Trends in U.S. Business Dynamism and the Innovation Landscape

UFUK AKCIGIT  
*University of Chicago*

SINA T. ATES  
*Federal Reserve Board*

CRAIG A. CHIKIS  
*University of Chicago*
Trends in U.S. Business Dynamism and the Innovation Landscape*†

Ufuk Akcigit  
University of Chicago  
NBER, CEPR

Sina T. Ates  
Federal Reserve Board

Craig A. Chikis  
University of Chicago

September 2023

Abstract

Business dynamism—the perpetual process of new firms forming, growing, shrinking, and dying—and the associated reallocation of factors toward more productive units is a fundamental source of aggregate productivity growth in a healthy economy. A variety of empirical regularities indicate that business dynamism in the United States has been slowing since the 1980s, and even more strikingly, since the 2000s. Our research implies that the competitive environment has been suffering from weaker technological diffusion in the economy. Hoarding patents and innovative resources, established market leaders have been entrenching their position in markets, contributing to a decline in overall innovativeness and dynamism of the economy. Concurrent with rising patent concentration, patent litigation has surged, with increased involvement of non-practicing entities more recently. A joint assessment of these trends indicates that strengthening criteria for patent grants and reducing uncertainty about their boundaries can go a long way in alleviating issues surrounding patents, innovativeness of firms, and business dynamism.

*The views and conclusions in this paper are solely the responsibility of the authors and should not be interpreted as the views of the Board of Governors of the Federal Reserve System or of any other person associated with the Federal Reserve System.

†E-mail addresses: uakcigit@uchicago.edu, sina.t.ates@frb.gov, and cachikis@uchicago.edu
1 Introduction

Businesses in the United States devote ever growing resources to research and development activities (Figure 1). Yet serious concerns about business dynamism in the U.S. economy—the perpetual process of entry, growth, downsizing, and exiting of firms—have been at the center of academic and policy discussions over the past decade or so.\(^1\) Indeed, an extensive set of empirical regularities, ranging from increased market concentration to a decline in the rate of new business formation, suggests that business dynamism has been slowing since the early 1980s in the U.S. economy (Akcigit and Ates, 2021).\(^2\)

![Figure 1: Innovative Resources](image)

*Sources: Inventor Employment History Database by Akcigit and Goldschlag (2023), NSF.*

At face value, expanding R&D resources would be expected to augur larger productivity gains. However, the aggregates mask important changes in the composition of the use of these resources, which limit productivity gains from these investments. For instance, Figure 2a shows that U.S. inventors are employed increasingly more by more established incumbents. Moreover, they appear to be utilized in less productive ways (compared to peers in younger firms), producing lower quality innovations with fewer citations, fewer citations per application, fewer

\(^1\) This process, which ensures reallocation of scarce resources such as labor and capital toward more productive uses, is integral to aggregate productivity growth and sustained long-run economic development (Foster et al., 2000).

\(^2\) We discuss the specific trends in detail in Section 2.
independent claims, and more self citations (Figure 2b). Concurrently, patent litigation has surged over the past several decades, with increased involvement of non-practicing entities more recently—a topic we discuss in detail in Section 5. These developments together with a slowdown in U.S. business dynamism are particularly worrying, suggesting that the business climate in the U.S. is moving away from a vibrant one in which firms strive to compete against and outperform rivals, boosting innovation, resource reallocation, and productivity growth, to a stalling one in which dominant players gain larger footprints, constraining competition and overall innovativeness. In this work, we summarize key findings from our recent research on factors that have been stymieing U.S. business dynamism and, within this context, discuss issues surrounding the landscape of patenting—an important gauge of firms’ innovative output.

In this paper we explore mechanisms that could have depressed U.S. business dynamism using an economic theory that explains the link between competition and firms’ incentives to make productive investments (innovations). The empirical trends suggest a marked slowdown in firms’ overall dynamism coinciding with a striking decline in competition and a notable increase in market concentration. Supported by a large body of empirical work, a canonical theory of economic growth postulates that close competition induces firms to invest in innovations with the goal of outstripping their close rivals. Based on this theory, our investigation indicates that the primary factor behind slowing business dynamism in the U.S. economy is increased distortions in the competitive environment that impair diffusion of knowledge and technologies from
market leaders to followers. In Akcigit and Ates (2022), we show that such declines in knowledge and technology diffusion can account for most empirical trends consistent with weaker business dynamism. It causes the technological gap between industry leaders and their competitors to widen, hurting firms’ incentives to improve and to compete. Laggard firms fall further behind the technological frontier and become more discouraged, as the chances of catching up with the industry leader diminish. Leading firms, in turn, relax their efforts, as competitive pressure to protect their advantages wanes. Finally, firm entry decreases, as startups observe higher barriers to competition.

Some of the trends we analyze may at first glance appear at odds with or unrelated to each other, and they have mostly been investigated in isolated contexts in previous work. By contrast, our study utilizes economic theory to propose a unifying framework that can rationalize this large set of empirical facts in a meaningful fashion. In addition, it helps us pin down the key and possibly common drivers behind these shifts.

Having established the theoretical background, we then turn our attention to patent data—the seminal representation of codified knowledge in the economy—and present empirical results consistent with a reduction in the intensity of knowledge diffusion. Based on Akcigit and Ates (2022), we show that patents are increasingly concentrated in more established firms, both via the production of new patents and purchases of existing ones from other firms. A similar trend is observed in patent litigation—embodied in an explosion of patent cases—with the involvement of non-practicing entities (NPEs) growing more prominent of late. The final section of this piece delves deeper into factors behind these trends and their impacts on firm performance and innovation. A key takeaway arises: a lack of transparency and clarity about patent boundaries and claims underlie multiple trends that we observe in the patenting environment. Improving the (re)examination and legal treatment of patents is imperative in order to overcome the challenges posed by these trends and revitalize U.S. firms’ innovativeness and competitiveness, which could then help restore business dynamism. In parallel, the increased dominance of established firms in the secondary market calls for a reexamination of the current regime to ensure the effective operation of this market to facilitate the diffusion of technologies and the efficient allocation of ideas.

One of the goals we pursue in our research agenda as well as in this piece is to provide a bird’s eye view of the business and innovation landscapes in the U.S. economy. To ensure that we construct a complete and coherent picture, we rely on a multitude of empirical trends and relationships. Given the breadth of the empirical facts of interest, we deem it appropriate to present a summary of them before we proceed further with the analysis. An (almost) exhaustive list is as follows.
Facts on Business Dynamism

1. The productivity growth rate has fallen, except for a temporary burst between the mid-1990s and the mid-2000s.
2. Market concentration has risen.
3. Average markups have increased.
4. Average profits have increased.
5. The labor share of output has gone down.
6. The rise in market concentration and the fall in labor share are positively associated.
7. The labor productivity gap between frontier and laggard firms has widened.
8. The firm entry rate and the share of young firms in economic activity have declined.
9. Job reallocation has slowed down.
10. The dispersion of firm growth has decreased.
11. The share of workers employed by small to medium-size firms has declined.
12. The rate of churn among top firms has decreased.

Facts on Inventor Landscape

13. The share of inventors in the total workforce as well the ratio of private R&D to GDP have increased.
14. The share of inventors employed by young firms has declined.
15. Inventors in young firms produce more impactful patents.
16. Entrepreneurship by inventors, who tend to found faster-growing startups, has dropped.

