The Compounding Effect of Experience in Innovation Dynamics

Saleh Zakerinia *

November, 2023

Please click here for the latest version

Abstract

This paper investigates how firms capitalize on their accumulated experience to drive impactful innovations, a vital aspect of corporate strategy and economic growth. Despite acknowledging that innovation and experience drive productivity, the specific ways that accumulated experience enhances a firm's innovation remain largely unexamined. This issue is particularly relevant in modern economies where dominant companies may exploit their vast experience to innovate and maintain market dominance. Motivated by this gap, the study constructs a dynamic framework to quantify how accumulating experience can enhance the scale, scope, and quality of innovation over time. The model captures how firms refine beliefs about unobserved ability through various signals, facilitating estimation of learning processes. The analysis applies this knowledge accumulation framework to innovations in artificial intelligence (AI), given the field's rapid emergence and policy significance. Structural estimation on patent data provides unique evidence on multi-dimensional learning effects in AI. The findings reveal that firms initially overestimate their innovation capabilities, but this perception adjusts downward with experience. Moreover, uncertainty about their capabilities decreases significantly over time. Counterfactuals assess prospective impacts of privacy and antitrust policies on AI innovation. The findings highlight the importance of shaping strategies and policies that promote innovation driven by experience across various industries.

Keywords: Innovation Dynamics; Experience Accumulation; AI Innovation; Privacy Policies ; Antitrust Policies

*Cornell University; `mz497@cornell.edu`. I am indebted to Nathan Yang, Chris Forman and Jura Liaukonyte for their mentorship and encouragement throughout this project. I thank Vithala Rao, Hong Luo, Sylvia Hristakeva, Emaad Manzoor, and Benjamin Leyden for invaluable comments and discussions. Any errors in this work are my own responsibility.

1 Introduction

How do firms leverage their accumulated experience to create more impactful innovations? This critical question sits at the forefront of corporate strategy, given that innovation is a key driver of long-term economic and productivity growth. Economists have long seen innovation, with its new technologies and methods, as crucial for the economy's future (Schumpeter, 1943). At the same time, companies often see immediate productivity gains by improving what they already do, drawing from their own learning and experience (Arrow, 1962). This learning, happening in every part of a company, helps to make work more efficient and cost-effective.

Yet, there is much we don't know about how learning affects a company's ability to innovate. This gap in our knowledge becomes even more important in today's economy where a few big companies dominate the market. These companies may be leveraging their extensive experience to foster innovation and maintain a competitive edge. For example, Samsung has successfully drawn on its extensive background in electronics to take the lead in emerging areas such as smartphone technology and digital displays.

Innovation operates differently from regular production. It involves knowledge that is dynamic, not stuck in one area, and can combine with ideas from different fields (Dyer et al., 2009). Tesla's application of its expertise and innovations from electric vehicles to advancements in energy storage and solar technology is a prime example of how experience in one sector can propel innovation in another.

The capacity for adaptation and creation in innovation transcends the constraints of physical resources that often limit traditional production. Experience impacts innovation in multifaceted ways: it contributes directly to new developments and indirectly through knowledge transfer across projects. This study delves into these nuanced effects, examining how a company's accrued experience informs its innovative capabilities. By unraveling the complexities of this influence, the research highlights the pivotal role of experience in driving successful innovation. Understanding this dynamic is essential, as it equips firms with the strategies to harness their experiential knowledge for impactful innovation, providing a competitive advantage in a rapidly evolving economy.

To delve deeper into this complex issue, this study will explore the growing field of Artificial Intelligence (AI) as a context. AI is a rapidly expanding sector where innovation is not only prevalent but pivotal. It stands out in the knowledge-intensive industrial landscape, with its transformative impact extending across the entire economy (Chui et al., 2023). The swift advancements in Large Language Models such as GPT-3 and BERT underscore the significant role of experiential learning in these technologies' evolution (Bommasani et al., 2021). The focus on AI allows me to contribute to the broader understanding of experience in innovation, while also offering timely insights that could inform policy in one of today's most dynamically advancing fields.

This paper introduces two significant advancements to the field: a novel framework that incorporates information acquisition into the learning process, enabling firms to navigate uncertainties in AI innovation and address endogeneity issues, a notable departure from traditional deterministic models. Secondly, it broadens the scope of organizational learning research by evaluating the quality and scope of a firm's innovation experience, moving beyond the conventional focus on cost-efficiency and filling a gap in existing literature. These contributions offer a more comprehensive understanding of firm growth and innovation dynamics.

I develop an empirical dynamic structural model to examine the impact of learning and various policies on AI innovations at the firm level. This model captures the dynamic nature of the decision-making process, where current innovation decisions have long-term implications for future states and beliefs regarding their ability to innovate in AI. The primary objective of this model is to quantitatively assess the influence of firms' experience in innovations on their future innovations through the process of learning. In particular, the model accounts for how firms learn about the underlying unobservable innovation ability as they accumulate experience and gather information from the depth and breadth of their portfolio of innovations. By incorporating these dynamics, the model provides insights into the mechanisms through which learning and various policies shape AI innovation outcomes at the firm level.

The model is dynamic in nature to capture the temporal aspects of firms' decision-making and learning processes. It recognizes that firm actions and beliefs in one period have implications for future states and beliefs. By incorporating this dynamic perspective, the model captures the feedback loop between firms' experiences, learning, and subsequent decision-making. This temporal dimension allows me to examine how firms' evolving beliefs and accumulated experience shape their performance and strategic actions over time. Ultimately, the dynamic nature of the model enables a nuanced analysis of the complex dynamics within the AI innovation landscape.

My model, inspired by Jovanovic and Nyarko (1996) and Ching, Erdem and Keane (2013), captures the dynamic learning process of firms in AI innovations. Initially, firms lack awareness of their ability in this domain and form beliefs about their ability based on limited information. These beliefs serve as their initial perceptions of their unobservable ability. However, as firms accumulate experience and receive signals related to the depth and breadth of their portfolio, they continually update and evolve their beliefs. This iterative learning process enables firms to enhance their understanding of the unobservable ability and make more informed decisions over time.

In addition to the firm-level dynamics, the model recognizes the presence of industry-level profitability in innovation, which is unobservable to econometricians but observable to firms. This variable intends to capture demand shocks as well as learning spillover from other firms. While I do not have direct data on industry-level innovation profitability, the model considers the serial correlation of industry innovation profitability over time. This recognition enables me to account for the broader industry context in which firms operate and learn. The industry-level dynamics, coupled with the firm-level learning, contribute to a comprehensive understanding of the evolution of innovation in AI.

The proposed model diverges from the conventional literature on learning-by-doing by explicitly incorporating the process of information acquisition, rather than relying on a deterministic function to describe productivity growth. This comprehensive framework analyzes how firms' experience and learning influence their performance and decision-making in the dynamic landscape of AI innovations. By considering both firm-level and industry-level dynamics, the model captures the multifaceted nature of learning in innovation, encompassing the accumulation of experience, active information acquisition, and strategic actions. This integrated approach provides insights into the complex interplay between firms and the evolving industry, shedding light on factors that drive success and competitive advantage. It facilitates informed analysis and decision-making, offering a holistic understanding of the innovation landscape in AI and its dynamics.

I estimate the dynamic parameters of the model by leveraging the Artificial Intelligence Patent Dataset (AIPD) (Giczy, Pairolero and Toole, 2022), which covers AI patents issued from 1976 to 2020. The AIPD identifies patents that incorporate eight different AI technology components, including machine learning, natural language processing, computer vision, speech, knowledge processing, AI hardware, evolutionary computation, and planning and control. To enhance the analysis, I merge this dataset with the Duke Innovation & Scientific Enterprises Research Network (DISCERN) dataset (Arora, Belenzon and Sheer, 2021), which connects innovation data to Compustat firms. Additionally, I incorporate the PatentsView forward citations dataset (PatentsView, 2022) and the KPSS dataset (Kogan et al., 2017) to capture the value of patents. This comprehensive approach allows me to identify AI patents at the firm level and assess their respective values. Furthermore, I calculate the time intervals between AI patents within each firm, using this as a proxy for the cost associated with patent development. By utilizing this rich dataset, I am able to estimate the proposed model.

I find that as firms gain more experience in AI innovation, several positive outcomes are observed. Firstly, the depth or value of their innovation increases, indicating that their AI technologies become more valuable and impactful. Secondly, the scope of their innovation expands, implying that firms are able to come up with a broader range of AI innovations. This implies that they are not limited to specific applications or narrow areas, but rather have the capability to venture into diverse domains within the AI landscape. This expansion in scope signifies the potential for firms to address various challenges and capitalize on new opportunities, further solidifying their position and potentially creating barriers to entry for competitors. Lastly, the cost of their innovation decreases, reflecting improved efficiency and cost-effectiveness in AI development. As firms accumulate experience and knowledge in AI innovation, they become more adept at optimizing their processes, streamlining operations, and reducing costs associated with developing AI technologies. This cost reduction indicates the ability to achieve higher levels of innovation at a lower investment, making AI development more accessible and economically viable for firms. Additionally, decreased costs may also contribute to increased competitiveness and market advantages for firms engaged in AI innovation.

The learning process results show that firms initially have optimistic, but uncertain, perceptions of their AI capabilities. As companies gain experience, their perception of their capabilities decreases. Additionally, their uncertainty about these capabilities diminishes quickly, leading to a clearer and more accurate understanding of their strengths. The study also reveals that in this learning process, the quantity of past patents is a more effective signal of capability than the diversity of these patents, suggesting that the scale of innovation is a more critical factor for learning than the scope. Furthermore, the analysis of the payoff function's estimated parameters shows that the value of granted patents exerts the most significant influence on per-period payoffs. This finding underscores the crucial importance of impactful and influential inventions in determining a firm's payoffs.

In the counterfactual simulation, the examination of merger and acquisition activities sheds light on their impact on firms' positions in AI innovations. The analysis reveals that when a firm acquires another firm, there is a notable increase in its holding of AI innovations. This acquisition-induced increase in AI holdings subsequently enhances the firm's experience in AI, resulting in improved depth and scope of their innovations, as well as reduced costs associated with AI development. These positive outcomes generate stronger signals, accelerate the learning process, and ultimately contribute to enhanced profitability. The findings underscore the potential benefits of mergers and acquisitions in bolstering firms' AI capabilities and their capacity to drive innovation in this rapidly evolving field. This perspective provides an interesting counterforce to arguments made in Cunningham, Ederer and Ma (2021) around killer acquisitions, in which incumbents acquire innovative targets solely to discontinue the target's innovations and preempt future competition. In contrast, the simulation here suggests that in the presence of learning effects, acquiring firms may have incentives to continue developing the target's innovations to enhance their own AI capabilities over time. While more research is needed, the counterfactual analysis indicates that with learning, mergers and acquisitions can strengthen AI innovation across firms rather than stifling it. In a parallel vein, the counterfactual simulation also investigates the consequences of antitrust policies aimed at dismantling big firms. While M&A activity showed potential for strengthening innovation through firm aggregation (Andrade, Mitchell and Stafford, 2001), break-ups display mirrored effects of weakened AI capabilities from forced fragmentation (Scott Morton et al., 2019). Specifically, undergoing a breakup diminishes a firm's AI experience, resulting in weaker signals, slowed learning, and reduced profitability (Giarratana and Fosfuri, 2007). This highlights potential unintended adverse impacts of antitrust measures on AI advancement. However, the symmetry between combining and splitting firms is not perfect. Knowledge accumulation in AI exhibits learning and increasing returns to scale, allowing aggregated entities to leverage shared data, resources, and talent to rapidly compound innovations (Agrawal, Gans and Goldfarb, 2018; Kaplan et al., 2020). Breaking up this ecosystem risks disrupting positive spillovers and slowing learning. But concentration also increases incentives for anti-competitive killer acquisitions (Cunningham, Ederer and Ma, 2021). Appropriately balancing these tensions is complex. Overall, the counterfactual analysis cautions antitrust policies could negatively impact AI development absent careful implementation attuned to innovation incentives.

Moreover, I conduct an examination of the impact of information privacy policies on AI innovations, aiming to address growing concerns over data privacy's influence on AI development and adoption. While some argue privacy regulations may slow AI progress (Chiou and Tucker, 2017; Goldfarb and Tucker, 2011), others suggest they incentivize privacy-preserving innovations (Jones and Tonetti, 2020). By analyzing the relationship between privacy policies and AI innovations through counterfactual simulation, this paper provides empirical insight into this complex issue. The findings shed light on how privacy regulations shape firm outcomes in AI, illuminating tradeoffs between competing forces like innovation, competition, and privacy. This exercise contributes to the literature on privacy and AI by investigating how privacy policies relate to positions in AI innovations. The results may inform forward-looking governance that fosters AI advancement while upholding consumer privacy.

