cross correlation of the two latter time series with the technological waves at different year gaps (Panel B). The time trend is derived by fitting a linear trend to the focal time series and taking its slope. The cross correlations are obtained by calculating the correlation coefficients of the detrended time series when the year gaps of the two series are respectively -3, -2, -1, 0, 1, 2, 3. The detrending process subtracts the linear trend from the original time series. The cross correlations capture not only the co-movement of the different time series, but also the relative timing of their movements. The first row of each panel shows the statistics for the whole sample; the subsequent rows are statistics by major industries according to the Standard Industrial Classification (SIC) code or technological fields according to the International Patent Classification (IPC).

Table 11: Time Trend and Cross Correlation

	Time Trend		Detrended Cross Correlation						
			Panel A. HHI						
	Tech Wave	HHI	$k = -3$	$k = -2$	$k = -1$	$k = 0$	$k = 1$	$k = 2$	$k = 3$
All	-0.002	0.007	-0.625	-0.720	-0.775	-0.736	-0.588	-0.366	-0.097
			Panel B. New-to-Incumbent Ratio						
	Tech Wave	N-to-I Ratio	$k = -3$	$k = -2$	$k = -1$	$k = 0$	$k = 1$	$k = 2$	$k = 3$
All	-0.002	-0.001	0.107	0.314	0.504	0.612	0.536	0.317	-0.001
Human Necessities	-0.001	-0.001	0.539	0.557	0.514	0.402	0.199	-0.063	-0.361
Performing Operations	-0.003	-0.001	0.117	0.210	0.283	0.301	0.191	0.017	-0.201
Chemistry; Metallurgy	-0.001	0.001	0.239	0.231	0.211	0.163	0.049	-0.128	-0.349
Textiles; Paper	-0.002	-0.001	0.458	0.451	0.462	0.417	0.371	0.270	0.189
Fixed Construction	-0.002	0	0.211	0.331	0.397	0.386	0.297	0.215	0.161
Mechanical Engineering	-0.001	0	-0.456	-0.505	-0.467	-0.358	-0.208	-0.061	0.059
Physics	-0.003	-0.001	-0.156	0.016	0.212	0.343	0.284	0.068	-0.207
Electricity	-0.003	-0.002	0.373	0.465	0.540	0.580	0.552	0.367	0.070

Notes: This table shows the trends of the technological waves, HHI, New-to-Incumbent ratio and the detrended cross correlations among them. The trend is derived by running linear regressions of the focal time series on year and taking the coefficient; the cross correlations are derived by computing the correlation coefficients at different year gaps of the detrended time series.

B.6 First-Stage Results and Instrument Relevance

Table 12 reports the first-stage results for the 2SLS regressions. Columns (1)–(2) relate the contemporaneous Novelty Index to its shift-share instrument in the HHI regression at the 6-digit NAICS-by-year level. Columns (3)–(4) and (5)–(6) report the corresponding first-stage relationships for the one-year and two-year lagged Novelty Index, respectively. Columns (7)–(8) show the first-stage relationship between the contemporaneous Novelty Index and its instrument in the patent-level regression with the new-firm indicator as the outcome. The last row reports the first-stage F-statistics assessing instrument relevance. Across all specifications, these statistics indicate a strong and statistically significant relationship between the instrument and the endogenous regressor.

Table 12: Relationship between HHI and Novelty Index at the 6-digit NAICS code

	HHI						New Firm Indicator	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Novelty (NAICS)	0.571***	0.736***						
	(0.157)	(0.0874)						
Novelty (NAICS, 1-year lag)			0.345*	0.719***				
			(0.183)	(0.0849)				
Novelty (NAICS, 2-year lag)					0.347*	0.722***		
					(0.184)	(0.0864)		
Novelty (IPC)							0.600***	0.707***
							(0.121)	(0.0999)
Observations	37500	37500	37500	37500	36500	36500	4032000	4032000
Industry/IPC FE	NO	YES	NO	YES	NO	YES	NO	YES
Year FE	NO	YES	NO	YES	NO	YES	NO	YES
F excl. instruments	13.21	70.85	3.54	71.86	3.56	69.92	24.53	49.99

Notes: Standard errors are clustered at the 6-digit NAICS-by-year level for columns (1)–(6), and at the 4-digit IPC-by-year level for columns (7)–(8). The inclusion of year and industry fixed effects, and the F-statistics assessing instrument relevance are shown in the last three rows. To comply with Census Bureau disclosure requirements, the number of observations is rounded to the nearest hundred. *** Significant at the 1 percent level; ** significant at the 5 percent level; * significant at the 10 percent level.

B.7 Alternative Measures of Market Concentration

The main text of this paper uses the Herfindahl-Hirschman Index to measure market concentration. It captures the whole distribution of firm sales in the economy, but the limitation is that it is based on only publicly listed firm. An alternative measure of market concentration is the share of sales by the top firms. This paper adopts the cleaned data series by Kwon, Ma and Zimmermann (2023) to calculate respectively the three-year moving average of the receipt share of the top 0.1% and 1% firms. The top shares are generated by the IRS data, which covers a more comprehensive set of firms. So, it can be used as a complement to the HHI measure in the paper. As displayed in Figure 15, the top shares exhibit increasing trends in general but with fluctuations. The peaks and troughs of the fluctuations appear nearly simultaneously with the HHI measured in this paper, showing the robustness of the market concentration patterns shown in the paper.