Facts on Patenting Landscape

17. The share of patents held by top firms—those already owning the largest patent stocks—has increased, while the share held by entrant firms has declined.
18. The share of patents reassigned to or acquired by top firms—those already owning the largest patent stocks—has risen.
19. The share of self-citations in patent claims (citations to earlier patents owned by the same firm) as well as the average claims length has increased since the early 2000s.
20. Higher concentration of patents in top firms is positively associated with higher market concentration, markups, and profits.
21. The share of litigated patents has increased.
22. A higher share of litigated patents in a sector is positively associated with higher market concentration, markups, and profits.
23. The share of litigation cases involving non-practicing entities has gone up.
2 U.S. Business Dynamism Landscape

In this section, we review the evidence for declining business dynamism in the United States.

*The productivity growth rate has fallen.* The decline in the productivity growth rate is widely documented; we plot its decline in Figure 3.\(^3\) Some have speculated that the decline in productivity growth reflects measurement error: time-intensive but otherwise free, new technologies would not show up in national accounts, even as they deliver substantial new surplus to end users. Syverson (2017) convincingly argues against such explanations and confirms the slow-down in productivity growth is all too real.\(^4\)

![Figure 3: TFP and its growth rate](image)

*Notes: Authors’ calculation based on nonfarm private business TFP data from BLS (retrieved from FRED, Federal Reserve Bank of St. Louis; [https://fred.stlouisfed.org/series/MFPNFBS](https://fred.stlouisfed.org/series/MFPNFBS), January 30, 2023).*

*Market concentration has increased.* Autor et al. (2017a, 2020) and Grullon et al. (2017) document rising within-industry market concentration. Figure 4a affirms this finding, plotting the average share of sales captured by the largest 4 and 20 firms within a given industry. Alternative measures of concentration, such as the average within industry Herfindahl-Hirschman Index (HHI), have also risen.\(^5\) Rising concentration is strongly, negatively correlated with declines in the labor share, as demonstrated conclusively by Autor et al. (2020) and Barkai (2020) (Fact 5).\(^6\)

---

\(^3\) For further evidence and implications see Fernald (2015); Fernald et al. (2017).

\(^4\) Aghion et al. (2019a) proposes an alternative mechanism for mismeasured and missing productivity growth, but their conclusions are broadly aligned with those of Syverson (2017).

\(^5\) See Council of Economic Advisers (2016) and OECD (2018a) for a thorough discussion. By contrast, notes by some participating delegations on the same subject doubt the notion of increased market concentration on the grounds of mismeasurement concerns and the lack of focus on relevant markets (OECD, 2018b,c).

\(^6\) For further studies on rising product market concentration and its aggregate implications, see Gutiérrez and Philippon (2017a,b) and Eggertsson et al. (2021), among others. In a similar vein, Azar et al. (2017) document concentration in the U.S. labor market using disaggregated data at the geographical-occupational level.
Concurrently, *markups have increased*. Though notoriously difficult to measure, there is a growing economic consensus that markups have risen in the United States—Figure 4b plots the rise in the average markup from 1980 to 2010. Hall (2018) and De Loecker et al. (2020) document rising markups; in particular, De Loecker et al. (2020) argue that the

---

7 Some recent work (e.g., Karabarbounis and Neiman, 2018 and Traina, 2018) disagrees with the empirical rise in markups, based on measurement concerns. In particular, these papers find that the alleged rise in markups is strikingly sensitive to the chosen measure of variable costs, being flat when practitioners exclude marketing and management costs.
increase in markups has taken place entirely in the right tail of the distribution: the median markup is unchanged, but the top decile of markups exploded between 1980 and present, concurrent with large reallocation toward high markup firms, capturing a flavor of the superstar effect of Autor et al. (2020). Building on this work, other papers claim that a rise in markups is a proxy for a rise in firm market power, which could explain various macroeconomic and asset pricing phenomena.  

Profit share of GDP has increased, too. Figure 4c shows that, concurrent with rising markups and a declining labor share, the profit share of GDP has risen. Gutiérrez and Philippon (2017b) provide reason to suspect such developments have macroeconomic consequences, documenting systematic underinvestment relative to firm profitability beginning in the early 2000s.

Meanwhile, the labor share has declined. To generations of economists, trained for decades on the facts of Kaldor (1961), a nonconstant labor share of output is heretical. Nevertheless, Figure 4d shows a trend decline in the labor share of output for the United States, starting in the 1980s (Elsby et al., 2013; Karabarbounis and Neiman, 2014; Piketty, 2014, Lawrence, 2015; Autor et al., 2017a, 2020; and Koh et al., 2020).

Market concentration and labor share are negatively associated. Figure 5a reproduces the findings of Autor et al. (2020), demonstrating a negative correlation between market concentration and the labor share across U.S. industries. Other recent research (e.g., Barkai, 2020 and Eggertsson et al., 2021) corroborates this observation.

Notably, the labor productivity gap between frontier and laggard firms has widened. A key fact that Akcigit and Ates (2022) find particularly instructive for identifying the primitive causes of declining business dynamism is the marked widening of the average within-industry productivity gap between frontier and laggard firms. Figure 5b shows that this gap—measured in terms of real value added per worker—has been widening (Andrews et al., 2015, 2016). These studies also demonstrate that aggregate productivity is weaker in industries that observe widening polarization between leaders and laggards. Complementing these findings, Decker et al. (2020) show that total factor productivity (TFP) dispersion across U.S. firms has risen.

---

8 For international evidence on rising markups, see Calligaris et al. (2018); De Loecker and Eeckhout (2018).
9 See Farhi and Gourio (2018); Barkai (2020); and Eggertsson et al. (2021) for links between rising market power, low risk-free interest rates, and risky asset prices.
10 For further evidence and implications, see Eggertsson et al. (2021).
11 Figure 5b reproduces the findings of Andrews et al. (2016), who present a cross-country comparison of the top five percent of firms with the highest labor productivity level (frontier) to the rest of firms (laggard). Although the Orbis database used in their study has a rather limited coverage of U.S. firms, in a complementary work, the authors claim that firms from advanced economies are well represented in the frontier group (Andrews et al., 2015).
Importantly, the firm entry rate and the economic share of young firms have declined. Small, young firms are key drivers of U.S. employment dynamics and employment growth (Haltiwanger et al., 2013). Thus, quite naturally, a widely debated symptom of declining business dynamism in the United States is the fall in new firm entry (Gourio et al., 2014, 2016) and in the share of young firms in employment (Furman and Orszag, 2018). Figures 5c and 5d illustrate these phenomena using Business Dynamics Statistics data. Numerous studies have documented that declining firm entry is linked to the recession in the right tail of the employment growth rate distribution (Criscuolo et al., 2015; Decker et al., 2016a,c; Bijnens and Konings, 2018).

Figures 5c and 5d illustrate these phenomena using Business Dynamics Statistics data. Numerous studies have documented that declining firm entry is linked to the recession in the right tail of the employment growth rate distribution (Criscuolo et al., 2015; Decker et al., 2016a,c; Bijnens and Konings, 2018).

Notes: Panel A and B reproduced from Autor et al. (2017b) and Andrews et al. (2016), respectively. Panel C and D show authors’ own calculations using the 2020 Business Dynamics Statistics Datasets from the U.S. Census Bureau. Young firms are five years old or younger.