Overall, these counterfactual simulations offer valuable insights into the potential effects of different policies and factors on firms' positions in AI innovations. They provide a deeper understanding of the complex dynamics and interdependencies within the AI landscape, helping policymakers and industry stakeholders make informed decisions and strategies.

This paper makes a contribution to the extensive literature on organizational learning, particularly focusing on the dynamics of learning in the context of innovation rather than traditional production settings. Numerous studies have confirmed the existence of learning-by-doing in specific production settings, highlighting its impact on factors such as knowledge acquisition, production costs, and product quality in traditional production environments. For instance, Epple, Argote and Devadas (1991) investigate organizational learning within different shifts of a car factory, while Irwin and Klenow (1994) document learning-by-doing in the semiconductor industry and examine its strategic use by firms. Benkard (2000,

2004) studies the role of learning-by-doing in reducing production costs of commercial aircraft, and Levitt, List and Syverson (2013) highlight how learning-by-doing improves the quality of production in an automobile assembly plant.

In contrast, as we move into an increasingly digital and knowledge-based economic landscape, understanding how firms learn to innovate becomes critically important. Most existing studies concentrate on learning in physical production settings, examining factors like capital accumulation and labor efficiency. These frameworks often lack the necessary nuance to capture the complexities and uncertainties involved in innovation within knowledge-based industries. Though Jain (2013) explores learning-by-doing in the specific context of biochemistry patents, the study is still limited to operational efficiencies and labor costs, rather than the broader implications of learning on innovation.

This paper distinguishes itself in two significant ways. First, it introduces a novel framework that incorporates information acquisition as a key element of the learning process. Unlike traditional models that rely on deterministic production functions, my approach allows firms to account for uncertainties associated with innovation in AI. By tracking trends in firms' AI patent portfolios, this study empirically explores how learning can mitigate such uncertainties, offering a more nuanced understanding of learning dynamics in innovation. This approach also effectively deal with the endogeneity issues commonly associated with traditional models, as discussed by Thompson (2012).

Secondly, this paper expands the scholarly conversation on organizational learning by moving beyond traditional metrics of cost-efficiency. Instead, it incorporates broader evaluative criteria such as the quality and breadth of a firm's innovation experience. By doing so, it builds on seminal works in the pharmaceutical sector that have examined scale and scope in drug discovery (Henderson and Cockburn, 1996; Cockburn and Henderson, 2001). Importantly, this research enhances both theoretical understandings of learning dynamics in knowledge-intensive industries and provides practical, policy-relevant insights. Such contributions are particularly vital in fast-paced, evolving fields like AI, where the study's frameworks and findings offer more nuanced tools for addressing issues around antitrust and privacy regulations. Additionally, this paper brings a methodological sophistication to bear by introducing an empirical framework that accounts for complexities introduced by unobserved state variable dynamics, especially those arising from multi-dimensional learning and knowledge production. Incorporating a learning mechanism into models where firms are inherently forward-looking presents significant challenges. This is non-trivial because forward-looking firms operating in a setting of multi-dimensional learning must navigate a more complicated landscape of strategic interactions and potential outcomes. My framework effectively addresses these complexities, providing a robust model for examining how learning in knowledge-intensive sectors influences firm behavior and market dynamics. This, in turn, enhances the paper's theoretical rigidity and empirical validity, offering a more holistic understanding of learning dynamics in innovation-driven industries.

This paper also contributes to the emerging literature that examines the effects of privacy policies on firms. A substantial body of research in this field focuses on how these policies restrict digital marketers' ability to target specific customer profiles and how firms can adapt their strategies in response to these restrictions (Ghosh, 2018; Aridor, Che and Salz, 2020; Godinho de Matos and Adjerid, 2021). Another strand of literature investigates the impact of privacy policies on market concentration, firms' investment decisions, and innovation activities. For instance, Jia, Jin and Wagman (2018) examine the short-term effects of GDPR on investment in new and emerging technology firms, while Peukert et al. (2021) document changes in websites and the web technology industry following the introduction of GDPR. Johnson, Shriver and Goldberg (2023) find that the implementation of the European Union's GDPR leads to a decrease in the usage of website vendors, indicating a reduction in third-party tracking and data collection. However, they also observe an increase in market concentration among technology vendors that offer support services to websites. In a relevant study to AI, Bessen et al. (2020) evaluate the impact of GDPR on AI startups, finding that training data and frequent model refreshes are crucial for startups relying on neural nets and ensemble learning algorithms. They also observe that firms with European customers are more likely to create new positions or reallocate resources to comply with GDPR.

Building upon this literature, my paper contributes by empirically modeling and estimating the effects of privacy policies on firms' ability to innovate in AI, the quality of their innovation, and the breadth of their AI innovations. Additionally, this study investigates the implications of these policies on profitability in various heterogenous firms. By examining the effects of privacy policies on firms' profitability across different contexts and considering the heterogeneity among firms, this research sheds light on the diverse outcomes and impacts of these policies in the AI landscape. Through a comprehensive analysis, this study provides valuable insights into the multifaceted relationship between privacy policies and firm-level outcomes in the context of AI innovations.

The paper is positioned within the literature that examines the relationship between digital capital and market concentration. Numerous studies have observed an increase in market concentration over recent decades and attribute it to the emergence of intangible assets, e.g. (Crouzet and Eberly, 2019; Tambe et al., 2020). AI, as a General Purpose Technology (GPT), necessitates substantial investments in firm-specific human and organizational capital to generate value (Brynjolfsson, Rock and Syverson, 2021). Consequently, firms that have accumulated significant intangible assets in AI may gain market power (Babina et al., 2021).

This paper argues that the learning mechanism serves as an additional avenue for firms to strengthen their position in AI innovation, leading to an augmentation of their intangible assets. This, in turn, can result in higher firm valuations and increased market power. By investigating the learning-by-doing process, the study contributes to understanding how firms leverage their accumulated knowledge to enhance their competitive advantage in the AI landscape.

2 Data

In this study, the data is sourced from various databases to capture the relevant information for analysis. The Compustat dataset offers comprehensive data on publicly traded companies in the US, including firm-level and industry-level information. To identify AI-related patents, the Artificial Intelligence Patent Dataset (AIPD) provided by the US Patent and Trademark Office (USPTO) is utilized (Giczy, Pairolero and Toole, 2022). This dataset utilizes natural language processing algorithms to categorize patents into eight subgroups related to AI. To enhance the dataset, it is merged with the Duke Innovation & Scientific Enterprises Research Network (DISCERN) dataset (Arora, Belenzon and Sheer, 2021), which connects innovation data to the firms in the Compustat database. Additionally, the Patentsview dataset, a comprehensive repository of patent applications filed with the USPTO since 1975, is merged with the dataset to access forward citations and proxy the value of patents (PatentsView, 2022). Moreover, the dataset is enriched with the economic importance measure developed by Kogan et al. (2017), which combines patent data with stock market responses to determine the economic significance of each innovation. Table 1 presents a overview of notable patented innovations in different categories of AI. The table showcases significant AI patents across eight different categories of AI.

Table 1: Notable patented innovations in AI

Patent Type	**Notable Patented Innovation**	**Patent Holder**	**Patent Number**
Knowledge Processing	Cognitive information processing system environment	Tecnotree	US11676042B2
Speech	Digital assistant providing whispered speech	Apple	US20170358301A1
AI Hardware	Hardware accelerated neural network subgraphs	Microsoft	US20190286972A1
Evolutionary Computation	Evolutionary search for robust solutions	Honda	US7783583B2
Natural Language Processing	Method to allow for question and answer system to dynamically return different responses based on roles	IBM	US10754969B2
Machine Learning	Communication optimizations for distributed machine learning	Intel	US20220245454A1
Computer Vision	Computer vision based security system using depth camera	Apple	US9396400B1
Planning/Control	System and method for realtime community information exchange	Google	US9275544B2

To assess the importance and depth of patents, forward citations and economic importance measures are employed. Forward citations, which refer to the number of times subsequent patents cite a particular patent, are used to capture the depth of knowledge and potential economic value of patents. The rationale behind using forward citations is that they often highlight the underlying knowledge on which subsequent inventions build (Hall, Jaffe and Trajtenberg, 2005). To address truncation issues, only forward citations received within the first five years are considered. To account for the correlation between forward citations and economic importance, the number of forward citations for portfolio of AI patents is calculated annually for each firm, serving as a metric for the firm's depth of knowledge. This measure is correlated with the economic importance measure developed by Kogan et al. (2017).

The breadth of knowledge is evaluated using the Herfindahl-Hirschman Index (HHI) metric, following the methodology outlined by Henderson and Cockburn (1996). This metric calculates the concentration of AI patents across eight subcategories and provides an indication of the firm's breadth of knowledge as shown in the following equation.

$$H = \sum_{j=1}^{8} (\frac{\text{number of AI patents in category j}}{\text{Total number of AI}})^2 \tag{1}$$

The diversity of the AI patent portfolio's value is quantified by an HHI calculation, which assesses the distribution of AI patent value among eight distinct categories. This measurement provides insight into the extent to which a company's AI value spans various AI categories, as illustrated in the equation below.

$$B = \sum_{j=1}^{8} (\frac{\text{value of AI patents in category j}}{\text{Total value of AI patents}})^2 \tag{2}$$

The speed of innovation is determined by calculating the average time between the filing dates of two successive AI patents for each year. This measure quantifies the pace at which firms innovate in the AI domain.

The dataset is constrained to the top 100 firms with the highest number of AI patents filed between 1990 and 2016. The choice of 1990 as the starting year is deliberate, as there was relatively limited AI activity before that point. 1990 marks the nascent stage of AI, where its practical applications began to materialize. Furthermore, during the 1990s, AI was in its infancy, and researchers and firms were just beginning to grasp its vast potential. The endpoint in 2016 was chosen due to data limitations; the last year of available data in the AIPD is 2020, and considering the typical delay in patent publication, 2016 provides a reasonable cutoff. It's important to note that patent processing can take several years, and the USPTO typically takes 18 months to publish applications. Table 2 presents the summary statistics of this dataset.

Table 2: Summary statistics

	mean	sd	min	max	count
AI patents	84.03	259.12	0	4,746	2,700
Machine learning patents	8.90	29.98	0	687	2,700
Evolutionary computation patents	2.60	10.86	0	241	2,700
Natural language processing patents	10.58	45.84	0	834	2,700
Speech patents	5.39	19.55	0	342	2,700
Vision patents	19.58	54.20	0	984	2,700
Knowledge processing patents	48.07	169.99	0	3,298	2,700
Planning and control patents	51.32	180.88	0	3,367	2,700
Hardware patents	35.78	135.67	0	2,544	2,700
Cumulative number of patents	700.74	2,418.85	0	45,068	2,700
Diversity of portfolio (HHI index)	3,008.31	1,546.51	1,378	10,000	2,213
AI Patent portfolio value ($M)	962.04	3,340.58	0	66,105	2,700
Diversity of portfolio value (HHI index)	3,130.29	1,747.29	1,383	10,000	1,967
Number of portfolio forward citations	1,359.95	4,358.22	0	77,001	2,700
Days intervals between AI patents	529.42	361.90	1	6,004	2,213

The table shows that on average, firms in the sample have a relatively large number of AI patents (mean 84.03) as well as a high cumulative number of patents overall (mean 700.74) which shows the prolificacy in AI innovation. There is also considerable variation across firms, with the maximum number of AI patents being 4,746 and maximum cumulative patents being 45,068. The diversity of patent portfolios as measured by the HHI index also varies substantially. Notably, the mean number of days between AI patents is 529, suggesting firms are patenting new AI inventions about every 1.5 years on average. Remarkably, the AI patent portfolio value demonstrates a considerable financial dimension, with an average value of $962.04 million per year. This highlights not only the quantitative strength in patent numbers but also the financial worth attributed to these AI innovations. The range is vast, with portfolio values in each year varying from zero to an impressive $66,105 million. The summary statistics of other key variables like number of patents in different categories, and forward citations are shown as well.

3 Descriptive evidence

In this section, I present compelling empirical evidence supporting the idea of learning within AI patenting. The primary objective is to convincingly demonstrate to the reader that as firms accumulate experience in AI innovations, their quality and quantity of innovation in this domain accelerates. To substantiate this claim, I present insightful data and leverage a reduced-form analysis.

