University of Nevada, Reno

Entrepreneurship in times of changing technology and labor markets

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Economics

by

Trevor D. McLemore

Dr. Frank. M. Fossen / Dissertation Advisor

May, 2025

THE GRADUATE SCHOOL

We recommend that the dissertation
prepared under our supervision by

Trevor D. McLemore

entitled

Entrepreneurship in times of changing technology and labor markets

be accepted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

Frank M. Fossen, Ph.D.
Advisor

Kym Pram, Ph.D.
Committee Member

Mark A. Pingle, Ph.D.
Committee Member

Garret Ridinger, Ph.D.
Committee Member

Ron Lembke, Ph.D.
Graduate School Representative

Markus Kemmelmeier, Ph.D., Dean
Graduate School

May 2025

Abstract

This dissertation investigates research questions in entrepreneurship, with a particular focus on the effects of recent and emerging technologies. Chapter One explores how entrepreneurship and market structure affect emerging technologies, while Chapters Two and Three examine the impact of these technologies on entrepreneurship. I incorporate Artificial Intelligence (AI) both as a methodological tool (Chapter One) and as a subject of inquiry (Chapter Three). The first chapter draws on data from the Google Play Store and applies Natural Language Processing techniques to measure the similarity of app descriptions, analyzing the relationship between market concentration and product differentiation. The second chapter uses data from the American Community Survey and the timing of Uber's entry into metropolitan areas to assess the impact of rideshare on taxi drivers' wages and labor supply. The third chapter reviews the emerging literature on AI and entrepreneurship.

The first chapter, "Do Apps Play Follow the Leader? Testing the Relationship between Market Power and Product Similarity with Language Models", coauthored with Kym Pram, examines the relationship between product similarity and market concentration. We employ Bidirectional Encoder Representations from Transformers (BERT) to embed product descriptions and use the Herfindahl-Hirschman Index (HHI) to capture market concentration. Our analysis reveals a robust U-shaped relationship that flattens for recently updated apps and apps where users interact. These findings suggest that the incentive for acquisition-driven market entry is the dominant mechanism only in markets characterized by high concentration.

The second chapter, "Revisiting the Uber Effect", is a solo-authored paper that replicates and extends a study by Berger et al. (2018) on whether Uber drivers displace conventional taxi drivers. The analysis leverages the timing of Uber's market entry as exogenous variation in a difference-in-differences approach that analyzes taxi driver wage, salary, and labor supply. I find an 8.5-9.8% decrease in hourly earnings among wage employed drivers, no significant effect on salary, and a 7.7-12.3% decrease in labor supply of wage employed drivers. I also identify data irregularities in the Berger et al. (2018) paper and find that the larger and more statistically significant results for wage and salary that they found were not robust.

The third chapter, "Artificial Intelligence and Entrepreneurship", coauthored with Frank M. Fossen and Alina Sorgner, reviews the literature on impacts of AI on entrepreneurship. It begins by clarifying definitions of AI to eliminate ambiguity and provide context to how various studies use the term. The chapter discusses theoretical frameworks and empirical evidence related to the adoption of AI technologies and how AI technologies affect entrepreneurial opportunities, decision making under uncertainty, entry barriers, and business performance. An original empirical analysis from the German Socio-economic Panel is introduced, showing that entrepreneurs demonstrate greater awareness and use of AI technologies than paid employees. We review indirect impacts of AI on entrepreneurship through the labor market, finding evidence suggesting that automation results in higher levels of necessity entrepreneurship while transformative technologies, those that do not necessarily displace workers, lead to higher levels of opportunity entrepreneurship. The entrepreneurial ecosystem literature suggests AI reshapes the importance and configuration of existing ecosystem elements and processes and may reduce the role of geography in entrepreneurial activity. We conclude with a discussion on regulation of AI, with a focus on developments within the European Union.

Contents

List of Tables

List of Figures

1 Do apps play follow the leader?

Testing the relationship between market power and product similarity with language models

Trevor McLemore (University of Nevada, Reno)

Kym Pram (University of Nevada, Reno)

1.1 Abstract

This paper studies the relationship between market concentration and product differentiation. In more concentrated markets, market entrants may try to imitate market leaders in order to be bought out, leading to less variety and innovation. We test this hypothesis using data from the Google Play app store in 2019. We calculate app similarity scores, using Bidirectional Encoder Representations from Transformers (BERT) embeddings of product descriptions, and market concentration is measured with the Herfindahl-Hirschman Index (HHI). We find a robust U-shaped relationship between market concentration and product similarity that flattens for recently updated apps and apps where users interact. This suggests that the buy-out incentive is the dominating mechanism only for markets with high concentration. We discuss various other potential mechanisms that may explain the opposite relationship in markets with low to moderate concentration. Understanding that an ideal range of competition exists, after which product differentiation decreases, is useful for merger policy.

1.2 Introduction

Currently, there is a case, Federal Trade Commission v. Meta Platforms, Inc., being held in the U.S. District Court for the District of Columbia. The Federal Trade Commission (FTC) alleges Meta illegally acquired Instagram (2012, $1B) and WhatsApp (2014, $22B). During the WhatsApp deal, Mark Zuckerberg, Meta's CEO, wrote in an email "it is better to buy than to compete". Post-merger price predictions, used in typical enforcement, are difficult to calculate with tech startups because they typically do not command a significant share of the market. The FTC argues that after the mergers, the platforms exhibited less data privacy with more buggy and expensive services to advertisers and deterred competitor entry. Meta argues that they compete in a wider range than claimed and that the FTC must prove that consumers would have had more (or better) options sooner without the acquisitions (Dave, 2025). By shedding light on the relationship between market concentration and product differentiation in the app market, our analysis contributes to informing such assessments. This case motivates analyzing concentration in the app market and shows the importance of considerations other than price to form a wholistic argument.

In our paper, we empirically test the effect of market concentration on product similarity using cross-sectional data from the catalog of 3.15M apps on the Google Play Store at the end of 2019. HHI is calculated for each of the 49 categories in the Google PlayStore to meausure market concentration. We use novel Natural Language Processing (NLP) techniques in the form of BERT embeddings that numerically capture semantics and vocabulary in the app descriptions to assign a (product) similarity score to each app. We regress HHI on similarity using a quadratic term to capture nonlinearity, testing whether similarity increases with market power. The tested mechanism is for new apps with features similar to existing apps to enter concentrated categories with the incentive to be acquired, which we only observe in categories that are already highly concentrated. We find a robust U-shaped relationship between market concentration and similarity that flattens for recently updated apps and apps where users interact.

Our research builds on the literature on industrial organization, specifically the relationship between market concentration and product differentiation, the literature on the app market, and the literature that uses natural language processing in the field of economics. The relationship between competition and innovation in product markets has been well studied, dating at least as far back as Joseph Schumpeter in the early 1900s (Gilbert et al., 2006). Empirical evidence shows positive effects of competition on innovation (Baily and Gersbach, 1995; Nickell, 1996; Blundell, Griffith, and van Reenen, 1999). Standard industrial organization theory predicts that innovation decreases with competition as innovators receive less reward from lower monopoly rents (Dasgupta and Stiglitz, 1980). However, more recent theoretical papers show positive (Aghion et al., 2001) or inverted U-shaped (Aghion et al., 2005) relationships, since most innovative activity occurs within industries of more than one firm and in firms already earning rents. Thus, rents to innovators may be a poor indicator of incentive to innovate.

The research mentioned above uses productivity, amount of money spent on Research and Development (R&D), patent or breakthrough technology count, or a combination of those metrics to capture innovation. Each of these methods has benefits and limitations. Productivity, measured through Solow residuals, is correlated with market power and has variance due to different levels of diffusion (Blundell et al., 1999). R&D is an input with innovation as the output goal; it can overlap in different firms, be spent narrowly or broadly, or have different levels of effectiveness. Innovations are not always patented, and not all patents are implemented as innovations, while counting breakthrough technologies as an alternative metric to patents is subjective. It is clear that productivity, R&D spending and patents should be correlated with what we think of as innovation, but from the shortfalls listed for each metric it should also be clear that, even together, they paint an incomplete picture of innovation.

Market entry is a measure of entrepreneurship that is highly affected by the competitive dynamics of the market. If entrepreneurs are predominantly copying existing products, that would negatively impact product differentiation and innovation. This paper focuses on the types of products that entrepreneurs develop. Phillips and Zhdanov (2013) show theoretically and empirically that a higher number of small firms leads to less innovation by large firms. A lower chance of beating all small firms to a specific innovation gives the incentive to shop innovations, that is to say, more acquisition targets, which is even more pronounced during periods of economic growth. Hoberg and Phillips (2010) show, using a bag-of-words approach in product descriptions of 10k reports, that firms are more likely to be acquired when the language describing their assets is similar to all other firms. A firm whose product is broadly similar likely owns assets that can be deployed in many different product markets. If, as the literature indicates, large firms are more likely to acquire firms than invest in their own R&D, acquisition targets are more likely to have products similar to all other firms, and there is an incentive to enter a market in order to be bought out, then product differentiation could be decreasing due to market power.

Rasmusen (1988) is a theoretical paper that shows that buyouts change the direction of credible threats and that an entrant will enter despite knowing that the duopoly price will be less than their average total costs if buyouts are possible. He acknowledges that the incumbent's capacity cost may already be sunk, but also adds the opportunity cost of not earning monopoly profits to the model, which is not a sunk cost. The joint profit of a monopolist is more than the combined profit of competitors, so there is an incentive to acquire the entrant; the entrant foresees the buyout. Dijk et al. (2024) find, when buyouts are allowed, that startups allocate more towards rival projects than they would otherwise, and this effect is robust to drastic innovations, multiple incumbent firms, and horizontal and vertical product differentiation. As large firms acquire entrants with broadly similar products, entrants have an incentive to enter a market with their product similar to existing products in order to be acquired, and thus we should indeed see product differentiation decreasing with market power.

Firms entering concentrated markets through the buyout channel (Rasmusen, 1988) can be harmful to welfare when these entrants make their assets similar to existing firms (Hoberg and Phillips, 2010), shift their investments to rival projects (Dijk et al., 2024), and cause large firms to innovate less (Phillips and Zhdanov, 2013), thus reducing product differentiation. A smaller selection of products has a negative impact on consumer welfare. There are also other effects and dynamics that make research into the interplay of product differentiation, market entry, and competition important. Merger enforcement usually relies on estimates of price effect immediately following a merger, but since startups typically do not have significant market share, have a lot of uncertainty in existing data, and may be complementary in some respects while substitutable in others, the typical analysis will usually suggest no potential harm. Antitrust laws prevent us from directly observing blatant examples of entry for buyout (Rasmusen, 1988). However, when startup acquisition is minimally regulated, Bryan and Hovenkamp (2020) show that over time the technological gap between leaders and less successful rivals expands, leading to additional concentration. Cunningham et al. (2021) show that conservative estimates are 5.3 - 7.4% of all acquisitions in their pharmaceutical industry sample are done simply to scrap the product, in other words, killer acquisitions. We see less R&D spending (Phillips and Zhdanov, 2013) and more acquisitions (Hoberg and Phillips, 2010) in concentrated markets, so there is no expectation that the market will become more innovative or self-correct from the pattern of powerful incumbents purchasing promising technology.

There may also be reasons to see effects in the opposite direction. When there are similar alternative firms, the target firms earn lower announcement returns (Hoberg and Phillips, 2010) reducing the incentive to enter the market in order to be bought out. Similarly, when a habitual acquirer becomes sufficiently dominant, it is less willing to pay for new technologies, reducing the returns and incentives for new firms to enter that market (Bryan and Hovenkamp, 2020). If large firms are able to successfully deter entry into the market as also suggested by Bryan and Hovenkamp (2020), we may also see an effect in the opposite direction, as similar firms would not compete. An example of this entry deterrence is if a marketplace requires significant network effects and there is a large incumbent that already has most of the market share.

The app literature is well studied, probably because the data for many transactions are captured virtually, providing large and high-quality data sets. Ghose and Han (2014) research consumer preference towards app characteristic effects (what makes users download more?); Roma and Ragaglia (2016) research app characteristic (paid, free, freemium, category, and platform) effects on revenue;

Wang et al. (2024) research entry deterrence effects of market consolidation from different types of mergers. Mergers across different categories of apps and those that acquired new technologies to enhance the functionality of their service had stronger entry deterring effects.

1.3 Further related literature

1.3.1 Natural language processing techniques

In this paper, we use natural language processing to estimate the similarities of textual app descriptions, with similarity being a proxy for product homogeneity. One of the earliest techniques in natural language processing is the bag-of-words model, appearing around the 1950s using word frequency to perform tasks such as classifying or searching for documents. Bag-of-words is still effectively used today having evolved with fine-grained modifications such as removing common words called 'stop words', reducing words to their root form through stemming or lemmatization, and using cross-document word count to normalize the contribution of each word to the model through TF-IDF (term frequency – inverse document frequency). Bag-of-words lacks the ability to capture the relation between words and the context of that order. This limitation is exemplified in the first reference to the term 'bag of words' that we could find, which was an influential article in linguistics, "language is not merely a bag of words, but a tool with particular properties which have been fashioned in the course of its use" (Harris, 1954, p 156). Language embeddings are vectorized forms of written passages that are trained using machine learning to capture elements of the language; in the transformer model that we use to generate embeddings, many of the semantics that a bag-of-words model would miss are captured. In this subsection, we walk through the development of the model we use.

Despite being the predominant language modeling method for many NLP tasks through 2014, Deep Neural Networks (DNNs) require fixed and known input and output lengths. Zero padding and truncation are workarounds for the varying length inputs. Sequence-to-sequence solved the problem of required fixed and known input and output lengths (Sutskever et al., 2014) by using Long Short-Term Memory (LSTM) to map an input sequence to a fixed dimension vector and then another LSTM to decode the vector. LSTMs are a subcategory of Recurrent Neural Networks (RNNs). RNNs capture order by making the output of a neuron part of the input at the next neuron, but sometimes stop learning due to vanishing gradients. LSTMs continue to learn by selectively remembering information through an input, output, and forget gate.

The transformer, the base framework for modern models including ours, kept the encoder-decoder architecture of sequence-to-sequence but replaced recurrent layers with multi-headed self-attention (Vaswani et al., 2023) increasing performance, reducing complexity per layer, adding parallelizability, and reducing path length between long-range dependencies in the network. Attention layers map a query vector, key vector, and value vector to an output vector. The weight assigned to each value is computed from the query and the key with a compatibility function, effectively telling the model which parts of the input are important.

In this paper, we use Bidirectional Encoder Representations from Transformers (BERT). BERT produced state-of-the-art results in 11 NLP tasks, including General Language Understanding and Evaluation (GLUE), which runs a series of tests, many of which capture semantics. BERT is a Masked Language Model (MLM), which means that in BERT's case during pre-training 15% of words are randomly replaced. 80% of replaced words are masked to be predicted, 10% are replaced by the correct word, and 10% are replaced by random words. Devlin et al. (2018) found that this process forced the model to keep a distributional context of every input layer instead of focusing only on masked words, which is what happens if all the changed words are masked. The MLM approach allows for the fusion of left and right context, so it can use information from the whole sentence surrounding the masked word. The model is also trained on Next Sentence Prediction (NSP) which 50% of the time places a random sentence from the corpus after sentence A and asks the model to predict when sentence B is actually the next sentence which further trains the model's representation of context. BERT is trained on the BooksCorpus (800M words) and Wikipedia (2.5B words). BERT has recently been used in the economic literature to estimate the premia associated with professional certificates (Bana, 2021) and to estimate hedonic prices from Amazon product descriptions for comparison with CPI (Bajari et al., 2023).

Both examples we gave for economic literature with BERT, used it with a fine-tuning approach. The fine-tuning approach is when a classification layer is added to the pre-trained model and all parameters are fine-tuned on a downstream task. In our paper, we do not have a classification task to fine-tune BERT on as there is no supervised data to generate the similarity score. Thus, we are interested in a feature-based approach which generates embeddings from the app descriptions using the pre-trained model. BERT is effective for both fine-tuning and feature-based approaches. From the original BERT paper, for a Named Entity Recognition (NER) task, which categorizes words, the feature-based approach had an F1 score of 96.1, the fine tuning 96.4, and ELMo, a high performing feature based embedding approach, 95.7 (Devlin et al., 2018). An F1 score of 100 represents perfect accuracy,

1.3.2 Market structure and product differentiation

In the Introduction, we write broadly about how the structure of markets affects innovation and how product differentiation fits into the innovation discussion. Here, we briefly discuss the relationship between market structure and product differentiation, which is an important theme in the field of economics that uses merger simulation or empirical evidence to provide insight into changes in product offerings and their welfare effects.

Mergers make a market more concentrated, but their effect on product differentiation is more complicated. An incentive exists not to compete against a firm's own product offerings, but does this mean 'killer acquisitions' (Cunningham et al., 2021)? In the radio industry, acquisitions are differentiated to avoid audience cannibalization at the firm level, which increases product differentiation. However, as each radio station differentiates their own products, they become more like their competitors, so the variety does not increase in aggregate (Sweeting, 2010). The consolidation of resources may, on the other hand, allow firms to offer additional products, so these and other competing factors make the effect of mergers on the product variety an empirical question (Berry and Waldfogel, 2001). Fan (2013) simulated a merger between two Minneapolis newspapers that the Justice Department blocked and found a projected decrease in content quality, local news ratio, and variety if the merger had been allowed to occur. Fan and Yang (2020) simulated a merger between Samsung and LG for their smartphone offerings and found a decrease in the number and variety of products. Wollmann (2018) shows evidence of expansion in the choice set in the commercial truck industry after a merger through the mechanism of increased mark-ups and a decreased output that attracts new product entrants to compete.

A recurring impetus for understanding the relationship between market structure and product differentiation is that price predictions from mergers can be biased when they treat product offerings as fixed (Mazzeo et al., 2018; Fan, 2013; Fan and Yang, 2020; Wollmann, 2018). Fan (2013) and Fan and Yang (2020) find that ignoring adjustment of product characteristics underestimates consumer surplus and overestimates producer surplus, while Wollmann (2018) finds that it overestimates price markups. Although our research does not analyze this topic from a multi-product firm merger perspective, our findings do provide general insights into the direction product differentiation may change due to market concentration.

1.4 Data and methods

1.4.1 Data

Table 1: Summary statistics

	Mean	σ	Min	Max
HHI	447.4	667.9	18.5	3592.4
similarity	0.586	0.050	0.153	0.714
weighted similarity	0.593	0.054	0.154	0.780
year created	2016.3	2.1	2011	2019
year of last update	2018.0	1.5	2011	2019
developer total apps	24.9	88.5	1	1141
screenshots	8.7	5.7	0	32
downloads	619.6k	25.4M	5k	5B
size (bytes)	19.4M	28.8M	3.2k	3.3B
number of ratings	11.1k	369.8k	0	101M
5 star rating	0.178	0.383	0	1
4 star rating	0.639	0.480	0	1
3 star rating	0.149	0.356	0	1
2 star rating	0.028	0.165	0	1
1 star rating	0.002	0.045	0	1
not rated	0.004	0.062	0	1
ads	0.688	0.464	0	1
digital purchases	0.217	0.412	0	1
not free	0.018	0.132	0	1
game	0.230	0.421	0	1
users interact	0.110	0.313	0	1
shares location	0.022	0.148	0	1
mature	0.030	0.170	0	1
violence	0.051	0.221	0	1

425,817 observations; Google Playstore 2019

We use a cross-sectional dataset that contains 3.15 million apps, all the apps in the Google Playstore at the end of 2019. We filter the dataset to only include apps in the English language to allow the use of an English language BERT model that provides more nuanced embeddings, but also to avoid the potential of duplicate apps for different languages. We remove apps with less than 5k downloads to capture fewer of the apps that are not built to be businesses (we use a variety of different cutoffs in the appendix with robust results). We remove apps created in 2008, 2009, and 2010 because the low number of apps created in these years could bias the results. None of the industry leaders with more than 5B downloads are included in those years and still offered in 2019. We remove apps that are labeled 'uncategorized' or do not have data for the creation date, the number of downloads, size, developer, ratings, or price because our regressions use this information and the number of apps missing these data is small. After cleaning the data, we have just over 425k apps remaining. Table

1 provides descriptive statistics of the apps in the sample.

Although our dataset is high quality with few missing values and a wide selection of reliable information attached to each app, cross-sectional data does have some shortcomings. Early unsuccessful apps may drop off the Playstore (developers pay to host apps), which biases the sample for older apps towards having been more successful. We do not believe that survival bias is a concern, since our focus is on whether apps become more similar in concentrated categories. If an app disappears due to lack of success, there is nothing to observe to copy. Since downloads are cumulative, we also see a bias toward a higher number of downloads in older apps, and we cannot tell whether the app is still a market leader or had a moment of popularity early in its release. This effect is pronounced, as apps created in 2011 have 18% of downloads, while apps created in 2019, despite being far more numerous, only have 2% of downloads (Figure 1). Again, our focus is on whether apps become more similar in concentrated categories, and since cumulative market share has a historical context, it may be better than observing current market share.

Figure 1: App quantity and cumulative downloads based on creation year

The dataset is not balanced temporally. Just over 1% of the apps increasing roughly linear to 20% in 2017 and then down to close to 14% in 2019 (Figure 1). The reason we see the small drop after 2017 is that we removed apps with less than 5k downloads, and as previously stated, downloads are cumulative. 2011 still has enough data to be informative in our sample, but we are unable to do an event study as most of the leading apps by download for each category were created in 2011 or 2012 and there are not enough pre 2011 data points to make comparisons.

All app categories have been the same since at least 2011, but are not balanced (Table 2). There are 49 categories. The category with the most apps, tools, comprises 8.1% of the total apps. The

Table 2: Category statistics ordered by market concentration

Category	HHI	Similarity	Downloads	n
travel and local	3,592	0.5885	880,186	9,681
video players	2,175	0.5883	1,967,187	5,606
social	1,791	0.5769	1,755,054	7,268
music and audio	1574	0.5588	700,106	18,278
news and magazines	1,213	0.5674	641,296	5,986
books and reference	1,134	0.5511	178,130	17,104
events	1,018	0.5682	67,346	537
libraries and demo	895	0.5706	145,868	20,071
communication	859	0.5817	3,500,001	7,516
maps and navigation	781	0.5914	408,957	4,765
tools	519	0.5852	1,145,969	34,688
beauty	485	0.5771	170,830	1,470
health and fitness	439	0.5880	259,758	9,940
game strategy	405	0.6102	1,024,648	2,684
game music	371	0.6029	786,425	1,165
game trivia	336	0.6074	333,092	2,010
productivity	332	0.5779	1,552,494	12,091
food and drink	301	0.5910	229,768	3,341
weather	271	0.5909	418,704	2,330
house and home	254	0.5785	188,274	1,214
parenting	245	0.5926	157,159	762
art and design	235	0.5898	146,059	3,599
personalization	218	0.6065	213,130	34,322
game word	203	0.6293	568,413	2,000
photography	193	0.6074	686,940	17,706
dating	178	0.6101	248,420	1,152
game sports	175	0.6096	1,309,969	3,595
game board	174	0.6071	476,317	3,101
sports	170	0.5757	159,756	6,611
game casino	169	0.6215	378,559	2,558
business	164	0.5808	147,583	11,814
game arcade	145	0.6112	1,016,698	9,622
game card	141	0.6070	361,936	3,246
entertainment	135	0.5657	348,978	30,456
auto and vehicles	135	0.5773	230,314	3,202
shopping	128	0.5951	654,687	5,582
comics	117	0.5985	290,996	2,174
game casual	99	0.6028	905,760	12,496
game racing	96	0.6230	1,389,793	5,215
game adventure	93	0.6066	381,343	6,249
finance	89	0.5947	193,572	12,355
lifestyle	78	0.5663	145,868	20,071
medical	78	0.5859	74,332	3,628
game educational	61	0.6050	465,694	4,950
game action	57	0.6051	1,104,808	9,199
game role playing	55	0.6015	586,359	3,171
game puzzle	42	0.6162	457,879	14,358
education	27	0.5636	90,526	31,868
game simulation	18	0.5930	586,412	12,402
average	447	0.5864	619,615	8,690

425,817 observations; Google Playstore 2019 category with the least apps, events, comprises only 0.1%. The events category still has enough data to be informative in our sample. The app developer chooses the category for their app, but there does not seem to be an advantageous reason to choose a category that does not fit. Labeling an app with a category that does not fit the app would make it more difficult to search and potentially annoy customers.

1.4.2 Similarity score construction

From the description of the app, which we use to generate similarity scores, we remove extra spaces, repeated punctuation, emoticons, symbols, and foreign characters. The main reason for cleaning in this manner is that BERT limits the input to 512 tokens. A token is a word or part of a word with 100 tokens equal to approximately 75 words (Bana, 2021). Symbols and extra punctuation use tokens at a much faster rate. After cleaning, for the very few descriptions that need truncation, we take the first 512 tokens to generate the embedding.

After cleaning the app descriptions, we run them through the BERT model which transforms the descriptions into embeddings, numerical vectors with two dimensions for each app[1]. The first dimension of the embeddings is the number of tokens from its description (variable and capped at 512). The second dimension of the embeddings is a fixed number, 768, based on the output architecture of the BERT model. There are many methods used to extract fixed features from the embeddings which all involve reducing the token dimension, which is the only variable dimension. Taking the last hidden layer, the second to last hidden layer, a weighted sum of the last four hidden layers, or concatenating the last four hidden layers are all mentioned in the original BERT paper (Devlin et al., 2018). The authors write that the best performing method in the feature-based approach concatenates the top four hidden layers of the pre-trained transformer. Bajari et al. (2023) also uses this method. We follow suit and thus bring the dimensions of our embeddings to 1 x (768*4). Thus, each row of our data has a 1 x 3072 numerical vector that represents the tokens (words and word fragments), positions, and sentence orders of each app description.

Once we have the embeddings (rows x 1 x 3072), we take the cosine similarity for each app compared to every other app in its category giving us square matrices for each category (rows in category x rows in category). We then take the mean across a matrix dimension to give us a vector of length equal to the number of rows in the category. Each value in this vector is the average similarity of

[1]The output gives three dimensions, but one dimension is simply the number of descriptions fed into the model.

that app description to the other app descriptions in its category. Since we do not want the app descriptions to be compared with themselves, we correct for the ones on the diagonal of the cosine similarity matrix with the equation:

$$\text{similarity} = \frac{\text{biased similarity} \times \text{number of apps in category - 1}}{\text{number of apps in category - 1}} \tag{1}$$

For robustness checks, we also take a version of similarity weighted by downloads where the magnitude of the self-comparison correction is much larger, especially for concentrated categories. Instead of taking the mean across a matrix dimensions, we multiply by the proportion of downloads of each app and then sum across that dimension giving us a vector equal to the number of rows in the category. The correction equation for weighted similarity is:

$$\text{weighted similarity} = \frac{\text{biased weighted similarity} \times \text{category downloads - app downloads}}{\text{category downloads} - \text{app downloads}} \tag{2}$$

1.4.3 Industry concentration score construction

The number of downloads for an app is given as a range in the dataset (0, 1+, 5+, 10+, 50+, 100+, 500+, . . . , 1B+, 5B+). Since we use the number of downloads to calculate market share, we must use a number instead of a range for the calculations. We use the minimum number of confirmed downloads to calculate the market share. The highest category of downloads is bounded below 10B.

The Herfindahl-Hirschman Index (HHI) is one of the most used indicators to detect anticompetitive behavior in industries. Evidence shows that higher HHI value indicates higher price-cost margin (Matsumoto et al., 2012). In Appelbaum (1982), which uses production theory to derive oligopoly market power, he notes that HHI is a special case of oligopoly power where the inverse demand elasticity equals one under Cournot behavior. Hoberg and Phillips (2010) is an example of a paper that uses HHI to classify industries into competitive and concentrated industries. As Figure 2 shows, markets are classified as moderately concentrated if their HHI is between 1k and 1800 and highly concentrated if their HHI is above 1800 (U.S. Department of Justice and the Federal Trade Commission, 2023). The HHI for a category c is defined as:

$$HHI_c = 10,000 * \sum_{i=1}^{N_c} s_i^2 \tag{3}$$

N_c: number of apps in category c

s_i: share of category downloads for app i

We calculate the market share of an app by taking its cumulative downloads and dividing by the cumulative downloads of the entire category where the app falls. By squaring the share of an app, which is less than one, before it is summed, we get small values for categories with lots of apps that are similar in size (competitive) and large values when nearly all downloads belong to a small number of apps (concentrated). We calculate an HHI for each category. For robustness, we also calculate a different concentration metric, which is the market share of the most downloaded app in each category.

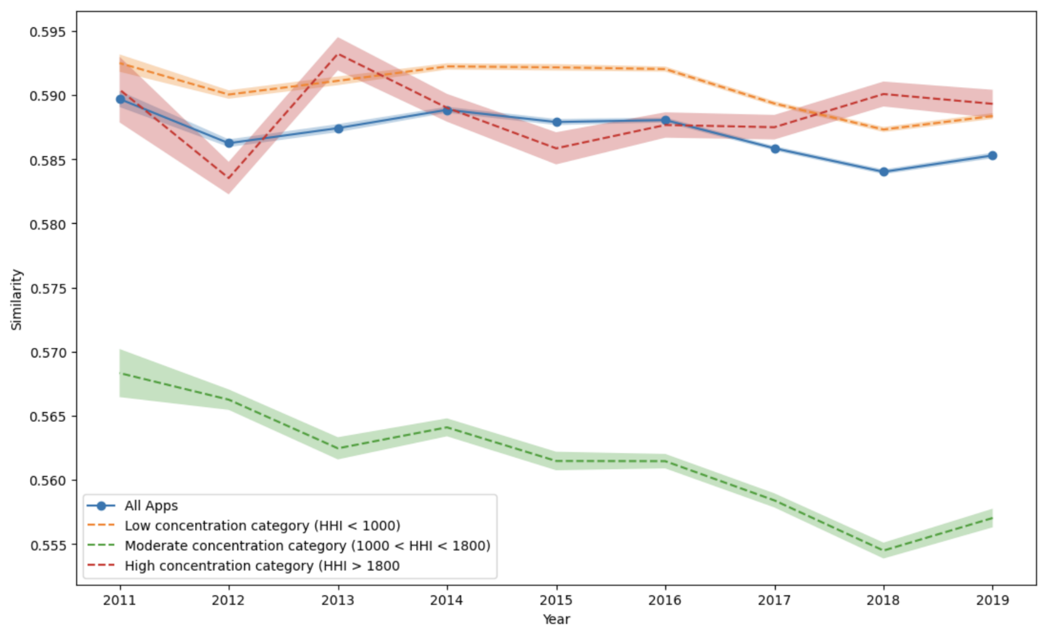


Figure 2: Similarity with standard errors based on app creation year

1.4.4 Methods

Our hypothesis is that new apps with features similar to existing apps enter concentrated categories with the incentive of being acquired. Our data appear non-linear when comparing similarity to market concentration which we take into account with a quadratic term. We test this hypothesis with the regression:

$$y_i = \alpha_1 HHI_i + \alpha_2 HHI_i^2 + \delta X_i + \xi_i \tag{4}$$

Equation (4) compares similarity and HHI with the assumption, based on observing the data when separated by concentration categories, that the relationship is nonlinear. y_i is the similarity score of each app to all other apps in its category. For robustness checks, we also look at similarity weighted by downloads. HHI is a continuous variable that measures the concentration of the category. For robustness checks, we also look at the download share of the top app in each category as an alternate evaluation of industry concentration. All control variables are measured at the app level. Control variables are introduced by tiers and include creation year, whether users interact, location sharing, info sharing, size, rating, editor's choice selection, ads, digital purchases, free, mature content, gambling, violence, adult language, suggestive themes, and drug references. The ability to control for app-level heterogeneity is the main motivation for running the regressions at the app level. Standard errors are clustered at the category level. We also run a robustness check using data aggregated to the category level.

We are also interested in whether the similarity of apps changes over time in categories with different levels of market concentration. We explore these interactions with the regression:

$$y_i = \alpha_1 HHI_i + \alpha_2 HHI_i^2 + \beta_1 HHI_i * t_i + \beta_2 HHI_i^2 * t_i + \beta_3 t_i + \delta X_i + \xi_i \tag{5}$$

Equation (5) compares similarity and HHI, allowing for nonlinearity, and interacting with t_i, a continuous variable that captures the year of app creation; 2011 is coded as 0, counting up to 8 for 2019. We run this regression with the creation year variable interacting with the HHI variable because we are interested in whether high HHI scores affect similarity differently as new apps enter that market. Our data contains only the most recent description for each app so we run the interaction with year of last update as well, coded 0 to 8. The final interaction regression run is with a users interact dummy variable replacing the creation year or update year variable. We run this regression with the users interact interacting with HHI because we are interested in whether high HHI scores affect the similarity variable differently in a market place where network effects may deter entry in concentrated categories.

The nonlinear relationship between similarity and market concentration can be captured with market concentration dummy variables instead of a quadratic term which we test with the regression:

$$y_i = \alpha_1 HHI_i^l + \alpha_2 HHI_i^h + \delta X_i + \xi_i \tag{6}$$

Equation (6) compares similarity and HHI as a piecewise constant function using dummy variables for different market concentrations. HHI_i^l is a dummy variable equal to 1 when HHI is below 1000, HHI_i^h is a dummy variable equal to 1 when HHI is above 1800, and HHI_i^m is the omitted base category representing the range 1000<HHI<1800. This regression is a robustness check testing whether there is a U-shaped relationship between HHI and similarity.