12 Bravo-Biosca et al. (2013) document this fact internationally. Akcigit et al. (2021a) discuss firm size and the growth of young firms in the context of managerial impediments to delegation in the developing world, broadly, and Indian manufacturing, specifically.

13 Karahan et al. (2019) argue the cause of the decline in firm entry is rooted in a demographics-induced contraction in labor supply, which began in the 1970s.

14 In particular, Gourio et al. (2016) estimates that, for the United States, there are 1.7 million fewer jobs than would have been obtained absent the trend decline in entry.
Job reallocation and churn have also gone down. Figure 6a exhibits a secular decline in the gross job reallocation rate (defined as the sum of job creation and destruction rates) in the United States. Davis and Haltiwanger (2014) provide evidence that this decline is international in scope, though more intense in the United States. The decline has been apparent in the retail trade and services sectors for several decades—due, in large part, to productivity-enhancing consolidation of activity in larger chains, often at the expense of smaller, so-called mom-and-pop shops—whereas, in the information sector, a pronounced decline started in the early 2000s.

Concurrently, the dispersion of firm growth rates has gone down. As the share of young, high-growth firms in economic activity has declined, the dispersion in firm growth rates has similarly declined (Figure 6b). This is particularly intriguing given the aforementioned rise in productivity dispersion across firms (Fact 7). The cause for declining dispersion in firm growth rates, as argued by Decker et al. (2016a), is the increasing share of the technology sector in economic activity. Using U.S. Census data, Decker et al. (2016a) show that the decline in firm growth rate dispersion accelerated in the post-2000 period due to the decline in young firm activity in high-tech sectors.

2.1 Additional Trends

While our recent work (Akcigit and Ates, 2021, 2022; Chikis et al., 2021) focused on these 10 observations following the debates in the academic literature, the additional trends listed above also seem to be consistent with a broad decline in market competition and business dynamism, especially through the lens of the economic theory we apply to interpret these dynamics. Echoing Fact (8), one such example is the decrease in the employment share of smaller firms (Figure 7a). Similarly, the share of inventors employed by young firms has been steadily declining (Figure 7b). On the one hand, this shift occurs in tandem with a broad decline in the rate of churn amongst the largest and most dominant firms—the probability of a top firm being replaced by a successful below-top competitor (Akcigit et al., 2021b).

15 For a thorough account of this phenomenon, see Decker et al. (2016a).
On the other hand, the rate of generating drastic ideas and inventions has declined in the economy (Park et al., 2023), as reflected in a reduction in the average creative content of U.S.-based patents (Kalyani, 2022). Our proposed theoretical explanation—the decline in knowledge diffusion—implies a widening technological gap between market leaders and followers, leading to an overall decline in innovativeness and dynamism, consistent with small and young firms capturing a smaller share of the economic pie, large firms becoming entrenched while hoarding innovative input, and the economy as a whole becoming less apt to generate breakthrough ideas.

Shedding light on the innovative landscape, Akcigit and Goldschlag (2023) provide further insights into the employment dynamics of U.S. inventors. They show that inventors are increasingly likely to be employed by mature and established firms, often to the detriment of inventors’ innovative productivity. In addition, inventors have become less entrepreneurial over time (Figure 8a). From the perspective of economic and business dynamism, this shift poses a serious challenge, as firms founded by inventors are more dynamic and grow considerably faster than peer firms set up by non-inventor founders (Figure 8b).

Notes: Panel A from authors’ calculations using BDS data maintained by the Census. Micro firms are defined as firms with less than 20 employees; mid-size firms are those with 20 to 99. Panel B reproduced from Akcigit and Goldschlag (2023). Young firms are defined as being five-years old or younger.

16 Akcigit and Goldschlag (2023) use event studies to analyze inventor activity around the time they switch jobs and compare the productivity of inventors that switch to more established firms to those with similar characteristics that switch to young ones. The results show that the number of patent applications by inventors drops after they join more established incumbents (relative to inventors with comparable characteristics who join young firms). Moreover, the citations to the patents for which inventors apply after switching to a mature incumbent firm are also lower relative to inventors hired by young firms, suggesting a deterioration in the quality of innovative output by inventors at incumbent firms. In addition, the share of self-citations of inventors hired by mature incumbents increases relative to inventors hired by young firms.

17 Please see the figure notes for the econometric specification underlying the analysis. Controlling for confounding factors, the specification identifies the difference in firm performance that can be attributed to the founder type.
Akcigit and Goldschlag (2023) also document that inventors are becoming less likely to switch jobs and are increasingly geographically concentrated. Patterns of increasing geographic concentration of innovative activity (Chattergoon and Kerr, 2021) and broader patterns of declining worker mobility pre-pandemic in the U.S. economy are consistent with declining knowledge diffusion across firms, aggravating the decline in business dynamism.

Figure 8: Inventor Entrepreneurship

Notes: Akcigit and Goldschlag (2023). Panel A reports estimates of $\beta_{k}$ from the equation $\text{Entrep}_{i,t} = \alpha + \sum_{2015}^{2011} \beta_{k} D_{k} + \psi_{i} + \epsilon_{i,t}$. For inventor $i$ in year $t$, $\text{Entrep}_{i,t}$ equals 1 if the inventor started a business in that year and zero otherwise. $D_{k}$ are year effects, and $\psi_{i}$ are person fixed effects. The mean of $\text{Entrep}_{i,t}$ in 2000 is 0.0064. Panel B shows estimates of $\beta_{a}$ from the equation $\ln(\text{FirmSize}_{f,t}) = \alpha + \sum_{10}^{1} \beta_{a} \text{FirmAge}_{f,t} + \psi_{f} + \epsilon_{f,t}$. For firm $f$ in year $t$, $\ln(\text{FirmSize}_{f,t})$ is the log of firm employment in year $t$, $\text{FirmAge}_{f,t}$ is the firm’s age in year $t$, and $\psi_{f}$ are industry-year fixed effects.

Finally, a phenomenon widely debated in both academic and policy circles has been the steady decline in real interest rates over the past several decades (Summers, 2014), as shown in Figure 9. Interestingly, this period of real interest rate decline overlaps with the numerous trends we discussed as signifying declining business dynamism. An overall decline in business dynamism could indeed have contributed to the fall in interest rates via deficient investment demand. Indeed, the theoretical mechanism that we explore in Section 4 implies a decline in innovative investment as competitive pressures weaken, resulting in lower aggregate growth and interest rates.

3 Proposed Mechanisms for Declining Business Dynamism

Having listed a set of rather dismal facts related to declining business dynamism in the United States, it is natural to ponder: what is behind such a decline, and are the causes exogenous and technological—in which case the scope for policy is perhaps limited or the outcome even efficient—or institutional—and thus subject to correction via policy? These questions, the causes and possible remedies, are subjects of active debate in the literature, which has tended to treat the assembled facts as separate, isolated phenomena. In the present section, we leverage the focused expertise of the profession in diagnosing the various factors behind declining business dynamism, largely thinking of the facts as distinct entities.
Koh et al. (2020) argues that the trend decline in the labor share documented by Karabarbounis and Neiman (2014), Autor et al. (2020), and others is entirely accounted for by the capitalization of intellectual property products.