Initially, I illustrate how a firm's innovations in AI accelerate corresponding to the cumulative number of

AI patents they hold. The left panel of Figure 1 visually represents this trend, depicting how the number of AI patent applications in each year for each firm increases exponentially with their cumulative number of filed AI patents. This logarithmic correlation signifies an exponential relationship between the filed AI patents and the cumulative number of AI patents. This relationship is a clear indicator of the escalating ability of firms to file patents as their cumulative experience in AI innovation grows.

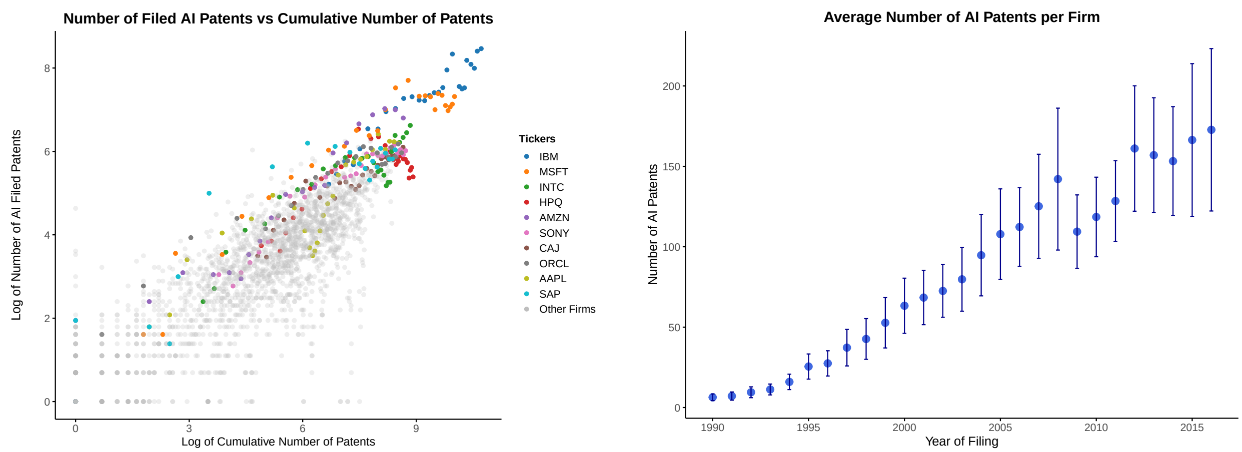

Figure 1: Number of AI patents and cumulative number of patents

Each dot in Figure 1 represents a firm-year value, with a clear emphasis on the top ten firms holding the most AI patents. For instance, it is evident that IBM, Microsoft, and Intel, the leading AI companies, exhibit a strong linear correlation between the log of filed AI patents and the cumulative number of patents. Additionally, the right panel of Figure 1 presents the average number of AI patents filed in each year across firms. A clear increasing trend in AI patent filings is observed, except for the years 2010, 2011, and 2012, primarily due to the recession. This, once again, signifies an exponential growth in filing AI patents, likely attributed to the accumulated experience of firms in the field of AI.

Next, I demonstrate that as firms accumulate more experience, the quality of their AI patent portfolio increasingly improves—an aspect of learning that has been scarcely explored in the literature. The impact of learning on quality is an underexplored dimension. As depicted in the left panel of Figure 2, a clear linear correlation exists between the logarithm of the value of AI patent portfolios held by firms and the logarithm of the cumulative number of AI patents they possess. This observation indicates a progressive enhancement in firms' ability to create a more valuable AI patent portfolio as their cumulative experience grows.

Similarly, the right panel of Figure 2 illustrates how the average value of a firm's patent portfolio has evolved over the years. A significant spike in the average value of the patent portfolio is observed in 2000, followed by a decline in subsequent years due to the dot-com bubble. Nonetheless, a consistent upward

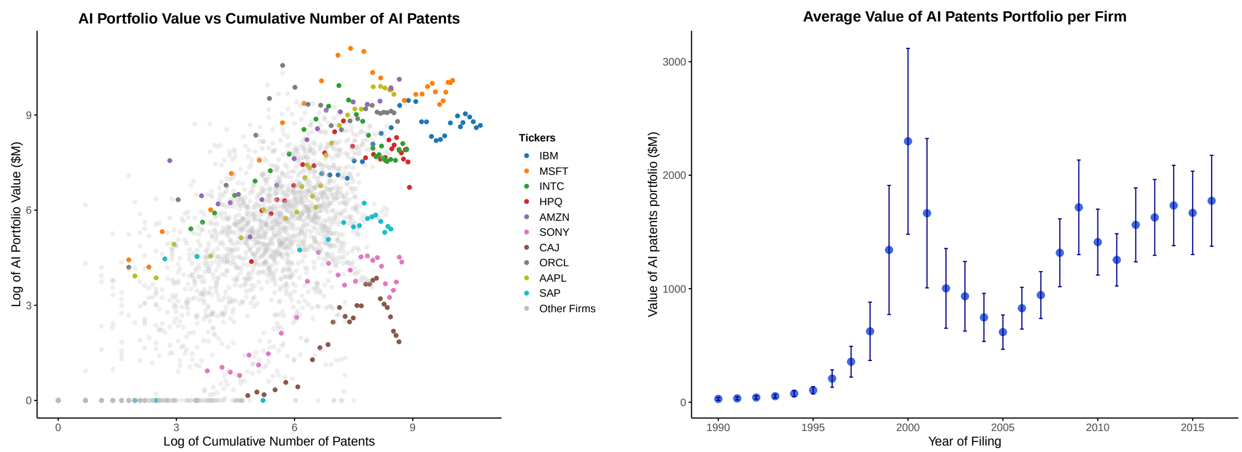

Figure 2: Value of AI patents and cumulative number of patents

trajectory in the average value reaffirms the increasing capability of firms with accumulated experience. Another crucial aspect of learning, which presents a novel perspective, involves examining how the diversity of firms' portfolios evolves as their experience accumulates. This facet of study is inherently intriguing as it offers insights into whether increased experience compels firms to specialize in a specific domain within AI or encourages them to broaden their expertise across various AI domains. This investigation holds significant implications because it can shed light on whether firms are merely diversifying their patent portfolios or if diversification coincides with the generation of high-quality patents across a wide spectrum of AI fields.

The left panel of Figure 3 depicts how the logarithm of the diversity index of AI portfolios, as defined in Equation 1, changes in response to the cumulative number of patents. A notable observation is the nearly linear negative correlation, signifying an exponential decrease in the diversity index as the cumulative number of AI patents grows. It's essential to note that a lower diversity index indicates a more diverse patent portfolio. Similarly, the right panel of Figure 3 showcases a decline with decreasing rate in the diversity index at a decreasing rate, which further underscores the role of learning in this process.

The left panel of Figure 4 delves into the diversity of the value of patent portfolios across different AI fields, as defined in Equation 2. This analysis examines how this diversity changes with the diversity index of firms' AI experience. The diversity index for firms' experience is defined as the HHI of all AI patents filed to date, in contrast to Equation 1, which calculates the HHI index of the portfolio for a specific year. A positive linear trend emerges, demonstrating that diversified accumulated experience leads to greater diversity in the value of the AI portfolio.

The right panel of Figure 4 delves into the shifting dynamics of the value-based diversity within patent portfolios over time. The noticeable decreasing trend, characterized by a diminishing rate, further confirms

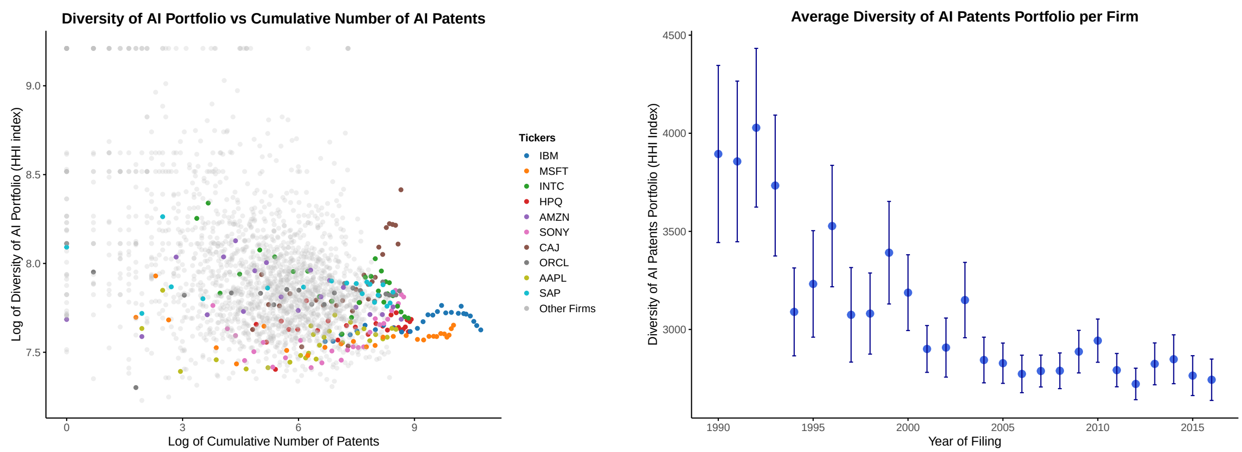

Figure 3: Diversity of AI patents and cumulative number of patents

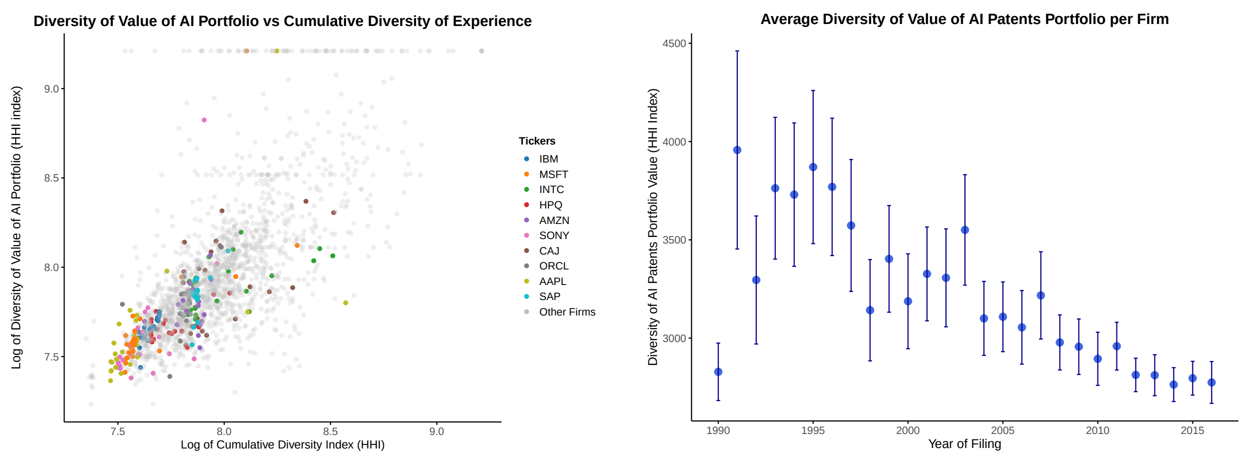

Figure 4: Value-diversity of AI patents and cumulative number of patents

that firms with more accumulated experience demonstrate an ability to enhance the value-based diversity within their patent portfolios. This suggests that as firms delve deeper into the AI domain and accumulate knowledge, they not only broaden their expertise but also strategically enrich the value facets of their patent portfolio. This nuanced understanding of how experience impacts the breadth and value of their portfolio is essential for formulating effective innovation strategies.

Shifting the focus to the cost of innovation, which is captured by the number of days between filing two successive AI patents, the left panel of Figure 5 illustrates that the total number of days spent on the entire AI portfolio per firm appears to increase as they accumulate more experience. However, it's essential to note that this growth in the total number of days spent on the AI portfolio can be attributed to the exponential increase in the total number of patents filed each year as firms gain more experience.