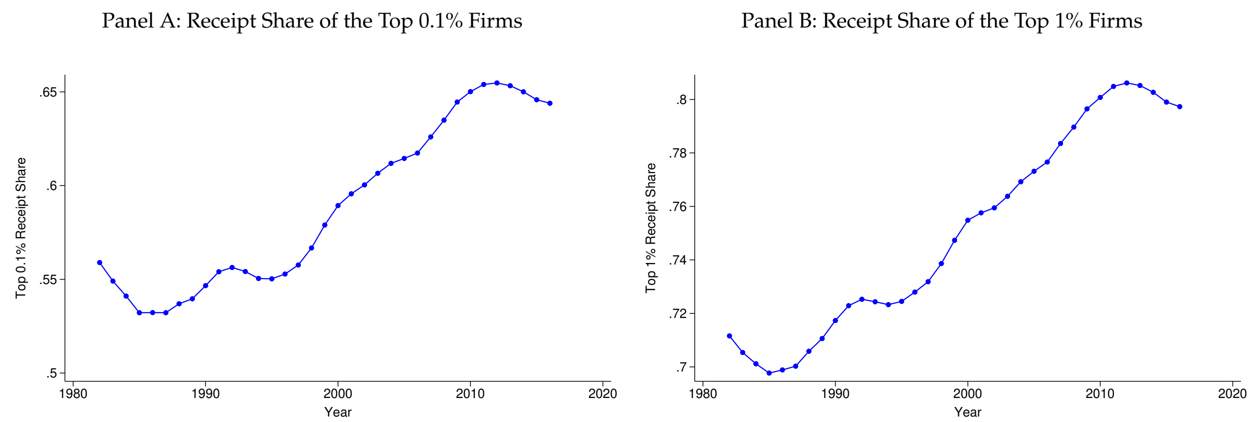

Figure 15: Receipt Shares by Top firms

Notes: This figure shows the three-year moving average of the receipt share of the top 0.1% (Panel A) and 1% firms (Panel B). The receipt shares are from the cleaned data series by Kwon, Ma and Zimmermann (2023), which is posted on `https://businessconcentration.com/`. The data source is the Statistics of Income (SOI) and the associated Corporation Source Book published annually by the IRS. Their statistics cover the whole population of US corporations.
Sources: `https://businessconcentration.com/`.

B.8 IPC-Level Relationship between Tech Waves and New-to-Incumbent Ratio

To assess the robustness of the relationship between idea allocation and technological waves, this paper compares the two trends by patent technological fields, categorized by the first digit of the patent IPC code. The IPC-level "Novelty" Index and "New-to-Incumbent Ratio" are computed using the same methodology as described in equations 1 and 5, with patent sets segregated according to their respective technology classes. Figure 16 illustrates that a positive correlation between idea allocation and technological waves is consistently observed across most technology classes. When a specific technology class experiences breakthroughs, there is an increase in the flow of ideas toward new startups. The contemporaneous correlation coefficients between the two curves are, respectively, 0.40 for Human Necessities, 0.30 for Performing Operations, 0.16 for Chemistry, 0.42 for Textiles, 0.39 for Fixed Constructions, -0.47 for Mechanical Engineering, 0.34 for Physics, and 0.58 for Electricity. The cross correlations when $k \in \{-3, -2, -1, 0, 1, 2, 3\}$ for each technological field are shown in Table 11 in Appendix B.5.

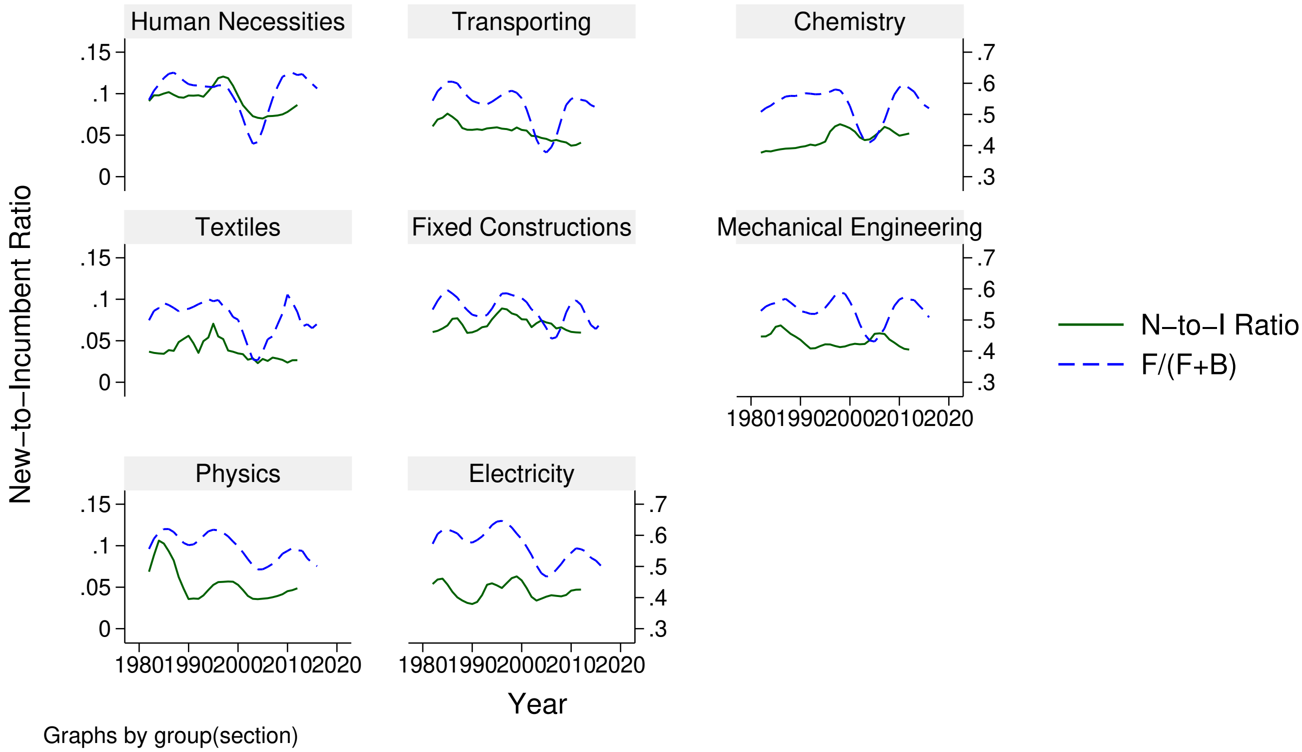

Figure 16: Technological Waves and Idea Allocation by Patent Technology Class

Notes: This figure shows the technological waves and the idea allocation between new and incumbent firms by patent technology class. The blue dashed curve, based on the methodology defined in this paper, calculates the relative ratio of forward citations to the sum of forward and backward citations. The green solid curve displays the "New-to-Incumbent Ratio" defined in the paper, capture where new ideas contribute their value. The two curves have different y-axises, shown respectively on the left and right.
Sources: Longitudinal Business Database (LBD) and USPTO patent and citation data.