1.5 Results

1.5.1 Quadratic regressions

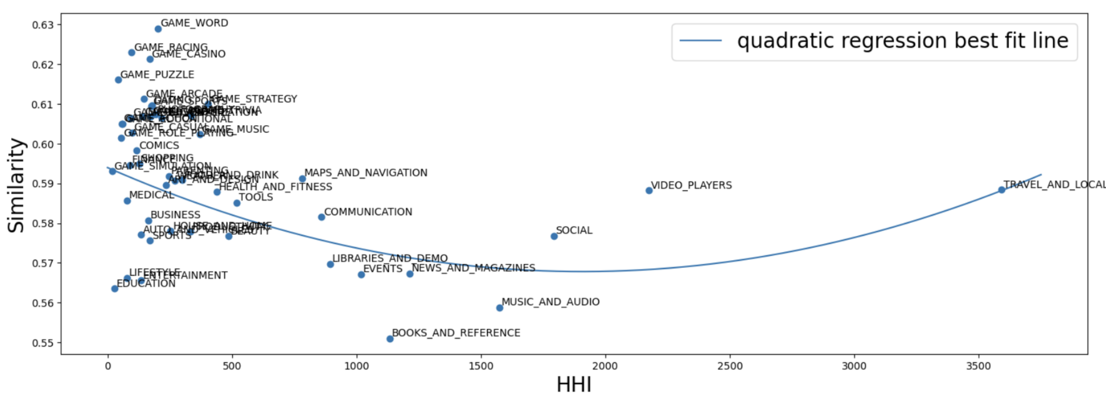


Figure 3: Similarity vs HHI

Notes: equation (4); function plots the regression predictions based on the app-level data

We observe a U-shaped relationship between similarity and HHI with significance in both HHI and HHI squared and a minimum similarity at the HHI value of 1,908 which is toward the beginning of the highly concentrated categories. Figure 3 shows the U-shaped relationship between similarity and HHI, as indicated by the significant coefficients of HHI and HHI squared at the 5% level in the quadratic regression reported in Table 3 (Column 1). The regression is robust to interacting with the year created (Column 2), update year (Column 3), or whether users interact on the app (Column 4); to representation as a piecewise linear function (Table 4); to category level regression (Table 5); to different download inclusion thresholds (Table 15); and to adding control variables (Table 16). If

Table 3: Quadratic regressions, outcome: similarity

	(1) Base	(2) Creation Year	(3) Update Year	(4) Users Interact
HHI/10k	-0.274**	-0.267***	-0.402***	-0.302**
	(0.122)	(0.094)	(0.136)	(0.127)
HHI/10k squared	0.717**	0.672***	0.929**	0.790**
	(0.315)	(0.231)	(0.365)	(0.329)
creation year		-0.000715		
		(0.000518)		
creation year x HHI/10k		-0.00191		
		(0.00966)		
creation year x HHI/10k squared		0.00996		
		(0.02746)		
update year			-0.000957**	
			(0.000423)	
update year x HHI/10k			0.0182***	
			(0.0062)	
update year x HHI/10k squared			-0.0304*	
			(0.0166)	
users interact				-0.00318
				(0.00285)
users interact x HHI/10k				0.167**
				(0.065)
users interact x HHI/10k squared				-0.451***
				(0.169)
Constant	0.594***	0.598***	0.601***	0.595***
	(0.006)	(0.005)	(0.007)	(0.007)
R-squared	0.0220	0.0229	0.0225	0.0232

425,817 observations; Google Playstore 2019
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
HHI is divided by 10k to scale the results

similarity is weighted by download share, the HHI_i^2 term falls just outside the significance range at the 10% level. Substituting the category leader download share for HHI increases standard errors, making the effect on similarity not significant, although the signs remain unchanged.

At 0 HHI, for every 100 points in which HHI increases, we see a decrease in similarity of 0.0027. The app that is least similar to all other apps in its category has a similarity score of 0.153 and the app that is most similar has a score of 0.714; there is a range of 0.561 between the most and least similar app. At a category level, the least similar category is books and references with a similarity score of 0.551 and the most similar category is word games with a score of 0.629 (see Table 2); the range between most and least similar categories is 0.078 so although those values appear small, they are still meaningful. The interpretation of the main result is that the apps become less similar as the HHI increases through the low and moderate category concentration range, but become more similar as HHI increases through the high category concentration range, suggesting there is an ideal range of competition that results in more differentiated products.

With quadratic regressions, there is always the risk that an outlier has a disproportionate weight. Travel and local is a category with an HHI of 3,592 making it much more concentrated than the next highest category, video players, with an HHI of 2,175 which is a much higher value than the category average HHI of 447. The travel and local category is one of 49 categories and consists of 9,681 apps, so although it contains more importance than a typical outlier, it does change the regression results. Figure 4 plots a linear regression with the travel and local category removed, which shows a decrease of 0.00153 similarity for every increase of 100 HHI, significant at the 5% level. With travel and local included, the linear regression is still negative, but not significant; just as with travel and local omitted, the quadratic has the same shape, but the coefficients are not significant. The categories, social and video players, also contribute to the U-shaped nature of the function, so the entire effect is not caused by a single outlier.

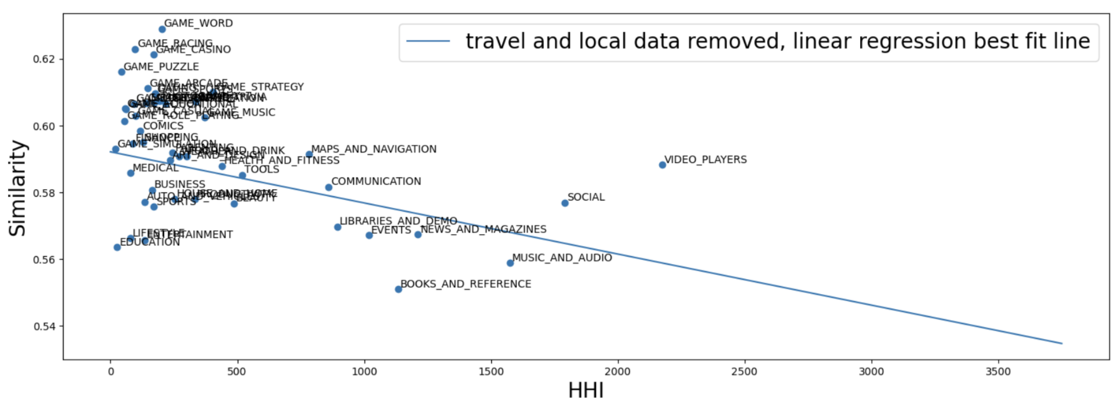


Figure 4: Similarity vs HHI - travel and local data removed

Note: function plots the regression predictions based on the app-level data

The year the app was created is not significant when treated as an interaction term (Table 3, Column 2). We observe the same robust U-shape as the main regression (Figure 3), but do not observe that similarity increases with HHI over time as we expected from our hypothesis, that entrants with similar products enter concentrated markets to be bought out. When we look at the creation year without interacting the variable with HHI, we do see a negative relationship with similarity at the 5% significance level, suggesting similarity decreases over time, but not more so in concentrated markets.

When we run the quadratic regression substituting the year the app was updated as the variable interacting with HHI, all variables in the regression are significant. The update year, without interaction, has a small negative value that is significant at the 5% level. The negative value suggests that when the apps are updated, the effect of them differentiating their descriptions from other apps in the category outweighs the effects in the other direction. The linear interaction term is positive, and the squared interaction term is negative, which means that the more recently an app was updated, the more the U-shaped relationship between similarity and HHI flattens, as shown in Figure 5. The mechanism is unclear.

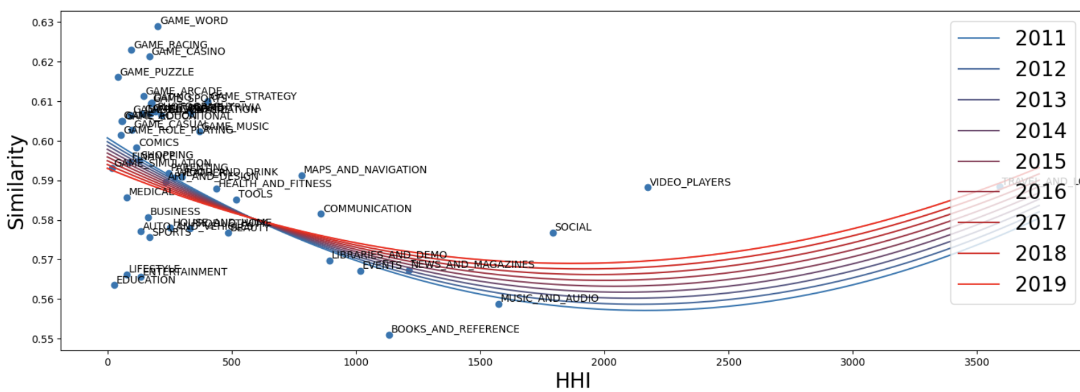


Figure 5: Similarity vs HHI and update year interaction

Notes: equation (5); function plots the regression predictions based on the app-level data

The interaction of the variable that captures whether users interact in the app, shows significance in the positive direction when interacting with HHI at the 5% level and in the negative direction when interacting with HHI squared at the 1% level. The direction of the interaction effects is opposite to the direction of the independent HHI terms, so we see a milder U-shaped function when users interact, as depicted in Figure 6. When the users interact variable is treated as a regular control, it does not show significance. A possible explanation of the interaction regression result is that positive network effects from a large user base are more difficult to achieve when powerful incumbents already benefit from those network effects, resulting in more frequent entry deterrence or early failure for entrants in concentrated markets.

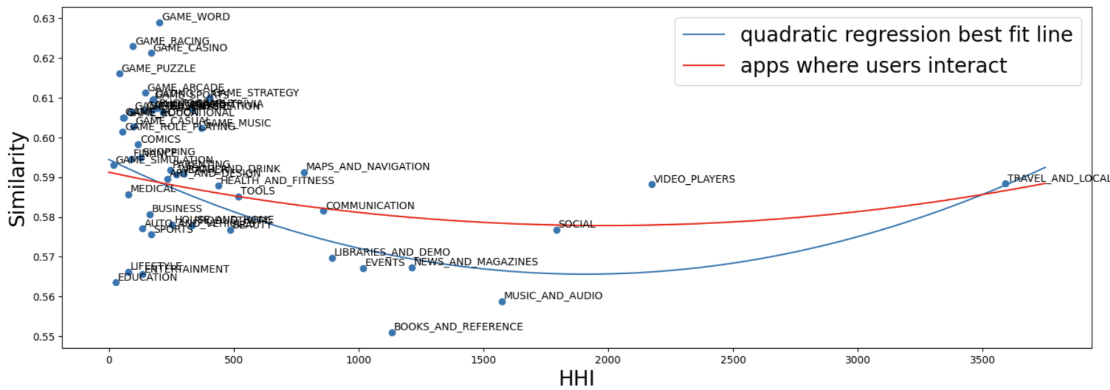


Figure 6: Similarity vs HHI and users interact interaction

Notes: equation (5); function plots the regression predictions based on the app-level data

1.5.2 Linear regression using market concentration dummy variables

Table 4: Linear regression with market concentration dummy variables, outcome: similarity

	Concentration Dummies
HHI low	0.030047***
	(0.005925)
HHI high	0.028482***
	(0.004150)
Constant	0.559957***
	(0.004150)
R-squared	0.0362

425,817 observations; Google Playstore 2019
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

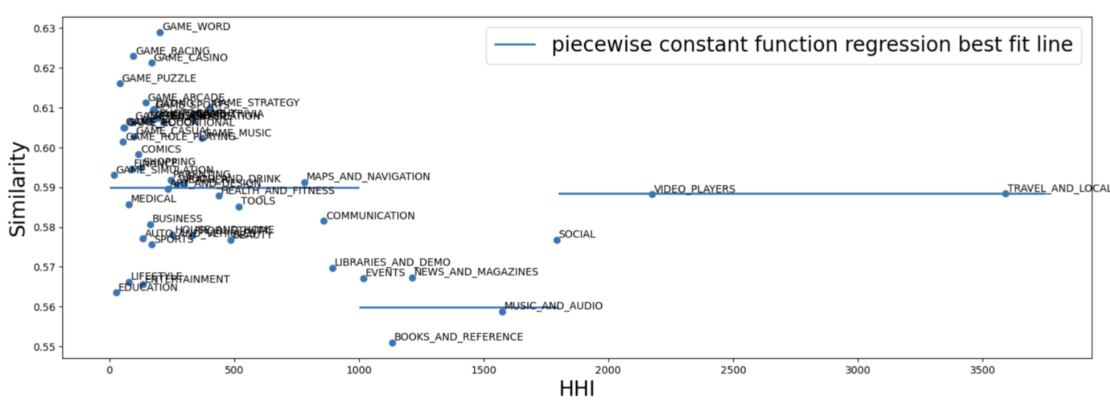


Figure 7: Similarity vs HHI piecewise constant

Notes: equation (6); function plots the regression predictions based on the app-level data

Table 4 shows the results from estimating eq. 6, using interval dummies of HHI as a robustness check. With the omitted base category as HHI_i^m, representing the range 1000<HHI<1800, it is clear from the relatively large, positive effects on similarity, significant at the 1% level for the low and high concentration categories, that we have a U-shaped function. This result can be seen in Figure 7.

1.5.3 Category level robustness check

We see the same relationship between similarity and HHI when we run the regression at the category level, but with slightly larger magnitudes and more significance (both HHI and HHI squared 1% significant, contrasted with 5% significant for both variables when run at the individual app level (Table 5).

Table 5: Quadratic regression at category level, outcome: similarity

	Category Level
HHI/10k	-0.338***
	(0.091)
HHI/10k squared	0.865***
	(0.300)
Constant	0.602***
	(0.003)
n	49
R-squared	0.2588

425,817 observations; Google Playstore 2019
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
HHI is divided by 10k to scale the results

1.6 Conclusion

This paper finds a robust U-shaped relationship between market concentration, measured by HHI, and product similarity, measured by a score derived from BERT embeddings of product descriptions, in the app market. We have not seen other literature that indicates, as our paper does, that there is an ideal range of competition that encourages product differentiation. Other literature suggests that there is an ideal range of competition that encourages innovation (Aghion et al., 2005); further research could look for ranges of market concentration that provide coinciding product differentiation and innovation improvements in different markets to inform merger policy. For recently updated apps and apps where users interact, the curve is flatter.

The results support the hypothesis that product differentiation, in terms of app dissimilarity, is lower in more concentrated markets. We only observe these results for markets with high concentration, and the opposite for markets with low to moderate concentration. Our results suggest that the incentive for apps to imitate the market leader hoping to be bought out is the dominant effect only in already concentrated markets. This result may be useful for considering cases such as FTC v. Meta, as it shows that product differentiation can decrease after mergers in already concentrated markets.

We used the BERT language model to estimate similarity of app descriptions, but it is not the only language model that may be relevant for the task. Sentence-BERT (SBERT) should be explored as a comparison to the BERT base model that we used. SBERT is a modification of BERT specifically designed for large-scale semantic similarity comparison. SBERT is BERT but fine-tuned on two datasets totaling 1M sentence pairs with labels of contradiction, entailment, or neutral (Reimers and Gurevych, 2019a). An example of SBERT in the economic literature is Carbonero et al. (2023a) matching work activities, from which measures of AI impact exist, to skills obtained from survey results in order to assess AI impacts on developing economies. There may also be more specialized models that perform better in measuring product differentiation. Al-Subaihin et al. (2019) explore many clustering techniques to categorize apps in the Google PlayStore, some of them custom. Improvements in language models provide numerous avenues in economic research for improving tasks such as crosswalking, using text as data, clustering text data, or replication of papers using early language modes, such as using BERT instead of bag-of-words for Hoberg and Phillips (2010).

2 Revisiting the Uber effect

Trevor McLemore (University of Nevada, Reno)

2.1 Abstract

Berger et al. (2018) examined whether Uber drivers displace conventional taxi drivers to answer the question of whether "gig work" displaces traditional work in a wide range of jobs. This paper uses the same methods and data set with seven additional years of data, through 2022, and additional Uber launches for more exogenous variation to better understand the effects found with data through 2015 when the rideshare industry was nascent. In my difference in differences and triple difference regressions, there was less evidence than the original paper showed that the diffusion of Uber reduced the earning potential of incumbent taxi drivers in the United States. Running regressions with the extended data range showed that Uber entry decreased the labor supply of incumbent taxi drivers, which was not a significant effect with the original data range. My inability to match the original authors' results using the same source dataset while limiting the data to their same time window, 2009-2015, suggests a lack of robustness in the original paper.

2.2 Introduction

Most people in the United States now have interacted with rideshare platforms, but relatively little is known about how they impact the taxi industry, or even a precise number of drivers. Census and tax data do not yet show significantly increasing numbers of gig workers outside of rideshare, but the number of rideshare gig workers has increased. The number of active Uber drivers in 2015, defined as providing at least four rides in that month, grew exponentially from less than 200,000 in January to 464,681 in December (Hall and Krueger, 2018). In Uber's Q2 2022 earnings, the CEO claimed that Uber had more than a million drivers in the US. From studying the rideshare industry, we can learn what may happen if other industries are similarly impacted in the future and be better prepared to understand the measurement of key metrics, taxes, and labor in those industries.

Using census data and press releases for Uber launches through 2015, Berger et al. (2018) showed a 10% decrease in the income of incumbent taxi drivers and chauffeurs in the metro areas where Uber was launched. The rideshare industry was still nascent in 2015. There was more than three times the number of Uber drivers by 2022 (based on estimates from Hall and Krueger (2018) and Uber's Q2 2022 earnings report). In this paper, I return to the methodologies of Berger et al. (2018) to update the data through 2022 so I can show if the impacts were lasting, if the increased intensity of market penetration strengthens the effects, and what impacts COVID had on these income dynamics. Uber remains a good proxy for the rideshare industry as it was a first mover in many metro areas, retains the highest market share, and many Uber drivers drive for multiple platforms.

There are potential negative labor impacts from the gig or contingent worker dynamic. From Muehlberger (2007), "If the self-employed person works only (or mainly) for one contractor in (partial) subordination, part of the entrepreneurial risk is transferred to the worker, while entrepreneurial possibilities are restricted" further finding that the dependent self-employed have weaker workplace rights under labor law, fewer social security benefits, and are often beyond trade union representation. The counterpoints that the rideshare dynamic offers flexibility, better earnings, and income smoothing are put forth by Hall and Krueger (2018). This paper offers insight into how these labor dynamics play out.

2.3 Literature

It is very difficult to give concrete information on contingent and gig work due to challenges surrounding the different sources of data that are needed to determine basic questions such as whether and where the contingent work force is growing and how that affects wages, market power, quality of life, and tax collection. Three approaches were commonly used in the research on rideshare and contingent workers. (1) Survey data, (2) analysis of tax filings, and (3) people affiliated with Uber publishing with proprietary data. The Current Population Survey (CPS), the American Community Survey (ACS), and other household surveys show contingent work has drifted downward since the mid-1990s, but tax filings provide strong evidence that non-employee work arrangements are increasing (Abraham et al., 2017). Although household surveys show a slight decrease in contingent work for the labor market as a whole, platform-based gig work is concentrated in certain occupations. Based on an analysis of 1099 returns from 2000 to 2016, platform-based driving work expanded dramatically without increasing the prevalence of other types of freelance work (Collins et al., 2019).

The Social Security Administration maintains a Master Earning File (MEF) database with all filled out W-2s. The Census Bureau receives the encrypted W-2s for each CPS respondent in an extract called the Detailed Earnings Report (DER). Abraham et al. (2017) states that for the 17 years they have the CPS and DER data, 65.4% of those with DER self-employment income do not report it on the CPS, and 51.1% of those with CPS self-employment income do not report it on the DER. As the gap between steady contingent workers in the CPS and growing contingent workers in the DER widens (Abraham et al., 2021), there is a temptation to assume that the tax data are more accurate. Garin et al. (2022) however cautions against deferring too heavily towards the tax data, finding that as much as 59% of the growth in self-employment rates can be attributed to changes in reporting behavior independent of changes in the nature of work driven by negative marginal tax rates in the EITC range. In 2017 the 1099-K reporting requirements substantially increased to 20,000 dollars, making it much more difficult to measure contingent work using tax data (Garin et al., 2023).

Hall and Krueger (2018) wrote an analysis of the labor market for Uber's driver partners in the United States that had findings such as that drivers are drawn to the platform due to flexibility and higher compensation that do not vary as much. The data is proprietary, and the paper acknowledges that Hall is affiliated with Uber and that Krueger worked as a consultant for Uber twice during the publication of the paper. The paper is clouded by an excoriating response from (Berg and Johnston, 2019, p. 64) that critiques its methodology, selective incompleteness, parroting Uber's corporate narrative, and incomplete labor market analysis ending with a quote, "there might emerge a privileged set of academic researchers presiding over private data from which they produce papers that cannot be critiqued or replicated…".

The 2009-2015 ACS is used by Berger et al. (2018) to analyze the effect of Uber on the income of taxi drivers and chauffeurs. They used press releases to identify the date and location of launches and the Google Trend search intensity to proxy Uber's penetration. This paper expands and updates this line of research which is important, as the number of active rideshare drivers has over tripled since their paper. The ACS intake is evenly distributed monthly and Uber's growth in this time period was exponential, so Berger et al. (2018) were working with an Uber driver impact on the ACS data below the mean of 200,000 and 464,681 drivers (Hall and Krueger, 2018) that had provided four rides in that month whether full or part time. There were at least 5 million individuals receiving information returns for platform gig work by 2021, nearly all from transportation platforms (Garin et al., 2023). Most Uber drivers work part-time which can be deduced from 68% and 66% of Uber drivers having other jobs in 2014 and 2015 respectively (Hall and Krueger, 2018). The tax return estimate likely includes most of these part-time drivers.

2.4 Data

This paper constructs its dataset from multiple sources:

1. 2009-2021 ACS individual-level annual data that is a 1-in-100 representative sample of the US population.

2. Berger et al. (2018) data on Uber launches in MSAs and the paper's Google Trends data used to measure Uber's market penetration. Both datasets are updated using press releases from entries into new markets and associated Google Trends data. Multiple news articles were used to verify that Buffalo, NY and Rochester, NY launched Uber in June 2017.

I use the personal consumption expenditure price index provided by the Bureau of Economic Analysis to adjust wage and earnings to 2022 values much like (Hall and Krueger, 2018) used the same index to adjust wage and earnings to 2015 values. I analyze the same 50 metro areas that were analyzed in the original paper based on population. However, I note that Salt Lake City, UT was used by the authors instead of Tucson, AZ or Honolulu, HI despite both metro areas being more populous from 2009 to 2015. By the end of 2017, Uber had launched in all of the 50 most populous MSAs.

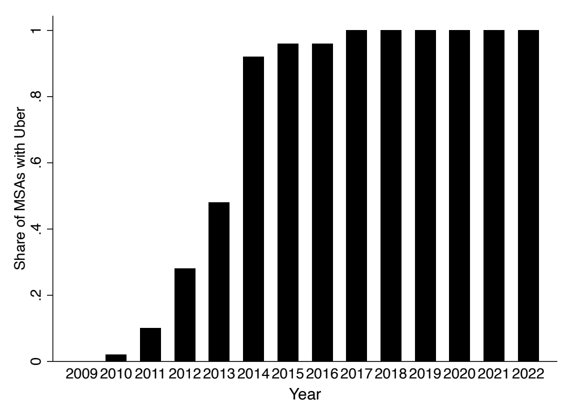


Figure 8: Share of MSAs with Uber

The trade-off for adding more MSAs is that the sample size gets small as metro areas are broken down into year groups, occupations, and wage-employed workers to the extent that Oklahoma City, OK, one of the less populous of the top 50 most populous MSAs, has 0 wage-employed taxi drivers, defined as receiving a salary from a for-profit company, in 2015 in the ACS data set. Berger et al. (2018) had data from 2015 for wage-employed taxi drivers and claimed that running the regressions for the full set of MSAs did not substantially change the results.

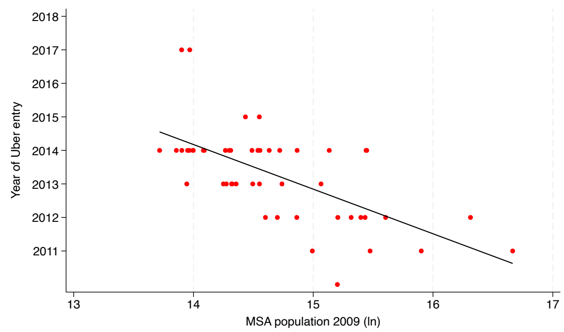


Figure 9: MSA population 2009 (ln)

The ACS data do not capture secondary employment or supplementary income, at least in a way that I could infer the secondary income came from gig work. Most Uber drivers continue full- or part-time employment after joining Uber (Hall and Krueger, 2018), so by sorting by occupation, I am capturing taxi drivers and chauffeurs. As a robustness check, I run additional regressions against the subset consisting of wage-employed drivers who are defined as an employee of a private for-profit company or business, or of an individual, for wages, salary, or commissions. I am determining the effect of ridesharing on the incumbent industry, so these robustness checks help me rule out that there are a significant number of Uber drivers remaining in the ACS.

The Berger et al. (2018) data on Uber launches in MSAs come from press releases verified across multiple news articles, so they are likely accurate. The Google Trends data were highly correlated with the number of Uber drivers per capita in an MSA (Angrist et al., 2017), but this correlation would not hold as strong as during the original publication, since general familiarity with Uber distorts the search queries centered around launching in a new market and thus will not be analyzed in this paper.

Uber does not introduce itself into markets randomly. Berger et al. (2018) points out that if Uber specifically targets locations that experienced differential trends in earnings or employment in the taxi sector, we have an issue with non-random selection. I repeat their regression with the year of Uber's MSA entry on MSA characteristics with an updated column 3 as I have more data to average for pre-Uber entry. They found that when a wide enough range of characteristics is used, the only remaining significant correlation is the MSA population, which accounts for roughly 40% of the variation. They also showed that the earnings in the metro area for all jobs were significant at the 10% level (I found significance at the 5% level), but the earnings of taxi drivers specifically showed no significance. I show the same results with respect to population and also show significance at the 5% level for Uber entering metro areas earlier when a greater share of the population is under 40 years of age. Despite slight differences, my regression reinforces the conclusion of Berger et al. (2018) that Uber's entry into MSAs based on earnings and employment in the taxi sector is plausibly random.

Table 6: What determined Uber's entry?

	Outcome: year of Uber entry		
	(1)	(2)	(3)
Share in taxi services	481.536	177.138	348.150
	(305.123)	(204.951)	(241.346)
Share female drivers	-1.681	-3.448**	1.824
	(2.224)	(1.616)	(3.404)
Share foreign-born drivers	-4.343***	-0.564	1.719
	(1.551)	(1.390)	(1.642)
Share low-educated drivers	0.276	1.648	7.352
	(1.840)	(1.306)	(4.522)
Share non-Hispanic white drivers	-0.139	-0.445	-0.325
	(1.118)	(1.034)	(1.032)
Share self-employed drivers	-0.146	0.773	1.244
	(1.266)	(0.971)	(2.275)
Mean earnings (ln)	-0.446	0.014	0.535
	(0.537)	(0.398)	(0.671)
Population (ln)		-1.188***	-0.960***
		(0.233)	(0.288)
Mean earnings (ln)		-3.226	-6.749**
		(2.683)	(2.692)
College share		-5.107	1.269
		(4.497)	(4.203)
Share aged < 40		-4.107	-8.583**
		(4.424)	(4.038)
Unemployment rate		4.764	-12.021
		(8.516)	(12.204)
Observations	50	50	50
R-squared	0.348	0.667	0.706

Notes: This table reports results from cross-sectional OLS regressions where I regress the year that Uber entered a MSA on a range of initial MSA characteristics measured in 2009 (columns 1 and 2) or averaged over the years prior to Uber's entry into a MSA (column 3). Robust standard errors reported in parentheses. Statistical significance is denoted by:*** p<0.01, ** p<0.05, * p<0.10.

2.5 Methods

This paper uses the methods of the Berger et al. (2018) paper with updated data through 2022.

Difference in differences of taxi earnings before and after Uber entry:

$$y_{it} = \alpha_i + \nu_t + \delta Uber_{it} + \gamma X_{it} + \xi_{it} \tag{7}$$

The regression compares changes in MSAs where Uber was introduced with areas where Uber was not introduced. y_{it} is the mean log earnings among taxi drivers in MSA i and year t. $Uber_{it}$ is a dummy variable that takes the value of 1 when Uber enters an MSA and for all subsequent years that Uber stays while taking the value of 0 for all preceding years. α_i controls for time invariant MSA fixed effects. ν_t controls for MSA time fixed effects. X_{it} controls for MSA and time variant covariates.

Berger et al. (2018) uses the log of Google Trends for the search term 'Uber' as a method to measure the intensity of adoption. The intensity of adoption is likely endogenous, but may provide further evidence of stronger trends in the MSAs where Uber is more heavily adopted. I have not compiled the Google Trends data for the expanded data set as the Google Trends data are affected by awareness of Uber, which is much stronger now than during the publication of the original paper.

Triple difference of taxi and other transportation occupation earnings before and after Uber entry:

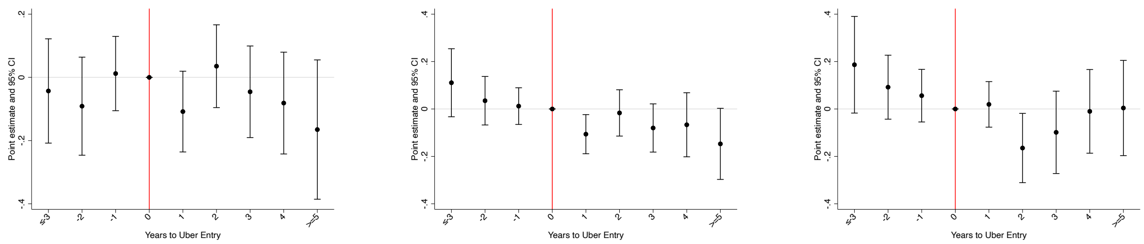
$$y_{it}^T - y_{it}^O = \alpha_i + \nu_t + \delta Uber_{it} + \gamma X_{it} + \xi_{it} \tag{8}$$

Although there is no evidence for it in this instance, changes in earnings may be correlated with unobserved time-varying factors that are also correlated with the introduction of Uber. Since unobservable time varying shocks are likely to affect other transportation services as well that are not competing with Uber, I check the results with a triple diff framework for other related driving occupations. y_{it}^T in equation 8 means taxi drivers and y_{it}^O means other transportation occupations.

2.6 Results

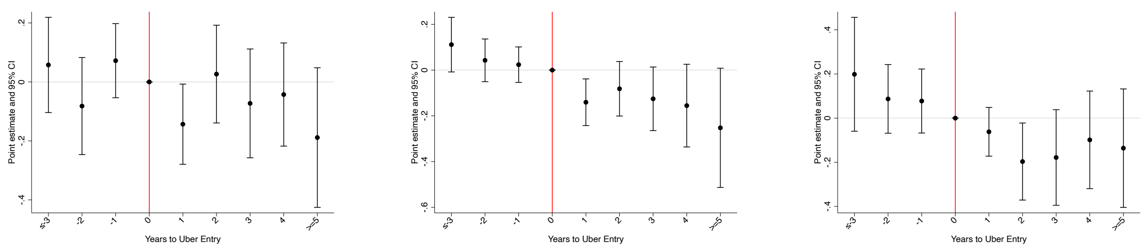
When constraining the data to 2009-2015, I did not get the same levels of significance as Berger et al. (2018). When observing wage and earnings effects for all drivers, I show magnitudes as small as one third the size of the original paper and only slightly lower magnitudes with slightly inflated standard errors when focusing on wage-employed drivers. In this results section, I will explain the results that I found for the 2009-2022 data.

2.6.1 Event study



(a) Mean earnings (b) Hrly earnings (c) Labor supply

Figure 10: All drivers



(a) Mean earnings (b) Hrly earnings (c) Labor supply

Figure 11: Wage-employed drivers

I employed an event study to determine whether treatment effects reflected preexisting differential trends instead of the effects of treatment. The event study I have conducted includes 4 additional years post treatment compared to Berger et al. (2018) and includes Buffalo, NY and Rochester, NY. I found no evidence of preexisting trends for the three years preceding Uber entry for earnings and a likely, but inconclusive, dip in earnings the year after treatment with a possible rebound in the second year before a slight downward trend over time. With hourly earnings, I see a possible, but inconclusive, preexisting downward trend with a dip the year after Uber entry and, similar to the total earnings, a slight rebound in the second year after entry followed by a slight downward trend that is inconclusive. For labor supply, I see a possible, but inconclusive, preexisting downward trend with a slight decline after Uber entry that looks like it steadies. The dip in labor supply after Uber entry is most significant after two years, while the dip in earnings and hourly wage are most significant after one year. I cannot completely rule out the possibility of preexisting trends for hourly earnings and labor supply, which may mean that these results are overstated.

2.6.2 Uber's impact on the earnings of taxi drivers

My results for the treatment effect on earnings after Uber entry were smaller in magnitude than the Berger et al. (2018) results with similar standard errors. I showed no significance for the effects of Uber entry on earnings (see Table 7) whereas the original paper showed a decrease in earnings between 11-18.5% with significance for all columns at either 5% or 1%.