To delve deeper into how experience contributes to the cost of innovation, the right panel of Figure 5

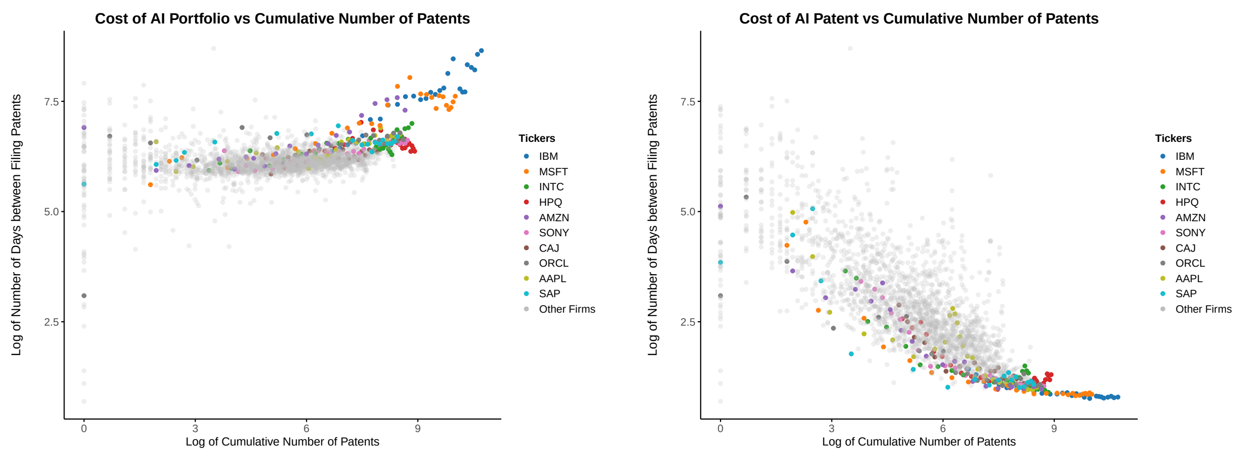

Figure 5: Cost of AI portfolio and patents vs cumulative number of patents

illustrates how the average cost of a single patent changes with the accumulation of experience, in contrast to the total AI portfolio. A discernible decreasing trend is evident, aligning with the hypothesis of learning in AI patent filing. This observation is in line with findings in the learning literature, as seen in studies such as Argote and Epple (1990); Benkard (2000); Levitt, List and Syverson (2013), which demonstrate that labor requirements decrease with the accumulation of experience. Therefore, our data supports the hypothesis of a decreasing cost of producing an AI patent with an increase in accumulated experience in AI.

In the following, I present a reduced-form model that examines the impact of cumulative experience and diversity of cumulative experience on various outcomes, including the number of AI innovations, diversity of AI innovations, average patent value, value-diversity of patents, and average patent cost. The model is described by the following linear regression equation, where y_{it} represents the mentioned output:

$$\log(n_{it}) = \beta_1 \log(e_{it}) + \beta_2 \log(eh_{it}) + \eta_i + \eta_t + \epsilon_{it} \tag{3}$$

Here, the variables $\log(e_{it})$ and $\log(eh_{it})$ correspond to the natural logarithms of the accumulated number of AI patents and the diversity of cumulative patent holdings of the firm, respectively. Additionally, η_i accounts for firm-specific fixed effects, η_t represents time dummies, and ϵ_{it} stands for the iid error term, capturing unobserved factors and random shocks affecting the investment decision.

The estimates of the firm's outputs are presented in Table 3. The results reveal a robust positive coefficient for log(cumulative number) in both log(number of filed patents) and log(mean value of patent), indicating that a 1% increase in the cumulative number of patents is associated with a 0.63% increase in the number of AI patents filed and a 0.37% increase in the average value of patents. Moreover, the increase in the

Table 3: Effect of experience and its diversity on various outcomes

VARIABLES	(1) log(# of filed patents)	(2) log(Diversity)	(3) log(Mean value)	(4) log(Value-diversity)	(5) log(Mean cost)
log(Cumulative number of patents)	0.626***	-0.0586***	0.369***	-0.111***	-0.586***
	(0.0501)	(0.0160)	(0.0822)	(0.0170)	(0.0470)
log(Diversity index of experience)	-0.623***	0.300***	-0.726**	0.934***	0.969***
	(0.223)	(0.0673)	(0.350)	(0.0808)	(0.198)
Constant	5.213***	5.830***	6.064*	1.145	-1.975
	(1.938)	(0.595)	(3.074)	(0.704)	(1.715)
Firm fixed effect	Yes	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes	Yes
Observations	2,221	2,151	2,221	1,967	2,151
R-squared	0.794	0.509	0.577	0.593	0.780

Standard errors are clustered at firm level
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

cumulative number of patents exhibits a negative effect on diversity, value-diversity, and the average cost of patents. It's important to note that diversity value, represented by the HHI index, operates inversely—higher diversity value indicates lower diversity. Consequently, firms with a more diverse cumulative patent holding have the capacity to file a more diverse portfolio of patents, including both a broad range and a varied value perspective. Also, as evidenced in Model (5), this increased diversity in cumulative patent holding decreases the average cost of patents by 0.6%, aligning with the findings of existing learning literature, particularly as reported in Jain (2013).

Another significant aspect is the impact of a more diverse patent holding on the outcomes of interest. This impact is captured by examining the coefficients of log(diversity index of experience). As shown in Table 3, a more diverse cumulative patent holding increases the number of patent holdings and the average value of patents. Additionally, a more diverse cumulative patent holding enhances diversity in both the diversity of the portfolio filed each year and the diversity value of granted patents. Furthermore, it decreases the average cost of patents.

The above evidence from raw data analysis and reduced-form models demonstrates learning within firms across various dimensions of AI patenting, including quantity and quality. However, reduced-form analysis alone cannot conclusively attribute these patterns to learning. A more structural approach is needed to rigorously model the learning process in AI innovation for several reasons.

First, as noted by Thompson (2012) , serial correlation complicates estimation of learning curves in reduced-form regressions. High serial correlation in productivity analyses makes identifying learning effects difficult, as discussed by Olley and Pakes (1996) . Without plausible instrumental variables for cumulative experience, like those used by Benkard (2004), reduced-form models struggle to distinguish learning from mere serial correlation.

Additionally, reduced-form analyses must assume where learning occurs, such as at the production input level as in Argote and Epple (1990). This significantly limits generalizability by imposing a fixed learning process. In contrast, a structural model can abstract away from these micro-level details and take a more flexible approach to modeling learning. By avoiding assumptions on the locus of learning, a structural approach maintains wider applicability.

Moreover, modeling the joint evolution of patent quantity and quality raises endogeneity concerns absent a structural approach. An omitted variable could drive both dimensions, biasing reduced-form inferences about learning. A structural model can formally account for this potential endogeneity by incorporating unobserved heterogeneity into the estimation. This avoids biases and ensures consistency when modeling the interrelated dynamics of patent quantity and quality.

Finally, reduced-form models cannot conduct counterfactual policy analysis or welfare assessment, which are essential to evaluate prospective impacts of polices such as privacy regulations. Understanding how these measures may shape innovation is crucial. Additionally, firms likely innovate with forward-looking motives, strategically investing in patents for future gains. But reduced-form analyses ignore these incentives. A structural model incorporating forward-looking objectives provides vital insights into policy impacts and strategic patenting behavior.

Given these limitations and the suggestive data patterns, a structural investigation of AI learning dynamics is crucial. The proposed model in the next section incorporates forward-looking behavior, Bayesian learning, and unobserved heterogeneity. This comprehensive framework estimates how experience shapes firms' AI innovations while accounting for forward-looking incentives and unobservables. Overall, the reduced-form evidence points to, but cannot definitively prove, learning in AI patenting. A structural approach is indispensable for rigorously quantifying learning, evaluating policies, and capturing forward-looking decision-making. This motivates the model developed in the next section to deliver robust insights into AI innovation and its drivers.

4 Model

In this section, I present a model that aims to quantify firms' innovation processes in AI. Each firm must decide the extent of their innovation, including both the number and breadth of innovations. These decisions are inherently dynamic, as they trigger a learning process that directly impacts the cost, quality, and scope of subsequent innovations. Here, 'learning' pertains to the dynamic nature of AI innovation investments, wherein production costs can decrease, and the quality and diversity of innovations can improve over time. To address questions surrounding the policy effects of antitrust and privacy regulations on AI innovations, I employ a structural model. This type of model is well-suited for studying the effects of these policies, which are yet to be implemented in the US and have the potential to significantly impact AI innovations. Through counterfactual scenarios, I assess the potential impact of these policies on AI innovations, providing insights into their implications for the field.

There are I firms, indexed by $i = 1, 2, ..., I$, that are actively involved in AI innovation. The level of AI innovation is measured using AI patents, which are categorized into eight distinct subcategories: knowledge processing, speech, AI hardware, evolutionary computation, natural language processing, machine learning, computer vision, and planning/control. At the beginning of each time period t, firms make two key decisions. First, they determine the total number of AI innovations they intend to pursue, denoted as n_{it}, which represents the number of patents they plan to file across all categories. Second, they decide the level of diversity in their portfolio of AI innovations, represented by h_{it}. This diversity is quantified using the Herfindahl-Hirschman Index (HHI), which ranges from 0 to 1. A value of $h_{it} = 1$ implies a focus on patents within a single category, while $h_{it} = 0$ indicates a highly diverse patent portfolio.

Based on its strategic decisions in previous years, firms receive a value of d_{it} for their patent portfolio in year t, measured by the number of forward citations or the monetary value of the patent portfolio. However, as there is a time lag between application and granting, the full benefits of patents are not realized in the same year they are filed. Rather, the benefits materialize a few years after filing an application. Thus, d_{it} in year t is associated with the filed patents in previous years not the current year. Additionally, I assess the breadth of portfolio value using the HHI index as b_{it}, reflecting the distribution of values across different categories, is quantified using the HHI index and indicated as b_{it}. This index shows the value-diversity among granted patents. Firms also incur costs c_{it} linked with developing the patent portfolio of AI innovations. These costs are realized when the firm files the patents. Consequently, in each year, the collective value of firm i's patent portfolio, its value-diversity index, and its associated cost are respectively represented as d_{it}, b_{it}, and c_{it}. Notably, d_{it} and b_{it} represent the value of patents and diversity of value that a firm filed in previous years but they realized in the current year, while c_{it} indicates the costs associated with the patents filed in the current year.

Firms' experience in AI innovation is measured by their cumulative number of AI innovations up to the previous period $t - 1$, denoted as e_{it}. This cumulative experience is calculated as the sum of the number of patents filed in previous periods:

$$e_{it} = \sum_{\tau=1}^{t-1} n_{i\tau} \tag{4}$$

Moreover, the cumulative diversity index of a firm's portfolio at time t, denoted as eh_{it}, approximates the diversity of the firm's cumulative patent holdings. It takes into account the number of patents filed in each period and their corresponding diversity levels:

$$eh_{it} = \sum_{\tau=1}^{t-1} \frac{n_{i\tau}}{e_{it}} h_{i\tau} \tag{5}$$

It's crucial to emphasize that firms don't immediately realize the full benefit of filed patents since the true value of patents is realized upon their grant, and the grant date is uncertain. The grant year is particularly significant as it grants the firm the right to fully utilize, extend, or otherwise leverage the patent's potential. Furthermore, filing patent applications serves purposes beyond merely securing intellectual property rights. These applications act as strategic tools, signaling a firm's commitment and capability to innovate in specific areas. They can enhance a firm's competitive positioning, attract partners or investors, and effectively communicate technological capability to the market. However, while filed applications serve signaling and strategic considerations, it's the ultimate granting of patents that provides the majority of benefits to firms. This is a critical juncture where the patent's legal protection is solidified, enabling the firm to capitalize on its innovation. This dual benefit structure generates a two-fold stream of payoffs for firms: one from the filing of an application, realized in the same year, and another from the stream of granted patents from previous filings.

Firms carefully determine the quantity and diversity of AI patents to file, aiming to maximize the potential benefits associated with these choices, depending on the specific state variables they face. The per-period payoff for firm i in period t can be described by the following equation:

$$U(n_{it}, h_{it}, d_{it}, b_{it}, c_{it}, z_t, q_{it}; \theta) = \theta_1 n_{it} + \theta_2 h_{it} + \theta_3 d_{it} + \theta_4 b_{it} - \theta_5 n_{it} c_{it} + v(q_i) + z_t + \epsilon_{it} \tag{6}$$

The payoff function factors various elements that directly contribute to the firm's profitability in each period. It takes into account the quantity of AI patents filed by the firm (n_{it}). More patents are likely to contribute directly to profits but also entail development costs. θ_1 captures this net effect on payoffs. The diversity of the patent portfolio is represented by h_{it}, with greater diversity across AI subfields allowing access to varied markets and resulting in higher payoffs. θ_2 reflects the potential benefits of diversification.

The value of granted patents in year t is denoted by d_{it}, where higher value patents represent more substantial innovations, directly enhancing profits. θ_3 quantifies this relationship. The value-diversity of the granted patent portfolio is indicated by b_{it}, reflecting that distributing value across categories reduces risk and broadens reach. θ_4 captures the advantages of balanced portfolios. c_{it} represents the marginal cost of innovation, reflecting the fact that developing patents requires investments, hence resulting in costs. θ_5 measures the contribution of costs in reducing the payoff function. With costs being implicit function of experience, this term is equivalent to the rate of learning as per Benkard (2004). Alongside these components, two additional terms influence a firm's profit from AI innovations: $v(q_i)$ and z_t. The term z_t signifies the profitability of AI innovation, assumed to be equal across firms but varying over time. It captures common shocks affecting the overall profitability of innovation in the industry, such as shifts in market demand, technological advancements, or competitive dynamics. Firms observe z_t, but it remains unobserved by the econometrician.