In doing so, however, we hope to illustrate a point we will lean on heavily in later sections: the drivers identified in the literature are all unified in that they represent forces that favor large, incumbent firms at the expense of small, young firms, leading, in the aggregate, to lower growth and innovation, higher market power, lower levels of gross reallocation, and a lower labor share. This will be the central idea on which our theory of declining business dynamism rests.

### 3.1 Declining Labor Share and Rising Market Concentration

The declining labor share is one of the central macroeconomic puzzles of the 21st century. Kaldor (1961) famously remarked upon the observed stability in factor income shares, but it is by now widely accepted that the labor share is declining, both in the United States and internationally. Numerous explanations have been proposed, ranging from technological (Karabarbounis and Neiman, 2014; Acemoglu and Restrepo, 2020); to institutional (Piketty, 2014; Gutiérrez and Philippon, 2017b; Kaymak and Schott, 2020); to structural (Elsby et al., 2013; Hopenhayn et al., 2022).

While mostly agnostic on the cause, Autor et al. (2020) has become the predominant paradigm for understanding the declining aggregate labor share. In their telling, some broad, exogenous impulse—improved search technologies for consumer products (Akerman et al., 2022) or globalization and the internationalization of product markets, for example—has raised the share of sales accruing to the most productive firms, who pay the lowest labor share of output. Notably, this “winner takes most” mechanism is also consistent with rising aggregate markups (Fact 3). The ideas of Autor et al. (2020) have been consistently validated and extended in concurrent and follow-on work (Diez et al., 2018; Bessen, 2020; Barkai, 2020; Eggertsson et al., 2021), and the superstar phenomenon has permeated multiple levels of economic discourse (Andrews et al., 2016; Van Reenen, 2018).

---

18 Koh et al. (2020) argues that the trend decline in the labor share documented by Karabarbounis and Neiman (2014), Autor et al. (2020), and others is entirely accounted for by the capitalization of intellectual property products.
One compelling explanation for rising concentration is technological in nature. Andrews et al. (2016), for example, stress that the increasing reliance on big—often proprietary—datasets and tacit knowledge is a key driver in the widening productivity gap between leading and lagging firms (Fact 7). Calligaris et al. (2018) show markups are higher in more digitally intensive sectors, and Bessen (2020) documents a robust, positive correlation between industry investment in information and communications technology (ICT) and the concentration of sales; Lashkari et al. (2018) demonstrate, empirically, that the intensity of investments in IT are increasing in firm size and rationalize this by arguing that the marginal return to investing in IT is increasing in size. Along a similar vein, Crouzet and Eberly (2018); Aghion et al. (2019b); and De Ridder (2021) emphasize the increasing importance of intangibles. Hsieh and Rossi-Hansberg (2022) also argue for the importance of technology in driving increasing national, sectoral sales concentration; their paper argues that ICT has allowed the geographic expansion of multiproduct service firms, which has lowered local concentration while increasing national concentration.

Technological explanations for rising concentration tend to obscure the role for policy, as their associated outcomes are often inherently efficient. In contrast, numerous studies argue that there are institutional factors at the heart of rising concentration (e.g., Philippon, 2019). Using cross-country data, Andrews et al. (2016) find that the productivity divergence between leading and lagging firms is largest in industries where pro-competition product market reforms were least extensive. Congruent with such a finding, there is a prominent but increasingly politicized strand of the academic legal literature that explores the competitive implications of antitrust law’s adoption of the Chicago-school, consumer welfare standard (Lynn, 2010; Baker, 2012; Khan, 2016). Khan (2016), in particular, argues that platform economics, in which companies prioritize growth over short-term profits, rewards firms for capturing enormous market share. Via the consumer welfare standard, prices remain low, so antitrust law has minimal bite, leading to large, super-firms that span numerous markets and have the heft to price their competitors out of business. Bessen (2016) argues that an increasing share of corporate valuations and profits can be ascribed to political rent-seeking, which the author causally links to more highly regulated markets. Haltiwanger et al. (2014) argue labor market regulations on hiring and firing explain cross-country variation in job reallocation. There is agreement in the literature that policy factors are likely at play, but little agreement on which and even less on potential remedies.

3.2 Labor Market Determinants of Declining Business Dynamism

3.2.1 Declining Firm Entry

There are two major schools of thought on what labor market developments may lie behind the decline in firm entry. One explanation is purely demographic. Karahan et al. (2019) note that the labor force growth rate began a secular decline in the 1970s.
In their model, long-run declines in the labor force growth rate, induced by the exhaustion of the baby boom population spike, are absorbed entirely along the entry margin. In other words: slower population growth necessitates slower new firm formation. Peters and Walsh (2021) and Hopenhayn et al. (2022) make variations of a similar argument.

A second school of thought links declining firm entry to forces constraining breakthrough innovations more broadly. Aforementioned studies such as Gordon (2016) and Bloom et al. (2020) take a pessimistic view on the prospect for major innovations to continue arriving at the frequency they did in the early 20th century. If entrants are responsible for more major innovations than incumbents as in Gort and Klepper (1982) and Acemoglu and Cao (2015), then it must be that the rate of entry is falling in response to the diminution in the stock of major ideas or dead-end duplication of effort (Akcigit and Liu, 2016).

### 3.2.2 Declining Job Reallocation

Focusing on job flows, Decker et al. (2020) argue that the culprit behind declining dynamism is the declining responsiveness of firms to shocks. They argue that such declining responsiveness likely reflects difficulties in the employment adjustment margin, which may depend on a variety of factors (see Decker et al., 2016b for a succinct overview). For instance, Davis and Haltiwanger (2014) suggest that lower worker fluidity may be a reflection of widespread occupational licensing practices or the inhibitory effects of employment protection regulations such as noncompete agreements.¹⁹

### 3.3 Patent Litigation and NPE activity

Finally, the period over which the U.S. economy has suffered from a decline in business dynamism and knowledge diffusion has concurred with a rise in patent litigation and NPE activity. We discuss these developments and the related literature in detail in Section 5. These shifts, driven chiefly by changes in the legal system regarding intellectual property and the examination standards of patents (Bessen and Meurer, 2008a; Jaffe and Lerner, 2004), could have possibly contributed to a tilting of the competitive environment against emerging competitors of large, established firms. We defer a further elaboration of this subject to Section 5.

---

¹⁹ Furman and Giuliano (2016) document that about a quarter of U.S. workers hold occupational licenses, a dramatic increase since the 1950s. As to the effect of non-compete laws, see Marx et al. (2009). Using a seemingly exogenous variation in noncompete laws in Michigan, the authors show the attenuating effect of such policies on labor mobility. White House (2016) highlights that noncompete contracts bind a sizable fraction of workers, even those without a four-year college degree and those earning less than $40,000, suggesting an abuse of the laws, possibly in ways harmful to job reallocation. Wessel (2018) provides a non-technical account of regulatory concerns regarding competition.
4 Basic Theory to Rationalize the Trends

Each of the trends described in Section 2 are subject to research dedicated to determine the exact magnitudes and possible causes, some of which we summarized in Section 3. Yet the question of whether this confluence of trends depicts a consistent and coherent picture of business dynamics and whether a joint analysis thereof could reveal underlying drivers of these symptoms has largely remained unanswered until recent work. To answer these questions, our research relies on economic theory that offers a unifying framework to jointly analyze the said trends and possible causes. Constructing a stylized model of the U.S. economy, Akcigit and Ates (2022) argue that the totality of empirical trends suggests a decline in U.S. business dynamism and attempt to diagnose the underlying factors behind this shift. A decline in knowledge diffusion from technologically leading to lagging firms emerges from the model as the main margin that explains declining business dynamism since the 1980s.