C Model and Proof

C.1 Starting up New Businesses

The partner's problem has the same form as the incumbent firm's, with $\tilde{q} = 0$ and the average innovation value being z_0 instead of $x_0(z_0, \tilde{q})$ (Equation 18).

$$\begin{aligned} &\max_a \left(\left(1-a\right)\lambda_0 e_I z_0 - T_I\right)\nu Q(t) \\ &\text{st}\, e_I = \arg\max\left\{u\left(c_I\left(a,0,T_I\right),e_I\right)\right\} \\ &u\left(c_I\left(a,T_I,0,Q(t)\right),e_I\right) \geq \bar{u}\left(z_0\right) \\ &\left(\left(1-a\right)\lambda_0 e_I z_0 - T_I\right)\nu Q(t) \geq 0 \end{aligned} \tag{41}$$

The partners are assumed to get zero profit due to competition.

The inventor decides her effort level by maximizing her utility, which yields:

$$e_I = \lambda_0 a z_0 - A a^2 \lambda_0 \mathbb{E}\left(z\left(z_0\right)^2\right) \tag{42}$$

The firm's problem in Equation 41 becomes

$$\begin{aligned} &\max_a \lambda_0 e_I z_0 - A a^2\left(\lambda_0 e_I \mathbb{E}\left(z\left(z_0\right)^2\right)\right) - \frac{1}{2}e_I^2 \\ &\text{st}\, e_I = \lambda_0 a z_0 - A a^2 \lambda_0 \mathbb{E}\left(\tilde{z}\left(z_0\right)^2\right) \end{aligned} \tag{43}$$

It gives the highest utility an inventor can obtain when working in a startup.

C.2 Production: Profit Maximization

The production sector features two types of firms: a representative final goods producer and intermediate goods producers. The final good producer assembles intermediate goods, denoted by j within the range $[0, N_F(t)]$, to produce final goods. The final goods producer chooses $\{y(t)_j\}_j$ to maximize its profit using the technology described in Section 3.2. The subscript t is omitted in this section whenever it does not cause a confusion. The final goods producer's problem can be written as:

$$\max_{\{y_j\}} \frac{1}{1-\beta}\int_0^{N_F} q_j^{\beta} y_j^{1-\beta} dj - \int_0^{N_F} y_j p_j dj. \tag{44}$$

The first-order condition

$$p_j = q_j^{\beta} y_j^{-\beta}$$

yields the demand function for goods produced by intermediate firms.

The intermediate goods are produced by their corresponding firm $j \in [0, N_F]$ using only labor $y_j = Ql_j$. Intermediate good producers engage in monopolistic competition, optimizing their profit by choosing l_j, p_j, y_j, given the wage level w:

$$\begin{aligned} &\max_{l_j,p_j,y_j} y_j p_j - w l_j. \\ &s.t.\ y_j = Q l_j \\ &\quad\ p_j = q_j^{\beta} y_j^{-\beta} \end{aligned} \tag{45}$$

The FOC yields:

$$y_j(t) = q_j(t)\left(\frac{Q(t)\,(1-\beta)}{w(t)}\right)^{\frac{1}{\beta}}, l_j(t) = \frac{y_j(t)}{Q(t)}, p_j(t) = \frac{w(t)}{Q(t)\,(1-\beta)}.$$

The labor market clears, which derives that

$$\frac{\int_0^{N_F} q_j \left(\frac{Q(1-\beta)}{w}\right)^{\frac{1}{\beta}} dj}{Q} = 1$$

Therefore, the wage is $w = (1-\beta)Q$.

C.3 Firm-led Innovation New

For firm-led innovation, plug the linear value function (Equation 33) into the firm choice problem specified in Equation 23

$$\begin{aligned} \Omega_F(z_0, q, t) &= \max_{e_F, T_F} \lambda_0 e_F \int \kappa x_F(z_0)\nu(t)N_F(t)q dx - \frac{1}{1+\delta}e_F^{\delta+1}\nu(t)N_F(t)q - T_F\nu(t)N_F(t)q \\ &= \max_{e_F, T_F}\left(\lambda_0 e_F \kappa x_{F0} - \frac{1}{1+\delta}e_F^{\delta+1} - T_F\right)\nu(t)N_F(t)q \end{aligned} \tag{46}$$

FOC yields:

$$e_F = (\lambda_0 \kappa x_{F0})^{\frac{1}{\delta}},$$

which is invariant to time and firm size. Equation 24 can be written as:

$$\lambda_0 e_F \kappa x_{F0}\nu(t) - \frac{1}{1+\delta}e_F^{\delta+1}\nu(t) = T_F\nu(t)$$

which yields:

$$\begin{aligned} T_F =& \lambda_0 e_F \kappa x_{F0} - \frac{1}{1+\delta} e_F^{\delta+1} \\ =& \frac{\delta}{1+\delta} (\lambda_0 \kappa x_{F0})^{\frac{1+\delta}{\delta}} \end{aligned} \tag{47}$$

which is also invariant to time and firm size. Inventor's utility in this case is:

$$U_I(T_F(z_0, q), 0) = T_F V(t|N_F(t)q) = \frac{\delta}{1+\delta} (\lambda_0 \kappa x_{F0})^{\frac{1+\delta}{\delta}} \nu(t) N_F(t) q,$$

which is linear in $\nu(t)N_F(t)$.