The ability of firm i in AI innovation, represented by q_i, plays a significant role in the objective function. It reflects the firm's potential to generate profits from AI innovations and is influenced by various factors such as technological infrastructure, expertise, and resource allocation. However, it is important to note that the true value of q_i is not known ex-ante to both the firms and the econometrician. Instead, firms rely on their prior knowledge and beliefs about their abilities, which are not precise measurements.

To gain a deeper understanding of their capabilities, firms can utilize signals that provide information about the unobservable variable q_i in each period. In this model, I consider the number of patents filed in the preceding period (n_{it-1}) and the scope of the filed patents portfolio in the prior period (h_{it-1}) as these signals. These variables serve as informative signals that provide valuable insights into the unobservable profitability potential (q_i) of firms in AI innovation. By incorporating these signals into the model, firms can gain valuable insights into their true abilities and refine their understanding of q_i over time.

The term $v(q_i)$ represents the utility that firm i derives from its ability in AI innovation. It reflects the potential advantages and benefits associated with higher capabilities, such as more efficient utilization of innovations or improved market positioning. However, since the true value of q_i is initially unknown, firms must actively participate in the innovation process, accumulate experience, and engage in a Bayesian learning process to update their beliefs and make more informed decisions based on their evolving understanding of their abilities.

4.1 Timing and transition

The model unfolds through a dynamic timeline, progressing in discrete time periods denoted by t. At the onset of each period, firm i possesses its accumulated stock of experience (e_{it}) and diversity of experience (eh_{it}) in AI patenting up to that point. Concurrently, the firm faces a specific marginal cost (c_{it}) of innovation determined by its current experience level.

Using the previous period's number of patents filed (n_{it-1}) and diversity choice (h_{it-1}) as decision variables, firm i shapes its beliefs concerning the current profitability (q_{it}) of AI innovations. These variables serve as informative signals, as firms with greater inherent capability can file more patents and cultivate a more diverse portfolio. It also takes into account the prevalent market trend (z_t) affecting industry profitability. Furthermore, the patents applied for in the previous period(s) are now granted, establishing the current portfolio value (d_{it}) and value-diversity (b_{it}). Armed with these insights and observations, the firm strategically decides the number of new AI patents to file (n_{it}) and the diversity (h_{it}) of its portfolio for that specific period.

Once these patenting decisions are concluded, the firm receives the payoff (U_{it}) for that period, determined by its choices of n_{it} and h_{it}, along with the other state variables.

At the end of period t and following the realization of outcomes, the firm's total experience (e_{it}) is updated based on the new patents (n_{it}) filed during that period, determined by Equation 4. Simultaneously, the diversity of cumulative experience (eh_{it}) is deterministically adjusted based on decision variables using Equation 5. Additionally, the market trend (z_{t+1}) evolves through an autoregressive process as shown in Equation 7 in preparation for the next period.

$$z_{t+1} = \rho z_t + \nu_t \tag{7}$$

Here, ρ represents the persistence parameter, signifying the degree to which the market trend maintains its direction from the current to the subsequent period. ν_t denotes an iid shock, capturing unforeseen and uncontrollable factors affecting the market trend.

Furthermore, the evolution of other state variables, including the value of the portfolio (d_{it}), diversity-value of the portfolio (b_{it}), and marginal cost (c_{it}), is subject to stochastic processes based on decision variables and other state variables. This evolution follows the state transition probabilities denoted by $f(s_{it+1}|s_{it}, a_{it})$, illustrating the likelihood of transitioning to state s_{it+1} in $t+1$ given the current state s_{it} and action a_{it}. The observed state variables are assumed to evolve according to:

$$\log(d_{it+1}) = \alpha_{0d} + \alpha_d \log(d_{i,t}) + \beta_d \log(e_{it}) + \eta_{1it} \tag{8}$$

$$\log(b_{it+1}) = \alpha_{0b} + \alpha_b \log(b_{i,t}) + \beta_b \log(e_{it}) + \delta_b \log(eh_{it}) + \zeta_b \log(e_{it}eh_{it}) + \eta_{2it} \tag{9}$$

$$\log(c_{it+1}) = \alpha_{0c} + \alpha_c \log(c_{i,t}) + \beta_c \log(e_{it}) + \eta_{3it} \tag{10}$$

Here, the persistence in value, breadth, and cost, denoted by αs, signifies institutional memory, reflecting the firm's ability to retain knowledge from past experiences. Additionally, the coefficients β, δ, and ζ represent how depth, breadth, and cost change with alterations in the size of experience and diversity of cumulative experience, akin to the learning-by-doing literature such as Benkard (2000). According to data variation, I posit that diversity of experience impacts only the breadth of knowledge. Finally, ηs denote the iid shocks.

The timeline then resets, commencing period $t+1$, wherein the firm again navigates patenting decisions, building upon updated knowledge, experience, portfolio characteristics, beliefs, and market conditions. This recursive and dynamic process persists as the firm accumulates experience, gains insights into profitability, and innovates progressively over time.

The dynamics of the problem arise from the fact that firms' decisions in one period affect their experience and subsequent state variables and decisions in the next period. This intertemporal link highlights the iterative nature of the decision-making process in the context of AI innovation.

As firms engage in the innovation process and accumulate experience over time, they gradually gain insights into their own abilities and refine their understanding of the profitability of AI innovations. The dynamic nature of the problem lies in the feedback loop between firms' decisions, their evolving experience, and their beliefs about profitability. This iterative learning process allows firms to adapt and adjust their strategies based on their evolving understanding of the industry and their own capabilities.

Overall, the accumulation of experience and the dynamic interplay between decision variables, state variables, and beliefs about profitability make this problem a dynamic one, requiring firms to continually update their strategies and adapt to changing market conditions in the field of AI innovation.

4.2 Learning

Firms engage in a dynamic learning process to update their beliefs about their unobserved AI profitability, denoted as q_i. At the beginning, the prior distribution for q_i is assumed to be identical across firms and follows a normal distribution $N(\mu_1, \sigma_1^2)$. This implies that prior to receiving any information, firms share the perception that the true value of their AI profitability is normally distributed, characterized by a mean of μ_1 and a variance of σ_1^2. Therefore, their initial information set is represented as $Q_1^i = \{\mu_1, \sigma_1\}$, capturing their common prior beliefs about AI profitability.

The learning process takes place in each time period t, allowing firms to update their beliefs based on observed signals. Firms acquire information through two distinct channels: the number of AI patents filed in the previous period (n_{it-1}) and the diversity of the patent portfolio from the previous period (h_{it-1}). This learning process enables firms to adjust their perceptions based on real-world feedback.

These channels provide noisy signals that inform firms about the true value of the unobserved AI profitability, denoted as $\tilde{q}_{it}^n$ and $\tilde{q}_{it}^h$. Each signal is obtained by adding a random variable to the firm's true profitability q_i.

For number of patnets, the signal is generated by adding a random variable f_{it}, following a normal distribution $N(0, \sigma_f^2)$:

$$\tilde{q}_{it}^n = q_i + f_{it} \tag{11}$$

Similarly, the diversity signal is obtained by adding a random variable g_{it}, following a normal distribution $N(0, \sigma_g^2)$:

$$\tilde{q}_{it}^h = q_i + g_{it} \tag{12}$$

These signals provide firms with valuable insights into the underlying profitability of their AI innovations. By incorporating these signals into their decision-making process, firms can update their beliefs regarding their own capabilities and make more informed decisions about their AI portfolio composition and strategic actions.

These two signals, stemming from the size and diversity of firms' AI portfolio, are informative about their AI profitability for several reasons. The signal derived from the portfolio's size captures the number of patents filed by a firm in the preceding period. A higher count of filed patents suggests that firms possess the competence to develop a substantial array of innovative AI technologies, signifying their proficiency in generating profits from their AI innovations. Furthermore, a larger patent count reflects the firm's commitment and investment in research and development, indicating a sustained dedication to AI innovation and a higher potential for reaping economic benefits. Similarly, the diversity signal reflects the breadth and variety of firms' AI portfolio, indicating their capability to address different market demands and adapt to technological advancements. Firms with a diverse portfolio are more likely to capture profitable opportunities in various domains and demonstrate their ability to effectively leverage their AI expertise. Collectively, these signals serve as informative proxies for firms' AI profitability, allowing them to update their beliefs and make strategic decisions to maximize their long-term payoffs in the dynamic landscape of AI innovation.

In the dynamic learning process, each firm updates its beliefs about the unobserved AI profitability. At each time period τ, firm i possesses an information set denoted as $Q^i_\tau = \{\mu_{i\tau}, \sigma_{i\tau}\}$. These information sets differ across firms due to the distinct signals they receive based on their unique abilities in AI innovation. The signals obtained from the size and scope of their AI patents in each period serve as individualized sources of information that shape their understanding of AI profitability. By incorporating these signals within a Bayesian framework, firms have the flexibility to adjust their beliefs and adapt their strategies accordingly. This dynamic learning process allows firms to refine their perceptions of their own abilities in AI innovation and make informed decisions to maximize their long-term payoffs in an ever-changing AI landscape.

In this framework, I leverage conjugate priors and signals, assuming that both the prior distributions for unobserved AI profitability and the noise in the signals follow a normal distribution. This choice of distributional assumptions allows for the derivation of formulas for Bayesian updating, facilitating firms' revision of their beliefs as new information is received (similar to DeGroot (2005)). By integrating the learning process and applying Bayesian updating principles, the perceived AI profitability for firm i at time t can be understood as the posterior belief that emerges after incorporating the signals and modifying the initial prior distribution. It can be expressed as:

$$q_{it} = \sigma^2_{it} \left[\frac{1}{\sigma^2_f} \tilde{q}^n_{it-1} n_{it-1} + \frac{1}{\sigma^2_g} \tilde{q}^h_{it-1} h_{it-1} + \frac{1}{\sigma^2_{it-1}} q_{it-1} \right] \tag{13}$$

Here, q_{it} represents the perceived AI profitability for firm i at time t, and σ^2_{it} denotes the variance of the posterior distribution, capturing the level of uncertainty surrounding the perceived AI profitability. The terms $\tilde{q}^n_{it-1}$ and $\tilde{q}^h_{it-1}$ represent the size and diversity signals obtained from the previous period, respectively. The variables n_{it-1} and h_{it-1} represent the corresponding size and diversity of the AI portfolio from the previous period, while q_{it-1} represents the belief about AI profitability from the previous period.

The variance σ_{it}^2 can be calculated as:

$$\sigma_{it}^2 = \left(\frac{1}{\sigma_f^2} n_{it-1} + \frac{1}{\sigma_g^2} h_{it-1} + \frac{1}{\sigma_{it-1}^2} \right)^{-1} \tag{14}$$

This equation represents the variability of signals generated by the learning process, incorporating the variability from the size and diversity signals, as well as the variability carried forward from the previous period. By updating their beliefs based on the observed signals and considering the level of uncertainty captured by σ_{it}^2, firms can make more informed decisions regarding their AI portfolio and navigate the dynamic landscape of AI innovation.

The equation for updating the posterior belief about AI profitability (q_{it}) incorporates signals from size ($\tilde{q}_{it-1}^n n_{it-1}$) and diversity ($\tilde{q}_{it-1}^h h_{it-1}$) of the AI portfolio. The weights assigned to these signals, determined by the inverse variances ($\frac{1}{\sigma_f^2}$, $\frac{1}{\sigma_g^2}$), represent the influence of each signal on the updated belief. Strong and precise signals from size and diversity have a significant impact on the posterior belief. These signals indicate a high level of profitability, resulting in a substantial adjustment in the belief (q_{it}). The weights assigned to these signals reflect the greater importance placed on these strong signals, leading to a stronger influence on the updated belief. As a result, the posterior belief becomes more aligned with the observed signals, reflecting a higher level of confidence in the profitability of AI innovation.

The strength of the signals also affects the variance of the posterior belief. When firms receive strong and precise signals, the uncertainty surrounding the belief is reduced, leading to a narrower standard error. The weight assigned to these strong signals contributes to a more accurate estimation of AI profitability, resulting in a higher degree of confidence in the updated belief. Conversely, weaker signals with larger variances introduce more uncertainty, leading to a wider standard error and a higher level of uncertainty in the updated belief.