Table 7: Earnings of taxi drivers after Uber's introduction, 2009–2022

	Outcome: mean ln earnings/wage income					
	Panel A. Mean earnings			Panel B. Mean wage income		
	(1)	(2)	(3)	(4)	(5)	(6)
	Sample: All drivers					
$Uber_{it}(=1)$	-0.037	-0.027	-0.053	-0.032	-0.019	-0.041
	(0.049)	(0.051)	(0.057)	(0.052)	(0.055)	(0.061)
Observations	700	700	700	700	700	700
	Sample: Wage-employed drivers					
$Uber_{it}(=1)$	-0.070	-0.060	-0.079	-0.039	-0.030	-0.051
	(0.057)	(0.062)	(0.067)	(0.058)	(0.064)	(0.069)
MSA and year FE?	Y	Y	Y	Y	Y	Y
Additional MSA controls?	N	Y	Y	N	Y	Y
MSA x linear time trend?	N	N	Y	N	N	Y
Observations	699	699	699	699	699	699

Notes: This table reports OLS estimates of Eq. (7) where the outcome is the mean ln earnings (panel A) and ln wage income (panel B) for all (upper panel) and wage-employed (lower panel) taxi drivers respectively. Additional MSA-level controls include mean ln earnings, the share of workers with a college degree, the female share of the labor force, ln MSA population, the unemployment rate, and the share of the labor force that falls in four age categories (16–25, 26–39, 40–54, and 55). Robust standard errors reported in parentheses. Statistical significance is denoted by:*** p<0.01, ** p<0.05, * p<0.10.

For Table 8, broken down into hourly earnings, my results show about a 6% drop in hourly earnings for the all driver panel at the 10% level for the most lean model that does not control for MSA-specific time trends or additional MSA controls consisting of mean ln earnings, the share of workers with a college degree, the female share of the labor force, ln MSA population, the unemployment rate, and the share of the labor force that falls in four age categories (16–25, 26–39, 40–54, and 55). For wage-employed drivers' hourly earnings, there is an 8.5-9.8% decrease at the 5-10% level. Berger et al. (2018) shows significant decreases in the 9.9-16.2% range in the all driver panel and significant decreases in the 11.8-12.7% range for wage-employed drivers mean hourly earnings. My findings suggest that the hourly earnings of taxi drivers decreased after Uber entry, but the magnitude and precision are not as high as in the original study.

Table 8: Hourly earnings of taxi drivers after Uber's introduction, 2009–2022

	Outcome: mean ln hourly earnings/wage income					
	Panel A. Mean hourly earnings			Panel B. Mean hourly wage income		
	(1)	(2)	(3)	(4)	(5)	(6)
	Sample: All drivers					
$Uber_{it}$(=1)	-0.058*	-0.050	-0.042	-0.060*	-0.054	-0.045
	(0.030)	(0.030)	(0.034)	(0.033)	(0.034)	(0.037)
Observations	700	700	700	700	700	700
	Sample: Wage-employed drivers					
$Uber_{it}$(=1)	-0.098**	-0.092**	-0.085*	-0.067	-0.061	-0.058
	(0.039)	(0.040)	(0.045)	(0.042)	(0.043)	(0.049)
MSA and year FE?	Y	Y	Y	Y	Y	Y
Additional MSA controls?	N	Y	Y	N	Y	Y
MSA x linear time trend?	N	N	Y	N	N	Y
Observations	699	699	699	699	699	699

Notes: This table reports OLS estimates of Eq. (7) where the outcome is the mean ln hourly earnings (panel A) and ln hourly wage income (panel B) for all (upper panel) and wage-employed (lower panel) taxi drivers, respectively. Additional MSA-level controls include mean ln earnings, the share of workers with a college degree, the female share of the labor force, ln MSA population, the unemployment rate, and the share of the labor force that falls in four age categories (16–25, 26–39, 40–54, and 55). Robust standard errors reported in parentheses. Statistical significance is denoted by:*** p<0.01, ** p<0.05, * p<0.10.

2.6.3 Uber's impact on the labor supply of taxi drivers

Table 9: Labor supply of taxi drivers after Uber's introduction, 2009–2022

	Outcome: ln labor supply					
	Panel A. All drivers			Panel B. Wage-employed drivers		
	(1)	(2)	(3)	(4)	(5)	(6)
$Uber_{it}$(=1)	-0.064	-0.067	-0.030	-0.123**	-0.119**	-0.077*
	(0.051)	(0.048)	(0.043)	(0.053)	(0.053)	(0.046)
MSA and year FE?	Y	Y	Y	Y	Y	Y
Additional MSA controls?	N	Y	Y	N	Y	Y
MSA x linear time trend?	N	N	Y	N	N	Y
Observations	700	700	700	699	699	699
R-squared	0.446	0.500	0.632	0.265	0.296	0.446

Notes: This table reports OLS estimates of Eq. (7) where the outcome is the ln labor supply of all (panel A) and wage-employed (panel B) taxi drivers respectively. Additional MSA-level controls include mean ln earnings, the share of workers with a college degree, the female share of the labor force, ln MSA population, the unemployment rate, and the share of the labor force that falls in four age categories (16–25, 26–39, 40–54, and 55). Robust standard errors reported in parentheses. Statistical significance is denoted by:*** p<0.01, ** p<0.05, * p<0.10.

For wage-employed drivers, I saw a 7.7-12.3% decrease in labor supply (see Table 9) at the 5-10% level of significance. Despite having the same signs and similar magnitudes, Berger et al. (2018) did not show significance for Uber entry on effecting taxi driver labor supply. Since I have data for more years after Uber entry, it could be a possible explanation that the labor supply of taxi drivers adjusts with a longer lag, but I need to do more research to make this claim as I do not get the same results as in the original paper when I restrict my data set to 2009-2015.

2.6.4 Uber's impact on the compositional changes among taxi drivers

I look at compositional changes in the taxi-driver labor pool to examine if more productive drivers are leaving to partner with Uber and thus describing some of the suggested decreases in earnings. I find that the share of foreign-born taxi drivers increased by 3.5% at the 10% level, but based on the small magnitude and imprecision, it is likely there is no significant change in the composition of the labor force that is describing the effect of Uber entry on taxi earnings.

Table 10: Compositional changes among taxi drivers after Uber's introduction, 2009–2022

	Aged < 40 (1)	Female (2)	Foreign-born (3)	No college (4)	Non-Hispanic white (5)	Self-employed (6)
$Uber_{it}(=1)$	-0.013 (0.019)	-0.014 (0.016)	0.035* (0.017)	-0.020 (0.014)	-0.010 (0.017)	0.026 (0.018)
MSA and year FE?	Y	Y	Y	Y	Y	Y
Additional MSA controls?	Y	Y	Y	Y	Y	Y
MSA x linear time trend?	Y	Y	Y	Y	Y	Y
Observations	700	700	700	700	700	700
R-squared	0.236	0.181	0.230	0.314	0.251	0.473

Notes: This table reports OLS estimates of Eq. (7) where the outcome is the share of taxi drivers that is aged below 40, female, foreign-born, not college educated, non-Hispanic white, and self-employed respectively. Additional MSA-level controls include mean ln earnings, the share of workers with a college degree, the female share of the labor force, ln MSA population, the unemployment rate, and the share of the labor force that falls in four age categories (16–25, 26–39, 40–54, and 55). Robust standard errors reported in parentheses. Statistical significance is denoted by:*** p<0.01, ** p<0.05, * p<0.10.

2.6.5 Triple difference and placebo tests

If my results reflect the effects of Uber entry rather than an omitted factor, I should not see earning in similar driving occupations that are unaffected by Uber services timed with Uber entry. In Table 11, I use Eq. (8) and subtract from the earnings of taxi drivers the earnings of the occupation codes, in turn, for bus drivers, locomotive operators, motor vehicle operators, tractor operators, and truck drivers. My results are of small magnitude with signs that flip back and forth and large standard errors. Since I did not originally find significant earnings, this robustness check was perhaps unnecessary. Berger et al. (2018) showed similar magnitudes and signs to their original findings for the decreased earnings of taxi drivers treated on Uber entry with decreased precision, but still showing significance at the 5% level for bus drivers and taxi drivers.

Table 11: Earnings of taxi drivers after Uber's introduction, 2009–2022: Triple-differences estimates

	Outcome: mean ln (hourly) earnings for taxi drivers – mean ln (hourly) earnings for occupation listed below				
	Bus drivers	Locomotive operators	Motor vehicle operators	Tractor operators	Truck drivers
	(1)	(2)	(3)	(4)	(5)
	Panel A. Mean earnings				
$Uber_{it}$(=1)	0.004	-0.088	0.100	-0.044	-0.042
	(0.067)	(0.100)	(0.173)	(0.076)	(0.057)
Observations	700	571	682	700	700
	Panel B. Mean hourly earnings				
$Uber_{it}$(=1)	-0.032	-0.119	0.040	-0.036	-0.041
	(0.042)	(0.091)	(0.104)	(0.043)	(0.034)
MSA and year FE?	Y	Y	Y	Y	Y
Additional MSA controls?	Y	Y	Y	Y	Y
MSA x linear time trend?	Y	Y	Y	Y	Y
Observations	700	571	682	700	700

Notes: This table reports OLS estimates of Eq. (8) where the outcome is the difference in the mean ln earnings (panel A) or ln hourly earnings (panel B) of taxi drivers and (1) "Bus and Ambulance Drivers and Attendants"; (2) "Locomotive Engineers and Operators"; (3) "Motor Vehicle Operators, All Other"; (4) "Industrial Truck and Tractor Operators"; and (5) "Driver/Sales Workers and Truck Drivers" respectively. Additional MSA-level controls include mean ln earnings, the share of workers with a college degree, the female share of the labor force, ln MSA population, the unemployment rate, and the share of the labor force that falls in four age categories (16–25, 26–39, 40–54, and 55). Robust standard errors reported in parentheses. Statistical significance is denoted by:*** p<0.01, ** p<0.05, * p<0.10.

Table 12: Placebo tests: earnings in other transport occupations after Uber's introduction, 2009–2022

	Outcome: mean ln (hourly) earnings					
	Taxi Drivers	Bus Drivers	Locomotive Operators	Motor Vehicle Operators	Tractor Operators	Truck Drivers
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A. Mean earnings					
$Uber_{it}$(=1)	-0.053	-0.057**	0.011	-0.149	-0.009	-0.011
	(0.057)	(0.025)	(0.076)	(0.144)	(0.046)	(0.019)
	Panel B. Mean hourly earnings					
$Uber_{it}$(=1)	-0.042	-0.010	0.073	-0.083	-0.006	-0.001
	(0.034)	(0.025)	(0.075)	(0.095)	(0.025)	(0.013)
MSA and year FE?	Y	Y	Y	Y	Y	Y
Additional MSA controls?	Y	Y	Y	Y	Y	Y
MSA x linear time trend?	Y	Y	Y	Y	Y	Y
Observations	700	700	571	682	700	700

Notes: This table reports OLS estimates of Eq. (7) where the outcome is the mean ln earnings (panel A) and ln hourly earnings (panel B) of taxi drivers and (2) "Bus and Ambulance Drivers and Attendants"; (3) "Locomotive Engineers and Operators"; (4) "Motor Vehicle Operators, All Other"; (5) "Industrial Truck and Tractor Operators"; and (6) "Driver/Sales Workers and Truck Drivers" respectively. Additional MSA-level controls include mean ln earnings, the share of workers with a college degree, the female share of the labor force, ln MSA population, the unemployment rate, and the share of the labor force that falls in four age categories (16–25, 26–39, 40–54, and 55). Robust standard errors reported in parentheses. Statistical significance is denoted by:*** p<0.01, ** p<0.05, * p<0.10.

Table 12 is a placebo test using similar logic to Table 11 in which I use regressions to find the treatment effects on the earnings of similar occupations. I still do not find significance for Uber entry on the earnings of taxi drivers, but do find a slight decrease (5.7%) in earnings of bus drivers significant at the 5% level. My results make it even less likely that Uber entry affects earnings. Berger et al. (2018) show significance of the decreased earnings for taxi drivers only.

2.7 Conclusion

This paper provides evidence that the Uber effect of a 10% decrease in taxi driver pay after Uber entry into an MSA, as reported by Berger et al. (2018), is not as large or robust. The effects had smaller magnitudes in almost all of the effects on earnings, but with similar standard errors when compared to the previous paper making most of the results insignificant. The impacts of Uber entry on labor supply were significant when using the extended data range, which was not observed in the more limited data range in the original paper. My results show many of the original findings right under the cusp of significance, even when limiting the data to match the original paper, 2009-2015. Two irregularities in the original paper's data stuck out. They used Salt Lake City, UT instead of Tucson, AZ or Honolulu, HI despite the higher populations in the latter two cities during their experiment. They had data for multiple wage employed taxi drivers in Oklahoma City, OK despite none existing in the American Community Survey. I conclude that the original paper's results were not robust.

3 Artificial intelligence and entrepreneurship

Frank M. Fossen (University of Nevada, Reno)

Trevor McLemore (University of Nevada, Reno)

Alina Sorgner (John Cabot University, Rome, Italy)

3.1 Abstract

Advances in artificial intelligence (AI) have brought the world to the threshold of significant new technological breakthroughs. These developments bring new opportunities and challenges to existing and potential entrepreneurs, raising pressing and promising research questions. We review emerging but fast-growing literature on impacts of AI on entrepreneurship, providing a resource for researchers in entrepreneurship and neighboring disciplines. We begin with a review of definitions of AI and show that ambiguity and broadness of definitions adopted in empirical studies may result in obscured evidence on impacts of AI on entrepreneurship. Against this background, we present and discuss existing theory and evidence on how AI technologies affect entrepreneurial opportunities and decision-making under uncertainty, the adoption of AI technologies by startups, entry barriers, and the performance of entrepreneurial businesses. We add an original empirical analysis of survey data from the German Socio-economic Panel revealing that entrepreneurs, particularly those with employees, are aware of and use AI technologies significantly more frequently than paid employees. Next, we discuss how AI may affect entrepreneurship indirectly through impacting local and sectoral labor markets. The reviewed evidence suggests that AI technologies that are designed to automate jobs are likely to result in a higher level of necessity entrepreneurship in a region, whereas AI technologies that transform jobs without necessarily displacing human workers increase the level of opportunity entrepreneurship. More generally, AI impacts regional entrepreneurship ecosystems (EE) in multiple ways by altering the importance of existing EE elements and processes, creating new ones, and potentially reducing the role of geography for entrepreneurship. Lastly, we address the question of how regulation of AI may affect the entrepreneurship landscape by focusing on the case of the European Union that has pioneered data protection and AI legislation. We conclude our survey by discussing implications for entrepreneurship research and policy.

3.2 Introduction

Artificial Intelligence (AI) transforms, destroys and creates human occupations and brings new opportunities and challenges to existing and potential entrepreneurs. AI may create new business opportunities for entrepreneurs (Shepherd and Majchrzak, 2022), but also push individuals into self-employment whose prior wage jobs are automated through the implementation of AI. AI startups such as DeepL are developing AI products that seem to be able to compete against tech giants like Amazon or Google (Weber et al., 2022). At the same time, other self-employed individuals and independent contractors feel threatened, such as writers in Hollywood, whose 2023 strike was partially motivated by concerns that studios may employ AI to create movie and television scripts.

In a Delphi study, Van Gelderen et al. (2021) asked 175 editors and Editorial Review Board members of the academic journals *Entrepreneurship Theory and Practice* (ETP) and *Journal of Business Venturing* what they think entrepreneurship will look like in 2030. AI is among the themes mentioned most often by the experts[2]. The authors conclude that one of the main future research questions in the field of entrepreneurship concerns how AI may affect the types of entrepreneurial ventures emerging and their performance. In particular, they raise the questions of how, why, and under what conditions AI might complement and support entrepreneurial activity or substitute for entrepreneurial judgement, perhaps even replacing the entrepreneur in the future[3]. While the literature on the AI-entrepreneurship nexus started before 2020, the global Covid-19 pandemic crisis substantially contributed to the spread of digital technologies and also spurred research on the impacts of AI on startups (Sorgner, 2023).

In this review, we discuss how extant literature in multiple disciplines including economics and management describes direct and indirect impacts of AI on entrepreneurship. We start our survey with a discussion of common definitions of AI and argue that wording matters for designing empirical research as well as government regulation of AI.

In terms of direct impacts, AI technologies have been identified as external enablers and facilitators of entrepreneurship (Davidsson et al., 2020; Obschonka and Audretsch, 2020; Chalmers et al., 2021; Davidsson and Sufyan, 2023). Innovative entrepreneurs discover and create new business opportuni-

[2]We note that AI is not a new phenomenon. A search for the terms "AI" and "artificial intelligence" using the Google Ngram Viewer tool and the corpus of English publications reveals that there was a significant increase in the interest in AI already during the 1960s, with a further surge in the 1980s culminating at around 1986. After that year, a strong decline in the trend followed that stabilized during the 1990s. A recent and still ongoing trend of an exponentially growing interest in AI started roughly after the year 2014.

[3]A call for paper proposals for a special issue of ETP on transformative AI and entrepreneurship is open until January 31, 2026 (https://www.entrepreneurship-ex-machina.org/).

ties using AI, and AI may reduce costs such as labor costs (through automation) and financing costs (through fintech services). Due to their prediction abilities, AI systems may help to resolve challenges of uncertainty and thereby create new possibilities of entrepreneurial action (Townsend and Hunt, 2019), but AI may also have fundamental limitations in dealing with uncertainty (Townsend et al., 2023). Recent literature reports that AI deployment has become increasingly important in the development of digital entrepreneurship, in identifying and acquiring knowledge, in customizing products and services as a competitive entrepreneurial strategy, and in managing product innovation (Mariani et al., 2023). Generative AI such as ChatGPT can support creative tasks such as pitching entrepreneurial business ideas to investors or generating business ideas (Short and Short, 2023; Boussioux et al., 2024). An important emerging theme is that AI plays a key role in (digital) entrepreneurial ecosystems (Acs et al., 2022; Wurth et al., 2023) by facilitating information sharing, creating and diffusing new products, and fostering innovation.

Due to the novelty of the phenomenon, most papers on the AI-entrepreneurship nexus have been conceptual, while few papers present evidence from surveys (e.g., Bessen et al., 2022; McElheran et al., 2024) or experiments (e.g., Otis et al., 2024). In this review, we supplement this literature by providing new survey evidence on how the self-employed use AI and are exposed to AI based on the German Socio-economic Panel (Fedorets et al., 2022).

Advances in AI may also have indirect impacts on entrepreneurship when AI affects the labor market more broadly. Some individuals are pushed into self-employment due to a lack of alternatives available to them. The number of such necessity entrepreneurs may increase if jobs are automated and workers are displaced by AI. We discuss various empirical measures of the impact of AI on occupations and continue by reviewing evidence of consequences of AI for the labor market and entry into different types of entrepreneurship (Fossen and Sorgner, 2021, 2022). As AI affects occupations, the impacts vary across regions depending on regional occupational structures, which creates challenges and opportunities for regional policymaking (Fossen et al., 2022). The rapid advances made in AI technologies, which have raised many concerns about the future of work, have led many countries and regions, most notably the European Union, to develop and implement sometimes very strict regulations of AI. We discuss likely impacts on the level and nature of AI entrepreneurship.

We conclude our review by deriving central questions for future entrepreneurship research. Do emerging government regulations of AI support entrepreneurs or keep them from achieving their goals? What are the right conditions – in terms of an entrepreneurial ecosystem – that allow entrepreneurs to put AI to beneficial use and to avoid potential harm to society (Baumol, 1996)? With our review we aim to take stock of what scholars have learned about the impact of AI on entrepreneurship, discuss potential policy implications, identify knowledge gaps, and provide avenues for further research in this rapidly growing research area.

The scope of this review is defined by the various impacts AI may have on entrepreneurship. Obschonka and Fisch (2022) identify two main areas of intersection between AI and entrepreneurship research: AI in entrepreneurship as a research topic and AI as a research method. We delve deeply into the first area, but do not cover potentials of AI to be used as new methods in entrepreneurship research. We refer the readers to Lévesque et al. (2022), who investigate this topic (see also Schwab and Zhang, 2019; Lamine et al., 2023). Implications of AI for entrepreneurship education are also outside the scope of this survey. Obschonka and Audretsch (2020) identify this as an important research area, and Winkler et al. (2023) offer significant inroads. Finally, we do not include AI in entrepreneurial finance in this review; Ferrati and Muffatto (2021) provide an extensive survey of emerging machine learning approaches in entrepreneurial finance (see also Bertoni et al., 2022).

This review is structured as follows. Section 3.3 reviews existing and the most widely used definitions of AI and provides an overview of various types of AI technologies and their applications. Section 3.4 reviews literature on the AI-entrepreneurship nexus that deals with the concepts of entrepreneurial uncertainty, opportunity, decision-making, and performance. Section 3.5 presents initial empirical evidence on AI adoption and usage by entrepreneurial businesses and complements it with our own novel analysis of survey data from the German Socio-economic Panel. Section 3.6 reviews evidence on indirect effects of AI on entrepreneurship through local labor markets. Section 3.7 discusses impacts of AI on entrepreneurship ecosystems. Section 3.8 presents current regulations of AI, with a focus on the EU, and discusses their potential impacts on the entrepreneurship landscape. Section 3.9 highlights insights from our review and discusses implications for entrepreneurship research and policy. Finally, Section 3.10 concludes the survey.

3.3 What can AI do?

Advances in AI have brought the world to the threshold of significant new technological breakthroughs. The number of AI patents granted globally has grown exponentially (Figure 12) according to Maslej et al. (2024)[4]. The number of newspaper articles published on AI exploded already before the release of OpenAI's popular chatbot ChatGPT (Alexopoulos and Cohen, 2018). This and other advances in Generative AI have further accelerated AI adoption: ChatGPT had more than 100 million users within weeks of its release (according to Bonney et al., 2024, citing the New York Times). Before we delve into potential impacts on entrepreneurship, this section lays out the background by discussing how AI can be defined, how AI works, and important applications of AI.

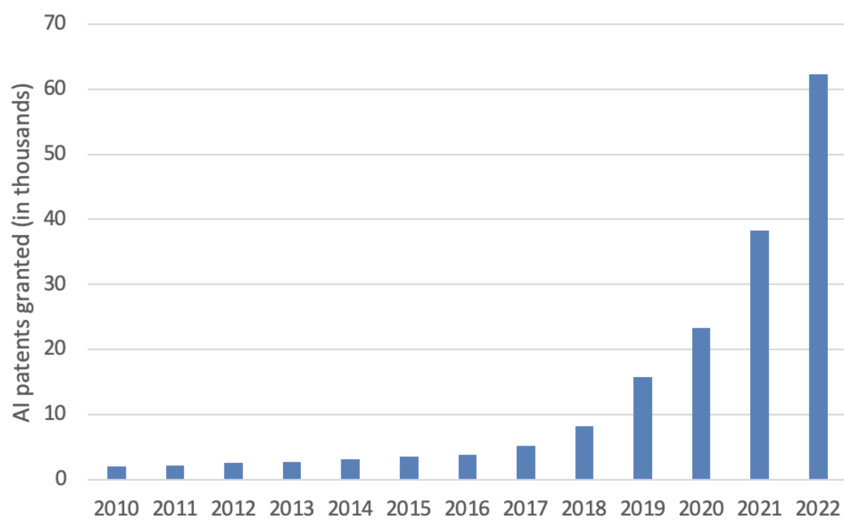


Figure 12: Global number of granted AI patents
Source: Adapted from Maslej et al. (2024), using data from the Center for Security and Emerging Technology.

3.3.1 Definitions of AI

AI is not easy to define, and this review is not going to provide a conclusive definition of AI. As stated by the European Commission Joint Resource Centre, "despite the increased interest in AI by the academia, industry and public institutions, there is no standard definition of what AI actually involves" (Samoili et al., 2021, 9). Most of the papers referenced in this review provide their own definition of AI, so we will mention a broad reputable definition, summarize different ways the definitions narrow, especially in the papers and documents referenced, and then highlight and break down specific AI subsystems that are commonly agreed upon in the definitions as key components in powering the technologies that have brought about the explosion in interest over the past decade.

According to the widely used definition provided by the Organisation for Economic Co-operation

[4]These patents do not necessarily translate to innovation, as many patents remain unused to strategically block potential competitors or left sleeping for other reasons (Torrisi et al., 2016).

and Development (OECD), an AI system is defined as a "machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy." (OECD, 2019, 7)[5]. In 2024, the OECD published a memorandum on an updated definition of an AI system, in which it is defined as "a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment." (OECD, 2024, 4). The OECD 2019 and 2024 definitions are rather broad, which the OECD 2024 memorandum acknowledges explicitly: "the updated definition of AI is inclusive and encompasses systems ranging from simple to complex. [...] When applied in practice, additional criteria may be needed to narrow or otherwise tailor the definition when used in a specific context" (OECD, 2024, 9).

The outputs of an AI system from the OECD 2019 definition, which are predictions, recommendations, or decisions, are different levels of automation. Prediction as an output could be, but does not need to be, a forecast such as the likelihood of default on a business loan. Prediction could also be whether an image contains a stop sign, or the next word a chatbot provides answering a query. A recommendation is the next level of automation that may involve a human decision maker like recommending a loan approval to a bank for a borrower with a low risk of default, but a recommendation does not need a human decision maker, for example in the case of recommending a movie to a viewer based on previous movies watched on a streaming service. A decision is the next level of automation and could involve the AI denying a loan with a high risk of default, or a self-driving car deciding to send a signal to an actuator to apply the brakes when the system's computer vision recognizes a stop sign at a certain distance with a high probability. If the machine's output is simply a prediction, and not a recommendation, decision, or action, it may not be an automat. Some definitions of AI specify a requisite amount of automation (High-level Expert Group on Artificial Intelligence, 2019; Boucher, 2020). Content was added as an output in the 2024 OECD definition to account for generative AI systems producing content such as text, image, audio, and video. Content as an output does not fall on the continuum of automation like the other outputs,

[5]The International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) have a substantially similar definition: "An AI system is an engineered system that generates outputs such as content, forecasts, recommendations or decisions for a given set of human-defined objectives. The engineered system can use various techniques and approaches related to artificial intelligence to develop a model to represent data, knowledge, processes, etc. which can be used to conduct tasks. AI systems are designed to operate with varying levels of automation." (International Organization for Standardization, 2022).

and many content creation algorithms are predictors of the next word, pixel, frequency, amplitude, or wave form. Thus, content can intersect with the other three outputs depending on the level of automation used in the AI system.

Continuous learning can be specified as a requisite for AI (High-level Expert Group on Artificial Intelligence, 2019; Samoili et al., 2021). Machine learning models can be trained and released without feedback mechanisms. With this definition constraint, models without feedback mechanisms would not be considered AI.

Performance measures are sometimes added to definitions such as normally requiring or exceeding human intelligence, reasoning, or predictive power under changing circumstances (Obschonka and Audretsch, 2020; Giuggioli and Pellegrini, 2023). Specifying performance differentiates more powerful AI systems from legacy technology included in the broad definitions, but as more systems meet the specified performance criteria, it may not narrow the definition much in practice going forward. Another type of AI definition uses novelty to narrow the definition in a way that has a similar effect as performance but will not necessarily expand what is considered an AI system over time. Consider, for example, (Tangredi and Galdorisi, 2021, 151-152): "Arguably, the term AI describes a human emotional response to new automation and is not a description of how that automation works [. . .]. What is and is not considered AI is always evolving and should be principally judged by whether a machine is doing something that, until recently, could only be done by human intelligence or couldn't be done at all". Novelty narrows the definition of AI in a way that updates over time, but also makes the definition vaguer and less stable as it is unclear how to determine how long, or until which event, a machine that imitated or outperformed humans in a new way should be considered AI.

None of the papers referenced have definitions explicitly requiring sensors and actuators because they explicitly or implicitly allow the possibility of an AI system interacting solely in a virtual environment. However, it is natural to think of sensors as part of an AI system. Definitions of AI usually acknowledge that AI systems use data to make their predictions. Humans receive data through our sensors (sight, hearing, smell, taste, touch). Machines, especially software in a virtual environment, can receive data in different ways. Data can be provided, input, scraped, etc., and sensors are but one method of getting data into an AI system. In a virtual system such as a Large Language Model, the model is initially trained on a large set of scraped or existing data and gets new data for continual learning from user or developer input and feedback without any sensors necessarily being involved. If not explicitly allowing for the possibility of virtual AI systems, it is easy to see language such as "interpret external data" and "achieve specific outcomes" (Shepherd and Majchrzak, 2022, 2) as suggestive of sensors and actuators being required. Other definitions exist that can appear even more suggestive that sensors and actuators are part of AI systems (High-level Expert Group on Artificial Intelligence, 2019; Boucher, 2020) with language like "capable of observing its environment [. . .] taking intelligent action" (Samoili et al., 2021, 9). Sensors and actuators can be part of AI systems (self-driving cars, robotics, Internet of Things), but there are alternative ways to access data and provide non-physical outputs, so we prefer definitions that do not make them a required component of AI.

Using human intelligence to define artificial intelligence is common (Townsend and Hunt, 2019; Obschonka and Audretsch, 2020; Chalmers et al., 2021; Shepherd and Majchrzak, 2022; Giuggioli and Pellegrini, 2023). Even if an AI definition does not directly refer to human intelligence, terms such as "intelligence", "behavior", "reason", "rational", "learn", "perceive", "interpret", "cognitive", "creativity", and "knowledge" appear frequently. Anthropomorphizing AI is a quick way to roughly associate the technologies that make up AI systems, but a problem of anthropomorphizing AI is that even in elements of AI that appear similar to human features, the actual functioning can be very different. The (International Organization for Standardization, 2022, 40-41) describes in their AI system functional overview: "AI systems do not under-stand; they need human design choices, engineering and oversight. The degree of oversight depends on the use case. At a minimum, oversight is typically present during training and validation". It also clarifies: "the knowledge concept is part of the data-information-knowledge hierarchy, according to which data can be used to produce information, and information can be used to produce knowledge. In the context of AI, these are purely technical, non-cognitive processes" (International Organization for Standardization, 2022, 18). Moreover, sensors may pick up non-visible light such as infrared or sounds outside of the human range of hearing and possibly combine it with data received from a cloud infrastructure, resulting in a very inhuman way to inform decisions.

3.3.2 How AI works

Describing how AI works is difficult to do systematically because along with AI not being clearly defined, many of the subcategories of AI overlap, interact, or have elements that are not considered AI. Demis Hassabis, the CEO of Google Deep Mind, said in a New York Times podcast interview that

AI is the "science of making machines smart, and then a subbranch of AI is machine learning, which is the kinds of AI systems that learn for themselves and learn directly from data and experience, and that's really what of course has powered the renaissance of AI in the last 10-15 years, and that's the sort of AI we work on today" (Klein, 2023). We will focus on describing machine learning and some of the subcategories of machine learning that are having big impacts on AI-associated innovation like deep learning and reinforcement learning. We will also describe machine reasoning and robotics along with some of their subcategories. Machine learning, machine reasoning, and robotics are not all the subcategories of AI, and a system containing elements of them does not necessarily make that system AI. Machine learning, machine reasoning, and robotics are also not distinct categories. For instance, a machine vision system could be a sensor for a drone, but that machine vision system is built with machine learning. This approach, however, should provide a basic understanding of the systems underlying most AI startups and the AI processes and technology that affect entrepreneurship. Cochran (2018) and the International Organization for Standardization (2022) provide a more comprehensive overview of many of these technologies.

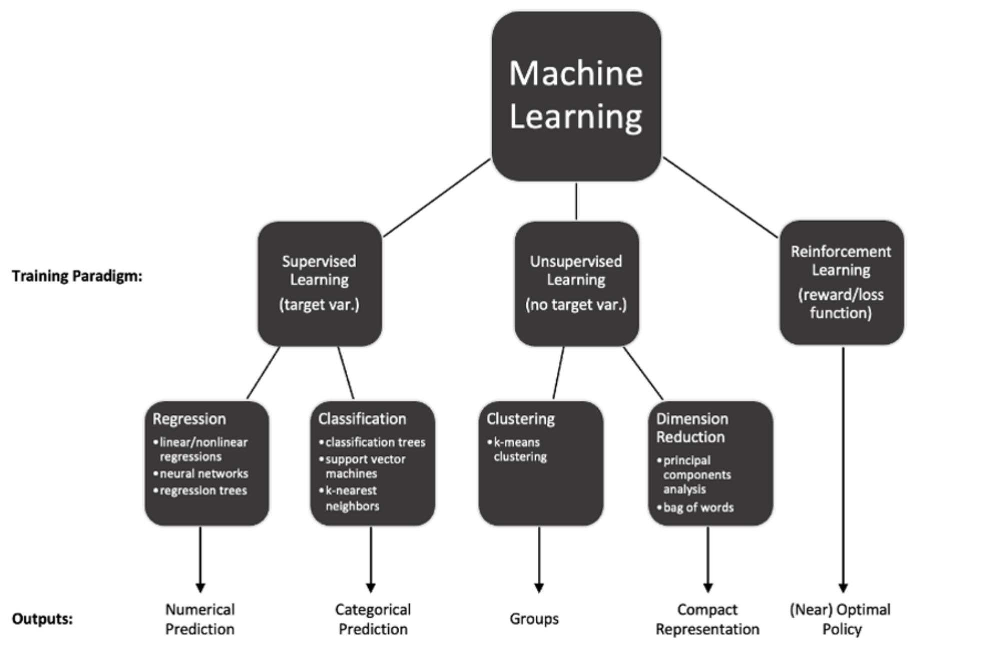


Figure 13: Types of machine learning

Source: Adapted from Cochran (2018). See also Emmert-Streib and Dehmer (2022).