C.4 Inventor-led Innovation

In the inventor-led case, the consumption is

$$\begin{aligned} & c_I(a, T_I, q, z_0, Q(t))dt \\ & = T_I V(t|Q(t))dt + adV(j, t|q, z_0), \\ & = T_I \nu(t) Q(t) dt \begin{cases} +a\tau\nu(t)qdt, & pr = 1 - \tau dt - \lambda_0 e_I dt \\ -a(1 - \tau dt - rdt)\nu(t)q, & pr = \tau dt \\ +a((1 - rdt)x(z_0, \tilde{q})Q(t) + \tau q dt)\nu(t), & pr = \lambda_0 e_I dt \end{cases} \end{aligned} \tag{48}$$

which is linear in $\nu(t)$. The expected consumption is

$$\mathbb{E}\left(c_I(a, T_I, q, z_0, Q(t))\right) = (a\lambda_0 e_I x_0\left(z_0, \tilde{q}\right) + T_I)\nu(t)Q(t),$$

and the associated variance is

$$\begin{aligned} \mathrm{Var}\left(c_I\left(a, T_I, \tilde{q}, z_0, Q(t)\right)dt\right) =& \mathbb{E}((c_I dt)^2) - \mathbb{E}(c_I dt)^2 \\ =& \mathbb{E}\left(a^2 \nu(t)^2 Q(t)^2 \left(\tilde{q}^2 \tau dt + x^2(z_0, \tilde{q})\lambda_0 e_I dt\right)\right) \\ & - \mathbb{E}\left(a\nu x_0(z_0, \tilde{q})Q(t)\lambda_0 e_I dt\right)^2 \\ =& a^2 \nu(t)^2 Q(t)^2 \left(\tilde{q}^2 \tau + \mathbb{E}(x^2(z_0, \tilde{q}))\lambda_0 e_I\right) dt \end{aligned}$$

The variance comes from two sources: non-innovation-related firm value and the R&D process. Both terms increases in firm size $\tilde{q}$, but the former one increases in a faster speed, implying that in larger firms, shocks unrelated to R&D are stronger. Hence, larger firms are subject to larger incentive problems and the equity held by the inventor provides a weaker incentive for R&D efforts. Upon reviewing all available contracts, an inventor determines her preferred firm $\tilde{q}$.

The instant utility is:

$$U_I(c_I(t), e_I(t)) = \Big((a\lambda_0 e_I x_0 \left(z_0, \tilde{q}\right) + T_I) - Aa^2 \left(\tilde{q}^2 \tau + \mathbb{E}(x^2(z_0, \tilde{q}))\lambda_0 e_I \right) - R(e_I) \Big) \nu(t) Q(t) \tag{49}$$

The inventor's problem is to maximize her utility:

$$\max_{e_I(t)} U_I(c_I(a, T_I, q, Q(t)), e_I(t)).$$

FOC yields:

$$e_I(t) = \left(a\lambda_0 x_0 \left(z_0, \tilde{q}\right) - Aa^2 \mathbb{E}(x^2(z_0, \tilde{q}))\lambda_0 \right)^{\frac{1}{\delta}} \tag{50}$$

which is invariant to $\nu(t)$.

The wage rate T_I is determined by the Betrand competition:

$$(1-a)\,\lambda_0 e_I \left[\int V(j, t|q + x(z_0, \tilde{q}(t))Q(t))dx - V(j, t|q) \right] - T_I V(t|Q(t)) = 0$$
$$T_I = (1-a)\lambda_0 e_I x_0(z_0, \tilde{q}(t))$$

which is invariant to $\nu(t)$. Therefore, the instant utility is linear in $\nu(t)Q(t)$:

$$\begin{aligned} U_I(c_I(t), e_I(t)) &= \Big((a\lambda_0 e_I x_0 \left(z_0, \tilde{q}\right) + T_I) - Aa^2 \left(\tilde{q}^2 \tau + \mathbb{E}(x^2(z_0, \tilde{q}))\lambda_0 e_I \right) - R(e_I) \Big) \nu(t) Q(t) \\ &= \left(\lambda_0 \left(\left(1 - \frac{a}{\delta+1} \right) x_0 \left(z_0, \tilde{q}\right) - \frac{\delta}{\delta+1} Aa^2 \mathbb{E}(x^2(z_0, \tilde{q})) \right) e_I - Aa^2 \tilde{q}^2 \tau \right) \nu(t) Q(t) \\ &= u_I(z_0, \tilde{q}) \nu(t) Q(t) \end{aligned} \tag{51}$$

where $u_I(z_0, \tilde{q}) = \left(\lambda_0 \left(\left(1 - \frac{a}{\delta+1}\right) x_0 \left(z_0, \tilde{q}\right) - \frac{\delta}{\delta+1} Aa^2 \mathbb{E}(x^2(z_0, \tilde{q})) \right) e_I - Aa^2 \tilde{q}^2 \tau \right)$, invariant to $\nu(t)$.

Combining the results from the previous section, the inventor's choice of firm, as specified in Equation (25), can be written as:

$$
\begin{aligned}
u\left(z_0, t\right) = & \max\{hU_I\left(c_I(z_0, q^*(z_0,t),t), e_I(z_0, q^*(z_0,t),t\right) + (1-h)\int_q U_I\left(T_F(z_0,q),0\right) f_t\left(\tilde{q}\right) d\tilde{q}, \\
& h_s U_I\left(c_I(z_0,0,t), e_I(z_0,0,t)\right) + (1-h_s)\int_q U_I(T_F(z_0,q),0) f_t\left(\tilde{q}\right) d\tilde{q}\} \\
= & \max\{hu_I\left(z_0,\tilde{q}^*\right)\nu(t)Q(t) + (1-h)\nu(t)Q(t)\int_q \frac{\delta}{1+\delta}(\lambda_0\kappa x_{F0})^{\frac{1+\delta}{\delta}} f_t\left(\tilde{q}\right) d\tilde{q}, \\
& h_s u_I\left(z_0,\tilde{q}^*\right)\nu(t)Q(t) + (1-h_s)\nu(t)Q(t)\int_q \frac{\delta}{1+\delta}(\lambda_0\kappa x_{F0})^{\frac{1+\delta}{\delta}} f_t\left(\tilde{q}\right) d\tilde{q}\} \\
= & \max\{hu_I\left(z_0,\tilde{q}^*\right)\nu(t)Q(t) + (1-h)\nu(t)Q(t)\frac{\delta}{1+\delta}(\lambda_0\kappa x_{F0})^{\frac{1+\delta}{\delta}}, \\
& h_s u_I\left(z_0,\tilde{q}^*\right)\nu(t)Q(t) + (1-h_s)\nu(t)Q(t)\frac{\delta}{1+\delta}(\lambda_0\kappa x_{F0})^{\frac{1+\delta}{\delta}}\}
\end{aligned}
\tag{52}
$$

Since $\nu(t)Q(t) > 0$, the decision rule is the same as:

$$
\begin{aligned}
u\left(z_0, t\right) = & \max\{hu_I\left(z_0,\tilde{q}^*\right) + (1-h)\frac{\delta}{1+\delta}(\lambda_0\kappa x_{F0})^{\frac{1+\delta}{\delta}}, \\
& h_s u_I\left(z_0,\tilde{q}^*\right) + (1-h_s)\frac{\delta}{1+\delta}(\lambda_0\kappa x_{F0})^{\frac{1+\delta}{\delta}}\}.
\end{aligned}
\tag{53}
$$

Thus, the decision rule $\tilde{q}^*(z_0)$ does not change over time.