By incorporating these signals into the learning process, firms can navigate the dynamic landscape of AI innovation more effectively. Strong signals provide reliable information that enables firms to make strategic decisions with greater confidence. They contribute to a more accurate estimation of AI profitability, reducing uncertainty and guiding firms towards profitable actions. In contrast, weaker signals introduce more uncertainty, requiring firms to exercise caution in decision-making and potentially seek additional information to reduce uncertainty and make more informed choices.

4.3 Objective function

In this model, firms adopt a forward-looking perspective, carefully considering the long-term implications of their actions. Their primary objective is to maximize the discounted stream of payoffs derived from AI innovations. Given the initial set of state variables s_{it} at time t, firm i aims to maximize the expected present discounted utility from AI innovation, considering idiosyncratic shocks and beliefs about profitability in the per-period payoff function. This objective can be expressed as:

$$E\left[\sum_{t=\tau}^{\infty} \beta^{t-\tau} U(a_{it}, s_{it}) \mid s_{i\tau}\right] \tag{15}$$

where β represents the discount factor. The expectation is taken over the firm's actions, beliefs on profitability, and private shocks in the current period, as well as the evolution of state variables, beliefs about profitability, and private shocks in future periods. The value function for firm i in state s_{it} can be formulated as a solution to the following Bellman equation, providing the expected present discounted value of profits obtained by firm i when choosing action a at state s (indices of time and firm are dropped):

$$V(s_i) = \max_{a_i} \left[E\left[u(s_i, a_i) + \beta E\left(V(s') \mid s, a\right) \mid s\right]\right] \tag{16}$$

It's important to note that the outer expectation is over current values of beliefs on profitability and private shocks, and the inner expectation is with respect to the state variables in the next period, s', conditional on the current state, s, and the actions of firm i. Additionally, the standard assumption is made that all firms possess correct beliefs regarding the state transition probabilities. The action-specific value function (ex-ante value function) is defined as:

$$V(s_i, a_i) = E\left[u(s_i, a_i) + \beta E\left(V(s') \mid s, a\right) \mid s\right] \tag{17}$$

where expectations are defined similarly to the above. Given the serial correlation of unobservable profitability through beliefs, the conventional assumptions proposed by Rust (1987) to solve the Bellman equation and estimate the model are not applicable. Instead, a method proposed by Norets (2009) is utilized to estimate the model.

5 Identification

In this section, I explore the conditions necessary for parameter identification in the proposed model. Drawing on a heuristic approach similar to that of Keane (2010) and Ching, Erdem and Keane (2013), I aim to determine the specific parameters based on the assumed model structure. Identification is crucial to ensure meaningful estimation and interpretation of the model parameters.

The identification in my model relies on the variations in firm actions and states related to AI patent development across different categories. These variations provide the primary source of identification, similar to other panel data models. However, my model introduces unobservable variables that are serially correlated, presenting challenges for identification. To address this, I employ a methodology that involves controlling for these unobservable states, following prior literature (e.g., Pakes (1986); Blevins, Khwaja and Yang (2018)). This approach, coupled with parametric distributional assumptions, allows for the identification of dynamics in the unobservable serially correlated components.

The identification process in the model relies on distinct projections to capture persistence and parameters related to learning. For the exogenous unobserved state z_t, its projection onto the lagged value z_{t-1} allows us to estimate the auto-covariance and determine the degree of persistence (ρ). This projection enables us to examine the temporal dynamics and assess the level of persistence exhibited by the state variable.

In contrast, the identification of parameters associated with learning, such as μ_1, σ_1, σ_f, and σ_g, involves projecting the posterior beliefs about the unobserved profitability q_{it} onto their lagged values q_{it-1}. It is important to emphasize that the evolution of q_{it} differs analytically from the evolution of z_t in significant ways. While z_t represents an exogenous unobserved state that remains unaffected by past decisions, q_{it} is directly influenced by the lagged size and scope of the AI portfolio, which, in turn, are influenced by the previous experience and diversity of experience. This distinction arises due to the Bayesian updating process, where firms incorporate observed signals and accumulated experience to revise their beliefs.

Furthermore, to ensure robust identification of these parameters, it is crucial that the data used for estimation contains sufficient variation in the size and breadth of AI portfolio. This variation allows for a comprehensive analysis of firms' decision-making processes and ensures that the model captures the complexities and heterogeneity of their AI innovation strategies accurately. The summary statistic in Table 2 provides evidence that the data used in the analysis exhibits sufficient variation in the depth, breadth, and cost of AI portfolio learning.

To identify the prior means and variance for the unobservable variable AI profitability (q_i), it is necessary to observe the initial size and scope of the portfolio as argued by Crawford and Shum (2005); Ching, Erdem and Keane (2013); Yang (2020). This condition is crucial because beliefs about AI profitability may converge towards complete information over time, resulting in the priors becoming negligible in the model. Without observing the initial periods when firms lacked complete information and faced uncertainty, it would be impossible to identify the priors and understand the learning process.

Fortunately, in the available data, this condition is satisfied as the dataset extends back to 1990, a crucial period when AI began to demonstrate its practical potential. During the 1990s, AI was in its early stages, and researchers and firms were just beginning to grasp the possibilities it held. Notable breakthroughs during this time, such as IBM's Deep Blue defeating world chess champion Garry Kasparov in 1997 (Campbell, Hoane Jr and Hsu, 2002), marked significant milestones in AI's journey towards practical applications. Moreover, advancements in speech recognition, handwriting recognition, and NLP during the 1990s showcased the power of data-driven approaches and their role in driving AI's practicality (Church and Rau, 1995). By including data from this initial period, the analysis captures the early stages of AI development and the learning process, ensuring that the priors are identifiable.

The observation of firm decisions and states over a long time period is another crucial factor for model identification. This characteristic has been highlighted in previous research on learning, such as the work of Ackerberg (2003); Ching, Erdem and Keane (2013); Dickstein (2018); Yang (2020). The variance in size and scope signals can be identified by examining how these measures change over time. Each observed time period for a given firm provides an opportunity to observe potential variations in these measures, allowing for the identification of parameters that govern information set transitions and the variance of signal noises associated with learning from size and scope.

Moreover, the changes in these signals over time contribute to shifting the information set across different firms, as discussed by Ching, Erdem and Keane (2013). With more than 27 years of observation available for each firm in my dataset, there are ample opportunities to observe variations over time as shown in Figures 1 and 3. This extensive time span provides valuable insights into the dynamics of firms' AI innovation strategies and allows for the identification of key parameters associated with learning and the evolution of the AI portfolio.

By leveraging the richness and duration of the data, the model can capture the complex dynamics of firms' learning processes in the context of AI innovation. The extensive observation period ensures that the model is well-identified and provides robust estimates of the parameters governing transitions, as well as the variances of signal noises associated with learning from size and scope. This strengthens the reliability of the findings and enhances our understanding of the factors driving firms' decision-making and learning in the field of AI innovation.

It is important to emphasize that the learning process in the model is influenced by the lagged values of size and scope of AI patents, which are themselves determined by the accumulated experience, diversity of experience and other observable and unobservable states in the previous periods.

A key exclusion restriction that aids identification of the preferences θ_3, θ_4, θ_5 is that the patent decisions n_{it} and h_{it} affect the per-period payoff but do not directly influence the state transitions. This provides distinct variation that identifies the preference parameters separately.

Additionally, while the states d_{it}, b_{it}, c_{it} enter both the per-period utility and transitions, functional form assumptions enable separate identification. For instance, lagged values of value, breadth and cost affect their transitions while having no impact on the current payoff. Nonlinear interaction terms like $\theta_5 n_{it} c_{it}$ also isolate effects on utility apart from the transition parameters.

In summary, the exclusion restrictions on patent decisions n_{it} and h_{it}, along with controls for unobserved heterogeneity and functional form choices, provide the identifying variation needed to estimate the preference and transition parameters distinctly, despite the state variables influencing both model components.

6 Estimation

In this section, I focus on estimating the parameters of the structural model introduced in the previous sections. The model encompasses four key parameter sets associated with the learning process, persistence in market trend, utility function preferences, state transition dynamics, and shock variances, which are collectively denoted as:

$$\Phi = (\mu_1, \sigma_1^2, \sigma_f^2, \sigma_g^2, \rho, \theta_1, \dots, \theta_5, \alpha_{0d}, \alpha_{0b}, \alpha_{0c}, \alpha_d, \alpha_b, \alpha_c, \beta_d, \beta_b, \beta_c, \delta_b, \zeta_b, \sigma_{\eta_1}^2, \sigma_{\eta_2}^2, \sigma_{\eta_3}^2) \tag{18}$$

The data we are working with comprises patenting decisions, and portfolio characteristics including its value, breadth and cost over a period. We denote the patent decision variables as $a_{it} = (n_{it}, h_{it})$, the observable state variables as $x_{it} = (d_{it}, b_{it}, c_{it})$, and the unobservable state variables as $y_{it} = (q_{it}, z_t)$.

Assuming independence of state variables across firms, the likelihood function is expressed as:

$$\begin{aligned} l(\{a, x, y\}_{t=1}^{T_i}, i \in \{1, 2, ..., I\}|\Phi) &= \prod_{i=1}^{I} p(a_{i1}, x_{i1}, y_{i1}, ..., a_{iT}, x_{iT}, y_{iT}|\Phi) \\ &= \prod_{i=1}^{I} \int p(a_{i1}, x_{i1}, y_{i1}, ..., a_{iT_i}, x_{iT_i}, y_{iT_i}|\Phi) dy_{i1} \dots dy_{iT_i} \end{aligned} \tag{19}$$

This joint density $p(a_{i1}, x_{i1}, y_{i1}, ..., a_{iT_i}, x_{iT_i}, y_{iT_i}|\Phi)$ can be further decomposed into the patent decision and state transition densities:

$$p(a_{i1}, x_{i1}, y_{i1}, ..., a_{iT_i}, x_{iT_i}, y_{iT_i}|\Phi) = \prod_{t=1}^{T} p(a_{it}|x_{it}, y_{it}; \Phi) f(x_{it}, y_{it}|x_{i,t-1}, y_{i,t-1}, a_{i,t-1}; \Phi) \tag{20}$$

Here, $f(.|.; \Phi)$ represents the state transition density, and $p(a_{it}|x_{it}, y_{it}; \Phi)$ is a choice probability conditional on all state variables.

Rust (1987) assumes that unobservable variables are iid, allowing for simplifications in high dimensional integration of Equation 20 through certain distributional assumptions. However, in our case, the presence of serially correlated and unobserved states makes it challenging to apply Rust's simplifications and introduces complexity to the estimation process.

The key challenge stems from the fact that firms base their decisions on the market profitability trend (z_t), which is not directly observable by the econometrician but is within the firms' knowledge. Moreover, firms engage in a learning process to evaluate their capability in AI innovations (q_i) using unobserved signals. In each period, q_{it} exhibits serial correlation as firms adjust their prior beliefs and update them using signals from previous actions. Thus, this variable is both serially correlated and endogenous. The endogeneity and serial persistence of the unobservables make estimation challenging, necessitating alternative approaches to evaluate the likelihood.

To tackle these obstacles, the literature presents various approaches. Bayesian methods, exemplified by Norets (2009); Imai, Jain and Ching (2009), offer one avenue. Additionally, Sequential Monte Carlo (SMC) methods, as demonstrated in Fernández-Villaverde and Rubio-Ramírez (2007); Blevins (2016); Gallant, Hong and Khwaja (2018), provide another effective solution.

The challenge with the Bayesian approach arises when dealing with a high-dimensional space featuring non-linear, non-Gaussian variables. Calculating marginal densities and utilizing Gibbs sampling for computing posterior distributions becomes impractical. On the contrary, the SMC method shines in such scenarios by presenting a practical and efficient means to compute posterior distributions. This is particularly beneficial when working with random variables that are non-linear, non-Gaussian, and high-dimensional. Given my encounter with a high-dimensional problem involving multiple unobservable variables, I opt for particle filtering to estimate my model which is a SMC method.

I can further decompose Equation 20 and write the likelihood function as:

$$l_I(\Phi) = \prod_{i=1}^{I} \prod_{t=1}^{T} \int p(a_{it}|x_{it}, y_{it}; \Phi) p(x_{it}|x_{i,t-1}, a_{i,t-1}; \Phi) p(y_{it}|y_{i,t-1}, a_{i,t-1}; \Phi) dy_{it} \tag{21}$$

This equality follows from decomposing observable and unobservable variables. Now, the observed data consists of three components after conditioning and integrating out the unobservable variables (y_{it}). For a given parameter set Φ, computing both $p(a_{it}|x_{it}, y_{it}; \Phi)$ and $p(x_{it}|x_{i,t-1}, a_{i,t-1}; \Phi)$ is feasible. However, Sequential Monte Carlo (SMC) helps us draw from $p(y_{it}|y_{i,t-1}, a_{i,t-1}; \Phi)$. I build on the sequential Monte Carlo approach of Blevins (2016) to estimate these components and draw from the unobservable distribution.