Machine learning is driven by data. It powers cutting edge technologies in fields such as natural language processing, machine translation, speech synthesis, speech recognition, fraud detection, machine vision, large language models, logistics, and many others. Machine learning tends to be able to adapt to new situations better than expert systems, which rely on encoded knowledge. The three machine learning training paradigms are supervised learning, unsupervised learning, and reinforcement learning (Figure 13). Systems exist that fall into more than one of these training paradigms like semi-supervised learning, which uses some data with a target variable and some data without, or reinforcement learning combined with deep learning (a subcategory of supervised learning or sometimes unsupervised learning), but the actual training paradigms have distinct characteristics. Emmert-Streib and Dehmer (2022) offer additional paradigms such as multi-label learning, one-class classification, positive-unlabeled learning, transfer learning, multi-task learning, and one-shot learning while at the same time showing from published article metadata that these paradigms are used to a much lesser extent than supervised learning, unsupervised learning, and reinforcement learning regardless of scientific community. Lupp (2023) links the different ML paradigms to the entrepreneurial decision logics of causation and effectuation (Packard et al., 2017). Lupp (2023) argues that supervised ML supports causation logic, whereas unsupervised or reinforcement ML support effectuation logic.

Supervised learning has a target variable, which could also be described as a dependent, response, or outcome variable. The goal of supervised learning is to minimize the difference between predicted and observed values when applied to new situations. Regressions are one of many forms of supervised learning. When training a model with supervised learning, data is typically split into training data, validation data, and test data with no intersection among the splits. Multiple models are trained on the training data and the results are compared using the validation data in a process called model selection. The predictive performance is one of the metrics used to select the model – others being cost, interpretability, time dependency, etc. After model selection, the validation data gets mixed with the training data so the selected model can train with more data. The model is then tested on the test data to estimate performance when fielded (predictive performance should not drop substantially from the validation data). After assessing the performance, the test data gets mixed in to train the model on all the available data. The process of validation and performance testing with split data helps prevent overfitting a model (the tendency for some nonlinear models to capture noise in data, appearing like an extremely good fit, and then performing poorly with new data). Not all cases use or require validation data. The target variable can be a number, leading to a numerical prediction, or a category, leading to a categorical prediction.

Unsupervised learning does not have a target variable. The goal of unsupervised learning is to identify the hidden or underlying structures in data. Subcategories of unsupervised learning include density methods, clustering methods, and dimension reduction. Density methods assume the data was generated by a parent distribution and estimate that parent distribution from the data. An ex-ample application would be creating a crime heat map for allocating policing resources. Clustering methods use a measure of similarity or dissimilarity to group similar items using algorithms. Clustering is the technique behind many streaming service recommendations, social media algorithms and ecommerce taglines like "bundles with this item" or "you might also like". Dimension reduction represents data in more compact forms with minimal loss of information. Dimension reduction can be used for file compression, formatting data to facilitate search or matching, or as preparation for a supervised learning algorithm such as ridge regression using principal component analysis (a dimension reduction algorithm) as a step (Cochran, 2018).

Reinforcement learning has a reward/loss function that iteratively improves the model. The model outputs provide feedback to improve performance. Examples often used with reinforcement learning are games such as pong, chess, go, and multi-player poker, starting with random moves and improving over time with points and wins used to algorithmically evaluate moves made. There is a trade-off in reinforcement learning between exploitation of known information and exploration of new policies that will reveal or update information.

Deep learning is a subcategory of neural networks, which is a subcategory of machine learning. It has become popular due to being better than most algorithms at finding structure with minimal input from an analyst and the now existing infrastructure for it to access the vast quantities of data it requires. Deep learning overlaps with the three machine learning training paradigms, being adapted for unsupervised (as emphasized by Shepherd and Majchrzak, 2022), but also supervised applications, and often applied in reinforcement learning. In artificial neural networks, input data are fed into a network of nodes (neurons) that perform transformations on that data and eventually emit output values that could be predictions or recommendations. The nodes in the middle of the artificial neural network are called hidden layers because they are not directly observable, each a function of the preceding layer and better able to predict the target variable than the original inputs. If there are multiple of these hidden layers, the network can be classified as deep learning. When training neural networks, the weights and biases transforming the inputs for the first iteration can be given randomly to calculate the outputs. A loss function (Mean Squared Error for example) is then calculated from the first iteration outputs that tells how bad the result is. Backpropagation is then performed which uses the chain rule to calculate gradients starting from the outputs and working back to the inputs. The weights are changed by some step size opposite the direction of the gradient to minimize the loss function in a process called gradient descent. The forward calculation and backpropagation happen iteratively until the loss function is at an acceptable level. In practice, the algorithms are more complex to consider efficiency, local minimums, overfitting, etc., but the process described is at the core of what the more complex algorithms do. Feature selection consists of choosing hyper-parameters such as learning rate, the loss function, the number of hidden layers, the number of nodes in each layer, defining the network paths, and for each node choosing the activation function (a function that typically squishes the calculated results from the weights and biases before providing the nodes output and can introduce nonlinearity). Deep learning is one of the least interpretable model types, so it is not well suited to public policy decision making or situations where an explanation of decisions is required (Cochran, 2018).

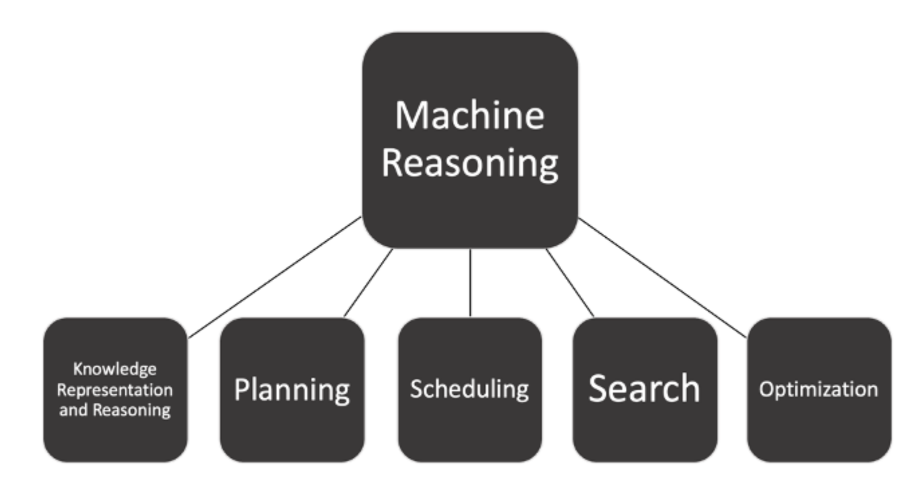


Figure 14: Machine reasoning types
Source: Own illustration adapted from High-level Expert Group on Artificial Intelligence (2019).

Machine reasoning (Figure 14) turns data into information into knowledge and uses that knowledge and preexisting knowledge to reason and plan. "Knowledge differs from information in that information is observed by the system, while knowledge is what the system retains from such observations. Knowledge is structured and organized; it abstracts away from the specificities of individual observations" (International Organization for Standardization, 2022, 18). The two main types of machine-readable knowledge are declarative, a statement about what something is, and procedural, a process of how to do something. While machine learning takes in data to form a prediction or recognize a pattern, most machine reasoning algorithms are logic based with rules encoded, i.e. the solution path in machine reasoning is typically independent of the data, so when the data is provided, a particular instance of the solution is the output (Cochran, 2018). An example of a complex machine reasoning model is Amazon's "mechanical sensei" which uses constraints in inventory, delivery lead times, capacity, and required delivery date to minimize shipping prices across their network. Upon a customer order, it charges and notifies the customer, hires the cheapest shipping provider that can meet the constraints, queues the item to be picked in a warehouse, updates the inventory, etc.

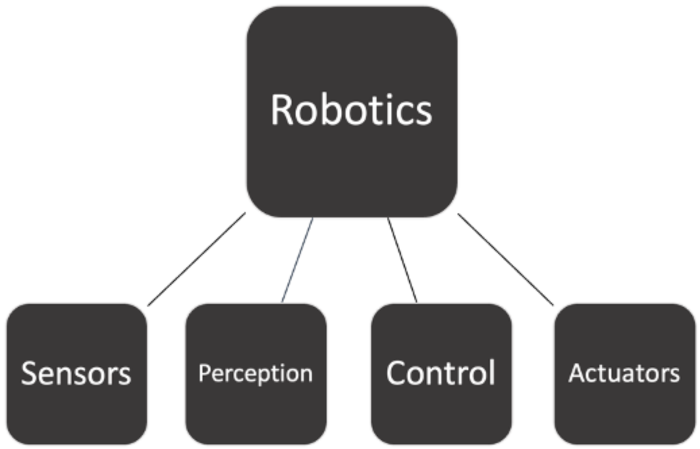


Figure 15: Aspects of robots
Source: Own illustration adapted from High-level Expert Group on Artificial Intelligence (2019).

A *robot* is an automation system composed of electronic, mechanical, firmware and software components that uses sensor inputs to control its activities and actuators to perform physical actions (International Organization for Standardization, 2022, 5, 27-28), as depicted in Figure 15. Machine learning is common in perception in robotics as neural networks are particularly good at many perception tasks such as audio processing and computer vision including object recognition. Industrial robots used for manufacturing are programmed to repeat actions precisely and without deviation so would likely not be considered AI. Service robots need to adapt to dynamic environments and thus often have machine learning components that may make them considered AI (International Organization for Standardization, 2022). There are synergies between machine learning and robotics that can create cycles that advance both fields. Machine learning can help design new hardware and robotics that robots can build. Machine learning is used for sensing, perception, and control in self-driving cars, and as that technology improves making more cars self-driving, more data is gathered, which further improves that technology.

3.3.3 AI applications

Due to varying definitions of AI and the breadth of its potential uses, we only highlight a few applications and focus on applications that most definitions agree are AI. Current AI systems are classified as narrow AI, defined as a "type of AI system that is focused on defined tasks to address a specific problem" (International Organization for Standardization, 2022, 4). What some AI companies are aggressively seeking, including OpenAI and Google DeepMind, is general AI, defined as a "type of AI system that addresses a broad range of tasks with a satisfactory level of performance" (International Organization for Standardization, 2022, 3). In the Fall of 2022, OpenAI launched ChatGPT, a Large Language Model, which brought AI technologies into the mainstream in a medium that users could interact with and explore[6]. Other AI applications less visible to the public, yet important, are also progressing, such as improving predictive maintenance, drug discovery, and fraud protection.

Large Language Models (LLMs) are deep learning neural networks trained on very large quantities of text data to predict the next word in a sequence of words and then tuned on smaller quantities of high-quality data so they can respond with the appropriate structure to queries. The accuracy of next word prediction in LLMs can be consistently predicted by the number of parameters and the number of tokens (amount of text) the model is trained on. The number of tokens in the best performing models, given a set training cost, is about 20 times the number of parameters (Hoffmann et al., 2022). Next word prediction is shown to be linked to improvements in abstraction, comprehension, vision, coding, mathematics, medicine, law, understanding of human motives and emotions, and more (Bubeck et al., 2023).

The trend in LLMs is towards multimodality and combination with search and tools. Generating text, images, audio, including speech and music, or videos is currently possible with varied stages of progress for each medium. Taking input from text, images, or audio is feasible. In reference to GPT-4 (an LLM by OpenAI), "It is able to reason about which tools it needs, effectively parse the output of these tools and respond appropriately (i.e., interact with them appropriately), all without any specialized training or fine-tuning" (Bubeck et al., 2023, 49) as long as the prompt specifies or expects it to use external tools. LLMs can use search engines and summarize search results, reference calculators, code with Python, etc. It is likely LLMs will eventually use other LLMs as external tools as well, as there is an ecosystem developing of LLMs trained and tuned with different specialties. Examples of specialty LLMs featured on OpenAI's GPT store include a website generator, a diagram

[6]GPT stands for Generative Pre-trained Transformer.

creator, a primer for learning things, a storage space to chat with personal pdfs, a text generator that writes research in the user's voice, and a travel guide. Even with referencing outside tools or more specialized LLMs when appropriate, LLMs are next word predictors and can fall short in tasks that require exploration, strategic lookahead, or where initial decisions play a pivotal role. Current research explores multiple reasoning paths and looking backwards and forwards to evaluate the best solutions and has shown positive results (Yao et al., 2024).

The examples of AI use mentioned throughout this section illustrate how broadly the technology is applied. Any situation where data is available and prediction or understanding structure is desired can be a suitable application for AI. Driving directions with real time traffic, credit or insurance risk, business plan evaluation, medical diagnoses, electrical grid optimization and management, swarm technologies for the defense industry, etc., it is possible to list almost endless potential applications. Using AI in an application does not always result in a finished product and often has humans, sometimes lots of humans, assisting in evaluating results and providing the feedback that many of the systems need.

3.4 The AI–entrepreneurship nexus

3.4.1 AI and uncertainty in entrepreneurship

AI has the potential to affect the core of entrepreneurship. As Townsend and Hunt (2019) point out, entrepreneurship theories are built around the fundamental question of how entrepreneurs deal with uncertainty. Entrepreneurs operate in environments where future possible states of the world and consequences of actions and their probabilities are unknown (Knight, 1921). For example, there is uncertainty about market demand for novel products and services, social resistance to innovation, and responses of competitors; at the same time, entrepreneurs are constrained by limited resources. The key capability of AI is prediction, which has the potential to reduce uncertainty (Lupp, 2023; Agrawal et al., 2024). AI can be used to search through a very large set of options to identify opportunities, for example for product design or the choice of market segments. Examples given by Townsend and Hunt (2019) include AI use in drug discovery (the startup Insilico Medicine) and AI-powered design tools (Autodesk or the startup Stitch Fix). Further, entrepreneurship theories often emphasize that entrepreneurs employ imagination and creative approaches to decision-making to identify opportunities under uncertainty (Kier and McMullen, 2018). Generative AI is starting to show signs of creativity – Townsend and Hunt (2019) mention Autodesk's generative design AI, or one could think of poems, images, videos and music generated by AI – an ability that was previously thought to be exclusively human domain. In the Delphi study by Van Gelderen et al. (2021), several entrepreneurship experts expected AI to augment the creative abilities of entrepreneurs by 2030.

The increasing capabilities of AI in prediction (and thus uncertainty reduction) and creativity raises the question whether further substantial advances in AI may change the core functions of the entrepreneur. For example, Foss and Klein (2012) theorize that entrepreneurial judgement is exercised in unstructured decision environments when no clear decision rules exist, and the exercise of judgement is a skilled activity accumulated through experiential learning. By facilitating systematic analysis of decision environments to identify resources and opportunities, AI may affect the degree to which human judgment is needed in entrepreneurial decision environments (Townsend and Hunt, 2019). AI may then not stop at helping entrepreneurs but might reduce the need in the economy for entrepreneurs to play the role of reducing uncertainty not only for themselves, but also for others. However, Townsend and Hunt (2019) also point out that, even if AI can help resolve what is possible, the human entrepreneur still needs to solve the fundamental problem of what is desirable according to the entrepreneur's goals, preferences, and objectives (e.g., McMullen and Shepherd, 2006).

Townsend et al. (2023) further elaborate that entrepreneurial decision environments, which are characterized by Knightian uncertainty, are inherently unpredictable and therefore a fundamental boundary even for enhanced AI. An entrepreneur is unable to determine what will be possible in the future not only due to a lack of data. Data is always about the past, so its usefulness is limited when the future deviates in important and unpredictable ways from the present and past. The uncertainty problem is also not only due to the entrepreneur's incapability of processing the data or a lack of tools to compute predictions. In situations where data or processing constraints are the bottlenecks to resolve uncertainty, AI has transformational potential, for example in drug discovery. In contrast, (Townsend et al., 2023, 11) argue that entrepreneurs face a decision environment that is indeterministic by nature, "neither analyzable nor predictable using statistical analysis and probabilistic reasoning". In their view, this type of Knightian uncertainty is unique to entrepreneurship and represents a boundary for the use of AI at the core of entrepreneurship. However, other researchers argue that entrepreneurs navigate uncertainty through the formulation of hypotheses, experimentation, and learning (e.g., Åstebro et al., 2014; Packard et al., 2017; Camuffo et al., 2020; Agrawal et al., 2021; Zellweger and Zenger, 2023). In such frameworks, AI may be able to play a more substantial role in uncertainty reduction for entrepreneurs, as AI could greatly increase the speed and reduce the cost of experimentation (Van Gelderen et al., 2021). The role of AI in entrepreneurial judgement and decision making under uncertainty is an important avenue for future research.

3.4.2 Entrepreneurial opportunities with AI

A process perspective on entrepreneurship (Shepherd et al., 2019) is useful to consider AI impacts on different stages of entrepreneurship (Chalmers et al., 2021; Obschonka and Fisch, 2022; Schiavone et al., 2022). Giuggioli and Pellegrini (2023) conduct a systematic literature review on AI as an enabler for entrepreneurs. They identify three clusters in the literature that align as sequential phases in an "AI-enabled entrepreneurial process": opportunity, decision-making, and performance[7]. The authors point out that these phases are compatible with the conceptual framework developed by Chalmers et al. (2021), who consider impacts of AI on antecedents of venture creation, firm-level activities (prospecting, organizational design, and exploiting), and outcomes of venture creation. In particular, opportunity relates to prospecting activities, decision-making to organizational design, and performance to exploiting activities. Gerling et al. (2022) present a critical review of the literature on AI in digital entrepreneurship from a socio-technical viewpoint. They classify their reviewed papers into three research perspectives: agency, processes, and outcomes. We structure the following subsections broadly along the three phases distinguished by Giuggioli and Pellegrini (2023), opportunity, decision-making, and performance.

AI as a technological breakthrough changes the business environment and has been described as an *external enabler* of entrepreneurial activities and success (Davidsson et al., 2020; Davidsson and Sufyan, 2023). In the opportunity phase, the first phase of the AI-enabled entrepreneurial process, the entrepreneur looks for ways to reshape and enhance traditional business model approaches with AI (Giuggioli and Pellegrini, 2023). AI can assist in discovering or co-creating opportunities (Lupp, 2023). Giuggioli and Pellegrini (2023) separate the research on this phase into two streams.

A first literature stream shows AI as an external enabler for new products and services or business models (Obschonka and Audretsch, 2020). AI will influence the decision to start a company and how entrepreneurs develop, design and scale (Chalmers et al., 2021). The AI revolution is predicted to be in full force during the 2030s having a greater impact than the industrial and digital revolutions combined in terms of speed of technological change, opening growth and profit opportunities and

[7] A fourth cluster, education and research, is not placed within this sequence and is outside the scope of our review.

increased competition from venture and crowdsourced startups (Madridakis, 2017). Many of the new opportunities for entrepreneurs with AI concern vertical integration, making the design of business models fundamental to get useful technologies to the marketplace (Garbuio and Lin, 2019).

A second stream focuses on the Internet of Things (IoT), which can overlap with AI technologies specifically around sensors and data analytics. A new entrepreneurial paradigm, enabled by the internet and digital technologies, reflects how digital technologies change technology entrepreneurship and a new venture creation process, and leverages innovation potential from large groups and dispersed individuals participating in entrepreneurship (Elia et al., 2020). In this context, AI may positively affect the economy, entrepreneurship development, and company opportunities (Mamedov et al., 2018). A subcategory of digital entrepreneurship, sensor-based entrepreneurship, uses passively sensed IoT data to provide products or services, changing how we interact with objects, each other, and how companies interact with their customers (Brown, 2017). IoT data also includes manufacturing process data, which enables an understanding of transaction costs for non-ownership services. This may give rise to entrepreneurial opportunities for new business models such as providing manufacturing assets, maintenance, operation, analytical services, and services targeted at the end user such as customization (Ehret and Wirtz, 2017).

The phase Giuggioli and Pellegrini (2023) refer to as “opportunity” is referred to as “prospecting” by Chalmers et al. (2021) and focuses on production of new venture ideas. Large datasets and learning algorithms that can see otherwise imperceptible patterns or make precise predictions can be turned towards entrepreneurial opportunity and exploitation (Cockburn et al., 2019). Chalmers et al. (2021) separate new AI ventures, based on form, function, and purpose, into three approaches for augmenting information search and idea production. The first approach is a subset of science and technology focused startups using AI to search across complex combinatorial problem spaces for technical solutions (Agrawal et al., 2019). The second approach uses sentiment analysis and Natural Language Processing to analyze online and social content (Gaspar et al., 2016). The third approach tests assumptions with AI such as predicting customer reaction to a feature or price change with existing data, which tends to be less biased and more generalizable than sourcing those reactions through customer engagement (Chalmers et al., 2021).

Shepherd and Majchrzak (2022) demonstrate how AI is transforming four sectors in the economy – customer service, financial, healthcare, and education – through applied examples explaining how those transformations affect entrepreneurship. For the customer service sector, they highlight facial recognition to respond to emotional reactions, predictive modeling for delivery speed, security through voiceprint, and multilingual chatbots. For example, a voice-enabled human-like avatar at an online store can answer questions, help navigate, highlight promotions, and make recommendations to increase customer engagement (Brown, 2021). Customer service entrepreneurial opportunities may include developing AI-based customer service products and using those products to identify customer needs, leading to additional entrepreneurial opportunities; improved venture speed, customization, quality, and reliability; and more secure payment (Shepherd and Majchrzak, 2022). For the financial sector, AI is enabling financial fraud detection and response, blockchain security improvements, financial investment identification, and customized risk assessments. The financial entrepreneurial opportunities emphasized are product creation and using these products as verticals in ventures for fraud protection, accessing debt capital, and increased financial security for customers. For the healthcare sector, AI is used for diagnosis, recommending treatments, predicting outcomes, and robotics. Healthcare entrepreneurial opportunities include products and services, but also ownership and management of healthcare businesses. For the education sector, AI is used for virtual teaching assistants, personalizing programs and interventions, checking plagiarism, and grading. Entrepreneurial opportunities in this sector comprise products and services, and more effectively identifying and exploiting opportunities, starting organizations, and growing ventures.

Shepherd and Majchrzak (2022) suggest several entrepreneurial opportunities to leverage AI or address the needs of others who implement AI. We are entering a feeling economy, differing from the mechanical economy (physical and repetitive focus) and thinking economy (process, decision, and knowledge updating focus), emphasizing interpersonal communication and empathy (Huang et al., 2019). Shepherd and Majchrzak (2022) present capitalizing on the feeling economy as an entrepreneurial opportunity. Early feeling economy AI includes monitoring attention and emotion of drivers in vehicles with autonomous features, chatbots responding to feelings detected, and robots with more anthropomorphic displays such as voice, touch, and emotion. Future feeling economy AI will use databases sourced from physical sensors and physiology-monitoring personal devices to improve communication by recognizing human-generated emotions and modeling the emotional activity in teams; the same data will also form stronger human-robot connections for rehabilitation and education (Franzoni et al., 2019). In line with these predictions, OpenAI released its AI voice assistant GPT-4o in May 2024, which analyzes the user's facial expressions and adapts its own tone accordingly. Another entrepreneurial opportunity mentioned by Shepherd and Majchrzak (2022), which may not be obvious, is to develop and monitor governance mechanisms for regulation. We discuss AI regulation further in Section 3.8. Noticing and matching entrepreneurial opportunities itself is changing with AI. Identified opportunities that are not exploited can be stored as data and taken up by interested entrepreneurs; startups can be matched with established organizations.

AI can support innovation management, not only in entrepreneurship, by offering more systematic approaches when traditional innovation management resources are overwhelmed and hampered by information processing constraints (Haefner et al., 2021). More specifically relating to digital entrepreneurship, AI supports identifying and acquiring knowledge, customizing products and services as a competitive entrepreneurial strategy, and managing product innovation, according to a systematic literature review by Mariani et al. (2023) on innovation and AI.

3.4.3 AI and entrepreneurial decision making

Chalmers et al. (2021) claim with support from Hoffman (2004) and Nambisan (2017) that foundational assumptions of mainstream organizational theories are shifting due to new practices enabled by digital technologies. To understand what the new practices mean for organizational design, they turn to a framework established by Burton et al. (2019) to capture dimensions, including organizational structure and decision systems, across which entrepreneurial firms are arranged. Based on case studies and survey data from senior managers and business executives, AI only moderately impacted venture structure at the time of publication, mostly supporting existing business, and AI was not yet used at scale (Ransbotham et al., 2018; Brock and von Wangenheim, 2019; Fountaine et al., 2019). When AI is deployed at scale, new forms of organizational structure will be created (Chalmers et al., 2021). AI can serve as a new form of input and process other inputs, reshaping the cost structure of a firm (Desai, 2019), including the cost of innovation as firms substitute away from labor-intensive research towards passively generated datasets and enhanced prediction algorithms (Cockburn et al., 2019). Huang and Rust (2018) predict how the distribution of tasks across a company evolves as technology improves, modeling service tasks with four types of intelligence (mechanical, analytical, intuitive, and empathetic; the order is from most developed with AI to least). AI initially augments jobs as it replaces individual tasks and may later lead to replacement of a job when it can do all a job's tasks. A more in-depth discussion of exposure of occupations to AI is detailed in Section 3.6. Firms employing AI are anticipated to create three new employee categories: trainers improve and add nuance to algorithms; explainers bridge the technical gap for business managers; sustainers manage the systems including ethics and compliance (Wilson et al., 2017; Daugherty et al., 2019).

Business models can give insight into organizational design of startups as they describe how startups intend to make a profit through a product or service. Weber et al. (2022) filter startups on Crunchbase that have AI technology as a core component of their product or service and perform a quantitative cluster analysis on the AI business models to determine patterns and how they differ from common information technology (IT) business models. Their cluster analysis confirmed AI widens the scope for applying IT, especially shifting IT applications towards the knowledge and service work domains both for augmentation and replacement (Coombs et al., 2020). Most AI startups focus on delivering complex AI technology that is otherwise difficult and costly to develop to their business customers (Jöhnk et al., 2021). The cluster analysis performed by Weber et al. (2022) revealed ethical aspects were not key characteristics of AI business model startups; data is an important element for value creation in most AI startups; data generates insights or detects anomalies in some AI startups and trains models that are embedded in products or services; and some AI startups form close relationships with industry partners to gain access to more exclusive data. Weber et al. (2022) reveal four archetypes from clustered patterns in the AI business models for how AI technology gets developed and delivered. The first archetype is AI-charged product/service provider startups, which offer a product or service that already has an embedded trained AI model. The second archetype is AI development facilitator startups, which provide a programmable interface or software development kit for a customizable solution. The third archetype is data analytics provider startups, which offer data analysis to support decision making. The fourth archetype is deep tech researcher startups, which develop research-based niche solutions.

Some initial studies suggest that generative language models can effectively assist entrepreneurs in creative tasks and lead to cost savings. This has been demonstrated for generating business ideas (Boussioux et al., 2024), elevator pitches, crowdfunding narratives and tweets (Short and Short, 2023), and for aiding in theory development (Davidsson and Sufyan, 2023). Generative models can quickly produce flexible content for entrepreneurs and are adept at style and content mimicry. Boussioux et al. (2024) investigate how human-AI collaboration using OpenAI's chatbot GPT-4 fares in a contest of business ideas. They launched a crowdsourcing challenge soliciting innovative business solutions and compared human submissions to the contest with solutions that GPT-4 created in collaboration with a human prompt engineer. The solutions were judged by 145 human evaluators who were not told which solutions were created by humans or in human-AI collaboration. The collaborative AI solutions were comparable to the human solutions in terms of creativity and provided higher value on average, but they did not completely match human ingenuity and diversity. By prompting the AI to mimic expert personas, the novelty ratings became statistically indistinguishable from the human submissions. It took months to collect the human submissions, but only hours to generate the AI solutions, suggesting that collaboration between humans and generative AI can lead to significant cost reductions for performing creative tasks.

In the organizational design framework Chalmers et al. (2021) use, established by Burton et al. (2019), decision systems are a component of organizational design. Giuggioli and Pellegrini (2023) focus on the decision system component in their organizational design equivalent section, calling the second phase of the AI enabled entrepreneurial process "decision-making"; the entrepreneur uses AI to assist in making predictions with available data. Giuggioli and Pellegrini (2023) separate the research on this phase into three streams. The first stream highlights benefits of more firm-generated data and improved analytics enhancing decision support for entrepreneurs. Machine learning can aid in business model validation through a hybrid intelligence decision support system that uses an iterative approach of interaction with relevant stakeholders and formal analysis (Dellermann et al., 2019). Future success of perceived opportunities can be predicted pre-startup by combining uncertainty factors with relevant datasets allowing decision makers to select favorable opportunities and identify high influence uncertainties (Tomy and Pardede, 2018). The second stream of studies centers on market analysis. Product introduction or discontinuation can be decided through forecasting produced by machine learning applied to historical sales transactional data (Ramesh et al., 2018). Fish and Ruby (2009) show an effective approach for startups expanding into export to screen foreign markets using self-organizing maps, a neural network clustering technique. The third research stream is about fundraising and crowdfunding. Deep learning, even using only basic project attributes such as category, funding target, and geographic location, can predict funding outcomes on Kickstarter data to a high degree of accuracy (Wang et al., 2020). Using text data and semantic analysis in crowdsource funding outcome predictions can improve the model and indicate what entrepreneurs should emphasize (Yuan et al., 2016; Wang et al., 2020). Machine learning has been used for facial feature detection to determine likelihood of crowdfunding success based on appearance of trustworthiness (Duan et al., 2020) and emotion (Raab et al., 2020). Machine learning has been combined with graph theory to determine how network position impacts the success rate of investors, which also gives insight into how startups should select venture capital providers (Glupker et al., 2019). For decision systems, Shrestha et al. (2019) propose a typology of: full human to AI delegation, hybrid AI-to-human sequential decision making, hybrid human-to-AI sequential decision making, and aggregated human-AI decision making. AI-to-human sequential decision making can optimize innovation strategies to source in-novation ideas, which is of interest to entrepreneurial firms as it reduces the cost of problem solving and of evaluating and selecting solutions.

3.4.4 Influence of AI on entrepreneurial performance and outcomes

After we have outlined how AI may lead to entrepreneurial opportunities and affect decisions to exploit them, the question arises how AI influences the performance of entrepreneurs and the outcomes of entrepreneurial activity. Relevant outcomes include business survival, pecuniary rewards such as earnings, and the wellbeing of the entrepreneur, which is partially determined by nonpecuniary rewards. Empirical research in economics on the effects of AI has so far focused more on job displacement through automation (see Section 3.6.1) than on productivity gains. Employment reductions hint at cost savings and thereby potentially higher profits for innovative entrepreneurs and their capital investors, at least in the short run. However, the extent of productivity gains through AI, their sustainability, and the distribution of the rewards are unclear; these are important research areas.

Chalmers et al. (2021) speculate that AI technologies might enable some entrepreneurs with high technological skills and venture capital firms to gain large financial returns with comparably little effort. On the other hand, they reckon that rewards might often be concentrated in large corporations who control critical amounts of capital and expertise. AI technologies may also favor large firms due to their ability to accumulate big amounts of data; Gerling et al. (2022) are concerned that powerful corporations may invade privacy to collect ever more data. These economies of scale for AI might lead to industry concentration and the rise of superstar firms (Autor et al., 2020).

A related discussion revolves around access to AI algorithms and models. Open-source technology would facilitate market entry for entrepreneurs, whereas proprietary algorithms may lead to oligopolistic or monopolistic tendencies. Open-source algorithms may form a scaffold for entrepreneurs upon which they can build their apps and tools, although open-source software may still come with conditions for use. Montes and Goertzel (2019) argue that decentralized and distributed AI can bring about more equitable development of AI. A counterargument against making AI open source is the concern that the open AI technologies could be misused by malicious actors. This may amplify risks of AI such as misinformation or bias. On the other hand, open-source technology can be scrutinized by a larger and independent community, which may decrease risks in a democratic way. OpenAI initially announced it would make its algorithms publicly available but has since retreated from that plan; Google and most other tech companies also keep their AI models closed. Meta made its trained generative AI text model Llama (the model's weights, evaluation code, and documentation) openly available, although not the training data and the code used to train it (Nolan, 2023).

The tension between opportunities for entrepreneurs due to the disruptiveness of AI and disadvantages for newly founded firms due to the particular economies of scale related to AI is an important avenue for entrepreneurship research. This research – we review initial empirical studies further below – should aim to inform government regulation of AI, antitrust policy toward big technology companies, and discussions on support programs for small businesses (such as certain exemptions from data protection requirements) or public investments in infrastructure (such as cloud services).

Some researchers have begun to examine the causal effects of AI on performance in entrepreneurship by running field experiments. Otis et al. (2024) randomly assigned 640 Kenyan entrepreneurs into a treatment and a control group. They gave the treated participants access to an AI mentor via their smartphones powered by the LLM GPT-4. The control group received a standard business guide instead. Access to the AI advice did not influence performance on average in the full sample. However, when splitting the sample by pre-treatment performance, the authors find interesting heterogeneity. Entrepreneurs in the treatment group with above-median performance increased their business performance by 15%, whereas those with below-median performance decreased their performance relative to the respective control groups. Further analysis revealed that entrepreneurs in both groups used the AI advisor, but low performers consulted it for more challenging tasks, which may have had adverse effects. Overall, the results point to the potential that AI might lead to an increasingly unequal distribution of rewards between high- and low-performing entrepreneurs.