C.5 Proof of Proposition 1 and Lemma 1

Proof. Production is linear in aggregate quality $Q(t)$:

$$Y(t) = \frac{Q(t)}{1-\beta} \quad \Rightarrow \quad \frac{\dot{Y}(t)}{Y(t)} = \frac{\dot{Q}(t)}{Q(t)} \equiv g.$$

Since $C_I(t)/Y(t)$ and $R_F(t)/Y(t)$ are linear functions of $\nu(t)$ and $Y(t) \propto Q(t)$, define

$$s_H(t) \equiv \frac{C_H(t)}{Y(t)}, \qquad s_I\nu(t) \equiv \frac{C_I(t)}{Y(t)}, \qquad s_F\nu(t) \equiv \frac{R_F(t)}{Y(t)},$$

so the resource constraint gives $s_H(t) = 1 - (s_I - s_F)\nu(t) \equiv S_H(\nu(t))$. Household optimization implies the Euler equation

$$\frac{\dot{C}_H(t)}{C_H(t)} = r(t) - \rho.$$

Because $C_H(t) = Y(t)s_H(t)$, we have

$$\frac{\dot{C}_H(t)}{C_H(t)} = \frac{\dot{Y}(t)}{Y(t)} + \frac{\dot{s}_H(t)}{s_H(t)} \quad \Rightarrow \quad r(t) = \rho + g + \frac{d}{dt}\ln s_H(t). \tag{54}$$

Plugging $V(j, t \mid q) = \nu(t)\, q$ into the HJB gives

$$\dot{\nu}(t) = (r(t) + \tau)\nu(t) - \beta.$$

Substitute $r(t)$ from (54) to obtain

$$\dot{\nu}(t)\Big(1 - \nu(t)\frac{S'_H(\nu(t))}{S_H(\nu(t))}\Big) = \big(\rho + \tau + g\big)\nu(t) - \beta, \tag{55}$$

Define

$$D(\nu) \equiv 1 - \nu\,\frac{S'_H(\nu)}{S_H(\nu)}, \qquad \Phi(\nu) \equiv \big(\rho + \tau + g\big)\nu - \beta. \tag{56}$$

According to (35) and (36), $s_I \geq 0$ and $s_F \geq 0$ (innovation becomes more attractive as ν rises). Hence

$$S'_H(\nu) = -(s_I + s_F) \leq 0, \qquad S_H(\nu) \in (0, 1) \quad \Rightarrow \quad D(\nu) \geq 1.$$

Equation (55) is therefore the autonomous scalar ODE

$$\dot{\nu}(t) = \frac{\Phi(\nu(t))}{D(\nu(t))}. \tag{57}$$

Because $g > 0$ and $\rho + \tau > 0$, so $\Phi'(\nu) = \rho + \tau + g > 0$. Since $\Phi(0) = -\beta < 0$, $\lim_{\nu\to\infty}\Phi(\nu) = +\infty$, hence there exists a unique $\nu^\star > 0$ with $\Phi(\nu^\star) = 0$.

Next, solving $\dot{\nu} - (r + \tau)\nu = -\beta$ by an integrating factor and imposing the transversality condition derive the present-value representation

$$\nu(t) = \int_t^\infty \beta \exp\Big(-\int_t^s \big(r(u) + \tau\big)\, du\Big)\, ds.$$

Since $r(u) = \rho + \dot{C}_H/C_H \geq \rho$, we have $r(u) + \tau \geq \rho + \tau$ and thus

$$0 < \nu(t) \leq \int_t^\infty \beta\, e^{-(\rho+\tau)(s-t)}\, ds = \frac{\beta}{\rho + \tau} \quad \text{for all } t. \tag{58}$$

From (57), $\text{sign}(\dot{\nu}) = \text{sign}(\Phi(\nu))$. If $\nu(t_0) > \nu^\star$ at some t_0, then $\Phi(\nu(t)) > 0$ thereafter and $\nu(\cdot)$ is strictly increasing, which contradicts the upper bound in (58). If $\nu(t_0) < \nu^\star$, then $\Phi(\nu(t)) < 0$ thereafter and $\nu(\cdot)$ is strictly decreasing and bounded below by 0, so it must converge to a limit $L \in [0, \nu(t_0))$ with $\Phi(L) \leq \Phi(\nu(t_0)) < 0$, which implies $\dot{\nu} < 0$ near L, a contradiction. Hence every admissible equilibrium path satisfies

$$\nu(t) \equiv \nu^\star \qquad \text{and} \qquad g(t) \equiv g^\star.$$

With ν constant, e_F and the inventor-side policies are time-invariant in Q-normalized units, so s_I and s_F are constant and $s_H = 1 - (s_I - s_F)\nu^*$ is constant. $\dot{s}_H = 0$ in (54) implies

$$r(t) = \rho + g^\star.$$

Because $C_H(t) = Y(t)s_H(t)$, $C_I(t) = Y(t)s_I(t)$, $R_F(t) = Y(t)s_F(t)$, and $Y(t) \propto Q(t)$, all aggregates grow at the common constant rate $g^\star$. The fixed point $\nu^\star$ is unique, so the competitive equilibrium is unique and coincides with the balanced-growth path. Besides, there is no transition dynamics. □

C.6 Closed-Form Model Solution

The firm's problem in Equation 18 can be rewritten as:

$$\begin{aligned} &\max_a \left(\lambda_0 e_I x_0(z_0, \tilde{q}) - Aa^2 \left(\tau \tilde{q}^2 + \lambda_0 x_0 \left(z_0, \tilde{q}\right)^2 \right) - \frac{1}{2} e_I^2 \right) \nu Q \\ &\text{st}\, e_I = \lambda_0 a x_0 \left(z_0, \tilde{q}\right) \end{aligned} \tag{59}$$