The model is a head-to-head, Schumpeterian innovation model. At the heart of its results is a dynamic patent race between two, finitely-lived firms (Aghion and Howitt, 1992; Aghion et al., 2001; Acemoglu and Akcigit, 2012). One firm is the leader and the other is the laggard; the leading firm has a production technology that allows it to produce the same variety as the laggard at lower cost, which means, for the duration that the leader has the better technology, it captures more of the market for its variety. Market leadership, however, probabilistically passes between the two firms based on endogenous investment decisions made by the two competing entities. Both firms are subject to firm death, or exit, upon arrival of an entrant possessing a better technology.

As the problem of the firm is dynamic, changes in the competitive environment of markets have multilayered, strategic implications. For example, higher competition reduces the persistence of market leadership; ceteris paribus, this makes the present value of being a leader lower, reducing innovative rents, and discouraging investment in innovation, both for leaders, who fret over more frequent replacement, and laggards, who are forward-looking, and know their labors will be less rewarded upon attainment of a less persistent market leadership. In contrast, higher competition, if focused in markets where leaders and laggards have broadly similar technologies, may spur higher investment in innovation: leaders, knowing that a successful string of innovations will deliver them to a less competitive promised land, will seek to “escape competition” by innovating intensively. In a like fashion, their competitors, the laggards, know the consequences of losing such a patent race are quite dire, as once leaders are sufficiently far ahead, it becomes increasingly difficult to ever attain market leadership themselves and earn positive profits.

The nature of the exercise of Akcigit and Ates (2022) is to calibrate the model to mimic key features of the 1980s U.S. economy and the transition path from that economy to the one prevailing in the mid 2010s (the latest period with consistently available data).
This terminal point is meaningful because it ushered in a period of extended stagnation in U.S. business dynamism that persisted up through the arrival of the coronavirus pandemic. The model allows for changes to key structural parameters such as: declines in corporate tax rates, increases in R&D subsidies, increases in entry costs, and declines in knowledge diffusion. Then, the authors assess, one-by-one, which of these channels can explain most of the evolution of the core facts on business dynamism presented above. They find that the overwhelmingly dominant force for the evolution of business dynamism in the United States between 1980 and 2010 is declining knowledge diffusion. It is not dispositive of other forces, but it is suggestive and consistent with numerous empirical trends—a keystone for the unified theory.

A reduction in knowledge diffusion works as follows. Over time, as knowledge diffusion declines, leaders’ technologies become more difficult to copy. On impact, this shift increases the value of being a leader and encourages those firms to innovate. Yet as they do so, this leads to a widening of the productivity gap between leaders and laggards (Fact 7). A wider gap allows leaders to capture a larger share of sales in their sectors (Facts 2 and 6); charge higher markups (Fact 3); earn higher profits (Fact 4); and pay lower wages (Fact 5). Falling behind and facing higher barriers to catch-up, follower firms become discouraged, losing their incentive to innovate and compete. While this discouragement effect concerns incumbents, it also trickles down to entrants. Entrants are forward-looking; upon entering, often as technological laggards, they realize that to increase their market share and earn larger profits they have to engage in a costly R&D race, now marred by lower technology diffusion. Thus, endogenously, entry declines as its ex-ante payoff is lower (Fact 8). As leaders—the established dominant players—consolidate their lead and intuit that their competitors—the laggards—are discouraged, they also choose a lower overall innovation rate, obtaining a lower level of gross reallocation (Fact 9). Thus, as the economy converges toward its new equilibrium balanced growth path (BGP), overall innovative exploration and productivity growth decline (Fact 1), as well as the equilibrium interest rate.

To sum up, the quantitative investigation of Akcigit and Ates (2022) underscores the importance of potential distortions in knowledge diffusion in explaining the decline in U.S. business dynamism. The next section presents novel empirical evidence on the symptoms of a decline in the intensity of knowledge diffusion in the United States, in line with the predictions of our analysis.

5 Lower Knowledge Diffusion: Symptoms and Causes

The previous section argues that a decline in knowledge diffusion is closely related to the decline in business dynamism, as summarized by the facts above, for the period 1980 to 2010. That leaves unanswered an important question: What is a reduction in knowledge diffusion, precisely? In the model, knowledge diffusion is a structural parameter that has no
obvious counterpart in the data. However, by leveraging micro-level data at the firm-and-inventor level, we can assess the case for diminished knowledge flows between the most productive and least productive firms and innovators, thereby revealing what precisely is meant by declining knowledge diffusion. Moreover, we can shed light, imperfectly, on whether firms are acting anti-competitively in protecting their intellectual property (IP). Lastly, we will delve into the growing prevalence of patent litigation over the past several decades, in addition to the increased activity of non-practicing entities more recently, and explore how they may relate to the puzzle.

5.1 Patent Concentration and Post-1980 Trends

The excessive accumulation of patents by a declining number of firms could potentially reduce inter-firm knowledge flows. The threat is particularly acute when innovations “stand on the shoulders of giants;” that is, innovators need to use the current frontier technologies in the design of new inventions. We will assess the case for the monopolization of patent holdings using patent application and reassignment data from the United States Patent and Trademark Office (USPTO).

As discussed in Section 2, the share of within-industry sales captured by the top 4 and 20 firms has increased. Figure 10a plots the share of patents held by the top 1 percent of innovating firms (as measured by their patent stocks), revealing a dramatic increase: The top 1 percent of firms by patent stock hold nearly 50 percent of all registered patents, as opposed to a little over 35 percent in 1980. Clearly, patent holding, like sales, has become more and more concentrated. Akin to Fact 8, the share of patents held by entrants—firms patenting for the first time—has declined from a peak of 9 percent to 4 percent over the same period (Figure 10b). Like the declining share of entrants in overall economic activity, entrants also seem to comprise a declining share of innovative activity.

\[ \text{Figure 10: Registry of Patents} \]

\[ \text{Source: Authors’ own calculation using U.S. Patent and Trademark Office data} \]
Serious ambiguities arise in interpreting the trends. The increasing concentration of patent holding could reflect predatory, anti-competitive forces. For example, firms may be seeking to build “patent thickets,” whereby firms acquire a critical mass of interrelated patents in a specific sector or technology cluster, which discourages competitors seeking to challenge the technological supremacy of the incumbent firm, either through fear of patent infringement litigation (Section 5.2) or the levying of exorbitant and overlapping licensing fees (Shapiro, 2001). They could also be acting rationally, finding it more efficient to assemble a trail of related patents to ward off litigation risk, or to patent even marginal innovations before someone else does. On the other hand, if the top innovators are simply innovating more intensively due to a distributionally focused positive “supply” shock in the stock of marketable ideas, one would certainly expect to see the trends in Figures 10a and 10b, but might have less to fear normatively. Ultimately, more high-quality innovation is good, regardless of who is undertaking it.

Patent data give us reason to believe the first force is at play, at least in part. First, a recent study by Argente et al. (2020) reveals that although market leaders innovate less intensively, when they innovate, they are more likely than their competitors to patent the innovation. This has the effect of reducing the number of follow-on innovations by their competitors, depressing the overall level.