One might argue that the generative AI used in this field experiment has not been sophisticated enough yet to aid effectively in a wide range of complex tasks in entrepreneurship. Does AI improve outcomes in more narrowly defined areas in the context of entrepreneurship such as judging business plans or angel investment? McKenzie and Sansone (2019) report that machine learning did not improve the prediction of performance outcomes (business survival, employment, sales, and profits three years later) in a business plan competition in Nigeria. They compared the performance of the ML approach using more than 500 variables to simple prediction models such as logit regressions using only a handful of ad-hoc predictor variables. However, human judges performed even worse than the simple prediction models, as their scores were uncorrelated with the outcomes. Blohm et al. (2022) compare the returns from investing via an angel investment platform between 255 human business angels and an ML algorithm. On average, the algorithm achieved higher investment performance, and only experienced human business angels outperformed the machine. Further analysis suggests that experienced human business angels are particularly successful when they are able to suppress three decision biases: overconfidence, loss aversion, and local bias, i.e., the tendency to make investments in close proximity to the investor's own location. Blohm et al. (2022) argue that ML algorithms are mostly unaffected by these biases, although more research is necessary, as bias might be introduced by the training data.

While the result of Blohm et al. (2022) seems promising for AI, the underwhelming results reported by Otis et al. (2024) and McKenzie and Sansone (2019) are consistent with the challenges brought by the large uncertainty in entrepreneurship discussed at the opening of this section. In contrast, generative AI has been shown to lead to measurable performance increases in more narrowly defined tasks such as writing tasks (Noy and Zhang, 2023) or customer support (Brynjolfsson et al., 2023), suggesting that AI might currently provide more useful support to more specialized employees than to entrepreneurs dealing with a broad range of challenging tasks in the face of uncertainty.

3.5 Initial evidence on AI in entrepreneurship

3.5.1 AI usage in entrepreneurial businesses

The development of AI technology and its implementation has accelerated rapidly in recent years. Therefore, evidence about the usage of AI in entrepreneurial businesses has only recently been emerging. McElheran et al. (2024) study early adoption of AI in a very large and high-quality sample of firm data in the United States. Their analysis is based on more than 440,000 employer businesses observed in two datasets provided by the U.S. Census Bureau, the Annual Business Survey (ABS) in 2018, which included a technology module, and the administrative Longitudinal Business Database (LBD). The analysis focuses especially on a subsample of 75,000 young firms not older than five years.

According to their results, only 6% of US firms used at least one of five AI technologies (automatedguided vehicles, machine learning, machine vision, natural language processing, or voice processing) in 2017. Rammer et al. (2022) report the same share of firms using AI in 2019 based on the Mannheim Innovation Panel (MIP), which is representative for firms in Germany in manufacturing and business-oriented services with at least 5 employees. This share of AI usage is likely lower than reported in studies such as Kazakova et al. (2020) for Europe or Chui and Malhotra (2018) because these studies do not use representative samples and oversample larger companies. Another reason may be that the LBD and the MIP asked for AI usage directly, but many people are not aware that they use AI technologies embedded in software, so indirect questions reveal higher rates of AI usage, as we show in the next section. The majority of firms using AI purchase commercially readymade applications rather than developing or customizing solutions in-house (Hoffreumon et al., 2024).

Despite the low adoption rates of AI, McElheran et al. (2024) find AI usage in every sector, supporting the view that AI has the potential to become a *general-purpose technology* (Goldfarb et al., 2023). In fact, more than half of the firms with more than 5000 employees report usage of at least one AI technology, and about a quarter of these indicate that they use AI intensively. Czarnitzki et al. (2023) also report that firms using AI are larger and grow faster than other firms based on the German MIP. The skewness of the size distribution of firms using AI may point at a disadvantage of small and entrepreneurial businesses. Potential reasons for the skewed distribution include high fixed costs and specialized skill requirements for AI implementation.

Among young firms up to five years old, McElheran et al. (2024) find that young owners with higher levels of education and experience are the most likely to report AI adoption. Owners using AI much more often state process innovation as their business strategy (39% in comparison to 20% among non-AI users) and that intellectual property is very important (40% versus 20%). Besides innovation, AI users also mention social entrepreneurship as an important motive for their entrepreneurial activity more often than owners not using AI. AI usage among young firms is positively associated with indicators of high-growth entrepreneurship such as venture capital funding (3% on average versus 1% among young firms not using AI), owning patents, and high capitalization at startup. Even controlling for these characteristics, AI adoption is found to be positively correlated with revenue growth among the young firms. As noted by the authors, this may suggest that AI usage supports firm performance, but the study cannot establish causality. For example, concerning venture capital, it is possible that startups using AI attract venture capital, or that venture capital firms push ventures to adopt AI. In any case, these descriptive results show that the relationship between AI use and high-growth entrepreneurship is a very promising avenue for future research. Among firms in general (not limited to young firms), evidence from Germany suggests that AI use is associated with innovations that are new to the world (not only to a regional or sectoral sub-market) as well as process innovations (Rammer et al., 2022) and productivity (Czarnitzki et al., 2023); the latter paper also attempts to establish causality by using instrumental variables methods.

McElheran et al. (2024) also document a pronounced *geographic disparity* in AI adoption among startups in the United States. Young firms using AI are concentrated in a small number of urban regions, including the Silicon Valley in California and the Research Triangle in North Carolina, suggesting that geographical proximity to leading academic centers and technology clusters is important (Kerr and Robert-Nicoud, 2020; Bessen et al., 2023). Local face-to-face interaction seems to be important for startup activity even in times of ubiquitous virtual communication (Fossen and Martin, 2018). As startup clusters tend to reinforce themselves with tech startups attracting other tech startups, this may lead to a geographical "AI divide" between super-star regions and other regions falling behind (McElheran et al., 2024). However, perhaps less obvious cities also emerge as centers of frequent AI-using startups, including Nashville, TN, San Antonio, TX, and Las Vegas, NV.

Due to the rapid technological progress in AI and accelerating adoption by firms, it is important to monitor the developments with a minimal delay to inform research and potential policy responses. Bonney et al. (2024) report results on AI use by firms in the United States based on the Business Trends and Outlook Survey (BTOS) covering September 2023 to February 2024. The BTOS is a bi-weekly survey provided by the U.S. Census Bureau, which is representative of U.S. employer businesses and includes questions on AI use until August 2024. A difference to the findings based on the ABS described above is that Bonney et al. (2024) report a more prevalent use of AI by small firms, leading to a U-shaped rather than increasing relationship of AI use with firm size. This might suggest that fixed costs for adoption of certain AI tools such as Generative AI have decreased, making these technologies more accessible for small and young firms.

3.5.2 AI startups and their entry barriers

The previous section reviewed evidence on the use of AI among young entrepreneurial businesses. Another empirical research question is how startups are actively involved in developing AI-enabled products and services and bringing them to market, and whether young, entrepreneurial firms are able to compete against big technology companies.

Fast-growing high-tech startups have played a central role in AI development. Some AI startups have received very large investments or were bought by Big Tech corporations. Some examples of such AI startups include OpenAI, the maker of ChatGPT, which received multiple billions of dollars in investments from Microsoft; Inflection, the developer of another LLM; MosaicML, a generative AI orchestration bought by the leading data storage company Databricks for US$1.3 billion in 2023; and Europe-based Mistral AI. Figure 16 presents the number of newly funded AI startups that received over US$ 1.5 million between 2013 and 2023 by geographic region. A strong increase is visible in particular in the United States, extending its lead over the other regions.

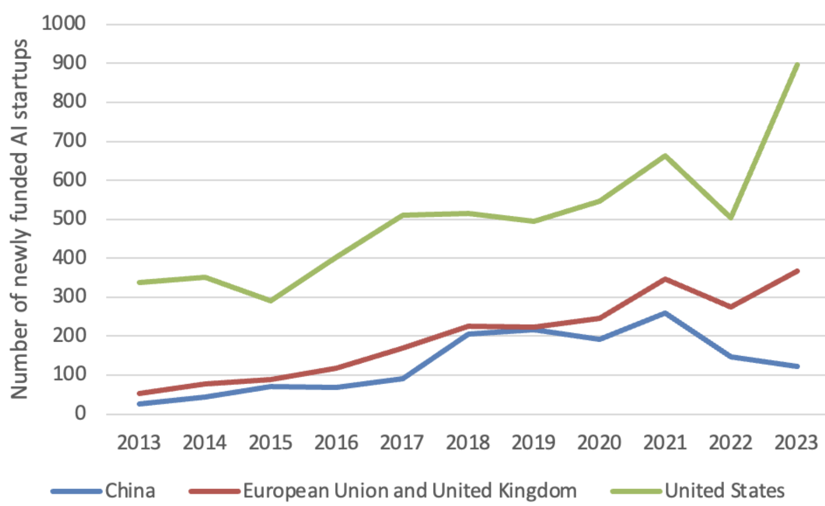


Figure 16: Number of newly funded AI startups by geographic area
Note: Number of newly funded AI startups that received over US$ 1.5 million in funding between 2013 and 2023.
Source: Adapted from Maslej et al. (2024), leveraging data from Quid.

Dinlersoz et al. (2024) analyze AI startups defined more inclusively, not limited to startups receiving large investments. They use administrative data on business applications from the U.S. Census Bureau's Business Formation Statistics to identify AI-related business applications over the period 2004-2023. The analysis shows that the annual number of new AI business applications was stable between 2004 and 2012, but began to accelerate thereafter, with a large upward jump in 2023, potentially related to the wide availability of generative AI. The likelihood of hiring others is found to be higher for AI-related businesses than for other businesses, suggesting that AI startups are beginning to contribute to overall business dynamism.

An emerging literature has started to collect survey evidence on the activities of AI startups. Bessen et al. (2023) administer five rounds of an online survey using Qualtrics. They recruit their survey participants mostly from startups labeled as AI firms in Crunchbase, a data provider reporting on more than 650,000 ventures and listing 28,000 startups related to AI in 2021 (Weber et al., 2022), and from additional data sources such as the Creative Destruction Lab, a startup incubator. Bessen et al. (2022) and Bessen et al. (2023) use subsets of these survey data, and the authors plan to collect further survey waves to continue the research.

Can resource-constrained AI startups compete with big technology companies such as Alpha-bet, Amazon, Meta and Microsoft in the AI arena or do large entry barriers tend to create and sustain oligopolies or monopolies? Firm entry, competition, innovation, and the future dynamism of the economy may depend on the answer to this question.

In Section 3.4.2, we discussed concerns in the conceptual literature on AI and entrepreneurship that rewards from productivity gains through AI could be concentrated among large firms. This prediction finds empirical support in Babina et al. (2024). Using a measure of AI investments at the firm level based on the skills of hired workers, they estimate a positive relationship between AI investments and firm growth triggered by product innovation. Importantly, they report that this positive effect of AI investments increases in the firm's initial size.

Advantages for large firms in AI development and usage may stem from several potential *entry barriers for AI startups*. The first potential entry barrier is that developing and refining AI algorithms requires training data (Stucke and Grunes, 2016; Bessen et al., 2022), and large firms with significant resources and stocks of customers may be in a better position to collect, acquire and maintain big data (Chalmers et al., 2021), for example in the form of large proprietary databases (Cockburn et al., 2019; Desai, 2019). Bessen et al. (2023) report that 80% of the AI startups in their sample use customer data and 63% data available from third parties, including publicly available data, suggesting that alternative data sources exist. There may be diminishing returns to the amount of big data beyond a certain point (Bajari et al., 2019), which would work against the emergence of unbreakable monopolies. However, it is likely that in some industries, the enormous data advantage of large incumbents makes it very hard for new competitors to enter the market, for example in the search engine market (Bessen et al., 2023). Concerning LLMs, Hoffmann et al. (2022) find that the performance of LLMs at the time of their study would have benefited more from additional training data rather than additional model parameters, highlighting the importance of the size of training data. However, they also speculate that scaling to ever larger data is only advantageous as long as the additional data is of high quality. (Gerling et al., 2022, 12) derive from their literature review the interesting research question asking how entrepreneurs "cope with the paradox of providing functions that require data to operate, but these data are only generated by the operation". We concur that further research is necessary to understand the relevance of data-related entry barriers for AI startups.

The second, related concern is that the training of AI models requires very significant IT infrastructure to store and process Big Data and to run the AI algorithms. The high fixed costs for capital investment may constitute a large entry barrier and lead to monopoly positions of big technology companies. However, Bessen et al. (2023) find that the AI startups in their sample use cloud computing effectively, a technology that may make hardware necessary for AI accessible without imposing prohibitively high fixed costs.

Third, it may be difficult for small, young firms to comply with data protection legislation and novel AI safety regulations. Firms of all sizes targeting the creation of similar AI products require similar data resources. Data regulations make it more difficult for firms to collect, store and analyze data, in particular personally identifiable or employment data. Large firms are better able to employ specialized legal staff to manage compliance with regulations, potentially putting small firms at a disadvantage. This is particularly a concern in the European Union (Bessen et al., 2023), which has stricter General Data Protection Regulation (GDPR) and passed the first AI safety law worldwide in 2023. Regulations requiring data sharing have been suggested as a potential remedy (Himel and Seamans, 2017). However, Bessen et al. (2022) find that AI startups using proprietary training data, as opposed to those using publicly available data, have a higher likelihood of successfully acquiring follow-on venture capital funding. A potential explanation is that startups and venture capital providers may be hesitant to invest in developing products using substitutable, not proprietary training data as such products may not be sufficiently innovative and differentiated to be protected from competition. Therefore, regulations that require data sharing may not necessarily help startups.

Rare skills are essential for AI startups. Gofman and Jin (2024) observe that universities that lost AI professors exhibit a reduced likelihood that their students establish AI startups and raise funding. They conclude that gaining knowledge from AI professors at universities is critical for successful AI startup activity among students. Scarcity of potential employees with the skills and talents necessary to develop AI products and services may be another barrier for AI startups, which may not be able to compete with big technology companies in terms of salaries and benefits they offer. Bessen et al. (2023) note that most startups in their sample develop their own software for most applications, suggesting that skilled developers are available to them, but future research should aim to provide direct evidence.

3.5.3 New survey evidence on AI exposure of entrepreneurs

In this section, we present a novel analysis of population survey data to shed light on AI usage and AI-related work tasks performed by entrepreneurs in comparison to paid employees. Within the entrepreneurs we distinguish between the solo self-employed (nonemployers) and employers. We also investigate whether these occupational groups share similar worries about how technological change may affect their work. The solo self-employed, who are the majority of entrepreneurs in Germany as well as the United States, were not included in many prior studies due to data limitations, for example in the studies by McElheran et al. (2024) and Bonney et al. (2024), because the ABS, the LBD and the BTOS provided by the U.S. Census Bureau only cover employer businesses.

For this original analysis we use the German Socio-economic Panel (SOEP), a large panel survey of individuals randomly drawn from the population in Germany (Goebel et al., 2019). Annual interviews started in 1984; since then, the sample has been extended several times. In addition to the SOEP Core, which is the main study, the SOEP Innovation Sample (SOEP-IS) has been running annually since 2011 (Richter and Schupp, 2015). While the SOEP Core and the SOEP-IS implement different questionnaires in separate samples, both include the same set of standard questions asked every year, for example questions on self-employment, and are therefore directly comparable. The SOEP Core mostly repeats the same questions in regular intervals (annually or less often, depending on the question) to allow longitudinal analysis. It covered about 30,000 respondents in 2020. In contrast, the SOEP-IS allows researchers to suggest their own set of questions (innovation module) to be administered in a smaller sample (about 5000 respondents in 2019). A SOEP survey committee chooses which proposed innovation modules will be included in the panel study. All SOEP data are released to all interested researchers worldwide (after an embargo period of usually one year for the SOEP-IS). The panel structure of the surveys allows connecting responses of the same respondent to various questions across different survey waves. The SOEP Core has frequently been used in entrepreneurship research (e.g., Nikolova, 2019; Caliendo et al., 2022; Sorgner and Wyrwich, 2022). More recently, entrepreneurship researchers have also introduced their own innovation modules in the SOEP-IS (e.g., Fossen and Neyse, 2024). Similar opportunities for researchers to introduce their own questionnaire modules within general population panel surveys are offered by the "Longitudinal Internet Studies for the Social Sciences" (LISS, see Scherpenzeel, 2011) in the Netherlands and the "Understanding Society Innovation Panel" in the UK (University of Essex, 2023), for example.

A subsample of the 2019 SOEP-IS included an innovation module on the topic of digitalization and AI (Fedorets et al., 2022). The 2020 SOEP Core repeated a subset of the questions in the larger sample. For our analysis of these data, we keep only currently working individuals (employees, solo self-employed individuals and employers) in our samples. The remaining sample from the SOEP Core contains 13,984-15,127 paid employees depending on the question (the variation is due to item non-response), 1053-1098 solo self-employed individuals, and 759-792 employers. The SOEP-IS digitalization module includes 676-698 paid employees, 36-39 solo self-employed, and 24-28 employers. Some results from the digitalization modules of the SOEP have been published (Giering et al., 2021; Giuntella et al., 2023), but these papers do not distinguish between entrepreneurs and paid employees. We use the larger and more recent 2020 SOEP Core sample to analyze the questions contained in both surveys and supplement the analysis with additional questions only available in the 2019 SOEP-IS.

The SOEP-IS inquired about AI use at work using one direct question and multiple indirect questions. The direct question asked for the use of "artificial intelligence or machine learning" explicitly. About 18% of the paid employees and employers and 14% of the solo self-employed answered that they work with these technologies (Figure 17). However, it is possible that respondents are not aware that they use AI because they are unsure what AI means. Therefore, the indirect questions asked for the use of more specific digital systems at work that are AI tools, but without mentioning the term AI: speech recognition, image recognition, text recognition, and automated answers to questions about specialized knowledge. Giering et al. (2021) observe that 45% of the respondents (pooling employees, solo self-employed and employers) answered that they used at least one of these digital tools weekly, and 37% even daily. The seemingly contradictory finding that these shares are much larger than the share of affirmative responses to the direct question on AI use suggests that many respondents are not aware that these digital tools, which are embedded in frequently used office software, are AI tools. For future research, this implies that indirect questions for AI use may lead to more accurate answers than direct questions; an alternative may be to explain the definition of AI to respondents before asking them about their usage of AI.

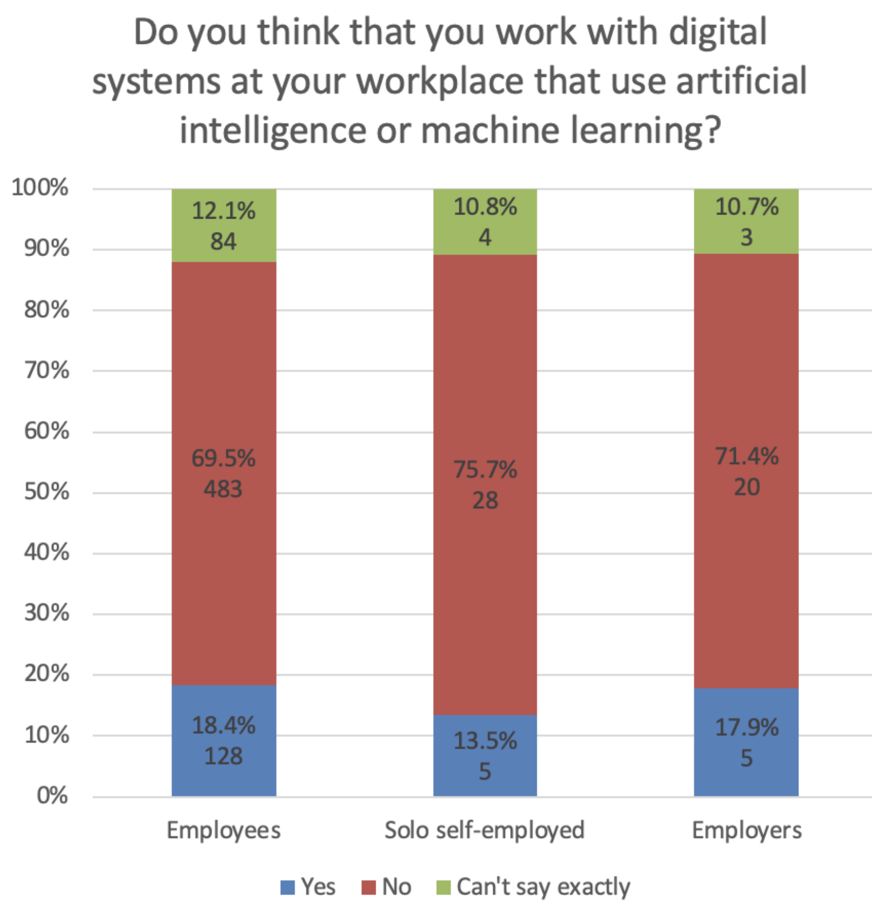


Figure 17: Direct question on AI use at work

Notes: The bars show the shares of answers given by respondents within the three occupations. Observation numbers are provided below the shares. Source: Own calculations based on the 2019 SOEP-IS.

Figure 18 shows the results from the indirect questions on AI use at work (first four questions) and, for comparison, one question on the use of a digital system that is not necessarily AI (processing information and data), based on the larger 2020 SOEP Core. Respondents were asked to indicate how often they work with each type of digital system on a scale from 1 (never) to 5 (multiple times a day). The figure shows mean values by occupation with 95% confidence intervals. The results indicate that most respondents never use each of the four AI systems, whereas more use digital processing of information and data. Comparing the occupational groups, one can see that employers use all the digital technologies, including the AI technologies, significantly more often than the other groups. Following behind the employers, solo self-employed individuals use AI tools processing language significantly more often than paid employees, whereas employees use non-AI digital systems processing information and data more often than the solo self-employed do.

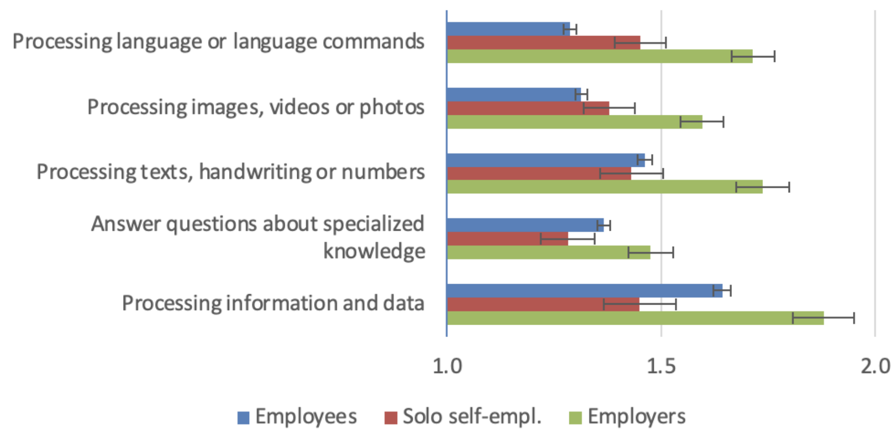


Figure 18: Frequency of work with digital systems

Notes: Respondents were asked to indicate how often they work with each type of digital system. The first four digital systems are AI systems, although they were not presented to respondents using this term. Scale: 1=Never, 2=Seldom, 3=Weekly, 4=Daily, 5=Multiple times a day. The x-axis is cut off at 2 for better readability. The figure shows mean values by occupation with 95% confidence intervals. Source: Own calculations based on the 2020 SOEP Core.

The SOEP-IS also asked respondents how often they perform work tasks in the same areas themselves. Figure 19 shows the mean responses by occupation and is structured like the previous figure. The first four work tasks are likely performed by AI, whereas the fifth task (processing and evaluating information and data records) could be performed by non-AI digital systems. The figure reveals that respondents perform work tasks that could in principle be automated much more often than the frequency they use digital systems for the same tasks. For example, on average, respondents perform text recognition more than weekly (Figure 19), but they use digital (AI) systems for this task between never and seldom (Figure 18). These results suggest that at the time of the surveys, AI systems were far from fully implemented in areas where it would be technologically feasible, and humans continued to perform these tasks. There are various potential reasons for why actual implementation of AI lags behind its potential uses (Giering et al., 2021). First, AI implementation can be costly and may economically not be viable for many of the tasks. Second, technological obstacles for certain purposes in practice are likely. Third, there may be social, legal and ethical reasons and labor policies hindering AI implementation. The first two barriers to AI implementation will almost certainly become lower in the near future as AI technologies become cheaper and more powerful, whereas the future development of the strength of the third barrier is harder to predict.

Comparing the occupational groups in Figure 19, an interesting result is that paid employees perform two of the tasks suitable for AI (recognizing and processing speech and images) more often than

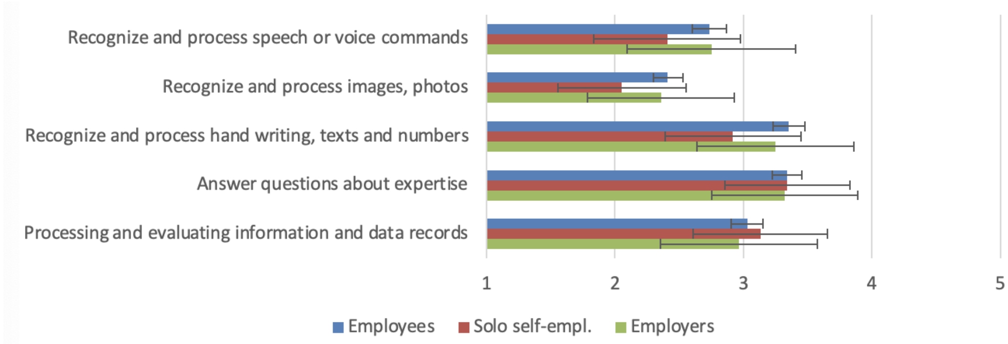


Figure 19: Frequency of work tasks suitable for AI

Notes: Respondents were asked to indicate how often they perform each work task. The first four work tasks are suitable for AI in principle. Scale: 1=Never, 2=Seldom, 3=Weekly, 4=Daily, 5=Multiple times a day. The figure shows mean values by occupation with 95% confidence intervals. Source: Own calculations based on the 2019 SOEP-IS.

the solo self-employed and about as often as employers (although the confidence intervals overlap in the small SOEP-IS sample), in stark contrast to the finding that employees use digital (AI) systems for speech and image recognition less often than both groups of entrepreneurs (Figure 18). This indicates that the gap between potential and actual use of AI is larger for employees than for entrepreneurs. Thus, there is more potential yet to be realized among employees by implementing AI. A potential reason is that employees may have been more resistant to AI implementation so far; entrepreneurs may be more willing to use AI to increase their productivity, as they are the claimants of the profits. The finding may suggest that larger transformations of their work are still ahead for employees, whereas solo self-employed and employers have already em-braced AI more in their work.

So far, we looked at the use of AI tools at work and the frequency of work tasks potentially suitable for AI. For comparison, Figure 20 shows the frequency of work with more traditional digital tools that do not necessarily involve AI, and Figure 21 reports the frequency of a wider spectrum of work tasks. It becomes apparent that digital tools like PCs and laptops are far more often used than AI tools. Only robots are used less often on average. Entrepreneurs (solo self-employed as well as employers) use laptops, notebooks, smartphones and tablets much more often for their work than employees, suggesting that their workplace is more mobile and flexible. Concerning general work tasks, it is interesting that searching, accessing, and providing information is among the most frequent work tasks of all respondents; these tasks are in principle suitable for AI, suggesting that AI will have an impact on most working individuals. Only communication with colleagues and superiors is even more frequent for employees. The differences in the frequencies of work tasks between occupational groups are plausible: Employers are most likely to give orders to other people and to develop work processes. They also most often write or evaluate texts. As writing tasks such as pitching business ideas can be supported by generative AI (Short and Short, 2023), the work of employers may be transformed strongly.

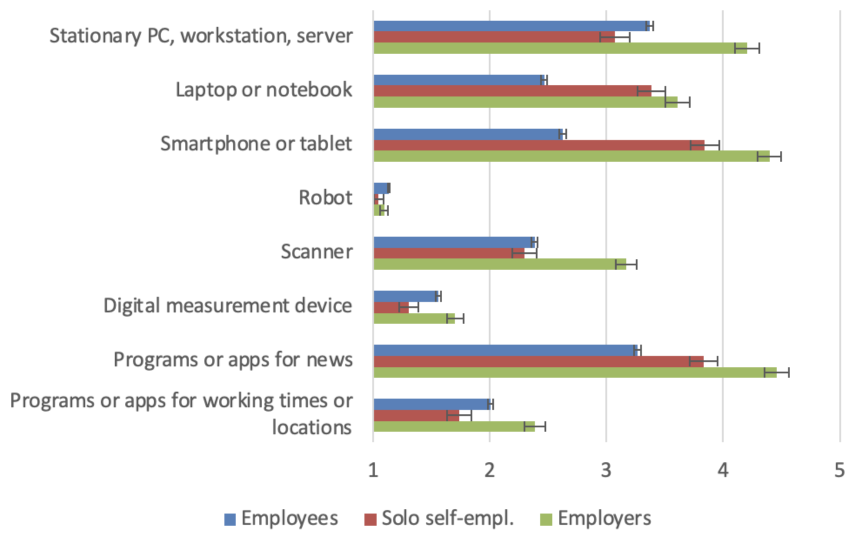


Figure 20: Frequency of work with digital (non-AI) tools
Notes: Respondents were asked to indicate how often they work with each type of digital (non-AI) tool. Scale: 1=Never, 2=Seldom, 3=Weekly, 4=Daily, 5=Multiple times a day. The figure shows means by occupation with 95% confidence intervals. Source: Own calculations based on the 2020 SOEP Core.

An often-voiced concern about AI at work is that AI systems can be used to monitor workers automatically or perform other supervising tasks that workers may experience as interfering with their self-determination and independence at work. Figure 22 shows that the self-reported frequencies of interactions with such automatic digital systems are low; the low implementation rates may be due to resistance of workers and awareness of managers.

The SOEP-IS also asked directly how often respondents felt they self-determine different aspects of their work (Figure 23). On average, respondents indicate that they often or always self-determine how they work, their work pace, and the order of their work. Learning new things and solving unforeseen problems is somewhat less often self-determined. The high self-determination is consistent with the low implementation of AI systems for monitoring and supervision. An interesting insight of Figure 23 is that entrepreneurs (both solo self-employed and employers) report significantly higher levels of self-determination than employees concerning how, how fast, and in which order they work. This is consistent with reports in the literature that autonomy at work is a major motivation for entrepreneurship (e.g., Benz and Frey, 2008), and it suggests that these expectations for indepen-

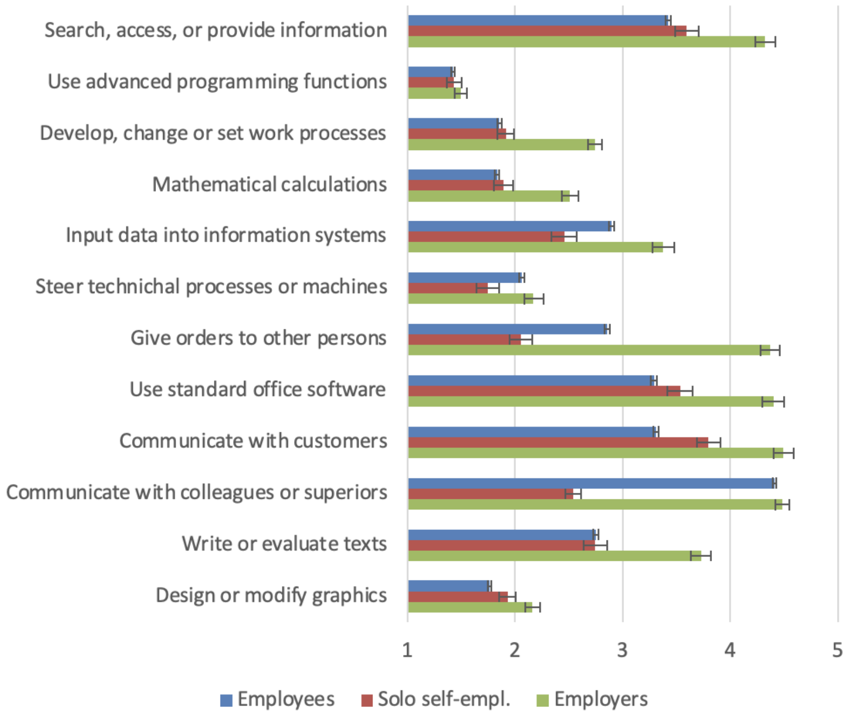


Figure 21: Frequency of general work tasks

Notes: Respondents were asked to indicate how often they perform each general work task. Scale: 1=Never, 2=Seldom, 3=Weekly, 4=Daily, 5=Multiple times a day. The figure shows means by occupation with 95% confidence intervals. Source: Own calculations based on the 2020 SOEP Core.