6.1 First-Stage Estimation

In the first stage, , I jointly estimate the transition equations, policy function conditional on unobservable levels, and the posterior distribution of unobservables using a sequential Monte Carlo method. Specifically, this stage entails recovering structural parameters linked to learning, the evolution of market trends, transition dynamics, and the reduced-form policy function. Subsequently, I will retrieve parameters related to the payoff function in the second stage.

Due to the interconnected nature of all three components in Equation 20, it is imperative to jointly maximize their log-likelihood. The reduced-form policy function is influenced by payoff-related observable states as well as unobservable variables. Transition densities are dependent on previous observable state variables and actions, which in turn are functions of unobservables. This interdependence extends to the transition of unobservables. The policy functions involve modeling two actions per period: the number of patents to file and the diversity of the portfolio. To estimate the number of patents, a two-stage model is employed due to the prevalence of zeros in the data. The first stage utilizes a probit model to differentiate between zero and non-zero choices, followed by an ordered probit model in the second stage, where the number of filed patents is discretized into ten categories. Similarly, for portfolio diversity choices, an ordered probit model is used, discretizing this variable into nine bins. The transitions of portfolio value, breadth, and cost are estimated using Equations 8, 9, and 10. For estimating the posterior of unobservables, particle filtering and the functional forms outlined in the learning section are applied. Particle filtering enables the drawing from the posterior distribution of unobservables and the calculation of the likelihood function. Details of the particle filtering procedure are elaborated subsequently. Particle filtering, a recursive algorithm, initiates by drawing a swarm of particles to approximate the initial distribution of unobservable variables. Subsequently, this swarm is used to draw the next period's swarm based on transition densities of latent variables, which is then filtered using sequential importance weights. The filter particles are employed to draw another swarm for the subsequent period, and this iterative process continues. Here's a step-by-step breakdown of the particle filtering procedure:

1. Initialization: Draw a swarm of particles y_{i0}^r from an initial distribution where $r = 1, ..., R$.

2. Prediction: Use the draw from the previous period to draw the distribution of $\hat{y}_{it}^r$ from $p(y_{it}|y_{i,t-1}, a_{i,t-1}; \Phi)$ for $t > 0$, leveraging the parametric assumptions imposed in the model equations.

3. Importance weight: Calculate the vector of importance weights for each particle $r = 1, ..., R$ as $w_t^r = p(a_{it}|a_{it-1}, x_{it}, \hat{y}_{it}^r; \Phi)$, calculating the probability of observing a_{it} given state x_{it} and drawn particle $\hat{y}_{it}^r$.

4. Resampling: Draw a new swarm of particles y_{it}^r with replacement from $\hat{y}_{it}^r$ using the importance sampling weights from the previous step as sample probabilities.

5. Proceed to the Next Time Period: Repeat the process from step 1 for the next time period.

Following Equation 21, the log-likelihood function can be expressed as:

$$L_I(\Phi) = \sum_{i=1}^{I} \sum_{t=1}^{T} \ln \int p(a_{it}|x_{it}, y_{it}; \Phi) p(x_{it}|x_{i,t-1}, a_{i,t-1}; \Phi) p(y_{it}|y_{i,t-1}, a_{i,t-1}; \Phi) dy_{it} \tag{22}$$

By utilizing particle filtering, we can approximate $p(y_{it}|y_{i,t-1}, a_{i,t-1}; \Phi)$ and rewrite the log-likelihood function as:

$$\hat{L}_{N,R}(\Phi) = \frac{1}{I} \sum_{i=1}^{I} \sum_{t=1}^{T} \ln \left[\frac{1}{R} \sum_{r=1}^{R} p(a_{it}|x_{it}, y_{it}; \Phi) p(x_{it}|x_{i,t-1}, a_{i,t-1}; \Phi) p(y_{it}|y_{i,t-1}, a_{i,t-1}; \Phi) \right] \tag{23}$$

This equation enables the maximization of the log-likelihood function for model estimation. Blevins (2016) demonstrate the consistency of the estimated parameters.

6.2 Second-Stage Estimation

In the second stage, I estimate payoff function parameters similarly to Bajari, Benkard and Levin (2007), assuming state variables adhere to a first-order Markov process, the data are generated by a single data generating process and this process is used by all firms in all periods.

This stage involves using previously estimated parameters for policy functions, transition densities, and unobservable distributions as fixed inputs in a forward simulation. Here, sequences of observable and unobservable variables are simulated to construct per-period future payoffs. The expected value of these payoffs is calculated by discounting, as shown below:

$$V_i(S_1, \hat{a}, \theta) = E[\sum_{\tau=1}^{\infty} \beta^{\tau-1} \Pi_i(S_\tau; \theta) | S_1, \hat{a})] = \frac{1}{K} \sum_{k=1}^{K} \sum_{\tau=1}^{T} \beta^{\tau-1} \Pi_i(s^k, \hat{a}^k; \theta) \tag{24}$$

In this equation, K simulated paths of length T are used, where each variable with subscript s represents a simulation s. The aim is to identify the parameter θ that rationalizes observed choice probabilities using revealed preference logic. This involves ensuring that the value function $V_i(S_1, \hat{a}; \theta)$ associated with θ exceeds any value function based on alternative strategies $\tilde{a}$, as represented in Equation 25:

$$g_b(S; \theta) = V_i(S_1, \hat{a}; \theta) - V_i(S_1, \tilde{a}; \theta) \tag{25}$$

In equilibrium, this equation should satisfy $g_b(S; \theta) \geq 0$, with b indicating indexed perturbations. The criterion for estimating θ is formulated as in Equation 26:

$$C(\theta) = \frac{1}{B} \sum_b (\min\{g_b(S'\theta), 0\}^2) \tag{26}$$

Finally, θ is estimated using a minimum distance estimator based on the above criterion.

7 Results

In the first stage, I employ 1,000 simulation draws for the particle filtering. Additionally, to assess firms' capabilities in AI innovations, I use a fifth-order polynomial in the payoff function with respect to perceived beliefs (q_{it}). This flexible approach effectively captures the potential nonlinear impact of this factor in the payoff function. For the second-stage estimation, I generate 500 paths, creating 500 random inequalities to compare various strategies against the optimal strategy. As discussed in the previous section, I perturb the coefficients of the estimated reduced-form policy function and forward simulate the entire path for each alternative strategy. This process is used to calculate an alternative value function and compare it with the optimal value function to obtain the structural parameters in the payoff function. Table 4 presents estimates of the main structural parameters in the first stage, including those governing transition densities, the persistence of market trends, and the learning process.

In the estimates of transition densities, a strong persistence is evident among all the state variables, as well as market trend over time. Additionally, the impact of experience on both the value and breadth of the

Table 4: First-stage results

Parameter	Estimate
Transition densities	
persistence in value (α_d)	0.755
effect of experience on value (β_d)	0.203
persistence in breadth (α_b)	0.150
effect of experience on breadth (β_b)	2.969
effect of diversity of experience on breadth (δ_b)	-2.956
effect of exp. X diversity of exp. on breadth (ζ_b)	0.151
persistence in cost (α_c)	0.721
effect of experience on cost (β_c)	-0.098
persistence in market trend (ρ)	0.745
Prior beliefs	
mean (μ_1)	0.882
variance (σ_1^2)	0.793
Signal variance	
number of past patents (σ_f^2)	0.119
diversity of past portfolio (σ_g^2)	0.281

portfolio is positive, while simultaneously reducing the cost of AI patent portfolio. An increase in the diversity of experience is also correlated with a broader diversity in the portfolio.

The learning process estimates offer some intriguing insights. The positive estimate for the prior suggests that firms generally have a favorable initial view of their AI innovation capabilities. However, there is significant uncertainty about these abilities, as indicated by the variance in prior beliefs. Furthermore, the data indicates that having a larger number of past patents gives firms a clearer signal about their AI innovation capabilities compared to the signal from a diverse portfolio, as demonstrated in Table 4. This suggests that learning from the sheer number of previous patent applications is more effective than learning from filing a diverse range of patents, showing that firms tend to learn more from scale as opposed to scope.

Table 5: Second-stage results

Parameter	Estimate
Payoff function	
number of patents (θ_1)	0.854
diversity of portfolio (θ_2)	0.612
value of granted patents (θ_3)	0.941
value diversity of granted patents (θ_4)	0.749
costs (θ_5)	0.889
Utility from perceived capability in AI innovation	
first degree	0.318
second degree	0.293
third degree	0.171
fourth degree	0.032
fifth degree	0.0992

Table 5 outlines the contributions of various factors to the payoff function. Notably, θ_3 emerges as the most influential factor, indicating that firms with a more valuable AI innovation portfolio experience a significant impact on their payoff. Following this, in descending order of importance, are the cost factor, the number of patents, the value diversity of granted patents, and the diversity of the portfolio. Additionally, it is evident that firms derive considerable utility from their perceived capability in AI innovations.

Next, I can take advantage of forward simulation to see how posterior beliefs evolve. The forward simulations are initialized using the values of states variables for each firm in the first year of data and using estimated parameters for transition densities and policy functions I can forward simulate the firms decisions in AI innovations, unobserved market trend as well as posterior beliefs about their capability in AI innovations. My forward simulations involve 500 paths with 27 length (for each year).

Next, I can leverage forward simulation to observe how posterior beliefs evolve over time. These forward simulations are initialized with the state variable values for each firm from the first year of data. By using the estimated parameters for transition densities and policy functions, I can simulate the firms' decisions regarding AI innovations and unobserved market trends, as well as their evolving posterior beliefs about their AI innovation capabilities. My forward simulations will consist of 500 paths, each spanning 27 years (corresponding to each year of data).

Table 6: Simulated beliefs statistics

Parameter	**Estimate**	**Std. dev.**
posterior mean (q_{it})	0.169	0.191
posterior std. dev. (σ_{it})	0.053	0.107

As illustrated in Table 6, the posterior mean is slightly positive, yet it is smaller than the prior mean. This indicates that firms initially perceive their ability to be higher, but this perception declines as they acquire more information about their underlying capabilities. Another notable result is the significantly reduced uncertainty surrounding the posterior belief compared to the prior mean, as indicated by the value of σ_{it}. This suggests that firms experience considerably less uncertainty regarding their posterior beliefs as opposed to their prior beliefs.

Figure 6 depicts the evolution of uncertainty surrounding posterior belief over time. It becomes clear that uncertainty about capability in AI innovation decreases quite rapidly and tends toward zero. This illustrates that firms rapidly learn about their underlying capabilities as they engage in innovations. These results are consistent with previous findings, such as those reported in Jain (2013).

Now, I am prepared to run counterfactual scenarios that examine the effects of relevant policies on AI innovations. Given the significant role of privacy policies in AI and the ongoing discussion about their impact on the development of AI technologies, my model is suited to evaluate the implications of such policies on AI innovation. The absence of such policy implementation in the United States, as per my

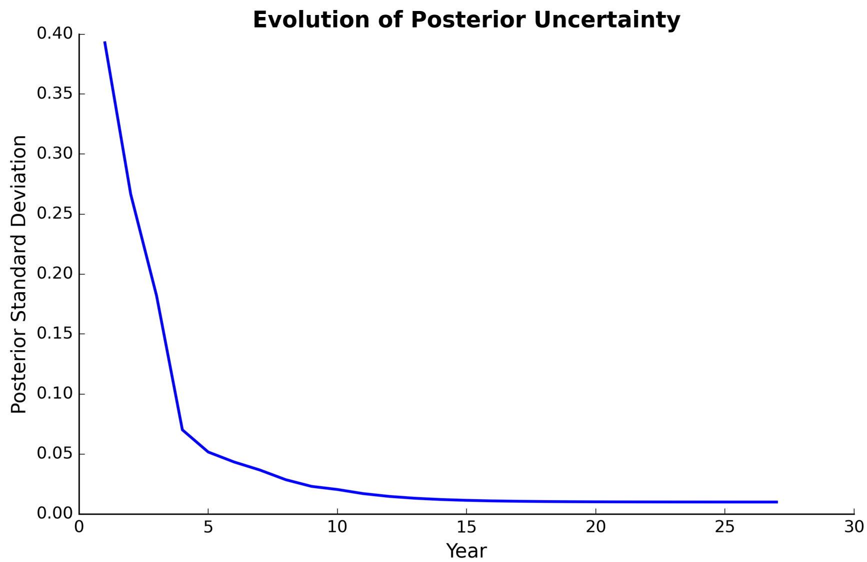

Figure 6: Evolution of Posterior Uncertainty

dataset, underscores the strength of my structural model in undertaking this task. This exercise contributes to the literature on the impact of information privacy policies on AI innovations, as referenced in works like (Chiou and Tucker, 2017) and (Jones and Tonetti, 2020).