Second, activity in patent reassignment data is similarly revealing. Figure 11a plots the share of transacted patents reassigned to the top percentile of firms (as measured by their patent stocks). This number has nearly doubled between 1980 and 2010. In a complementary vein, the share of transacted patents reassigned to “small business concerns” declined by roughly one-third.20

Third, even at the inventor and geographic level, patents are becoming more concentrated. Akcigit and Goldschlag (2023) document an increasing share of inventors being employed in top, incumbent firms. In a similar vein, Chattergoon and Kerr (2021) show that the spatial concentration of patents has risen substantially since 1980, especially in software. Thus, the preponderance of evidence points toward increasing concentration of innovative activity that is, at least partially, the result of strategic behavior of large, incumbent firms.

Fourth, evidence from pairwise citations and claims data suggests a declining generality and quality of patents. Akcigit and Kerr (2018) posit firms undertake two distinct forms of innovation. Internal innovations are those that seek to build upon and improve a firm’s existing suite of products. External innovations are those that relate to fields that are new to firm and are thus more explorative in nature, offering a broader and larger

---

20 The designation as a “small business concern” derives from the USPTO’s U.S. Patent Grant Maintenance Fee Events database, which records information on patent renewals.
innovative content. Patent claims are used to describe the contributions of the invention that merit award of monopolistic exclusivity. More marginal inventions that make incremental advancements on previous innovations tend to have longer claims reflecting the need to clarify their novelty vis-à-vis older patents. Figure 12a shows that the share of self-citations has increased between 1980 and 2010, pointing to a rise in the share of incremental innovation, which, on average, have a narrower focus and offer more limited technological advancement. Figure 12b tells a more ambiguous story, but it is unequivocal that there has been a sharp uptick in claim length beginning in 2000.

Figure 11: Reassignment of Patents
Source: Authors’ own calculation using U.S. Patent and Trademark Office data

Figure 12: Self-Citation and Claim Length Patterns
Source: Authors’ own calculation using U.S. Patent and Trademark Office data

---

\[21\] In the data, a patent is internal if 50% or more of its citations are of patents owned by the firm itself, so called “self-citations.” Otherwise, it is external.
Having documented the strategic use and increased concentration of patents, we next explore the relationship between this patent-level evidence and firm- and sector-level evidence of declining business dynamism. To do so, we run the following regression specification in the spirit of Autor et al. (2020):

$$\Delta Y_{st} = \beta \Delta PatTop5_{st} + \gamma Npat_{st-1} + \psi t + \epsilon_{st},$$ (1)

where $s$ and $t$ stand for sector and time period, respectively. This equation links changes in numerous market-based outcomes ($Y_{st}$) to the (contemporaneous) change in patent concentration ($PatTop5_{st}$) measured by the share of patents held by the top five percent of firms by sales at the sector level. We match the patent data from the USPTO database with data on public firms from Compustat, which provides the information on market-based outcomes. The specification controls for period fixed effects ($\psi t$) and the number of patents assigned in a sector, lagged to alleviate simultaneity issues (please see the notes below Table 1 for other details). The results are summarized in Table 1. We observe highly positive, statistically significant correlations between the change in the share of patents held by the top 5 percent of firms and the change in HHI, markups, and profit share. Indeed, a one standard-deviation change in the patent share of the top 5 percent firms is associated with up to 10 percent of the variation in the market outcome variables. While not causal, these results are broadly consistent with the mechanism inferred from the model of Akcigit and Ates (2022).

<table>
<thead>
<tr>
<th>Table 1: Patent Concentration and Dynamism Indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>$\Delta$ patent share of top 5% firms</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>$\Delta$ HHI</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>$\Delta$ Markups</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>$\Delta$ Profit share</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Observations</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>R-square</td>
</tr>
</tbody>
</table>

Notes: The regressions show the correlation between contemporaneous changes in patent litigation intensity and changes in indicators of market power. Dependent variable at the top of each column, and $\Delta$ refers to changes. The share of litigated patents refers to the ratio of the number of patents subject to litigation in a given year to the total patent stock at the sector (4-digit NAICS) level. The period covered runs from 2003 to 2016, and the regression includes year fixed effects. In all regressions, clustered standard errors at the sector level in parentheses. (** p < 0.01).
5.2 Litigation Explosion

In the post-1980 period, a parallel trend concerning patents in the U.S. has been the dramatic increase in the number of patent cases filed, which some authors have dubbed the “patent litigation explosion” (Bessen and Meurer, 2013). As can be gleaned from Figure 13, the annual number of litigation cases filed per 100 granted patents rises from about 1.2 in the early 1990s to an average of about 1.5 between 1995 and 2010, before rising again to more than 1.8 between 2010 and 2015 and only receding marginally since then. Next, we look closer at the potential forces behind the striking trend in patent litigation.

As Figure 13 shows, while the number of cases involving patents expanded precipitously, so did the number of patents granted, which has motivated a technology-based argument that posits that the former trend may simply reflect the larger pool of patents in circulation. While there may be merit to that argument, the number of cases per patents granted has also risen. Similarly, Bessen and Meurer (2013) argue that in the period up to 2000, the hazard rate of being subject to a lawsuit involving patents has increased notably, even after controlling for the size of a firm’s patent portfolio, and even more so for smaller firms. The authors argue that the rise in cases was widespread, having been observed in several sectors, casting doubt on technology-based explanations.

![Figure 13: Litigation](source: USPTO, Judicial Facts and Figures)

In explaining the rise in litigation cases, several studies put the emphasis on changes in legal and patenting practices. A notable change that preceded the pickup in litigation is the 1982 creation of the Court of Appeals for the Federal Circuit, a centralized appellate court, with the aim of unifying and strengthening the judicial treatment of patent rights in the
United States. The court’s rulings have broadly been regarded as reflecting a pro-patent shift in the judicial treatment of patents, with significant increases in the plaintiff success rates (Lerner, 1995). Jaffe and Lerner (2006) highlight that their survey of practitioners suggests a substantial increase in the willingness of patentees to file a suit, together with a diminution of the incentives of defendants to fight. Such evidence is in accordance with the interpretation of increased litigation hazard rates by Bessen and Meurer (2013). In addition, to be upheld, the fact that an infringement claim requires only a preponderance of evidence tilts the legal field in favor of the patentee, especially considering that the presumption of validity—the presumption that a granted patent is valid—raises the bar for proving a patent invalid (it requires “clear and convincing evidence”).

But broader concerns underlying this upward litigation trend are issues relating to the overall standards of patents (see Gallini, 2002 for a review). Scholars have argued that a decline in the bar for patent grants has enabled the surge in litigation, allowing patent holders to enlarge and enforce their portfolios with ill-defined or lower quality patents while also benefiting from the pro-patent shift in the legal realm. Jaffe and Lerner (2006) raise concerns about the nonobviousness of business method patents, for example, in light of the prior art, which is difficult to detect, as most such methods have not been considered to be patented until quite recently. Similarly, the treatment of software patents has attracted substantial criticism based on the perceived easing in their enablement; legal criteria concerning their novelty and applicability appear to have loosened (Burk and Lemley, 2002; Bessen and Hunt, 2004). As such, the interplay between the underlying technological change—with the software industry growing in importance—and the pro-patent, pro-grant shift in the legal treatment of patents has paved the way for the rise in patent litigation.