Putting the expression of e_I into the maximization problem and taking the FOC with regard to the equity share, a, derives,

$$\begin{aligned} a^* &= \frac{\lambda_0^2 x_0 \left(z_0, \tilde{q}\right)^2 \nu^2}{\lambda_0^2 x_0 \left(z_0, \tilde{q}\right)^2 \nu^2 + 2A \left(\tau \tilde{q}^2 \nu^2 + \lambda_0 x_0 \left(z_0, \tilde{q}\right)^2 \nu^2 \right)} \\ &= \frac{1}{1 + 2\frac{A}{\lambda_0} \left(\frac{\tau \tilde{q}^2}{\lambda_0 x_0 (z_0, \tilde{q})^2} + 1 \right)} \end{aligned} \tag{60}$$

Upon reviewing all contracts, an inventor with idea quality z_0 chooses which firm $\tilde{q}$

to work for by maximizing her instant utility:

$$
\begin{aligned}
\max_{\tilde{q}} u\left(c_I, e_I\right) = & \mathbb{E}\left(c_I\right) - \frac{A}{\nu Q(t)} \mathrm{Var}\left(c_I\right) - R\left(e_I\right) \nu Q(t) \\
\mathrm{st}\, a = & \, a^*\left(\tilde{q}\right) \\
T_I = & \, T_I^*\left(\tilde{q}\right)
\end{aligned}
\tag{61}
$$

Putting the expression of the optimal equity level, $a^*(\tilde{q})$, and $T_I^*(\tilde{q}) = (1-a^*)\,\lambda_0 e_I x_0\left(z_0, \tilde{q}\right)$ into the maximization problem and solving the first-order condition,

$$
\frac{\partial x_0\left(z_0, \tilde{q}\right)}{\partial \tilde{q}} = \frac{2A\tau}{4A\tau \frac{\tilde{q}}{x_0(z_0,\tilde{q})} + \frac{2A\lambda_0 + \lambda_0^2}{\tilde{q}/x_0(z_0,\tilde{q})}},
\tag{62}
$$

derive the optimal firm size,

$$
\tilde{q}^* = \left(\frac{\left(2A\lambda_0 + \lambda_0^2\right)\left(\gamma\left(z_0\right)\right)^2 b}{2A\tau q_0^{2b}\left(1-2b\right)} \right)^{\frac{1}{2-2b}}.
$$

The left-hand side and right-hand side of the first-order condition when $b < 0.5$ and $\eta = -1$ are shown in Figure 17.

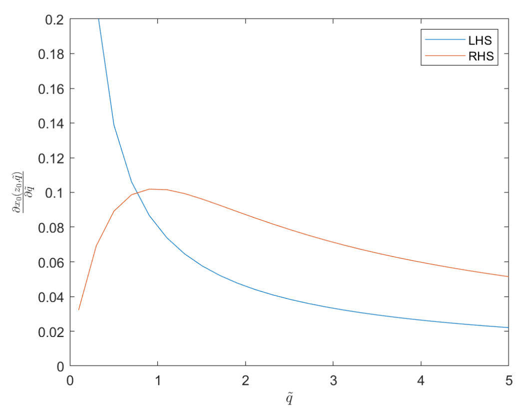

Figure 17: FOC Condition

Notes: This figure shows respectively the left-hand side (the blue curve) and the right-hand side (the red curve) of Equation 62. When $b < 0.5$, there exists a unique intersection.

C.7 Proof for Proposition 2

Proof.

$$\frac{\partial \tilde{q}^*}{\partial z_0} = \frac{\tilde{q}^*}{\gamma(z_0)(1-b)} \frac{\partial(\gamma(z_0))}{\partial z_0}$$
$$= \frac{\tilde{q}^*}{\gamma(z_0)(1-b)} \left(z_0^\eta + B^\eta\right)^{\frac{1}{\eta}-1} z_0^{\eta-1}$$

where

$$\tilde{q}^* = \left(\frac{\left(2A\lambda_0 + \lambda_0^2\right)(\gamma(z_0))^2 b}{2A\tau q_0^{2b}(1-2b)}\right)^{\frac{1}{2-2b}}.$$

When $b < 0.5$, the optimal size increases in the inventor's idea quality, z_0. □

C.8 Proof for Proposition 3

Proof. The highest expected utility $u_N(z_0)$ an inventor z_0 can obtain when working in a new firm is

$$u_N(z_0) = u\left(c_I(z_0, \tilde{q}), e_I(z_0, \tilde{q})\right)$$
$$= \frac{1}{2}\frac{\lambda_0}{\lambda_0 + 2A}\lambda_0^2 z_0^2 \nu Q,$$

which is quadratic in the idea quality $z0$. However, the highest expected utility $u_I(z_0)$ an inventor z_0 can obtain when working in an incumbent firm depends on B:

$$u_I(z_0) = u\left(c_I(z_0, \tilde{q}^*), e_I(z_0, \tilde{q}^*)\right)$$
$$= \frac{1}{2}\lambda_0^2 \frac{\left(\frac{2A\lambda_0 + \lambda_0^2}{2(1-2b)A\tau} b\right)^{\frac{b}{1-b}}}{\left(1 + \frac{2A(1-b)+\lambda_0 b}{\lambda_0(1-2b)}\right) q_0^{\frac{2b}{1-b}}} \gamma(z_0)^{\frac{2}{1-b}} \nu Q$$
$$= \frac{1}{2}\lambda_0^2 \hat{q}_0 \gamma(z_0)^{\frac{2}{1-b}} \nu Q,$$

where $\hat{q}_0$ is a parameter ($\hat{q}_0 = \frac{\left(\frac{2A\lambda_0 + \lambda_0^2}{2(1-2b)A\tau} b\right)^{\frac{b}{1-b}}}{\left(1 + \frac{2A(1-b)+\lambda_0 b}{\lambda_0(1-2b)}\right) q_0^{\frac{2b}{1-b}}}$). The derivative is proportional to $\gamma(z_0)^{\frac{1+b}{1-b}} \left(\frac{\gamma(z_0)}{z_0}\right)^{1-\eta}$. When $\eta < 0$, the change of $u_I(z_0)$ in z_0 is much smaller than the change of $u_N(z_0)$.