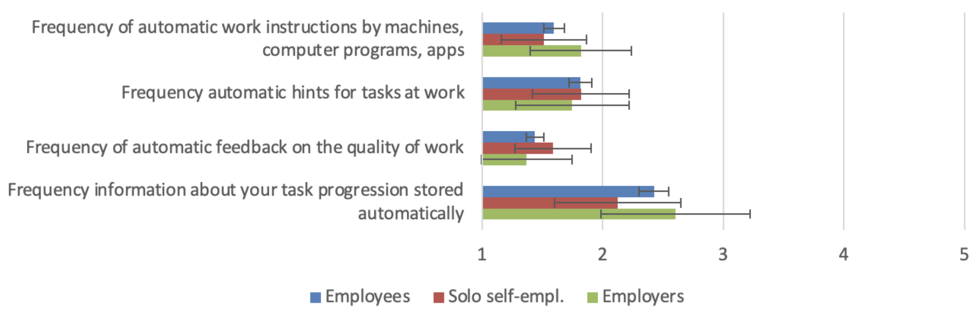


Figure 22: Frequency of automatic interactions

Notes: Respondents were asked to indicate how often they interact with each automatic digital system. Scale: 1=Never, 2=Seldom, 3=Weekly, 4=Daily, 5=Multiple times a day. The figure shows means by occupation with 95% confidence intervals. Source: Own calculations based on the 2019 SOEP-IS.

dence are indeed met in entrepreneurship. Concerning learning new things and solving unforeseen problems, the difference between entrepreneurs and employees is smaller, which seems plausible as unforeseen problems (and the need to learn new things to deal with these problems) often arise exogenously for individuals in both occupations.

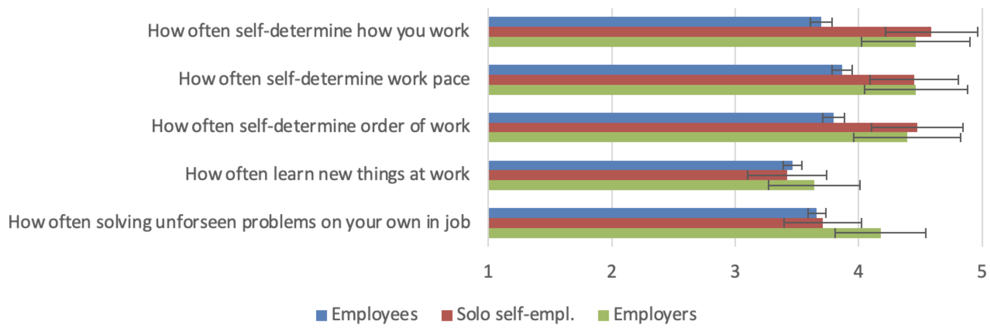


Figure 23: Self-determination at work

Notes: Respondents were asked to indicate how often they self-determine different aspects of their work. Scale: 1=Never, 2=Rarely, 3=Sometimes, 4=Often, 5=Always. The figure shows means by occupation with 95% confidence intervals. Source: Own calculations based on the 2019 SOEP-IS.

Although we find that respondents feel they currently work in a self-determined way and that implementation rates of AI are low, people might be concerned that this could change, especially given the public debate about potential displacement of workers due to AI in the near future (Frey and Osborne, 2017). Are respondents worried about how technological change might impact their work? The 2019 SOEP-IS asked respondents about various concerns they might have, and answers were given on a 3-point scale (1=no worries, 2=some worries, 3=big worries). Figure 24 shows that worries about technological progress were low in Germany at the time of the survey, with average responses closer to "no worries" than "some worries". The solo self-employed are significantly less concerned than paid employees that their professional and personal lives are not well aligned. This is consistent with reports in the literature that the self-employed enjoy larger flexibility concerning the location and timing of their work than employees, which is a major motivation for self-employment, especially among women due to child-care responsibilities (e.g., Georgellis and Wall, 2005).

Figure 24: Worries about technological change

Notes: Respondents were asked about different concerns they might have relating to technological progress. Scale: 1=No worries, 2=Some worries, 3=Big worries. The x-axis is cut off at 2 for better readability. The figure shows means by occupation with 95% confidence intervals. Source: Own calculations based on the 2019 SOEP-IS.

3.6 AI and the future of work: indirect impacts on entrepreneurship

3.6.1 How does AI affect labor markets?

There is widespread concern that AI will accelerate automation and displace workers from their jobs, potentially leading to a large increase in unemployment or underemployment. This concern has manifested in union-led protests and strikes seeking to protect workers from automation through AI; the strike of Hollywood writers in 2023 is an example. Unlike prior waves of automation technologies, AI has the potential to automate many tasks performed in jobs that require high skills previously presumed to be out of reach for machines, such as creativity. However, after initial studies suggested that automation of a high percentage of jobs will be technologically feasible in the near future (Frey and Osborne, 2017), subsequent studies emphasized that only certain tasks within most occupations can be automated (Arntz et al., 2017; Dengler and Matthes, 2018), likely leading to a reorganization of many occupations rather than their complete displacement (Brynjolfsson et al., 2018).

At the same time, researchers have emphasized that waves of automation due to technological innovation not only substitute for labor, but also complement labor. These complementarities may increase productivity of labor and output, which may in turn lead to higher demand for labor and raise earnings (Autor, 2015). Like other technologies, AI also leads to the creation of new tasks and new jobs (Acemoglu and Restrepo, 2018a), so the net effect on employment is ambiguous.

Lane et al. (2023) document results from a survey of more than 5000 workers and more than 2000 firms in several developed countries. More than two thirds of employers report that AI has already automated tasks that workers used to do, and around half of them that AI has created tasks that were not previously done by workers. Acemoglu et al. (2022) analyze online job vacancies in the United States from 2010-2018 collected by Burning Glass. They report that AI-related vacancies have been increasing strongly, and that establishments adopting AI reduce hiring in non-AI positions at the same time.

An understanding of the impacts of AI on the labor market is important to derive potential implications for entrepreneurship. If AI leads to increased unemployment and underemployment, this may induce displaced workers who struggle to find employment to become self-employed as a last resort, which is also called *necessity entrepreneurship* (Congregado et al., 2012; Boeri et al., 2020; Fairlie and Fossen, 2020; Henley, 2021). In the Delphi study by Van Gelderen et al. (2021), a minority of entrepreneurship experts expressed the belief that job loss due to AI automation will force many individuals to become necessity entrepreneurs by 2030. On the other hand, AI leads to potentially large opportunities for entrepreneurs who use new technologies for process and product innovation. For example, relating to an earlier technological invention, Shane (2000) shows how the introduction of 3D printing technology created opportunities for growth-oriented entrepreneurship ranging from printing out organs to applications in the construction sector. We will return to implications for entrepreneurship in the next subsection after developing a better understanding of general labor market effects of AI.

The impact of AI is likely to vary widely across different occupations, depending on whether substitution or complementarities between AI and human labor dominate, and to what extent occupations can be reorganized to incorporate new tasks. Fossen and Sorgner (2019) distinguish between *destructive AI* from the perspective of workers, threatening to displace workers, and *transformative AI*, making workers more productive in reorganized jobs if workers are willing and able to adapt. Depending on how much occupations are affected by transformative or destructive AI or both, the space of occupations can be divided into four segments: rising star occupations are those affected by transformative, but not destructive AI, potentially leading to increased earnings; occupations in machine terrain are affected by both destructive and transformative AI, making it crucial for workers to keep up with the changes at work through training; occupations in human terrain are not affected by either form of AI; and collapsing occupations are becoming obsolete due to complete automation through destructive AI.[8]

Acemoglu and Restrepo (2018b, 2019) model *automation* as capital taking over tasks previously performed by human labor. If it is cheaper to perform tasks by employing capital rather than labor, automation will result in displacement effects, that is, a decrease in wages and employment (see also Aghion et al., 2019). However, the displacement effect may be mitigated or even over-compensated due to productivity gains brought by automation, increasing the demand for workers performing existing tasks that cannot be automated, and due to the creation of new tasks for human workers. This may then lead to a labor-reinstating effect[9]. In the following, we show more formally how automation technology enabled by AI may change labor demand and wages, based on Acemoglu and Restrepo (2018a)[10]. Aggregate output Y can be represented as

[8] Casas and Román (2023) use a similar classification of occupations.

[9] Gries and Naudé (2022) build on this approach, but argue that the services provided specifically by AI, in contrast to other automation technologies, can be modeled as abilities rather than tasks or skills.

[10] We represent the outline of the formal model as in Fossen and Sorgner (2022).

$$\ln Y = \int_{N-1}^{N} \ln y(x)\, dx, \tag{9}$$

where $y(x)$ denotes the output of task x. Tasks are indexed by x and normalized to lie between $N-1$ and N. Each task can be produced by human labor, $l(x)$, or by machines, $m(x)$, if it can be automated. In particular,

$$y(x) = \begin{cases} \gamma_L(x)l(x) + \gamma_M(x) & \text{if } x \in [N-1, I] \\ \gamma_L(x)l(x) & \text{if } x \in (I, N], \end{cases} \tag{10}$$

where $\gamma_L(x)$ is the productivity of labor in task x and $\gamma_M(x)$ is the productivity of capital in automated tasks. Parameter I denotes the range of tasks that can be automated. Tasks with $x \leq I$ are automated whereas tasks with $x > I$ are not. Automation is modeled as the expansion of the range of tasks that can be performed by machines, that is, as an increase in I. Acemoglu and Restrepo (2018a) derive that, as the range of automated tasks increases, the wage W will change as follows:

$$\frac{d \ln W}{dI} = \frac{d \ln (N-I)}{dI} + \frac{d \ln (Y/L)}{dI}, \tag{11}$$

where the first term on the right-hand side of the equation describes the displacement effect due to automation, which is always negative, and the second term is a productivity effect, which is always positive. It follows from Eq. 11 that automation technologies that are labor displacing but besides that not very productive ("so-so" technologies) will reduce wages and labor demand. In turn, brilliant automation technologies that substantially improve labor productivity can increase wages and labor demand.

$$\frac{d \ln W}{dN} = \left[\ln \left(\frac{R}{\gamma_M(N-1)} \right) - \ln \left(\frac{W}{\gamma_L(N)} \right) \right] + \frac{1}{N-1} > 0. \tag{12}$$

R is the cost (or rental rate) of machines. Both terms on the right-hand side are positive. The first term (in square brackets) is a productivity effect and the second the reinstatement effect. Thus, technologies that expand the total amount of tasks with labor at a comparative advantage will increase wages and labor demand. In addition, deepening of automation, that is, displacement of previous automation technologies by more productive ones, will not displace workers (under the assumption of no change in I), but will tend to increase wages and labor demand. Finally, automation through AI may also lead to capital accumulation and thereby raise the overall demand for human labor. In sum, the overall theoretical impact of AI automation technologies on wages and labor demand are ambiguous, as they depend on whether the labor displacement effect or the productivity and labor-reinstating effects are stronger, making this an empirical question.

Therefore, Fossen and Sorgner (2022) set out to analyze empirically whether exposure of occupations to AI leads to changes in wages and employment. They use microdata representative for the US population for 2011-2018 from the Current Population Survey and its Annual Social and Economic supplement as provided by the U.S. Census Bureau. Based on these data, they estimate the associations between exposure measures of occupations to AI with two sets of labor market outcomes at the individual level: year-to-year changes in individual wages and probabilities of transitions out of the current paid employment. In particular, they consider transitions into non-employment (unemployment or non-participation), switching between different wage and salary jobs, and transitions into self-employment.

Several *measures* of impacts of digitalization and AI on occupations have been developed in the literature. The computerization probabilities (CP) of occupations estimated by Frey and Osborne (2017) and the suitability of occupations for machine learning (SML) scores provided by Brynjolfsson and Mitchell (2017a) and Brynjolfsson et al. (2018) are based on expert judgements. The AI Occupational Exposure (AIOE) scores presented by Felten et al. (2018, 2021) instead utilize past progress in nine categories of AI as tracked by the Electronic Frontier Foundation. These progress measures are then linked to abilities used in occupations as described in the occupation database O*NET provided by the U.S. Department of Labor. Webb (2020) measures occupational exposure to AI, to non-AI software, and to robots by assessing textual similarities between descriptions of patents and job tasks in O*NET. Paolillo et al. (2022) develop another measure of exposure of jobs to potentially AI-enhanced robots, the automation risk index. They compare robotic abilities as described in the European H2020 Robotics Multi-Annual Roadmap to human abilities as defined in O*NET. While this approach is similar to that of Frey and Osborne (2017), Paolillo et al. (2022) use a finer matching based on more abilities.

Felten et al. (2023) and Eloundou et al. (2023) focus on exposure of occupations to generative language models such as GPTs specifically. The resulting ability-level score by Felten et al. (2023)

measuring exposure to LLMs is very highly correlated (correlation coefficient: 0.979) with the more general AIOE score provided in Felten et al. (2021). Felten et al. (2023) report that telemarketers followed by post-secondary humanities teachers are the occupations most exposed to LLMs, and the most exposed industries are legal services followed by the securities, commodities, and investment industries.

Fossen and Sorgner (2022) use four of the measures of occupational exposure to AI: the CP, the SML as well as the within-occupation standard deviation of these SML scores, and the AIOE. In their regressions, they control for a wide range of individual and occupation-level characteristics that may be correlated with AI exposure and may also affect the labor market outcomes, such as education, age, gender and prior income as well as offshoreability of the occupation (Blinder and Krueger, 2013). They also account for year, industry, and occupation category effects. The results show that effects of AI exposure were already detectable in the US labor market in the period under analysis. A larger CP or SML score of an occupation is associated with slower individual wage growth as well as a higher probability of entry into non-employment. In contrast, a higher AIOE score is associated with stronger wage growth and a lower probability of becoming non-employed. These results are consistent with the interpretation that CP and SML capture destructive AI, whereas the AIOE scores capture transformative AI[11].

Fossen et al. (2022) extend the individual-level analysis of wage changes using the alternative measures of exposure of occupations to AI, to non-AI software, and to robots, as provided by Webb (2020). Fossen et al. (2022) find that occupational exposure to AI is associated with accelerated wage growth, whereas non-AI software and robots are associated with slower wage growth. Thus, Webb (2020) measure of occupational exposure to AI may capture transformative AI, like the AIOE scores. These econometric results contrast with simulation results reported by Webb (2020). For his simulations, Webb (2020) assumes that the relationship of wage changes with AI exposure will be negative, like the relationship between exposure to robots and software was in the past. Based on this assumption, he predicts that AI exposure will decrease wages at the 90th percentile relative to the 10th percentile in the future. However, this assumption may be questionable because Fossen et al. (2022) find that the association of wage changes with AI exposure is positive and thus has the opposite sign in comparison to the association of wage changes with exposure to robots and

[11] For example, physicians and surgeons have a very low CP score, as these occupations will not be completely automized in the near future. At the same time, these occupations have a high AIOE score, as physicians and surgeons increasingly use new AI tools for diagnostics and other tasks, transforming their occupation.

software. Further research is necessary, especially because AI implementation in firms was still at an early stage in the data analyzed so far, and impacts are expected to become much stronger. Özgul et al. (2024) use the same measure of AI exposure by Webb (2020) to estimate impacts on the German labor market. They report heterogeneous effects by skill level and occupations. Future research should continue to consider effect heterogeneity; Fossen and Sorgner (2022) and Fossen et al. (2022) provide initial evidence of differential impacts by education, gender, age, industry, and other characteristics.

3.6.2 Opportunity and necessity entrepreneurship

Given the widespread prediction that AI will fundamentally impact the labor market and the initial evidence reviewed above supporting this claim, it is straightforward to expect that the transformation of the labor market will also impact entrepreneurship, as an individual's decision to engage in entrepreneurship in part depends on the valuation of alternative employment options. A prerequisite for a meaningful empirical analysis is a measurable definition of different types of entrepreneurship, which we discuss in this section.

Entrepreneurship is a multifaceted concept, and definitions and measurements in empirical studies should take into account the pronounced heterogeneity among entrepreneurs (Congregado, 2007). Researchers using general population data often use self-employment as an observable, broad and inclusive measure of entrepreneurship, where the self-employed include nonemployers as well as employers and those working alone as well as those working with partners in a partner-ship business (e.g., Fairlie, 2013; Congregado et al., 2024). Self-employment incorporates key elements of most conceptual definitions of entrepreneurship, such as more autonomy and a larger income risk in comparison to wage and salary employment. However, not all self-employed innovate, and many self-employed do not have any intention to grow their business (Hurst and Pugsley, 2011). Entrepreneurs who innovate and grow have a larger impact on the economy, as they generate positive external effects through innovation and may create jobs, and for this reason, they are often the focus of academic interest and public policy.

This motivates the need to distinguish between types of entrepreneurs within the self-employed. Most researchers use dichotomous distinctions and differentiate between two types of entrepreneurs with different labels, which to a large extend capture similar concepts, such as opportunity versus necessity entrepreneurship (e.g., Van der Zwan et al., 2016; Fairlie and Fossen, 2020) or pull- versus push-entrepreneurship (e.g., Storey, 1991; Ritsilä and Tervo, 2002). Opportunity entrepreneurs become entrepreneurs because they see opportunities superior to paid employment, they are pulled into entrepreneurship by conditions they see as favorable. Necessity entrepreneurs are those who become self-employed due to a lack of alternatives. They cannot find acceptable paid employment, so they are pushed into self-employment as a last resort to make a living[12].

Different empirical strategies have been suggested in the literature to operationalize the distinction. Although some scholars call for moving beyond dichotomous classifications (Caliendo and Kritikos, 2019; Dencker et al., 2021) and a more fine-grained view of necessity entrepreneurship (O'Donnell et al., 2024), such approaches are currently often limited by data availability. The Global Entrepreneurship Monitor, which has been used frequently in entrepreneurship research (Bosma, 2013), asks survey respondents directly for their motivation to become entrepreneurs. A disadvantage is that the answer is subjective and might be influenced by the success of the business after its launch. Fairlie and Fossen (2020) suggest classifying the self-employed based on their labor market status prior to entry into self-employment. If they were paid employees (or not in the labor force, for example, in education) before, they are labeled opportunity entrepreneurs, and if they come out of unemployment, they are labeled necessity entrepreneurs[13]. Those who are registered as unemployed are actively seeking employment by definition, so if they become self-employed, this is likely due to a lack of alternatives. In contrast, somebody who quits a paid job to become self-employed has an alternative, but likely perceives a better opportunity in entrepreneurship. This operational definition of necessity versus opportunity entrepreneurs has the advantages that it is consistent with standard entrepreneurship models, that it is objective, that it is determined before the launch of the business, and that it is readily observable in many micro datasets, either based on repeated observation of an individual in panel data or on retrospective questions (Fairlie and Fossen, 2020).

Of course, this is not the only possible operationalization of the different types of entrepreneurs, and other approaches are also useful depending on the research question and the data at hand. Some papers distinguish between nonemployers and employers, arguing that employers have shown growth ambition and have a larger impact on the economy by creating at least one job for others (e.g., Congregado et al., 2012; Coad et al., 2017; Fairlie and Miranda, 2017; Caliendo et al., 2022; Nikolova et al., 2023). Other studies use the legal form of the business entrepreneurs are running as a separating device. Levine and Rubinstein (2017) show that the self-employed who own an incorporated

[12] This has also been called a "refugee effect" (Thurik et al., 2008).

[13] Block and Sandner (2009), Block and Wagner (2010) and Fossen and Buettner (2013) use similar approaches.

business in the United States are more likely to innovate than those who own an unincorporated business, and Åstebro and Tåg (2017) find that entrepreneurs with an incorporated business in Sweden are more likely to create jobs. The classification based on incorporation status is useful within a certain country and limited period, but less appropriate when making comparisons across countries or over longer periods, because incentives to incorporate differ widely across countries and often change over time, for example due to reforms of tax legislation. Within single countries, Fairlie and Fossen (2020) document a large correlation between their classification of opportunity versus necessity entrepreneurs based on the prior employment status with classifications based on being an employer or incorporation status. Moreover, Fossen (2021) finds that the entry rate into unincorporated self-employment increased during the Great Recession in the United States, when unemployment was high, but not into incorporated self-employment, providing further evidence of the association of unincorporated and incorporated self-employment with necessity and opportunity entrepreneurship, respectively.

3.6.3 Occupational impacts of AI and transitions into entrepreneurship

Fossen and Sorgner (2021) investigate empirically whether early impacts of AI on occupations are already influencing transitions into entrepreneurship. The study highlights the importance of nuanced analysis distinguishing between different types of entrepreneurs as well as between different AI impacts on occupations, i.e., destructive and transformative. The paper separates between the self-employed with incorporated and with unincorporated businesses, following Levine and Rubinstein (2017). Fossen and Sorgner (2021) hypothesize opposite effects of new digitalization and AI technologies on unincorporated and incorporated entrepreneurship. They predict that employees working in occupations that are exposed to destructive digitalization (from the perspective of workers) are more likely to transition to unincorporated entrepreneurship, because they may become necessity entrepreneurs as an adjustment mechanism when they face the risk of displacement from their paid jobs. However, these individuals at risk of losing their jobs due to AI automation may lack skills or finance necessary to start an incorporated business. Different effects are expected in occupations that are affected by transformative digitalization, where employees become more productive due to complementarities of human labor with new AI technologies. As employees in these occupations potentially experience productivity and wage growth, they will be less likely to become self-employed with small-scale, unincorporated businesses because their opportunity costs of doing so increase. At the same time, employees exposed to transformative digitalization in their occupation may discover new opportunities to start growth-oriented, incorporated businesses making use of advances in AI in their area of work.

The empirical analysis is based on individual-level rotating panel data from the Current Population Survey provided by the U.S. Census Bureau from January 2011 to October 2018, thus, referring to a period of early implementation of AI. The authors use multinomial logit models to estimate the choice of paid employees to transition to incorporated or unincorporated entrepreneurship from one month to the next; alternative outcomes are becoming unemployed or switching between different wage occupations, while remaining in the current job is the base category. The main explanatory variable is one of the measures of exposure of occupations to digitalization and AI that we introduced in Section 3.6.1: CP (Frey and Osborne, 2017), the mean and within-occupation standard deviation of SML (Brynjolfsson and Mitchell, 2017a; Brynjolfsson et al., 2018) and AIOE (Felten et al., 2018, 2021).

The empirical results presented by Fossen and Sorgner (2021) support the hypotheses, although only for certain groups of individuals in some cases. Employees in occupations that are subject to destructive digitalization and AI are more likely to transition to self-employment with unincorporated businesses, suggesting that they have an increased probability of becoming necessity entrepreneurs as they face the risk of displacement from their jobs. Interestingly, this effect is found most persistently for high-skilled individuals, not for low-skilled individuals, who are more likely to become unemployed instead. The option of starting a business as an adjustment mechanism in the face of destructive impacts of AI on one's job does not seem to be viable for many low-skilled workers. The effect of destructive digitalization on transitioning to incorporated entrepreneurship is negative for low-skilled individuals, who may lack the skills necessary to start an incorporated business when their work experience is devalued due to automation. The effects of transformative digitalization of one's occupation go in the opposite directions. Transformative new digital and AI technologies decrease the probability of entry into unincorporated self-employment, plausibly reflecting higher opportunity costs of self-employment due to increased productivity and wages in paid employment. However, for employees with a high-school degree and older employees, an increased probability of starting an incorporated business is detected in occupations exposed to transformative digitalization, suggesting that advances in AI indeed lead to new business opportunities for growth-oriented entrepreneurship. Analyzing data from 31 European countries, Bachmann et al. (2024) find that advances in AI are correlated with higher probabilities of transitioning from paid employment to solo self-employment and vice versa, thus, higher worker mobility. Future research might consider utilizing measurements of AI impacts at the task or work-activity level rather than at the occupation level for a more fine-grained analysis.

3.6.4 AI and entrepreneurship in regional labor markets

Given the evidence that AI impacts on occupations have started affecting labor markets in general and transitions into entrepreneurship in particular, one can expect that regions with a large share of workers employed in occupations exposed to AI will see stronger effects on their local labor market and entrepreneurial activity than less exposed regions. This may lead to new regional disparities and high pressure on some regions to adjust. Fossen et al. (2022) analyze heterogeneous impacts of AI on entrepreneurship in states and counties within the United States due to the local occupational structures. The analysis is based on individual-level data from the American Community Survey provided by the U.S. Census Bureau, which is well suited for regional analysis. The study documents pronounced regional differences in the impact of digitalization and AI on incorporated and unincorporated entrepreneurship. Entrepreneurs in metropolitan areas are most likely to benefit from advances in AI (for example, in the Bay Area in California, extending via Sacramento to Reno in Nevada), but heterogeneity between different urban regions is large. The findings suggest that AI exposure through the regional occupational structure is an important element of regional entrepreneurial ecosystems, which we will discuss in detail in the next section.

Most impact measures of digitalization and AI on occupations (e.g., CP, SML, AIOE, and Webb's measure) were developed for the United States and reflect tasks performed and skills used in occupations there. Naturally, there is high interest in analyzing AI impacts on labor markets and entrepreneurship in other countries as well. Some studies use the same AI impact measures for the same occupations in other countries, employing crosswalks of occupation codes if necessary (e.g., Casas and Román, 2023; Bachmann et al., 2024). This approach will capture differences in AI impacts across countries as far as they are determined by differences in the occupational structure, i.e., the prevalence of certain occupations in different countries. However, a limitation of this approach is that the same occupation may be comprised of different tasks in different countries and require different skills (Arntz et al., 2017; Carbonero et al., 2023b). For example, an important part of the occupation of craftspeople in Germany is teaching, because they instruct apprentices, whereas teaching crafts is performed by teachers in schools in other countries. Differences in contents of occupations are particularly stark between developed and developing countries. For instance, a large share of a farmer's work in a developing country may be manual field labor, whereas a farmer's workday in the United States is filled to a larger extend with accounting work. Therefore, new AI technologies may have varying effects on occupations in different countries, as AI tools may be suitable for tasks that are important in an occupation in one country, but not in another country. A potential solution for research could be to develop new AI impact scores for other countries, but this is often prohibitively expensive and difficult, especially in resource-constrained countries. Another approach consists of regressing AI impact scores in the country that they were developed for on job characteristics or characteristics of individual workers, and then using the estimated model to predict the AI impact scores in a different country, based on characteristics observed there (Arntz et al., 2017). However, this approach still requires a crosswalk of occupations, and the precision achievable with this method seems limited.

Therefore, Carbonero et al. (2023b) develop an alternative method to translate occupational AI impact scores developed for one country to another country. They suggest employing individual-level surveys of workers' skills available in both countries, such as the World Bank's Skills Measurement Program (STEP) for developing countries or the Survey of Adult Skills (PIAAC) for OECD countries. A crucial step in this approach is to find semantic similarities between the textual descriptions of work activities in the original country, for which AI impact scores are available, and the textual descriptions of the workers' skills elicited in the survey. Some studies rely on manual linking. For example, Georgieff and Hyee (2022) manually match abilities in job descriptions in O*NET for the United States to skills in PIAAC to use the AIOE scores developed for the US by Felten et al. (2018) for 23 other OECD countries. However, a manual assessment of similarities is costly and subjective. Carbonero et al. (2023b) attempt to overcome this limitation by applying the automated semantic textual similarity matching technique SBERT (Reimers and Gurevych, 2019b). The constructed matrix of similarity then allows reweighting the AI impact scores based on the skills that workers in a particular occupation use in different countries, as observed in the workers' skills survey. Carbonero et al. (2023b) illustrate the method by translating the SML scores (Brynjolfsson and Mitchell, 2017a; Brynjolfsson et al., 2018) from the United States to the developing countries Lao PDR and Viet Nam. Brynjolfsson et al. (2024) are also working on a project named WorldSML to present their SML scores for other countries, it will be very interesting to compare the approaches and results.

3.7 AI and entrepreneurial ecosystems

The concept of *entrepreneurial ecosystems* (EE) is an attempt to describe the emergence and evolution of entrepreneurship within a specific context in a systemic way (Spigel, 2017; Stam and Spigel, 2018; Wurth et al., 2022, 2023). At the core of the EE concept is the premise that entrepreneurship is undertaken by individuals based on the structure of incentives, such as payoffs. Individual entrepreneurial behavior is further affected by a complex institutional framework that might facilitate or hinder entrepreneurship. This framework consists of multiple elements, such as formal regulations, physical infrastructure, talent supply, networks, finance, or entrepreneurship culture. Moreover, all elements of EE are mutually interdependent, so that it is not merely entrepreneurship that is affected by the environment, but the context itself can change in response to evolving entrepreneurial activities. The output of a successful EE is productive entrepreneurship that fosters economic development and growth. In this sense, innovative and growth-oriented startups, e.g., 'gazelles' or 'unicorns', are frequently considered as a proxy for productive entrepreneurship.

While EE can be defined at different regional levels (Malecki, 2018), such as nations (Acs et al., 2015), regions within countries (Stam, 2015), or cities (Audretsch and Belitski, 2017), the role of geography in EE has frequently been emphasized in various studies (Stam and Spigel, 2018). The main rationale for studying EE at the local level is that at least some of the EE elements are determined at a rather narrowly defined regional level (often within countries). While many formal institutions can be defined at a national level (e.g., labor market regulations) or a supranational level (e.g., the General Data Protection Regulation that is defined at the EU level), entrepreneurial culture or natural conditions for doing business (e.g., risk of natural disasters, proximity to natural resources) tend to be local.

There are several ways in which AI may affect EE, which can broadly be divided into three categories: effects of AI on EE elements, effects of AI on EE outputs, and effects of AI on EE processes, links, and feedback mechanisms (see Table 13 for an overview). First, new digital technologies, such as AI, may *affect specific elements of EE*. For instance, AI can alter the structure of incentives and change the opportunity cost of setting up a business by impacting local job markets (Fossen et al., 2022) or reducing the economic cost of running a business (Goldfarb and Tucker, 2019). In addition, universities with departments specialized in AI may become a new source of knowledge spillovers that might affect the nature of entrepreneurial opportunities in an EE. One concern is, however, that, currently, Big Tech companies dominate AI research (including basic research in computer science),

Table 13: Effects of AI on entrepreneurial ecosystems

Effects on EE elements	Effects on EE outputs	Effects on EE processes, links, and feedback mechanisms
• Effects on existing EE elements by modifying them and/or changing their relative importance within an EE; • Creation of new EE elements.	• Effects on the quality of entrepreneurship, for instance, in terms of productivity, job creation, and effects on regional economic performance (e.g., economic development, economic growth, economic and social inequalities); • Effects on novel or rare types of entrepreneurship, such as unicorns and digital startups.	• Effects on links between EE elements and/or between EE elements and EE outputs; • Effects on long-term persistence of regional entrepreneurship by (potentially) reducing the role of geography for certain types of entrepreneurship.

such that there are worries about the viability of publicly funded research in this field (Jacobides et al., 2021). Colombelli et al. (2024) provide evidence that the presence of digital knowledge spillovers in a NUTS3[14] region, measured by the share of graduates in Information, Communication and Technology (ICT) topics and the stock of ICT patent applications per inhabitant, as well as the regional digital skill endowment are positively associated with the number of digital innovative start-ups in a region. In addition, AI may also reduce the importance of some previously relevant EE elements, such as professional services, by promoting disintermediation and reducing the power of intermediaries in value chains (Autio et al., 2018).

Moreover, AI may lead to the *emergence of new EE elements*, such as access to digital infrastructure, big data, cloud providers and cybersecurity expertise by entrepreneurs. These new EE elements may play a key role in competitiveness and sustainability of an EE. For instance, access to larger databases reduces the cost of training AI models and increases the accuracy of predictions, thereby facilitating data-driven decision-making in startups. AI ecosystems may themselves *represent a novel element of an EE* that interacts with other established elements, facilitates them, and is affected by them. The evolution of AI ecosystems varies widely across geographic areas, which is due to differences in institutional, political, academic, and cultural backgrounds. Regarding the latter, for example, while 86% of end users of AI in China generally trust AI-driven decisions, only 45% of European users and

[14]The nomenclature of territorial units for statistics (NUTS) is a geographical system, according to which the territory of the European Union is divided into regions at three hierarchical levels: NUTS 1, 2 and 3, moving from larger to smaller territorial units.

39% of American users do so (Jacobides et al., 2021, 423). Moreover, evolution of AI ecosystems may affect institutional frameworks that regulate their usage and their further development (e.g., the approved EU AI Act will restrict employment of specific types of AI systems).

Second, AI may *affect the output of EE*, which is productive entrepreneurship. It is still an open question in how far AI affects productivity and job creation in startups. Despite the rapid recent developments in AI, the corresponding productivity gains appear rather modest (Furman and Seamans, 2019), with only 11% of companies deploying AI reporting a significant benefit from it (Ransbotham et al., 2020). Brynjolfsson et al. (2019) speculate that this is due to lack of complementary investments into commercialization of innovative business ideas in this field. In terms of job creation, one could consider unicorns as an example. While unicorns are defined by the estimated value of the company, which should be at least 1 billion U.S. $, rather than by its size, there are examples of unicorns that would formally fall within a category of Small and Medium Sized Enterprises (SMEs). Also, AI has facilitated algorithmic management practices on digital platforms (Parent-Rocheleau and Parker, 2022), which accelerated the rise of gig work that would be classified as self-employment (not necessarily entrepreneurship in the Schumpeterian sense). In addition, there is a risk of destructive digital entrepreneurship that may lead to so-called digital dystopias, such as digital platform capitalism, the surveillance state, digital divides, cybercrime (Naudé, 2023), and underground entrepreneurship (Goel et al., 2015), among others.