How does litigation activity relate to business dynamism? To address this question, we proceed as in the previous section, examining the relationship between contemporaneous changes in the share of litigated patents in a sector and market-based outcomes. Specifically, we run the following regression specification:

\[
\Delta Y_t = \beta \Delta \text{Litshare}_t + \psi_t + \epsilon_t, \tag{2}
\]

where \(Y_t\) is defined as in equation (1), and \(\text{Litshare}_t\) denotes the share of patents among all patents in a sector being subject to litigation in a given year (obtained from the USPTO).\(^{22}\)

\(^{22}\)To match patents to specific sectors, we utilize the probabilistic crosswalk designed by Goldschlag and Tabarrok (2018). The procedure is able to assign 88 percent of unique litigated patents between 2003 and 2016 (the period for which the identification numbers of litigated patents are available). Moreover, in 93 percent of the cases, it can match at least one of the involved patents.
As shown in Table 2, our analysis reveals that a higher rate of litigated patents in an industry is positively correlated with higher levels of market power indicators in the post-2000 period, implying a similar magnitude of variation being driven by changes in patent litigation rates as observed in the analysis of top firms’ patent share. This correlation in the post-2000 period is particularly interesting in light of the aforementioned evidence that the decline in business dynamism accelerated after 2000, especially in some high-tech sectors.

<table>
<thead>
<tr>
<th>Table 2: Litigation Activity and Dynamism Indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Δ share of litigated patents</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>(1)</td>
</tr>
<tr>
<td>(2)</td>
</tr>
<tr>
<td>(3)</td>
</tr>
<tr>
<td>∆HHI</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Observations</td>
</tr>
<tr>
<td>R-square</td>
</tr>
</tbody>
</table>

Notes: The regressions show the correlation between contemporaneous changes in patent litigation intensity and changes in indicators of market power. Dependent variable at the top of each column, and ∆ refers to changes. The share of litigated patents refers to the ratio of the number of patents subject to litigation in a given year to the total patent stock at the sector (4-digit NAICS) level. The period covered runs from 2003 to 2016, and the regression includes year fixed effects. In all regressions, clustered standard errors at the sector level in parentheses. (*** p < 0.01).

When interpreted through the theory we proposed in Section 4, these results are indicative of lower innovation intensity and dynamism across U.S. firms. Previous work also highlighted the negative effects of litigation hazard on firm performance. Bessen and Meurer (2008a) argue that the preponderance of litigation reflects inadvertent infringement (up until the last decade or so, at least), which would weigh on firms’ R&D efforts by increasing the risks of legal disputes for potential innovations. In addition, Bessen and Meurer (2008b) document substantial private costs of patent litigation to alleged infringers, with a notable rise in the litigation hazard per R&D dollar spent over the final decades of the 20th century. Lerner (1995) finds that new biotechnology firms avoid patenting in technology subclasses where they are more likely to face costly litigation, even if those subclasses offer attractive awards, and Lanjouw and Lerner (2001) observe that preliminary injunctions by large firms depress small firms’ R&D activities. Similar concerns are raised by studies that focus on the activity of non-practicing entities. These NPEs have accounted for a growing share of patent litigation in the past decade, as will be discussed in Section 5.3. Putting the pieces together, the distortions arising under the current regime are undermining the incentive to innovate itself.

23 Data availability restricts the period of analysis to the 2000s.
Interestingly, the recent pickup in litigation cases is by no means the first or strongest wave of its kind. Indeed, many of today’s debates and arguments are a repeat of history—perhaps unsurprisingly—as exemplified in this quote:

_There are a considerable number of patents issued annually from the Patent Office which are of no force or value except for blackmailing and for interfering with the business of parties competing with their owners... [These patents] do not cover practical machines, but contain principles upon which other more practical inventors have built, and which are infringed by the other patent devices, and are good for nothing except to be bought and speculated upon by those who are justly called patent sharks._

---

J. H. Raymond, Secretary and Treasurer of the Western Railroad Association

In fact, litigation activity was more prevalent from the mid- to late 19th century when measured against the patents in force. Beauchamp (2021) estimates that the number of patent cases per 1000 patents in force was about 30 in 1850 and still around 5 by 1870. For comparison to the recent surge, the same measure peaked at about 2.7 in 2013. Still, the ratio dropped sharply toward the end of the century with the help of certain legal changes, which may cast light on potential remedies to today’s challenges.

Beauchamp (2021) highlights two notable changes that paved the way for the relative abatement of litigiousness in the 19th century. First, legal action in the 1880s limited the practice of term extensions and reissues, hindering excessively lengthy protection of patents and their use for infringement claims. Second, a more gradual shift was the clearer demarcation of patent claims, clarifying the breadth and scope of the patent and decreasing the chances of overlapping claims. These observations appear highly relevant for today’s challenges considering the aforementioned issues. Accordingly Jaffe and Lerner (2004) and Bessen and Meurer (2008a) both urge practitioners to improve the notice of patents. Moreover, the common practice of “continuing applications,” which allows one to extend the patent claim over time, resonates with the problems seen more than a century ago, likely inviting a similar remedy.

Of note, Beauchamp (2021) acknowledges that technological changes and the ensuing shifts in economic organization likely contributed to the retrenchment of litigation activity more than a century ago. Over time, horizontal and vertical integration caused production to consolidate in larger firms at the expense of many smaller ones, leaving fewer participants in the production chain.

---

24 Testimony before the Committee on Patents of the US Congress (US Senate 1878, pp. 123, 230).
Among other things, this shift in the production landscape made collective action in settlement negotiations easier and, more importantly, reduced the number of disputes concerning the ownership of the economic surplus captured via patents, as a larger share of patents remained within a large organization instead of being dispersed across many small ones. The case today is to some extent reminiscent of the earlier period; for instance, Hall and Ziedonis (2001) document the rise of the many specialized, small design firms in the software industry. That said, while the shift toward larger organizations may have helped reduce litigation activity, it also implied higher market concentration. As our theoretical analysis underscores, this rise in concentration is a cause for concern in and of itself. Today, solving the litigation problem through consolidation would only further exacerbate the dynamism problem. Suffice it to say that there is clear historical precedent for new waves of technological change and industrial organization prompting the need to revisit and adapt certain tenets of the nation’s intellectual property protection regime.

Finally, the increase in litigation activity has more recently been accompanied by a surge in suits brought by non-practicing entities, the subject of the next section.

5.3 NPE activity

The share of cases brought by NPEs in total patent cases has dramatically increased in the past decade (Figure 14a), which has attracted the attention of both policymakers and researchers. Indeed, a 2013 White House report titled “Patent Assertion and U.S. Innovation” focuses specifically on patent assertion entities (PAEs)—NPEs that acquire patents chiefly to collect royalties, rather than to facilitate knowledge transfer or technology commercialization, and, most crucially, comprise the largest group among NPEs. In this section, we look closer at recent developments concerning NPE activity.

To start, the surge in NPE activity around the turn of the past decade begs an examination as to the determinants of such a dramatic shift. In a recent study (Miller, 2018), Stanford Law School researchers examined a large random sample of litigation lawsuits since 2000 by carefully differentiating between different types of litigants. The findings suggested that most of the pickup in litigation cases around 2010 is attributable to PAEs. Interestingly, however, the study shows that the number of defenders per lawsuit decreased dramatically in PAE cases, while Cotropia et al. (2014) observes that the number of unique plaintiffs has not increased between 2010 and 2012.