An inventor decides whether to join a startup by comparing the incumbent-startup utility $u_I(z_0) - u_N(z_0)$ and zero. The utility gap decreases in z_0, meaning that a larger share of inventors would choose incumbent firms when the entire distribution of z_0 shifts to the right. $\square$

C.9 Proof of Proposition 4

Proof. An inventor decides whether to launch a startup by comparing the highest utility offered by incumbents $u_I(z_0)$ and startups $u_N(z_0)$. When $\frac{\gamma(z_0)^{\frac{1}{1-b}}}{z_0} < \left(\frac{\lambda_0}{\hat{q}_0(\lambda_0+2A)}\right)^{\frac{1}{2}}$, $u_I(z_0) < u_N(z_0)$, inventor chooses to initiate a startup. When $\eta < 0$, if $b < \frac{\min\left(z_0^{-\eta}\right)}{\min\left(z_0^{-\eta}\right)+\max(B^{-\eta})}$, $b - \frac{B^\eta}{z_0^\eta + B^\eta} < 0$ always holds:

$$\frac{\partial\left(\gamma(z_0)^{\frac{1}{1-b}} z_0^{-1}\right)}{\partial z_0} = \frac{\gamma(z_0)^{\frac{1}{1-b}}}{z_0^2}\frac{1}{1-b}\left(b - \frac{B^\eta}{z_0^\eta + B^\eta}\right) < 0,$$

since $\gamma(z_0) = \left(z_0^\eta + B^\eta\right)^{\frac{1}{\eta}}$. $\gamma(z_0)^{\frac{1}{1-b}} z_0^{-1}$ monotonically decreases in z_0, when holding B constant. It implies there exists a cutoff $\bar{z}_0(B)$, when $z_0 > \bar{z}_0(B)$,

$$\frac{\gamma(z_0)^{\frac{1}{1-b}}}{z_0} < \left(\frac{\lambda_0}{\hat{q}_0(\lambda_0 + 2A)}\right)^{\frac{1}{2}}$$

always holds, and hence $u_I(z_0) < u_N(z_0)$, inventors opts for new businesses instead of incumbent firms. $\square$

C.10 Full Model

The firm's problem in Equation 18 becomes

$$\begin{aligned}&\max_a \left(\lambda_0 e_I x_0(z_0,\tilde{q}) - Aa^2\left(\tau\tilde{q}^2 + \lambda_0 e_I \mathbb{E}\left(x(z_0,\tilde{q})^2\right)\right) - \frac{1}{2}e_I^2\right)\nu Q \\ &\text{st}\, e_I = \lambda_0 a x_0(z_0,\tilde{q}) - Aa^2\lambda_0\mathbb{E}\left(x(z_0,\tilde{q})^2\right)\end{aligned} \tag{63}$$

FOC yields the optimal a solves a cubic equation [31]:

$$FOC = 2A^2\lambda_0^2 E(x^2)^2 a^3 - \left(\lambda_0^2 E(x)^2 + 2A\lambda_0^2 E(x)E(x^2) + 2A\tau q^2\right) a + \lambda_0^2 E(x)^2 = 0$$

Characterize the optimal allocation a^*. The solutions to the FOC is $a_1 \leq a_2 \leq a_3$. When $a \to -\infty, f(a) < 0$, and $a = 0, f(a) > 0$. Therefore, there is at least one solution satisfying $a_1 < 0$. Consider the SOC. It can be written in this form: $f'(a) = C_1 a^2 + C_2, C_1 = 6A^2\lambda_0^2 E(x(z_0,\tilde{q})^2)^2 > 0, C_2 = \left(\lambda_0^2 E(x(z_0,\tilde{q}))^2 + 2A\lambda_0^2 E(x(z_0,\tilde{q}))E(x(z_0,\tilde{q})^2) + 2A\tau\tilde{q}^2\right) > 0$. Obviously, there exists a cutoff $\bar{a} > 0$, such that $SOC > 0$ when $a \in (-\infty, -\bar{a}) \bigcup (\bar{a}, \infty)$, and $SOC \leq 0$ when $a \in [-\bar{a}, \bar{a}]$ Therefore, there is one and only one solution $a_1 < 0$, and there can be either two positive solutions $a_2, a_3 > 0$ (including the case where $a_2 = a_3$), or two solutions with a imagine imaginary. We are interested in the case where $a^* \in (0,1)$ and $SOC < 0$, which is a_2 (the smaller one between the two positive solutions), if exists (easy to show that SOC at a_3 is always positive). Consider four cases:

1. $a_2 \in [0,1], a_3 > 1$. This is the optimal allocation. $a^* = a_2$.
2. $a_2 \in [0,1], a_3 < 1$. a_2 is an interior solution and $a = 1$ is a boundary case that may or may not be dominated by a_2. Therefore, we need to compare the case where $a = 1$ with a_2 to determine a^*.
3. $a_2 \notin R$, meaning that $FOC > 0, \forall a > 0$. $a^* = 1$.
4. $a_2 > 1$, $FOC > 0, \forall a \in [0,1]$. $a^* = 1$.