Third, elements within an EE are highly interdependent (Leendertse et al., 2022), and *AI may affect processes within an EE* by influencing links and feedback mechanisms between EE elements as well as between EE elements and EE outputs. It has often been stressed that well-functioning EE create self-perpetuating effects on future levels of entrepreneurship, that is, currently successful EE are likely to create more productive entrepreneurship in the future. This also implies that it is challenging to promote entrepreneurship in regions with inefficient EE that result in low levels of high-quality entrepreneurship. This effect has been observed in various studies of long-term persistence of regional entrepreneurship (Fritsch et al., 2019; Fritsch and Wyrwich, 2023). A frequently suggested policy implication to foster entrepreneurship in regions with less successful EE is to focus on EE elements that are more persistent over time, such as local entrepreneurship culture, which appears to be one of the main drivers of long-term persistence of entrepreneurship. It is yet to be investigated how far AI will affect the self-perpetuating effect entrepreneurship has had on EE, but some preliminary evidence on the role of regions in AI-adopting behavior of firms suggests that it could potentially be highly disruptive.

For instance, a study of AI technology adoption in more than 380,000 firms in Germany, Austria, and Switzerland found that the pattern of AI adoption is highly clustered, thus, indicating the presence of regional hotspots associated with production of AI knowledge (Dahlke et al., 2024). This is consistent with the findings by McElheran et al. (2024) for the United States reported above. Dahlke et al. (2024) found AI adopters to be highly interconnected within the AI knowledge network, whereby cognitive proximity to the source of AI knowledge appears to play a key role[15]. A surprising finding was that the transmission of AI knowledge is not affected by geographic proximity. Thus, even though the production of knowledge continues to be local, AI knowledge could potentially spill over to very distant locations. This could imply a reduced dependency of AI-adopting startups on geographic proximity to EE, but a strong dependency on cognitive proximity to EE. Although the AI knowledge base is strong in certain places, knowledge spillovers from AI are yet very limited (Cetindamar et al., 2020). This raises important questions concerning barriers for commercialization of AI knowledge and what could be done to facilitate these processes.

While the reduced dependence of AI-adopting business ventures on geographic distance may come as little surprise, it is not straightforward whether AI is likely to have a positive or a negative effect on the overall importance of EE in fostering productive entrepreneurship. This is because AI may increase the importance of some EE elements (e.g., universities as a source of knowledge spillovers and talent) and decrease the importance of other EE elements (e.g., dependence on intermediaries and local social networks), as discussed above. While geographic location played a key role in traditional EE, it is an open question in how far geography matters in an increasingly digitized world. Desai (2019) raises the research question asking if AI will diminish the geographical dimension of entrepreneurship due to reduced reliance of entrepreneurs on local human capital and other resources and local opportunities. Recent literature on digital affordances seems to suggest that the dependence of startups on regional clusters as sources of entrepreneurial opportunities has decreased, since startups are now enabled by digital technologies to access opportunities far beyond regional clusters (Autio et al., 2018). This preliminary evidence calls for re-considering the concept of EE, and particularly, the role of geography in building successful EE.

So far, a main limitation of the EE framework was that it did not systematically account for the role of technological transformation (Song, 2019; Subramaniam et al., 2019; Ferreira et al., 2023). As a

[15]Cognitive proximity refers to the similarity of cooperating firms' knowledge bases.

result, some researchers consider the impact of digital transformation on EE to be so profound and disruptive, such that a new concept of *digital entrepreneurial ecosystems* (DEE) has been put forward to better understand entrepreneurship in the digital world from the perspective of *platform-based ecosystems*. According to (Sussan and Acs, 2017, 63), "the digital entrepreneurial ecosystem is the matching of digital customers (users and agents) on platforms in digital space through the creative use of digital ecosystem governance and business ecosystem management to create matchmaker value and social utility by reducing transactions cost." This definition relies on four key elements that make up a DEE, namely, digital infrastructure governance, digital user citizenship, digital entrepreneurship, and digital marketplace. The digital infrastructure governance refers to a set of rules and technological standards for digital usage that are related to entrepreneurial activities. The digital user citizenship element describes participation of citizens in a digital environment and implies that a DEE relies on the availability and the quality of digital users, for instance, in terms of their digital literacy. Digital marketplaces include, for instance, social network-based businesses, e-commerce, e-health, e-education, and e-government. Lastly, digital entrepreneurship refers to entrepreneurial activities that optimize the utilization and reconfiguration of digital infrastructure in the form of new systems, new platforms, and new networks (Nambisan, 2017; Berger et al., 2021).

This approach has been criticized and revised by Song (2019), who proposed, for instance, to extend the notion of digital entrepreneurship in the DEE framework by Sussan and Acs (2017) to include digital technology entrepreneurship that refers to "entrepreneurial opportunities based on ICT related to IoT, data security, connectivity solutions, cloud platforms, networking software, management solutions, smart home, and so on" (Song, 2019, 575). Overall, the modified DEE framework proposed by Song (2019) includes (1) digital user citizenship; (2) digital technology entrepreneurship; (3) digital multi-sided platforms; and (4) digital infrastructure governance concept (this element remains without modification). It should be noted that the concept of DEE and the frameworks described above are not exclusively related to AI technologies, but they are designed to include a broader set of digital technologies. Moreover, by explicitly referring to digital marketplaces and multi-sided platforms, their focus is more on digital platform economies rather than on digital transformation within a local EE. In sum, the DEE concept seems to complement rather than substitute the concept of EE, while potential overlaps are still to be investigated. For instance, several studies proposed a notion of DEE that seeks to integrate both frameworks more strongly (see, e.g., Du et al., 2018; Elia et al., 2020; Keyhani et al., 2022).

While the literature on DEE is still in its emerging state, a recent study offering a systematic literature review of DEE suggests a DEE typology, thereby recognizing that digital technologies may fulfill different roles in different types of DEE (Bejjani et al., 2023). The authors distinguish between four types of DEEs depending on the degree of autonomy in governance and the degree of collaboration within the ecosystem. For instance, in the "marketplace ecosystem" (low autonomy, high collaboration), the main role of digital technologies consists in mediating interactions, and the ecosystems emerge around multi-sided platforms. In turn, in the "chat room ecosystem" (high autonomy, high collaboration), the role of digital technologies is to facilitate entrepreneurial opportunities around entrepreneurial clusters. This approach appears promising, as it provides a framework, within which similarities and distinct characteristics of EE and DEE can be studied. It could be a useful extension to this approach to focus on the dominant type of digital technologies in each DEE type, as the approach does not explicitly mention the role of AI.

It is still an open question what can be considered as an appropriate output of a DEE. Sussan and Acs (2017) see a sustainable ecosystem as the outcome of the DEE. Recent studies have suggested using unicorns to measure the performance of DEEs, as they usually rely on digital technologies (Torres and Godinho, 2022; Venancio et al., 2023). Torres and Godinho (2022) also show that all elements of DEEs are necessary to produce digitally enabled unicorns, while only a few of them are relevant for a high level of new business formation. This suggests that DEEs play a more important role in fostering high-quality (digital) entrepreneurship rather than in increasing the pure quantity of new businesses.

In general, there is a striking tendency in the above-mentioned literature not to focus on AI specifically, but rather to assume that it is a general-purpose digital technology that has disruptive effects on EE. In their literature review, Mariani et al. (2023) identify only a handful of studies examining AI in relation to entrepreneurial ecosystems. One potential reason behind this significant research gap is that AI entrepreneurial ecosystems represent an *emerging* entrepreneurial ecosystem. While established EEs likely have most key elements materially visible, which makes them more easily measurable using established metrics, emerging EEs are not yet mature enough to display their distinct material elements, so that non-conventional approaches, such as cultural mapping, are needed to correctly identify them (Hannigan et al., 2022). In response to this challenge, Jacobides et al. (2021) provide a deeper insight into the evolution of AI ecosystems that consist of the complex interdependency of multiple actors, including developers, manufacturers and users of AI. By distinguishing among AI enablement, AI production, and AI consumption, the authors were able to show that AI production is dominated by a small number of Big Tech firms who set the trends in the evolution of AI. In terms of AI adoption, only a small share of firms can access high-quality data and technology. While (national) AI ecosystems are deeply rooted in a country's cultural and political backgrounds, a few companies are building global AI ecosystems. By doing so, they create benefits not only for themselves, but also for their complementors in the segment.

Digitalization of economies creates the need for new tools to measure digitally transformed EE to inform policy makers. Several measures of DEEs have been put forward in various studies. One of such measures, the *European Index of Digital Entrepreneurship Systems* (EIDES), was proposed by Autio et al. (2018) to support policies in the domain of digital innovation and start-ups. The EIDES is designed to measure both physical and digital conditions for entrepreneurship in EU 28 countries, and it closely resembles the traditional concept of EE. It consists of three sets of conditions: i) the general framework conditions including, for instance, culture, formal institutions, regulation, taxation, market conditions, and physical infrastructure; ii) the systemic framework conditions including human capital, knowledge creation and dissemination, finance, and networking and support, and iii) digitalization conditions that provide the digital context for both above-mentioned framework conditions. While the general framework conditions apply to all stages of entrepreneurship, the systemic framework conditions are divided into three areas corresponding to three stages of the entrepreneurial process, namely, stand-up, start-up, and scale-up stages. In addition, the digitalization conditions provide an additional layer that emphasizes the role of digital transformation in the general and systemic framework conditions. For instance, the rule of law, which is one of the general framework conditions, is augmented by the digital condition of e-government and freedom of the net. In turn, entrepreneurial culture, measured by social desirability and general acceptance of entrepreneurship in the population, is co-shaped by the use of the internet by individuals and businesses (Autio et al., 2018, 24). The overall value of the EIDES accounts for all these aspects of digitally transformed EE. Based on the distribution of the EIDES, EU countries are divided into four groups in terms of their DEE performance, such as the leaders (e.g., Denmark, Finland, Sweden), the followers (e.g., Austria, Belgium, Estonia, France, Germany), catchers-up (e.g., Cyprus, Czech Republic, Portugal), and laggards (e.g., Bulgaria, Croatia, Greece, Hungary, Italy, Poland). It should be noted, however, that the EIDES was not designed to measure the impact of AI technologies specifically on EE, as there is no explicit measure of AI in the digitalization conditions framework. In addition, the EIDES has the same shortcoming as other, more conventional measures of EEs, such as the one provided by the Global Entrepreneurship Monitor, namely, that it cannot be calculated at a narrower regional level defined within countries. Lastly, the EIDES is aimed at capturing the augmenting effects of digital technologies on EE elements. It does not account for other effects, such as creation of new EE elements that were discussed above.

Another example of measuring DEEs that comes closer to the notion of platform-based eco-systems is the Digital Platform Economy (DPE) index that was proposed by Acs et al. (2022). As in the case of the EIDES, the DPE index applies at the country-level. It was developed using the frameworks of Sussan and Acs (2017) and Song (2019). In doing so, it measures various aspects of digital technology infrastructure, digital user citizenship, digital multi-sided platforms, and digital technology entrepreneurship, each of them representing a sub-index and a building block for calculating the final DPE index. The structure of the DPE index allows identification of potential bottlenecks in a country's platform-based ecosystem and development of corrective policy measures. The study by Acs et al. (2022) shows that the digital platform economy is positively correlated with economic development. Moreover, there exists a significant gap between most European countries and the DPE top performing countries, which are the US and the UK. This gap is likely due to the EU's institutional setup that supports small business ownership rather than fast-growing digital businesses, such as unicorns. The authors call for the need to rebalance the EU's digital entrepreneurial ecosystem policy to "promote technology innovation, platform companies and create sustainable platform economy" (Acs et al., 2022, 111).

3.8 AI regulation and entrepreneurship

In March 2021, the European Commission announced, in its communication entitled "2030 Digital Compass: the European Way for the Digital Decade", an ambitious target, namely, that it wants 75% of European enterprises to use cloud computing services, big data and AI by the year 2030[16]. Just a month later, in April 2021, the European Commission published a draft of its proposed Artificial Intelligence Act (EU AI Act), the world's first and most restrictive and comprehensive regulation of AI. In December 2023, the European Parliament and the Council of the EU reached a political agreement on the AI Act. It is expected that it will be fully applicable by 2025, while the Commission encourages, with the support of the AI Pact[17], an early transition to complying

[16]EUR-Lex - 52021DC0118 - EN - EUR-Lex (europa.eu)
[17]https://digital-strategy.ec.europa.eu/en/policies/ai-pact

behaviors of AI developers and businesses from Europe and beyond in advance of the legal deadline. According to some estimates, the EU AI Act will result in a significant increase in the cost of AI adoption for European businesses (the estimated cost for the European economy is €31 billion over the next five years) and a reduction of AI investments by almost 20% (Mueller, 2021).

AI systems in the US are implicitly regulated by common law such as tort and contract law and statutory and regulatory obligations on organizations (Cuellar et al., 2024). In October 2023, US President Biden signed the Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence (Kang and Sanger, 2023). While the US President has broad powers to regulate AI use in the federal government, the President's reach into the private sector is limited, requiring congressional action. The executive order centers on safety and security mandates such as establishing best practices for watermarking AI generated photos, videos, and audio to lessen dangers of "deep fakes" that can swing elections or swindle consumers, and requiring reports of foreign customers to the federal government from cloud service providers. It also invokes the Defense Production Act to require companies to test their systems to ensure they cannot be used to produce biological or nuclear weapons and to report the findings to the federal government. According to Axios (Heath, 2024), "rapid AI innovation and a federal regulatory vacuum have given state legislatures the impetus to generate a six-fold increase in AI draft legislation compared to a year ago" with nearly half of the draft legislation addressing deep fakes[18]. Analyzing the variance in awareness of AI across states provides insight into regulatory monitoring, guardrails, and entrepreneurship incentivization by individual states. Goel and Nelson (2024a) index Google internet searches for general AI awareness and show, ceteris paribus, more prosperous states and states with more economic regulation having higher levels of awareness of AI and ChatGPT normalized by internet users.

The focus of this section will be on the European approach to the regulation of AI, since the EU has pioneered the legislation around AI, and it can be considered as one of the most important players in the field of AI and data privacy. At the same time, it should be stressed that the European vision of AI and its regulatory approach to AI ecosystems differs substantially from approaches employed in other parts of the world, most notably, in the U.S. and China (Saheb and Saheb, 2023; Timoteo et al., 2021; Comunale and Manera, 2024). Taking availability of and access to training datasets as an

[18]Other relevant proposals of AI regulation in the U.S. include the Algorithmic Accountability Act and domain-specific regulations proposed by federal regulators such as the Food and Drug Administration (FDA), the National Highway Traffic and Safety Administration (NHTSA), and the Federal Trade Commission (FTC) (for details, see Cuellar et al., 2024).

example, the U.S. approach puts strong emphasis on data privacy[19], while in China citizen's data are viewed as a public good, which makes the training data availability, also considering its population size, China's key advantage (Jacobides et al., 2021). Beraja et al. (2023) argue that the Chinese government's use of facial recognition AI for social control benefits AI firms that receive government contracts and thereby stimulates AI innovation by Chinese firms. The European approach is, in turn, strongly based on ethical considerations reflecting European values and a general tendency to avoid risks. While it is beyond the scope of the present research to compare regulatory approaches in various regions, it should be noted that the U.S. is currently a global leader in the number of unicorns (the global share of 51.5%), followed by China (19.4%) and India (6.1%). The countries of the European Union have together produced about 6.7% of the global number of unicorns (Figure 25). It is another ambitious target of the EU's Digital Strategy to double the number of unicorns in the EU by 2030[20].

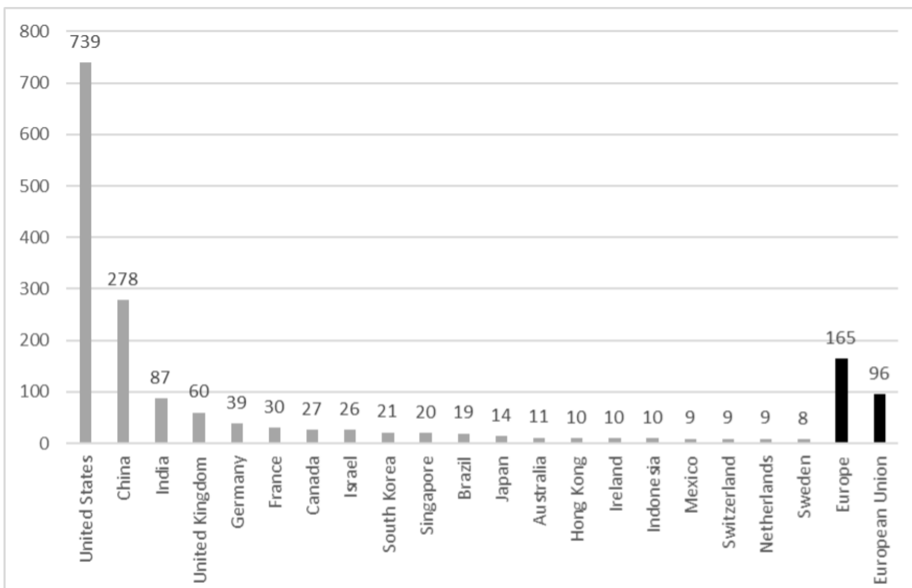


Figure 25: Number of unicorns worldwide as of February 2024, by country

Note: A unicorn company is a privately owned startup that has a current valuation of one billion U.S. dollars or over. Once a company has gone public (IPO) or has been acquired it is no longer termed as a unicorn. Source: Authors' own representation based on data provided by CrunchBase and Statista.

In what follows, we shall offer an overview of the European approach to AI legislation and discuss two key documents regulating AI and data privacy in the EU - the AI Act and the General Data

[19]For instance, the California Consumer Protection Act (CCPA) is a state statute that was introduced in California in 2020 to protect consumer privacy rights. Other than the General Data Protection Regulation that is applied in all regions in the EU, the CCPA only applies to California residents.

[20]EUR-Lex - 52021DC0118 - EN - EUR-Lex (europa.eu)

Protection Regulation - in terms of their potential effects on entrepreneurship, as the empirical evidence is yet to be produced. Finally, we will present and discuss the EU's strategy to boost startups and innovation in trustworthy AI, as presented in the recent 2024 Communication.

The EU's vision of itself is that of a leader in 'ethical' or 'human-centric' AI (Stix, 2020). The narrative of the need to build trustworthy AI is strikingly pronounced in all EU communication pertaining to its Digital Strategy[21]. For instance, the High-Level Expert Group on AI that was established in 2018 had as one of its objectives to propose to the Commission 'Draft Ethics Guidelines for Trustworthy AI', a document, in which it is explicitly stressed that "a human-centric approach to AI is needed, forcing us to keep in mind that the development and use of AI should not be seen as a means in itself, but as having the goal to increase human well-being." The European Digital SME Alliance (2021) considers the EU's vision for ethical AI as Europe's advantage in global competition. This vision has led the EU to implement a series of regulatory measures to ensure that the development and employment of AI tools will be aligned with the vision of 'ethical AI', while at the same time the objectives outlined in the Digital Strategy are achieved. As we will see later, this is likely to lead to difficult trade-offs that might have adverse effects on the entrepreneurship landscape in Europe. (Comunale and Manera, 2024, 47) refer to trade-offs that regulators face between safety and attracting AI investments, as companies may choose locations where policies are looser.

In the AI Act, an 'artificial intelligence system' (AI system) is defined as "software that is developed with one or more of the techniques and approaches listed in Annex I[22] and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with." This definition is very similar to the OECD definition given and discussed in Section 3.3.1. In addition, AI systems are classified following a proportionate risk-based approach, differentiating between uses of AI that create i) unacceptable risk, ii) high risk, and iii) low or minimal risk. Unacceptable AI practices are considered particularly harmful and prohibited. These include, for instance, the use of 'real-time' remote biometric identification systems in publicly accessible spaces; exploiting vulnerabilities of a specific group of

[21] Within the Digital Strategy, the Digital Decade Policy Programme 2030 was established in 2022 that guides all actions related to digitalization under the slogan "The EU is pursuing a human-centric, sustainable vision for digital society throughout the digital decade to empower citizens and businesses." For more information about the Digital Decade, see Europe's Digital Decade — Shaping Europe's digital future (europa.eu).

[22] These techniques and approaches include: a) machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning; b) logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems; c) statistical approaches, Bayesian estimation, search and optimization methods.

persons due to their age, physical or mental disability; and classification of the trustworthiness of natural persons based on their social behavior. In turn, "high-risk" AI systems that pose significant risks to the health and safety or fundamental rights of persons include eight areas: 1) biometric identification and categorization of natural persons; 2) management and operation of critical infrastructure; 3) education and vocational training (for instance, used to determine access of individuals to educational institutions); 4) employment, workers management and *access to self-employment* (for instance, used for recruitment, monitoring and evaluating performance); 5) access to and enjoyment of essential private and public services and benefits; 6) law enforcement; 7) migration, asylum, and border control management; and 8) administration of justice and democratic processes. High-risk AI systems are subject to specific requirements, and risk management systems must be established in relation to them. For non-high risk AI systems, there is an obligation to disclose that the human is interacting with an AI system and label deep fakes when not used for legitimate purposes.

Several measures were established to support "small-scale providers", which are defined as micro or small enterprises[23]. For instance, they are provided with priority access to so-called AI regulatory sandboxes that are established to support innovation and reduce the regulatory burden by means of providing a controlled environment that facilitates the development, testing and validation of AI systems. In addition, specific awareness raising activities tailored to the needs of small-scale providers and establishment of a dedicated channel for communication to provide guidance and respond to queries about the implementation of the regulation will be established. Moreover, small and micro-enterprises enjoy an exemption from the regulation: they can put into service for their own use AI systems for the purpose of creditworthiness assessment and credit scoring.

The regulation outlined in the AI Act can affect the entrepreneurship landscape in Europe in several ways. First, the regulation will give rise to *new compliance costs* that might be unbearable for those entrepreneurs who need to engage professional services to stay compliant with the regulation. Non-compliance of an AI system with any requirements or obligations under the AI Act is subject to administrative fines of up to €20 million or, if the offender is a company, up to 4% of its total worldwide annual turnover for the preceding financial year, whichever is higher. Under certain circumstances, administrative fines are up to €30 million or up to 6% of an offending company's total worldwide turnover. A study for the U.S. found that providing managers with information

[23] In the Commission Recommendation 2003/361/EC, a small enterprise is defined as an enterprise which employs fewer than 50 persons and whose annual turnover and/or annual balance sheet total does not exceed €10 million. In turn, a microenterprise is defined as an enterprise which employs fewer than 10 persons and whose annual turnover and/or annual balance sheet total does not exceed €2 million.

about AI regulation had a negative effect on their intention to adopt AI technologies by investing into AI training of current employees and purchasing AI packages from external vendors, while it had a positive effect on managers' perception of the importance of ethical issues related to AI (Cuellar et al., 2024). This effect was largely independent of the type of regulation design observed, including broad proposals of general AI regulation, such as the Algorithmic Accountability Act, domain-specific regulations that apply to businesses operating in specific sectors, and data privacy regulation treatment, such as the CCPA. The study further finds that the negative impact of AI regulation on AI adoption intent is stronger for small rather than for large firms.

Second, the regulation could increase the *general uncertainty* associated with the rather broad definition of AI adopted in the AI Act. This uncertainty is further complicated by the classification of AI systems in different risk categories and unprecise definition of 'harmful approaches'. It appears also important to consider that the risks associated with an AI system depend on the sector, in which an AI system is intended to be employed, rather than the AI system itself, which is merely a software that employs statistical data analysis techniques. Institutional uncertainties provide powerful incentives for evasive entrepreneurship, and, in some cases, for entrepreneurial exit (Bylund and McCaffrey, 2017). In addition, institutional uncertainty may diverge entrepreneurial activities into unproductive or even destructive channels (Baumol, 1996). Institutional uncertainty related to the AI Act could result in startups being more reluctant to invest into and adopt AI technologies classified as 'high-risk', and more risk averse entrepreneurs could potentially be more reluctant to adopt even 'non-high risk' AI technologies. Startups could be more sensitive to this type of uncertainty than established big companies due to lower levels of resources available to them. Uncertainty could furthermore decrease the attractiveness of the EU for investors and international entrepreneurs. Lastly, it is difficult to judge how far the definition of AI systems adopted in the AI Act can be considered as future-proof, as the technology is still in its emerging state.

Third, AI systems that are used to access self-employment are explicitly mentioned in the AI Act as high-risk AI systems. It is not further specified what cases exactly fall under the regulation. One could speculate that these purposes include predicting the success of a business venture in advance, deciding whether to grant a business loan, or granting subsidies and public support to startups. If these cases are affected by the AI Act, the regulation can lead to *information asymmetries* that are considered a notable barrier for market entry, growth, and survival. Information asymmetries may impede access to financial capital for startups that lack established business networks, bank relationships, reputation, and a track record, and startups might therefore struggle more than larger and older companies to access formal financing (Klapper and Love, 2011). This could result in adverse effects of regulation on entrepreneurs in Europe.

Conversely, regulations may also serve as *profitable sources of entrepreneurial opportunities*. For instance, consumer protection regulations put an additional burden on entrepreneurs and create the need for compliance with the safety procedures and consumer health norms, but some entrepreneurs may perceive this as an opportunity for offering new compliance services (Jacquemin and Janssen, 2015). Regulation of AI is not an exception in this respect. For instance, with the introduction of LLM technologies, many universities revised their anti-plagiarism and anti-cheating policies to address potential misuse of ChatGPT and similar AI tools for cheating. Goel and Nelson (2024b) provide some evidence that anti-plagiarism and anti-cheating policies are positively correlated with resources that facilitate cheating behaviors. This hints toward the presence of entrepreneurial opportunities arising from anti-plagiarism policies that are being actively pursued by entrepreneurially minded individuals.

It is highly important for AI-adopting startups to have affordable access to AI infrastructure, such as cloud services and supercomputers, AI talent, and training datasets to train AI models, as we discussed in Section 3.5.2. According to a study of barriers to AI adoption by firms in OECD countries, the cost and a lack of skills are currently greater barriers to AI adoption than government regulation: "25% of employers in finance and 19% in manufacturing said that government regulation was a barrier, compared to 53% and 58% that said that high costs of the technology were a barrier, and 41% and 43% that said that lack of relevant skills was a barrier" (Lane et al., 2023, 14). However, this study also found that government regulation was a stronger barrier for AI adoption in EU rather than in non-EU OECD countries (Lane et al., 2023, 95). The EU's 2024 "Communication on boosting startups and innovation in trustworthy artificial intelligence" lays out the key actions to help startups in the EU to engage with AI within the framework of the European approach to AI[24]. It is planned to establish "AI Factories", which are open ecosystems formed around a network of European public supercomputers that can be accessed by AI startups to train large-scale AI models. It is expected that access to supercomputers will provide European AI startups with a substantial competitive advantage by accelerating the training of large AI models from 6-9 months on average to just a few weeks and achieving substantial cost benefits compared to using commercial

[24]Communication on boosting startups and innovation in trustworthy artificial intelligence — Shaping Europe's digital future (europa.eu)

cloud services[25]. In addition, "AI Factories" will take on an important role in AI entrepreneurship ecosystems by means of collaborating with startups, universities and research centers, as well as key industrial sectors. "AI Factories" will be financially supported mainly, but not exclusively, by public funds.

Moreover, the EU intends to improve the availability of and access to quality data in key sectors (currently 14), such as health, media, mobility, environment, and manufacturing. Access to training data is an important input for AI startups to develop AI-enabled products (Bessen et al., 2022) and, without an appropriate regulation, there is a risk of market tipping (monopolization) that inhibits innovation in this area (Graef and Prüfer, 2021). This objective shall be achieved with the creation of Common European Data Spaces that will provide controlled access to high-quality data in these sectors[26]. This appears to be a highly relevant initiative for supporting digital entrepreneurs in the EU, since access to private data is heavily restricted, most prominently, by the General Data Protection Regulation (GDPR).

The GDPR entered into force in May 2016, and it has been applied since May 2018 to protect personal data of individuals who are in the EU. It applies also if the processing of data occurs outside of the EU, thus affecting EU- as well as non-EU-based businesses. The GDPR requires an informed opt-in consent for collected data, while it generally prohibits processing of specific categories of personal data, such as data on ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, health and genetic data, among others. Furthermore, individuals are granted several rights, for instance, the right to obtain information on the personal data collected and processed, the right to rectification of inaccurate personal data, the right of erasure of personal data, the right of restriction of processing of personal data, and the right of data portability. Several studies have analyzed the effects of the GDPR on businesses and startups. Early studies of short-run effects of the regulation report a negative effect on investment into new and emerging technology firms in the EU, relative to US-based firms (Jia et al., 2018). Moreover, the GDPR has placed a significant burden particularly on small- and medium-sized enterprises, which may adversely affect their development and competitiveness compared to larger firms that can allocate dedicated resources to manage the GDPR-related issues (Wilkinson, 2018). In fact, AI startups with customers in Europe were found to be more likely to create a new position (69% of respondents) or to reallocate firm resources (63% of respondents) to stay compliant with the GDPR, while almost 75% of responding firms have deleted

[25] Access to Our Supercomputers - European Commission (europa.eu)
[26] Common European Data Spaces — Shaping Europe's digital future (europa.eu)

data due to the GDPR, which might affect their performance in the future (Bessen et al., 2020). Semi-structured interviews with 15 tech startups based in the UK revealed that they face several challenges in meeting their data protection obligations, particularly, due to lack of clarity if their technology could be reconciled with the GDPR and whether aspects of the GDPR applied to them (Norval et al., 2021). According to estimates reported in (Mueller, 2021, 9), thirty-four percent of large enterprises spent more than €1 million to implement GDPR. Moreover, larger businesses face higher costs, and U.S. companies with more than 500 employees spent up to $10 million each on GDPR compliance and up to $150 billion in total. Despite substantial fines for non-compliance with the GDPR, some firms were found to turn into multi-violators who are not affected by the amount of the fines imposed on them, developing a "fine is a price" attitude (Méndez-Suárez et al., 2023).

Furthermore, the 2024 Communication announced further supporting initiatives, such as strengthening the EU's generative AI talent pool, thus, directly addressing the concern that lack of AI skills is one of the main barriers to AI adoption. In addition, it was announced that the EU intends to encourage public and private investments, including venture capital, in AI startups and scale-ups. Lastly, the initiative to support the development of trustworthy AI algorithms appears to be a promising one to help reduce the burden of AI regulation, and particularly, the discussed AI Act and the GDPR, on entrepreneurs. Table 14 summarizes the AI regulations and policy initiatives in the EU discussed in this section.

3.9 Discussion

Our review reveals that AI has the potential to make a profound impact on virtually all aspects of entrepreneurship, albeit the research on this topic is still in its infancy. In this section, we summarize some of the key insights resulting from our literature survey and our own research on AI and entrepreneurship, identify main trends in this emerging field, and formulate implications for entrepreneurship research and policy.

It should be noted first that emerging literature on AI and entrepreneurship is often overlapping with a much broader and more established literature on digital entrepreneurship. AI is frequently considered as just a subset of the many digital technologies in the literature, and, given that the field is still in its emerging state, the term AI is frequently used as a rather generic term, without an explicit reference to how it is defined or measured. Being a novel field, the inclusion of AI is considered a promising direction for the future of digital entrepreneurship research (Kollmann et al., 2022). Thus,

Table 14: Summary of relevant AI regulations and policy initiatives in the EU

Regulation/ policy initiative	Key objectives	Expected effect of regulation / policy initiative on entrepreneurship
AI Act (agreed upon in December 2023; expected to be fully applicable by mid-2026)	• Ensure that AI systems placed on the EU market and used are safe and respect existing law on fundamental rights and EU values; • Ensure legal certainty to facilitate investment and innovation in AI; • Enhance governance and effective enforcement of existing law on fundamental rights and safety requirements applicable to AI systems; • Facilitate the development of a single market for lawful, safe and trustworthy AI applications and prevent market fragmentation.	• Increased cost of compliance with the regulation. • Increased uncertainty, potentially leading to lower AI adoption rates. • Information asymmetries, potentially leading to limited access to formal sources of financing. • Decreased attractiveness of the EU for international (digital) entrepreneurs and investors, potentially leading to reductions in the availability of venture capital. • New entrepreneurial opportunities directly arising from regulation (e.g., compliance services).
General Data Protection Regulation (GDPR) (came into force in May 2016; fully applicable since May 2018)	• Protection of personal data of individuals who are in the EU.	• Increased cost of compliance with the regulation. • Lower availability of and difficult access to training datasets to train AI models, potentially leading to lower AI adoption rates and lower firm performance. • Increased uncertainty, potentially leading to lower AI adoption rates or evasive entrepreneurship. • Cross-border impact on firms with customers in the EU. • New entrepreneurial opportunities directly arising from regulation (e.g., compliance services).
Communication on boosting startups and innovation in trustworthy AI (January 2024)	• Creation of a thriving startup and innovation ecosystem for trustworthy AI. • Establishment of supporting initiatives, such as "AI Factories" and "Common European Data Spaces". • Strengthening the EU's generative AI talent pool. • Encouraging public and private investments in AI startups and scale-ups. • Supporting the development of trustworthy AI algorithms.	• Potentially substantial competitive advantage for AI startups due to public access to supercomputers and high-quality training datasets in key sectors. • Better availability of AI talent and venture capital for AI startups. • Reduced burden of AI regulation (AI Act and GDPR) on AI startups.

we decided to start our review by providing an overview of various existing definitions of AI, which turned out to be a challenging task in itself, given the variety of AI technologies and approaches to regulating AI. Particularly definitions that are used in regulation of AI tend to be very broad, which is probably the result of a difficult trade-off between the need to define the subject of regulation and the goal of ensuring that the definition is "future-proof", so that regulation will not become obsolete soon, given that AI is an emerging and fast developing field. One of the key insights that we learned from reviewing AI definitions is that we will unlikely understand the full range of implications that AI has for entrepreneurship when research continues to adopt unclear definitions of AI and not distinguish between different types of AI technologies. Theoretical entrepreneurship research will likely fail to make clear-cut predictions, and empirical research will likely produce ambiguous results. At the same time, we acknowledge that in the absence of comprehensive frameworks to measure AI development, implementation, and impacts, it will be difficult to achieve this goal. Thus, it appears important to develop comprehensive frame-works and corresponding inventories that would make it possible to collect high-quality data on AI to study its impacts on entrepreneurial businesses.