Since information privacy policies directly limit firms' abilities to access and retain data, and my model does not explicitly account for data (due to the lack of data on data), I have made certain assumptions for this counterfactual analysis. I hypothesize that implementing a privacy policy will affect the cost of innovation, as it compels firms to devise innovative solutions to overcome data scarcity, which likely incurs additional costs. Therefore, I simulate this counterfactual by modeling an increase in the cost of innovation across firms.

8 Concluding remarks

This paper develops an empirical, dynamic structural model to examine the impact of learning and experience on innovation in the context of AI innovation at the firm level. Motivated by the exponential growth and nascent stage of AI, along with emerging policy debates, the model aims to provide insights into how experience accumulation and regulations may shape the evolving AI landscape. The analysis centers on formally modeling the learning process in AI patenting, whereby firms refine beliefs about unobserved profitability based on signals from their patent portfolios. This dynamic Bayesian learning framework combined with forward-looking optimization enables the quantification of how experience in AI patents enhances innovation scale, scope, quality, and cost efficiency over time. The model's structural nature facilitates counterfactual simulations to evaluate prospective impacts of privacy and antitrust policies on firm-level AI innovation, providing perspective on key tradeoffs.

The empirical results using comprehensive AI patent data uncover learning effects that increase patent value, expand technological breadth, and reduce innovation costs as AI experience grows. Furthermore, the results indicate that companies initially have an overly optimistic view of their AI innovation capabilities, which becomes more realistic over time. Initially, there is also a high level of uncertainty about their capabilities, but this uncertainty quickly reduces with increased experience. Overall, this research contributes a rigorous analytical framework for understanding the role of experience in AI innovations. By illuminating the central role of learning and experience in progressing AI technologies, while also assessing policy tradeoffs, the study offers insights to inform strategies and governance for stakeholders across industry, government, and research seeking to realize AI's immense economic potential in a socially responsible manner.

References

Ackerberg, Daniel A. 2003. "Advertising, learning, and consumer choice in experience good markets: an empirical examination." *International Economic Review*, 44(3): 1007–1040.

Agrawal, Ajay, Joshua Gans, and Avi Goldfarb. 2018. *Prediction machines: the simple economics of artificial intelligence.* Harvard Business Press.

Andrade, Gregor, Mark Mitchell, and Erik Stafford. 2001. "New evidence and perspectives on mergers." *Journal of economic perspectives*, 15(2): 103–120.

Argote, Linda, and Dennis Epple. 1990. "Learning curves in manufacturing." *Science*, 247(4945): 920–924.

Aridor, Guy, Yeon-Koo Che, and Tobias Salz. 2020. "The economic consequences of data privacy regulation: Empirical evidence from gdpr." *NBER working paper*, , (w26900).

Arora, Ashish, Sharon Belenzon, and Lia Sheer. 2021. "Knowledge spillovers and corporate investment in scientific research." *American Economic Review*, 111(3): 871–98.

Arrow, Kenneth J. 1962. "The economic learning implications of by doing." *The Review of Economic Studies*, 29(3): 155–173.

Babina, Tania, Anastassia Fedyk, Alex He, and James Hodson. 2021. "Artificial intelligence, firm growth, and product innovation." *Firm Growth, and Product Innovation (November 9, 2021).*

Bajari, Patrick, C Lanier Benkard, and Jonathan Levin. 2007. "Estimating dynamic models of imperfect competition." *Econometrica*, 75(5): 1331–1370.

Benkard, C Lanier. 2000. "Learning and forgetting: The dynamics of aircraft production." *American Economic Review*, 90(4): 1034–1054.

Benkard, C Lanier. 2004. "A dynamic analysis of the market for wide-bodied commercial aircraft." *The Review of Economic Studies*, 71(3): 581–611.

Bessen, James E, Stephen Michael Impink, Lydia Reichensperger, and Robert Seamans. 2020. "GDPR and the Importance of Data to AI Startups." *NYU Stern School of Business.*

Blevins, Jason R. 2016. "Sequential Monte Carlo methods for estimating dynamic microeconomic models." *Journal of Applied Econometrics*, 31(5): 773–804.

Blevins, Jason R, Ahmed Khwaja, and Nathan Yang. 2018. "Firm expansion, size spillovers, and market dominance in retail chain dynamics." *Management Science*, 64(9): 4070–4093.

Bommasani, Rishi, Drew A Hudson, Ehsan Adeli, Russ Altman, Simran Arora, Sydney von Arx, Michael S Bernstein, Jeannette Bohg, Antoine Bosselut, Emma Brunskill, et al. 2021. "On the opportunities and risks of foundation models." *arXiv preprint arXiv:2108.07258*.

Brynjolfsson, Erik, Daniel Rock, and Chad Syverson. 2021. "The productivity J-curve: How intangibles complement general purpose technologies." *American Economic Journal: Macroeconomics*, 13(1): 333–372.

Campbell, Murray, A Joseph Hoane Jr, and Feng-hsiung Hsu. 2002. "Deep blue." *Artificial intelligence*, 134(1-2): 57–83.

Ching, Andrew T, Tülin Erdem, and Michael P Keane. 2013. "Learning models: An assessment of progress, challenges, and new developments." *Marketing Science*, 32(6): 913–938.

Chiou, Lesley, and Catherine Tucker. 2017. "Search engines and data retention: Implications for privacy and antitrust." National Bureau of Economic Research.

Chui, Michael, Lareina Yee, Bryce Hall, and Alex Singla. 2023. "The state of AI in 2023: Generative AI's breakout year."

Church, Kenneth W, and Lisa F Rau. 1995. "Commercial applications of natural language processing." *Communications of the ACM*, 38(11): 71–79.

Cockburn, Iain M, and Rebecca M Henderson. 2001. "Scale and scope in drug development: unpacking the advantages of size in pharmaceutical research." *Journal of health economics*, 20(6): 1033–1057.

Crawford, Gregory S, and Matthew Shum. 2005. "Uncertainty and learning in pharmaceutical demand." *Econometrica*, 73(4): 1137–1173.

Crouzet, Nicolas, and Janice C Eberly. 2019. "Understanding weak capital investment: The role of market concentration and intangibles." National Bureau of Economic Research.

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

DeGroot, Morris H. 2005. *Optimal statistical decisions.* John Wiley & Sons.

Dickstein, Michael J. 2018. "Efficient provision of experience goods: Evidence from antidepressant choice." *Working paper.*

Dyer, Jeffrey H, Hal B Gregersen, Clayton M Christensen, et al. 2009. "The innovator's DNA." *Harvard business review*, 87(12): 60–67.

Epple, Dennis, Linda Argote, and Rukmini Devadas. 1991. "Organizational learning curves: A method for investigating intra-plant transfer of knowledge acquired through learning by doing." *Organization science*, 2(1): 58–70.

Fernández-Villaverde, Jesús, and Juan F Rubio-Ramírez. 2007. "Estimating macroeconomic models: A likelihood approach." *The Review of Economic Studies*, 74(4): 1059–1087.

Gallant, A Ronald, Han Hong, and Ahmed Khwaja. 2018. "The Dynamic Spillovers of Entry: An Application to the Generic Drug Industry." *Management Science*, 64(3).

Ghosh, Dipayan. 2018. "How GDPR will transform digital marketing." *Harvard Business Review Digital Articles*, 2–4.

Giarratana, Marco S, and Andrea Fosfuri. 2007. "Product strategies and survival in Schumpeterian environments: Evidence from the US security software industry." *Organization Studies*, 28(6): 909–929.

Giczy, Alexander V, Nicholas A Pairolero, and Andrew A Toole. 2022. "Identifying artificial intelligence (AI) invention: A novel AI patent dataset." *The Journal of Technology Transfer*, 47(2): 476–505.

Godinho de Matos, Miguel, and Idris Adjerid. 2021. "Consumer consent and firm targeting after GDPR: The case of a large telecom provider." *Management Science.*

Goldfarb, Avi, and Catherine E Tucker. 2011. "Privacy regulation and online advertising." *Management science*, 57(1): 57–71.

Hall, Bronwyn H, Adam Jaffe, and Manuel Trajtenberg. 2005. "Market value and patent citations." *RAND Journal of economics*, 16–38.

Henderson, Rebecca, and Iain Cockburn. 1996. "Scale, scope, and spillovers: the determinants of research productivity in drug discovery." *The RAND Journal of Economics*, 27(1): 32–59.

Imai, Susumu, Neelam Jain, and Andrew Ching. 2009. "Bayesian estimation of dynamic discrete choice models." *Econometrica*, 77(6): 1865–1899.

Irwin, Douglas A, and Peter J Klenow. 1994. "Learning-by-doing spillovers in the semiconductor industry." *Journal of political Economy*, 102(6): 1200–1227.

Jain, Amit. 2013. "Learning by doing and the locus of innovative capability in biotechnology research." *Organization Science*, 24(6): 1683–1700.

Jia, Jian, Ginger Zhe Jin, and Liad Wagman. 2018. "The short-run effects of GDPR on technology venture investment." National Bureau of Economic Research.

Johnson, Garrett A, Scott K Shriver, and Samuel G Goldberg. 2023. "Privacy and market concentration: intended and unintended consequences of the GDPR." *Management Science*.

Jones, Charles I, and Christopher Tonetti. 2020. "Nonrivalry and the Economics of Data." *American Economic Review*, 110(9): 2819–2858.

Jovanovic, Boyan, and Yaw Nyarko. 1996. "Learning by Doing and the Choice of Technology." *Econometrica*, 64(6): 1299–1310.

Kaplan, Jared, Sam McCandlish, Tom Henighan, Tom B Brown, Benjamin Chess, Rewon Child, Scott Gray, Alec Radford, Jeffrey Wu, and Dario Amodei. 2020. "Scaling laws for neural language models." *arXiv preprint arXiv:2001.08361*.

Keane, Michael P. 2010. "Structural vs. atheoretic approaches to econometrics." *Journal of Econometrics*, 156(1): 3–20.

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.

Levitt, Steven D, John A List, and Chad Syverson. 2013. "Toward an understanding of learning by doing: Evidence from an automobile assembly plant." *Journal of political Economy*, 121(4): 643–681.

Norets, Andriy. 2009. "Inference in dynamic discrete choice models with serially orrelated unobserved state variables." *Econometrica*, 77(5): 1665–1682.

Olley, Steven, and Ariel Pakes. 1996. "The dynamics of productivity in the telecommunications equipment industry." *Econometrica*, 64(6): 1263–1297.

Pakes, Ariel. 1986. "Patents as Options: Some Estimates of the Value of Holding European Patent." *Econometrica*, 54(4): 755–784.

PatentsView. 2022. "Forward Citation Dataset." *[Data set] PatentsView.* https://www.patentsview.org/download/.

Peukert, Christian, Stefan Bechtold, Michail Batikas, and Tobias Kretschmer. 2021. "Regulatory Spillovers and Data Governance: Evidence from the GDPR." *SSRN Electronic Journal*.

Rust, John. 1987. "Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher." *Econometrica: Journal of the Econometric Society*, 999–1033.

Schumpeter, Joseph A. 1943. *Capitalism, socialism and democracy.* routledge.

Scott Morton, Fiona, Pascal Bouvier, Ariel Ezrachi, Bruno Jullien, Roberta Katz, Gene Kimmelman, A Douglas Melamed, and Jamie Morgenstern. 2019. "Committee for the study of digital platforms: Market structure and antitrust subcommittee report." *Chicago: Stigler Center for the Study of the Economy and the State, University of Chicago Booth School of Business*, 36.

Tambe, Prasanna, Lorin Hitt, Daniel Rock, and Erik Brynjolfsson. 2020. "Digital capital and superstar firms." National Bureau of Economic Research.

Thompson, Peter. 2012. "The relationship between unit cost and cumulative quantity and the evidence for organizational learning-by-doing." *Journal of Economic perspectives*, 26(3): 203–24.

Yang, Nathan. 2020. "Learning in retail entry." *International Journal of Research in Marketing*, 37(2): 336–355.