Given the contracts, inventor chooses which firm $\tilde{q}$ to work for by maximizing her utility. In each firm, her optimal effort level in given in Equation **??**.

$$\begin{aligned}
&\max_{\tilde{q}} u\left(c_I\left(a, T_I, q, Q(t)\right), e_I\right) \\
&\quad = \left(a\lambda_0 e_I x_0\left(z_0,\tilde{q}\right) + T_I - Aa^2\left(\tilde{q}^2\tau + \mathbb{E}(x^2(z_0,\tilde{q}))\lambda_0 e_I\right) - R\left(e_I\right)\right)\nu Q(t) \\
\text{st}\, e_I &= \lambda_0 a\left(z_0,\tilde{q}\right) x_0\left(z_0,\tilde{q}\right) - Aa\left(z_0,\tilde{q}\right)^2 \lambda_0 \mathbb{E}\left(x(z_0,\tilde{q})\left(z_0,\tilde{q}\right)^2\right)
\end{aligned}$$

[31] In startup, the decision rule becomes:

$$FOC = 2A^2 E(x^2)^2 a^3 - \left(E(x)^2 + 2AE(x)E(x^2)\right) a + E(x)^2 = 0$$

$$FOC = 2A^2 k_z^2 z^2 a^3 - (1 + 2Ak_z z)\, a + 1 = 0$$

$$
\begin{aligned}
g = \frac{\dot{Q}(t)}{Q(t)} = & \int_{z_0 \in \{z_0 | \tilde{q}^* > 0\}} \left(h\lambda_0 e_I\left(z_0, \tilde{q}^*\right) x_0\left(z_0, \tilde{q}^*\right) + (1-h)\,\lambda_0 e_F \kappa z_0\right)\, d\Psi\left(z_0\right) \\
& + \int_{z_0 \in \{z_0 | \tilde{q}^* = 0\}} \left(h_s \lambda_0 e_I\left(z_0, \tilde{q}^* = 0\right) z_0 + (1-h_s)\,\lambda_0 e_F \kappa z_0\right)\, d\Psi\left(z_0\right)
\end{aligned}
\tag{64}
$$

The firm-level innovation arrival rate can be written as:

$$
\begin{aligned}
\lambda_q\left(\tilde{q}\right) N_F \tilde{f}\left(\tilde{q}\right) d\tilde{q} = & h\lambda_0 e_I\left(z_0^*, \tilde{q}\right) \psi\left(z_0^*\right) dz_0^* + (1-h)\,\tilde{f}\left(\tilde{q}\right) d\tilde{q} \int_{z_0 \in \{z_0 | \tilde{q}^*(z_0) > 0\}} \lambda_0 e_I\left(z_0, \tilde{q}\right) \psi\left(z_0\right) dz_0 \\
& + (1-h_s)\,\tilde{f}\left(\tilde{q}\right) d\tilde{q} \int_{z_0 \in \{z_0 | \tilde{q}^*(z_0) = 0\}} \lambda_0 e_I\left(z_0, \tilde{q}\right) \psi\left(z_0\right) dz_0
\end{aligned}
\tag{65}
$$

where z_0^* is the inventor whose optimal choice is $\tilde{q}$.[32]

[32]If an inventor z_0 works in a firm $\tilde{q}$ when the novelty index is γ, the utility level is:

$$
u\left(z_0, \tilde{q}\right) = \left(\lambda_0 e_I x_0\left(z_0, \tilde{q}\right) - a^2 A\left(\lambda_0 e_I k x_0^2\left(z_0, \tilde{q}\right) + \tau \tilde{q}^2\right) - e_I^2/2\right) \nu Q
$$

Take derivative with respect to $x_0\left(z_0, \tilde{q}\right)$ yields:

$$
\begin{aligned}
\frac{\mathrm{d}u/(\nu Q)}{\mathrm{d}x_0} = & \frac{1}{\nu Q}\left(\frac{\partial u}{\partial x_0} + \frac{\partial u}{\partial e_I}\frac{\partial e_I}{\partial x_0} + \frac{\partial u}{\partial a}\frac{\partial a}{\partial x_0}\right) \\
= & \lambda_0 e_I - 2a^2 A \lambda_0 e_I k x_0\left(z_0, \tilde{q}\right) + (1-a)\,\lambda_0 x_0\left(z_0, \tilde{q}\right) \frac{\partial e_I}{\partial x_0} \\
= & \lambda_0^2 a x_0\left(z_0, \tilde{q}\right)\left(1 - 2a^2 A k x_0\left(z_0, \tilde{q}\right)\right)\left(1 - aAkx_0\left(z_0, \tilde{q}\right)\right) \\
& + (1-a)\, a \lambda_0^2 \left(1 - 2Aakx_0\left(z_0, \tilde{q}\right)\right) \\
= & \lambda_0^2 a x_0\left(z_0, \tilde{q}\right)\left(2 - 3aAkx_0\left(z_0, \tilde{q}\right) + 2a^3 A^2 k^2 x_0\left(z_0, \tilde{q}\right)^2 - a\right) \\
= & \lambda_0^2 a x_0\left(z_0, \tilde{q}\right)\left[(1-a) + \left(1 - aAkx_0\left(z_0, \tilde{q}\right)\right)\right. \\
& \left. - 2aAkx_0\left(z_0, \tilde{q}\right)\left(1 - a^2 A k x_0\left(z_0, \tilde{q}\right)\right)\right]
\end{aligned}
$$

As long as $aAkx_0\left(z_0, \tilde{q}\right) < 1$, the derivative is positive and the utility increases in x_0, and hence it increases in γ.

D Quantification Details

D.1 Distribution: matching data with model

The distributional data are from the BDS, which provides information on the number of firms and total employment in each size bin, capped at 10,000 employees per firm. We use a spliced log-normal distribution with a Pareto tail to parametrically represent the firm size distribution in the initial period. Specifically, the distribution is assumed to be log-normal when $q < \bar{q}$ and Pareto otherwise. The cutoff is set at 1,000 employees.[33] The parameters are estimated using GMM. Moments are divided into two sets:

1. cumulative density related, $g_{cdf,i}$, and
2. average employment related, $g_{emp,j}$.

The first type of moments measures the cumulative density in each bin, while the second type matches the average employment of firms within each bin. With 10 bins, this yields 19 moments to match 4 parameters.

Numerically, the estimated distribution is used as the initial distribution. The main challenge is converting from a per-employee measure to the model measure, which is expressed in terms of Q. We proceed with the following algorithm:

1. Calculate the measure of firms, N_F, in the initial period.
2. Construct a grid of employment levels covering both the log-normal and Pareto parts of the distribution.
3. Calculate the CDF from the parametric distribution.
4. Using the CDF, compute the mass of firms in each bin, ensuring that the masses sum to the total measure of firms.
5. Calculate Q based on the grids and the mass allocation.
6. Normalize the grid using Q.

The normalized grid, together with the mass from step 4, is then used as the initial distribution.

[33] We test alternative cutoffs, which do not materially affect the results.