---

25 A distinction between NPEs and PAEs is to be made. NPE is a broader concept that includes universities, technology development companies or even individual inventors who monetize on others’ use of their patents. In our discussion, we use PAEs to refer to entities whose main goal is to create patent portfolios to obtain licensing fees, as adopted by many scholars (Miller, 2018).
These observations point to a unique policy change that likely explains most of the surge in cases: the introduction of the America Invents Act (AIA) in 2011, which limited the scope for joining accused infringers in an action (the joinder provision). Consequently, separate actions need to be taken against different alleged infringers, inflating the number of lawsuits. Still, by looking at distinct defendant-lawsuit pairs (focusing on the number of defendants rather than lawsuits), Miller (2018) finds that there has been a notable increase in distinct patent disputes between the years 2000 and 2015, as shown in Figure 14b. Critically, most of the rise in total pairs in the early 2010s are accounted for by increased PAE activity (even with some relative growth preceding the AIA).

In theory, NPEs can help facilitate a better allocation of patents, acting as middlemen in the secondary market and improving its efficiency, which in turn bolsters innovation and economic growth (Akcigit et al., 2016). Firms that fail to commercialize their inventions may monetize them by selling them to NPEs, who can then license these patents on a larger scale to firms or inventors that need the underlying technology for their own innovative endeavor, thereby enabling the match between the patent and its best use. However, recent work casts doubt on the fulfillment of this premise in general. In their survey, Feldman and Lemley (2015) document that innovation based on NPE-licensed technology is a rare phenomenon. Based on a survey covering more than 300 venture capital (VC) firms and venture-backed startups, Chien (2014) reports that the option to monetize on patents via NPEs is not an incentive for innovation, echoing Feldman (2013), who documents that few startups sell their patents to NPEs. Moreover, most VCs do not consider NPEs as a viable exit option for failed portfolio companies.

---

26 For instance, Abrams et al. (2019) document that the probability of a patent being sold to a NPE increases if the technology covered by the patent is technologically more distant from the core business of the firm.
Recent studies also suggest that business strategies of NPEs are not in general conducive to innovative activity. Chien (2014) reports that relatively small firms comprise a large target group: 66 percent of unique defendants in cases involving PAEs earn less than $100mn in annual revenue. Chien (2014) also observes strategic litigation behavior: targeted startups are usually caught off guard at vulnerable times such as “the eve of a funding or acquisition event.” Cohen et al. (2019) echo this finding, documenting opportunistic NPE behavior in a larger sample. They also document that NPEs in general target firms with large cash holdings or positive cash shocks, a distinct motive when compared to other litigant types and one used exclusively by a certain type of NPE, namely the “patent aggregators.”

Recent evidence also suggests that opportunistic NPEs utilize patents of lower quality or disputable nature. While earlier work based on small samples of NPE patent portfolios did not find an indication of NPEs making use of lower-quality patents (e.g., Shrestha, 2010), studies based on larger samples challenge this finding. For instance, Feng and Jaravel (2020) document that a large part of NPE portfolios consist of patents granted by more lenient examiners; Miller (2013) finds that about 60 percent of NPE patents have at least one invalid claim; and Abrams et al. (2019) document that NPEs acquire more litigation-prone patents. In addition, Cohen et al. (2019) show that, on average, NPEs repeatedly assert lower-quality patents (see also Love, 2013), and increasingly so near the expiration date. Evidence suggests that NPEs engage in opportunistic forum-shopping, as well, with one district court eventually hearing more than 25 percent of all patent cases in the United States—symbolic of growing distortions in the system (Anderson and Gugliuzza, 2021).

As could be expected, these practices impose substantial monetary costs on defendants in litigation cases brought by NPEs, especially for smaller firms, with losses from NPE suits totaling over 25 percent of annual industrial R&D investment (Bessen and Meurer, 2013). In a more systematic examination of causal impact, utilizing a difference-in-differences specification, Cohen et al. (2019) compare defendants that lose a case (or agree to settle) with those that prevail and show that the “loser” group, following the decision, decreases their innovative effort considerably relative to their counterparts in the “winning” group, with losing companies reducing their research and development by $163 million in the two years after the litigation (Cohen et al., 2019). In addition, Abrams et al. (2019) estimate that the acquisition of a patent by an NPE causes innovation by downstream innovators to decline, as the risk of a follow-up innovation by these entities becoming the subject of an infringement case rises when the patent of the original innovation is acquired by an NPE.
6 Conclusion and Policy Discussion

To summarize, our investigation starts with the premise that sustained growth in aggregate productivity necessitates a healthy degree of business dynamism, which has been losing its pace in the United States over the past several decades. The theoretical investigations in our broad research agenda hint at factors that tilt the competitive playing field in favor of established firms, with a prominent contributing factor being a slowdown in knowledge diffusion in the U.S. economy. Patent concentration, which can affect diffusion, has risen over the past several decades with a concurrent surge in patent litigation cases and an increased involvement of NPEs in these cases. While this surge reflects in part technical and legal factors, academic research and our empirical findings indicate that the rise in litigation and NPE activity is plausibly associated with a worsening business climate for inventive activity, with the largest burdens being shouldered by smaller and upstart firms. Several recent studies also provide convincing analysis documenting causal relationships (e.g., Cohen et al., 2019), though the exact contribution of these shifts to declining business dynamism remains yet to be addressed systematically.

Where do our theoretical and empirical findings leave us as to business dynamism, knowledge diffusion, and patents? A direct implication is that to enhance dynamism, policies should focus on enhancing the diffusion of knowledge and technologies from the frontier firms to the rest even as they safeguard the incentive for all types of firms to innovate. To do so, it is imperative to address issues surrounding the excesses and abuses of the patent system that may stifle the transfer of knowledge. For instance, Galasso and Schankerman (2015) find that invalidation of a patent leads to a 50 percent increase in subsequent citations to the focal patent, especially when the invalidated patent was held by a large firm. Concerns regarding business method and software patents also suggest that the issues may be even more egregious for the expanding digital economy, considering that digitally-intensive sectors are more prone to consolidation of market power and litigation. Fortunately, potential solutions to alleviate increased litigation and NPE activity would likely ameliorate broader issues related to patents.

As the discussion in Sections 5.2 and 5.3 suggest, the main issue is to reduce uncertainty surrounding the validity of patents and their claims. Enhancing the definition of patents’ boundaries is a priority agreed upon by prominent scholars (Jaffe and Lerner, 2004; Bessen and Meurer, 2008a; among others). This requires tighter pre-grant examination and post-grant reexamination processes. Considering the burden these practices may entail for the patent office, Jaffe and Lerner (2006) propose a selective reexamination process that focuses on the most prominent patents.

---

27 Miller (2018) estimate that about 80 percent of litigation cases in software involve NPEs.
The 2011 America Invents Act indeed takes steps in such a direction, establishing new post-grant review procedures. This provides a less expensive and faster way to challenge the validity of a patent, thereby alleviating, at least in part, the disadvantageous legal position of defendants in litigation suits.

While our discussion emphasizes problems surrounding the patent system, it is worthwhile to mention that the U.S. system of patent examination was introduced in 1836 and has since been the reference system for many countries around the world. Measures that improve the patenting landscape would certainly help rekindle the U.S. model, which has democratized innovation and enabled the unprecedented technological progress witnessed over the past two centuries (Khan, 2005).
References