A conceptual literature on AI in entrepreneurship has emerged, highlighting potential roles of AI in entrepreneurship theories. As uncertainty is a core concept in some of the most prominent theories on entrepreneurship, an interesting tension is whether AI, as a prediction device, can re-duce this uncertainty, or whether entrepreneurial uncertainty about the future is of such unpredictable nature that it poses a fundamental limitation even to enhanced AI using any amount of data about the past (Townsend et al., 2023). Further theory development in this area could consider how AI might facilitate entrepreneurial experimentation to reduce uncertainty (e.g., Zellweger and Zenger, 2023) by reducing the cost and time needed for experimentation (Van Gelderen et al., 2021).

A useful approach in some conceptual papers and literature surveys has been to describe how AI can affect various stages of the entrepreneurial process (Chalmers et al., 2021) such as opportunities (their discovery or creation and exploitation), decision making, and performance (Giuggioli and Pellegrini, 2023). AI is often seen as an external enabler of entrepreneurship (e.g., Davidsson and Sufyan, 2023). Researchers have provided examples of how AI affects entrepreneurship by transforming various sectors of the economy such as customer service, the financial sector, healthcare, and education (Shepherd and Majchrzak, 2022). Technological change is often seen as a source of opportunities for entrepreneurs as the agents of innovation (Shane, 2000), and the more radically AI transforms the environment in which businesses operate, the more entrepreneurs may emerge as the agents of disruptive change.

However, the literature also discusses that rewards from productivity gains through AI innovation may be distributed very unequally. Large firms controlling critical amounts of capital, data and expertise may have advantages in AI development, leading to industry concentration and dominating superstar firms, an environment that is detrimental to new firm formation (Chalmers et al., 2021; Gerling et al., 2022).

The emerging empirical literature indeed reports that large firms are more likely to adopt AI than smaller firms (McElheran et al., 2024) and that AI investments lead to stronger growth in initially larger firms (Babina et al., 2024). Therefore, an important empirical research topic going forward is to study potential entry barriers to AI development and usage. Such barriers may hinder entrepreneurship and favor oligopolies or monopolies, decreasing competition and potentially innovation. Entry barriers for AI startups may emerge due to the need for very large amounts of training data; proprietary AI models and algorithms; the need for significant capital investment in IT infrastructure; scarcity of talent; and the difficulty of coping with data protection and AI regulation. Research should investigate mitigating factors to inform discussions on potential policy interventions. For example, capital requirements can be reduced by cloud services, and open-source AI models may facilitate a democratization of AI. AI tools are accessible in rural areas through the internet, and LLMs can reduce language barriers, which may contribute to leveling the playing field in terms of geographical divides (Serrano, 2024). There also seems to be awareness among regulators that small businesses may need support to facilitate compliance with regulations.

More generally, the empirical literature on AI impacts on entrepreneurship has struggled to keep pace with the rapid technological development and implementation of AI. Comprehensive studies such as McElheran et al. (2024) and Bonney et al. (2024) report that only a small share of firms utilize AI. Our own analysis of survey data from Germany reveals that individuals are often unaware that they use AI, partially due to uncertainty around the definition of AI mentioned above. It may therefore be beneficial in future collections of survey data to ask specifically for the use of certain applications that involve AI rather than asking for AI use directly. Carefully thinking about the way to elicit AI use in surveys may become even more important going forward as AI is increasingly integrated in everyday applications such as web search, e-mail programs and office suites.

A research challenge in the empirical literature is to identify causal effects of AI on entrepreneurial outcomes. Randomized field experiments (Otis et al., 2024) and comparisons between the performance of humans and AI in entrepreneurial tasks (McKenzie and Sansone, 2019; Blohm et al., 2022) have yielded mixed results, which is consistent with the notion that uncertainty makes entrepreneurship a difficult application for AI. More research is necessary as AI becomes more powerful, and research may find that AI increases performance in specific areas relevant for entrepreneurship.

In addition to these direct effects of AI on entrepreneurship, AI indirectly affects entrepreneurship through its broader impact on the labor market. As AI automates tasks and displaces certain workers from their jobs, it may push individuals into necessity entrepreneurship. Other employees become more productive and may receive higher wages due to complementarities between AI and human workers, changing the opportunity cost of entrepreneurship. At the same time, the technological change enables new business ideas and new forms of opportunity entrepreneurship. The impacts of different types of new digital technologies and AI on workers' occupations and wages and, as a result, their subsequent transitions into different types of entrepreneurship are already observable in the data (e.g., Fossen and Sorgner, 2021, 2022). As AI technologies are further evolving and AI implementation in firms accelerates, it is important to monitor the developments in the labor market to understand implications for entrepreneurship. As we discussed in this review, future research should carefully consider appropriate measures of AI impacts on occupations and work tasks, especially in international and regional contexts, and of different types of entrepreneurs affected in differential ways.

Literature on entrepreneurial ecosystems has recently attempted to systematically account for impacts of digital transformation on entrepreneurship ecosystems, although, in most cases, it does not explicitly focus on AI. We could distinguish between two strands in this emerging literature. One strand in the literature, which has its roots in regional studies, tries to describe how existing entrepreneurship ecosystem elements and outputs are affected by digital technologies, while another strand in this literature takes a platform-based approach to digital transformation of entrepreneurship. The term "digital entrepreneurship ecosystems" is commonly applied in the studies following the platform-based approach, while it is less common in studies that conceptualize entrepreneurship ecosystems as a (regional) network of interdependent elements, such as institutions, knowledge, finance, talent, or culture, which together result in productive entrepreneurship that leads to economic growth. In general, our finding is that this latter strand in the literature has not yet fully conceptualized the effects of AI on entrepreneurship ecosystems. To fill in this gap, we propose to focus on the effects of AI on i) entrepreneurship ecosystem elements, ii) entrepreneurship ecosystem outputs, and iii) entrepreneurship ecosystem processes that include interdependencies between the elements and between the elements and outputs. So far, particularly the latter aspect has not been sufficiently investigated, although it is among the most promising aspects in terms of policy implications. For instance, we discussed how long-term persistence of regional entrepreneurship is the result of persistent processes within an entrepreneurship ecosystem that can potentially be disturbed or facilitated by AI. Moreover, it appears promising to study the complementarities between two established and related concepts of AI ecosystems and entrepreneurship ecosystems. The main challenge of this approach is that measuring an AI entrepreneurship ecosystem is difficult due to it being an emerging ecosystem, such that novel, unconventional approaches that do not merely rely on well-established indicators from official statistics need to be proposed.

In turn, literature on digital entrepreneurship ecosystems that employs a platform-based approach does not necessarily distinguish between different types of digital technologies or between AI and other digital technologies, but it mainly focuses on digital platforms and digital marketplaces. In addition, the focus on the role of the region for digital entrepreneurship ecosystems is substantially less prominent in this literature, thus, further distinguishing it from the other strand in the literature on entrepreneurship ecosystems where the role of the region continues to be pivotal. In sum, the two distinct developments in the literature on digitization of entrepreneurship ecosystems created a certain level of confusion due to partly overlapping terminology that is used in both literature strands, and AI is not explicitly conceptualized in the entrepreneurship ecosystems literature, while literature on AI entrepreneurship ecosystems is in its embryonic stage.

Another fascinating emerging topic regards the role of AI regulation in entrepreneurship. We show that regions and countries differ substantially in their vision for AI and, therefore, their approach to regulating AI. While it is beyond the scope of the present research to conduct a comparative analysis of regulatory approaches in different regions, we decided to focus primarily on the European approach, also known as "ethical AI" or "human-centric AI" approach, since the EU with its recent AI Act and the GDPR has pioneered the legislation around AI. While the AI Act will likely affect the type and the level of entrepreneurial activities in the EU, it is too early to see its effects. Therefore, we discuss the potential impacts AI regulation could have on entrepreneurship, including impacts on the cost of doing business, entrepreneurial uncertainty, information asymmetries, and the attractiveness of the EU for AI startups and investors. We have also reviewed supportive measures that the EU has initiated to promote AI entrepreneurship and digital unicorns and to help AI startups stay compliant with the restrictive regulation. Future research that compares approaches to regulating AI and their impacts on the level and type of entrepreneurship activities appears to be particularly promising, since there is a general agreement among regulators in different regions that promoting AI entrepreneurship is a key strategic objective. Moreover, the stark competition for

AI talent and AI startups incentivizes different regions to create unique competitive advantages to become an attractive location for AI entrepreneurs. While the EU's claimed comparative advantage lies in its "trustworthy" approach to AI, it is still to be researched what approach will turn out to be most efficient.

In this regard, one of the promising areas for future research would be to investigate the complex relationship between trust, AI, and entrepreneurship. Several dimensions, such as transparency, reliability, tangibility, or responsiveness, were found to positively influence human trust. These dimensions may be present to varying degrees in different types of AI, such as robotic or embedded AI (Glikson and Woolley, 2020). Trust is also crucial to entrepreneurs, as it is indispensable in uncertain environments (Caliendo et al., 2012; Mickiewicz and Rebmann, 2020). It would be an interesting and relevant research question to investigate how trust impacts on entrepreneurs' decision to adopt AI, as AI usage rates appear to be still rather low (see Section 3.5.1). In addition, it appears important to also investigate the moderating role of institutions, such as AI regulation, in the relationships between trust, AI and entrepreneurship.

Substantial risks and ethical concerns regarding AI have been widely discussed. These include bias in AI output reflecting bias in the training data or human trainers (Desai, 2019; Bubeck et al., 2023), potential violation of intellectual property rights when human creative works are used to train AI (Robertson, 2024), untransparent AI decisions (Keding, 2021) and lack of fact validation, displacement of human workers and loss of autonomy at work (Shepherd and Majchrzak, 2022), and potential loss of human control more generally. Moreover, malicious human actors could misuse and weaponize AI, for example to spread disinformation, for cyberattacks, or physical weapon design (Harris et al., 2024). These risks and ethical concerns have motivated some of the government regulations discussed above. In the context of entrepreneurship, Shepherd and Majchrzak (2022) highlight a potential accountability gap if entrepreneurs can shift blame for bad outcomes to AI's involvement in the decision making. If AI shields entrepreneurs from responsibility, this may induce some entrepreneurs to engage in activities that are harmful to society to gain a personal benefit. Conversely, AI can also help entrepreneurs to spot mistakes sooner (Desai, 2019). Future research should address which factors enable the use of AI for entrepreneurship that is productive and contributes to sustainable development (Gupta et al., 2023; Bickley et al., 2024) while preventing destructive outcomes for society. In particular, the optimal design of regulation in the context of AI and entrepreneurship is an important research area, as Baumol (1996) points out that good institutions can help channel entrepreneurial energy away from destructive and toward productive use, but overregulation could also impede productive entrepreneurship.

3.10 Conclusion

AI has begun to transform entrepreneurship and its environment. Research on the AI-entrepreneurship nexus is still in its infancy, characterized by a large number of AI definitions, prevalence of conceptual over empirical studies, an ongoing search for reliable and precise measurement tools, taxonomies, and novel datasets, attempts to reconsider the role of traditional determinants of entrepreneurship, as well as the role of entrepreneurship for the well-being of individuals, economies, and societies in the age of AI. We can state with certainty that the research on AI and entrepreneurship is in its early 'entrepreneurial' stage, and it is yet difficult to foresee how it will likely evolve in the future. Trying to do so at this moment, when AI technologies themselves are developing so fast, would be - by far - a too ambitious task. Nevertheless, several crucial research avenues emerged from our survey. What are the boundaries of AI's capability of reducing uncertainty in entrepreneurship? How does AI affect the types of entrepreneurial ventures emerging and their performance? Will AI displace many workers from their jobs and induce them to become necessity entrepreneurs? How does AI impact on entrepreneurs' well-being in terms of their physical health, their mental health, and their job and life satisfaction? How does AI interact with entrepreneurial creativity? What is the role of AI in entrepreneurial ecosystems and how can it be included in measurements to inform policy? Does AI challenge the role of geography in entrepreneurial ecosystems? How should institutions be designed to help productive AI startups access necessary resources including training datasets and supercomputers to enter the market and succeed? At the same time, do regulations requiring open-source algorithms or data sharing provide inclusive access to AI for entrepreneurs or make it more difficult to protect innovations and access financing? What is a good AI entrepreneurship policy? Addressing these and other research questions we raised throughout this review will provide much-needed guidance for entrepreneurship policy and practitioners in their attempts to channel entrepreneurial activity around AI into productive use while preventing potential harm to society.

References

Abraham, K., J. Haltiwanger, K. Sandusky, and J. Spletzer (2017). Measuring the gig economy: Current knowledge and open issues. *Measuring and Accounting for Innovation in the 21st Century*.

Abraham, K. G., J. C. Haltiwanger, C. Hou, K. Sandusky, and J. R. Spletzer (2021). Reconciling survey and administrative measures of self-employment. *Journal of Labor Economics 39*(4), 825–860.

Acemoglu, D., D. Autor, J. Hazell, and P. Restrepo (2022). Artificial intelligence and jobs: Evidence from online vacancies. *Journal of Labor Economics 40*(S1), S293–S340.

Acemoglu, D. and P. Restrepo (2018a). Artificial intelligence, automation, and work. In A. K. Agrawal, J. Gans, and A. Goldfarb (Eds.), *The Economics of Artificial Intelligence: An Agenda*. NBER. https://www.nber.org/chapters/c14027.

Acemoglu, D. and P. Restrepo (2018b). The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review 108*(6), 1488–1542.

Acemoglu, D. and P. Restrepo (2019). Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives 33*(2), 3–30.

Acs, Z. J., E. Autio, and L. Szerb (2015). National systems of entrepreneurship: Measurement issues and policy implications. In *Global Entrepreneurship, Institutions and Incentives*, pp. 523–541.

Acs, Z. J., L. Szerb, A. Song, E. Lafuente, and E. Komlosi (2022). Measuring the digital platform economy. In M. Keyhani, T. Kollmann, A. Ashjari, A. Sorgner, and C. E. Hull (Eds.), *Handbook of Digital Entrepreneurship*. Edward Elgar Publishing.

Aghion, P., N. Bloom, R. Blundell, R. Griffith, and P. Howitt (2005). Competition and innovation: An inverted-U relationship. *The Quarterly Journal of Economics 120*(2), 701–728.

Aghion, P., C. Harris, P. Howitt, and J. Vickers (2001). Competition, imitation and growth with step-by-step innovation. *The Review of Economic Studies 68*(3), 467–492.

Aghion, P., B. Jones, and C. Jones (2019). Artificial intelligence and economic growth. In A. Agrawal, J. Gans, and A. Goldfarb (Eds.), *The Economics of Artificial Intelligence: An Agenda*, pp. 237–290. University of Chicago Press.

Agrawal, A., J. S. Gans, and A. Goldfarb (2019). Artificial intelligence: The ambiguous labor market impact of automating prediction. *Journal of Economic Perspectives 33*(2), 31–50.

Agrawal, A., J. S. Gans, and S. Stern (2021). Enabling entrepreneurial choice. *Management Science 67*(9), 5510–5524.

Agrawal, A., J. McHale, and A. Oettl (2024). Artificial intelligence and scientific discovery: A model of prioritized search. *Research Policy 53*(5), 104989.

Al-Subaihin, A., F. Sarro, S. Black, and L. Capra (2019). Empirical comparison of text-based mobile apps similarity measurement techniques. *Empirical Software Engineering : An International Journal 24*(6), 3290–3315.

Alexopoulos, M. and J. Cohen (2018). Canadian productivity growth, secular stagnation, and technological change. *International Productivity Monitor* (35), 113–137.

Angrist, J., S. Caldwell, and J. Hall (2017). Uber vs taxi: A driver's eye view. *NBER Working Papers* (w23891).

Appelbaum, E. (1982). The estimation of the degree of oligopoly power. *Journal of Econometrics 19*(2), 287–299.

Arntz, M., T. Gregory, and U. Zierahn (2017). Revisiting the risk of automation. *Economics Letters 159*, 157–160.

Åstebro, T., H. Herz, R. Nanda, and R. A. Weber (2014). Seeking the roots of entrepreneurship: Insights from behavioral economics. *Journal of Economic Perspectives 28*(3), 49–70.

Åstebro, T. and J. Tåg (2017). Gross, net, and new job creation by entrepreneurs. *Journal of Business Venturing Insights 8*, 64–70.

Audretsch, D. B. and M. Belitski (2017). Entrepreneurial ecosystems in cities: Establishing the framework conditions. *The Journal of Technology Transfer 42*, 1030–1051.

Autio, E., L. Szerb, E. Komlósi, and M. Tiszberger (2018). The European Index of Digital Entrepreneurship Systems. Technical Report EUR 29309 EN, Publications Office of the European Union, Luxembourg. 2018, ISBN 978-92-79-91303-7, doi:10.2760/39256, JRC112439.

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

Autor, D. H. (2015). Why are there still so many jobs? The history and future of workplace automation. *Journal of Economic Perspectives 29*(3), 3–30.

Babina, T., A. Fedyk, A. He, and J. Hodson (2024). Artificial intelligence, firm growth, and product innovation. *Journal of Financial Economics 151*, 103745.

Bachmann, R., M. Gonschor, S. Milasi, and A. Mitra (2024). Technological progress and the dynamics of self-employment: Worker-level evidence for Europe. *International Labour Review*. Forthcoming.

Baily, M. and H. Gersbach (1995). Efficiency in manufacturing and the need for global competition. *Brookings Papers on Economic Activity*, 307–347.

Bajari, P., Z. Cen, V. Chernozhukov, M. Manukonda, S. Vijaykumar, J. Wang, R. Huerta, J. Li, L. Leng, G. Monokroussos, and S. Wan (2023). Hedonic prices and quality adjusted price indices powered by AI. *arXiv:2305.00044v1*.

Bajari, P., V. Chernozhukov, A. Hortaçsu, and J. Suzuki (2019). The impact of big data on firm performance: An empirical investigation. *AEA Papers and Proceedings 109*, 33–37.

Bana, S. H. (2021). job2vec: Using language models to understand wage premia. (Working Paper). Stanford Digital Economy lab. https://conference.iza.org/conference˙files/DATA˙2021/bana˙s26582.pdf.

Baumol, W. J. (1996). Entrepreneurship: Productive, unproductive, and destructive. *Journal of Business Venturing 11*(1), 3–22.

Bejjani, M., L. Göcke, and M. Menter (2023). Digital entrepreneurial ecosystems: A systematic literature review. *Technological Forecasting & Social Change 189*, 122372.

Benz, M. and B. S. Frey (2008). Being independent is a great thing: Subjective evaluations of self-employment and hierarchy. *Economica 75*(298), 362–383.

Beraja, M., A. Kao, D. Y. Yang, and N. Yuchtman (2023). AI-tocracy. *The Quarterly Journal of Economics 138*(3), 1349–1402.

Berg, J. and H. Johnston (2019). Too good to be true? A comment on Hall and Krueger's analysis of the labor market for Uber's driver-partners. *ILR Review 72*(1), 39–68.

Berger, E. S., F. Von Briel, P. Davidsson, and A. Kuckertz (2021). Digital or not–The future of entrepreneurship and innovation: Introduction to the special issue. *Journal of Business Research 125*, 436–442.

Berger, T., C. Chen, and C. B. Frey (2018). Drivers of disruption? Estimating the Uber effect. *European Economic Review 110*, 197–210.

Berry, S. T. and J. Waldfogel (2001). Do mergers increase product variety? Evidence from radio broadcasting. *The Quarterly Journal of Economics 116*(3), 1009–1025.

Bertoni, F., S. Bonini, V. Capizzi, M. G. Colombo, and S. Manigart (2022). Digitization in the market for entrepreneurial finance: Innovative business models and new financing channels. *Entrepreneurship Theory and Practice 46*(5), 1120–1135.

Bessen, J. E., I. Cockburn, and J. Hunt (2023). Is distance from innovation a barrier to the adoption of artificial intelligence? Working Paper, https://www.jenniferhunt.org/uploads/ai-diffusion-230829-lores.pdf, accessed 3/18/2024.

Bessen, J. E., S. M. Impink, L. Reichensperger, and R. Seamans (2020). GDPR and the importance of data to AI startups. Working paper available at SSRN: https://ssrn.com/abstract=3576714.

Bessen, J. E., S. M. Impink, L. Reichensperger, and R. Seamans (2022). The role of data for AI startup Growth. *Research Policy 51*(5), 104513.

Bessen, J. E., S. M. Impink, L. Reichensperger, and R. Seamans (2023). The business of AI startups. Working paper available at SSRN: https://ssrn.com/abstract=3293275.

Bessen, J. E., S. M. Impink, and R. Seamans (2023). The role of ethical principles in AI startups. Working paper available at SSRN: https://ssrn.com/abstract=4378280.

Bickley, S. J., A. Macintyre, and B. Torgler (2024). Artificial intelligence and big data in sustainable entrepreneurship. *Journal of Economic Surveys*. https://doi.org/10.1111/joes.12611.

Blinder, A. S. and A. B. Krueger (2013). Alternative measures of offshorability: A survey approach. *Journal of Labor Economics 31*(S1), S97–S128.

Block, J. and P. Sandner (2009). Necessity and opportunity entrepreneurs and their duration in Self-Employment: Evidence from German micro data. *Journal of Industry, Competition and Trade 9*, 117–137.

Block, J. and M. Wagner (2010). Necessity and opportunity entrepreneurs in Germany: Characteristics and earnings differentials. *Schmalenbach Business Review 62*(2), 154–174.

Blohm, I., T. Antretter, C. Sirén, D. Grichnik, and J. Wincent (2022). It's a people's game, isn't it?! A comparison between the investment returns of business angels and machine learning algorithms. *Entrepreneurship Theory and Practice 46*(4), 1054–1091.

Blundell, R., R. Griffith, and J. van Reenen (1999). Market share, market value and innovation in a panel of british manufacturing firms. *The Review of Economic Studies 66*(3), 529–554.

Boeri, T., G. Giupponi, A. B. Krueger, and S. Machin (2020). Solo self-employment and alternative work arrangements: A cross-country perspective on the changing composition of jobs. *Journal of Economic Perspectives 34*(1), 170–195.

Bonney, K., C. Breaux, C. Buffington, E. Dinlersoz, L. S. Foster, N. Goldschlag, J. C. Haltiwanger, Z. Kroff, and K. Savage (2024). Tracking firm use of AI in real time: A snapshot from the Business Trends and Outlook Survey. NBER Working Paper w32319, National Bureau of Economic Research.

Bosma, N. (2013). The Global Entrepreneurship Monitor (GEM) and its impact on entrepreneurship research. *Foundations and Trends in Entrepreneurship 9*(2), 143–248.

Boucher, P. N. (2020). Artificial intelligence: How does it work, why does it matter, and what can we do about it?

Boussioux, L., J. N. Lane, M. Zhang, V. Jacimovic, and K. R. Lakhani (2024). Generative AI and creative problem solving. Technical Report 24-005, Harvard Business School Technology & Operations Mgt. Unit. Working Paper No. 24-005, https://doi.org/10.2139/ssrn.4533642.

Brock, J. K. U. and F. von Wangenheim (2019). Demystifying AI: What digital transformation leaders can teach you about realistic artificial intelligence. *California Management Review 61*(4), 110–134.

Brown, P. (2021). Improving customer interactions with AI-based avatars. https://electronics360.globalspec.com/article/17568/improving-customer-interactions-with-ai-based-avatars.

Brown, T. E. (2017). Sensor-based entrepreneurship: A framework for developing new products and services. *Business Horizons 60*(6), 819–830.

Bryan, K. A. and E. Hovenkamp (2020). Startup acquisitions, error costs, and antitrust policy. *The University of Chicago law review 87*(2), 331–356.

Brynjolfsson, E., D. Li, and L. R. Raymond (2023). Generative AI at work. Technical report, NBER Working Paper w31161, National Bureau of Economic Research.

Brynjolfsson, E. and T. Mitchell (2017a). What can machine learning do? Workforce implications. *Science 358*(6370), 1530–1534.

Brynjolfsson, E. and T. Mitchell (2017b). What can Machine Learning do? Workforce implications. *Science 358*(6370), 1530–1534.

Brynjolfsson, E., T. Mitchell, I. Rahwan, S. H. Bana, D. Rock, and M. Frank (2024). Suitability for Machine Learning Rubric (WorldSML). https://digitaleconomy.stanford.edu/research/suitability-for-machine-learning-rubric-worldsml/, Accessed April 21, 2024.

Brynjolfsson, E., T. Mitchell, and D. Rock (2018). What can machines learn and what does it mean for occupations and the economy? In *AEA papers and proceedings*, Volume 108, pp. 43–47. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Brynjolfsson, E., D. Rock, and C. Syverson (2019). Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics. In A. Agrawal, J. Gans, and A. Goldfarb (Eds.), *The Economics of Artificial Intelligence: An Agenda*, pp. 23–60. Chicago: University of Chicago Press.

Bubeck, S., V. Chandrasekaran, R. Eldan, J. Gehrke, E. Horvitz, E. Kamar, and Y. Zhang (2023). Sparks of Artificial General Intelligence: Early experiments with GPT-4. arXiv preprint arXiv:2303.12712.

Burton, M. D., M. G. Colombo, C. Rossi-Lamastra, and N. Wasserman (2019). The organizational design of entrepreneurial ventures. *Strategic Entrepreneurship Journal 13*(3), 243–255. https://doi.org/10. 1002/sej.1332.

Bylund, P. L. and M. McCaffrey (2017). A theory of entrepreneurship and institutional uncertainty. *Journal of Business Venturing 32*, 461–475.

Caliendo, M., F. Fossen, and A. Kritikos (2012). Trust, positive reciprocity, and negative reciprocity: Do these traits impact entrepreneurial dynamics? *Journal of Economic Psychology 33*(2), 394–409.

Caliendo, M., F. M. Fossen, and A. S. Kritikos (2022). Personality characteristics and the decision to hire. *Industrial and Corporate Change 31*(3), 736–761.

Caliendo, M. and A. S. Kritikos (2019). "I want to, but I also need to": Start-ups resulting from opportunity and necessity. In E. E. Lehmann and M. Keilbach (Eds.), *From Industrial Organization to Entrepreneurship: A Tribute to David B. Audretsch*, pp. 247–265. New York: Springer.

Camuffo, A., A. Cordova, A. Gambardella, and C. Spina (2020). A scientific approach to entrepreneurial decision making: Evidence from a randomized control trial. *Management Science 66*(2), 564–586.

Carbonero, F., J. Davies, E. Ernst, F. M. Fossen, D. Samaan, and A. Sorgner (2023a). The impact of artificial intelligence on labor markets in developing countries: A new method with an illustration for Lao PDR and urban Viet Nam. *Journal of Evolutionary Economics 33*(3), 707–736.

Carbonero, F., J. Davies, E. Ernst, F. M. Fossen, D. Samaan, and A. Sorgner (2023b). The impact of artificial intelligence on labor markets in developing countries: A new method with an illustration for Lao PDR and urban Viet Nam. *Journal of Evolutionary Economics 33*(3), 707–736.

Casas, P. and C. Román (2023). Early retired or automatized? Evidence from the survey of health, ageing and retirement in Europe. *The Journal of the Economics of Ageing 24*, 100443.

Cetindamar, D., T. Lamers, and Y. Zhang (2020). Exploring the knowledge spillovers of a technology in an entrepreneurial ecosystem—The case of artificial intelligence in Sydney. *Thunderbird International Business Review 62*, 457–474.

Chalmers, D., N. G. MacKenzie, and S. Carter (2021). Artificial intelligence and entrepreneurship: Implications for venture creation in the fourth industrial revolution. *Entrepreneurship Theory and Practice 45*(5), 1028–1053.

Chui, M. and S. Malhotra (2018). AI adoption advances, but foundational barriers remain. McKinsey and Company. https://www.mckinsey.com/featured-insights/artificial-intelligence/ai-adoption-advances-but-foundational-barriers-remain.

Coad, A., K. Nielsen, and B. Timmermans (2017). My first employee: An empirical investigation. *Small Business Economics 48*, 25–45.

Cochran, J. J. (Ed.) (2018). *Informs Analytics Body of Knowledge*. John Wiley & Sons.

Cockburn, I., R. Henderson, and S. Stern (2019). The impact of artificial intelligence on innovation: An exploratory analysis. In A. Agrawal, J. Gans, and A. Goldfarb (Eds.), *The Economics of Artificial Intelligence: An Agenda*, pp. 115–148. University of Chicago Press. https://doi.org/10.7208/9780226613475-006.

Collins, B., A. Garin, E. Jackson, D. Koustas, and M. Payne (2019). Is gig work replacing traditional employment? evidence from two decades of tax returns. *Unpublished paper, IRS SOI Joint Statistical Research Program*.

Colombelli, A., E. Paolucci, E. Raguseo, and G. Elia (2024). The creation of digital innovative start-ups: The role of digital knowledge spillovers and digital skill endowment. *Small Business Economics 62*, 917–937.

Comunale, M. and A. Manera (2024). The economic impacts and the regulation of AI: A review of the academic literature and policy actions. IMF Working Paper No. 2024/65.

Congregado, E. (2007). *Measuring Entrepreneurship: Building a Statistical System*, Volume 16. Springer Science & Business Media.

Congregado, E., F. M. Fossen, N. Rubino, and D. Troncoso (2024). Long-term effects of shocks on new opportunity and necessity entrepreneurship. IZA Discussion Paper 16930. https://docs.iza.org/dp16930.pdf.

Congregado, E., A. Golpe, and A. Van Stel (2012). The 'recession-push' hypothesis reconsidered. *International Entrepreneurship and Management Journal 8*, 325–342.

Congregado, E., A. A. Golpe, and S. C. Parker (2012). The dynamics of entrepreneurship: Hysteresis, business cycles and government policy. *Empirical Economics 43*, 1239–1261.

Coombs, C., D. Hislop, S. K. Taneva, and S. Barnard (2020). The strategic impacts of intelligent automation for knowledge and service work: An interdisciplinary review. *Journal of Strategic Information Systems 29*(4), 101600.

Cuellar, M. F., B. Larsen, Y. S. Lee, and M. Webb (2024). Does information about AI regulation change manager evaluation of ethical concerns and intent to adopt AI? *The Journal of Law, Economics, and Organization 40*, 34–75.

Cunningham, C., F. Ederer, and S. Ma (2021). Killer acquisitions. *The Journal of Political Economy 129*(3), 649–702.

Czarnitzki, D., G. P. Fernández, and C. Rammer (2023). Artificial intelligence and firm-level productivity. *Journal of Economic Behavior & Organization 211*, 188–205.

Dahlke, J., M. Beck, J. Kinne, D. Lenz, R. Dehghan, M. Wörter, and B. Ebersberger (2024). Epidemic effects in the diffusion of emerging digital technologies: Evidence from artificial intelligence adoption. *Research Policy 53*(2), 104917.

Dasgupta, P. and J. Stiglitz (1980). Industrial structure and the nature of innovative activity. *The Economic Journal 90*(358), 266–293.

Daugherty, P. R., H. J. Wilson, and P. Michelman (2019). Revisiting the jobs artificial intelligence will create. *MIT Sloan Management Review 60*(4), 0_1–0_8.

Dave, P. (April 13, 2025). FTC v. Meta Trial: The future of WhatsApp and Instagram is at stake. *Wired*. https://www.wired.com/story/meta-ftc-trial-begins-instagram-whatsapp.

Davidsson, P., J. Recker, and F. Von Briel (2020). External enablement of new venture creation: A framework. *Academy of Management Perspectives 34*(3), 311–332.

Davidsson, P. and M. Sufyan (2023). What does AI think of AI as an External Enabler (EE) of entrepreneurship? An assessment through and of the EE framework. *Journal of Business Venturing Insights 20*, e00413.

Dellermann, D., N. Lipusch, P. Ebel, and J. M. Leimeister (2019). Design principles for a hybrid intelligence decision support system for business model validation. *Electronic Markets 29*(3), 423–441.

Dencker, J. C., S. Bacq, M. Gruber, and M. Haas (2021). Reconceptualizing necessity entrepreneurship: A contextualized framework of entrepreneurial processes under the condition of basic needs. *Academy of Management Review 46*(1), 60–79.

Dengler, K. and B. Matthes (2018). The impacts of digital transformation on the labour market: Substitution potentials of occupations in Germany. *Technological Forecasting and Social Change 137*, 304–316.

Desai, S. (2019). Artificial intelligence and entrepreneurship: Some thoughts for entrepreneurship researchers. In D. B. Audretsch, E. E. Lehmann, and A. Link (Eds.), *A Research Agenda for Entrepreneurship and Innovation*, pp. 197–207. Edward Elgar Publishing.

Devlin, J., M.-W. Chang, K. Lee, and K. Toutanova (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv:1810.04805v2*.

Dijk, E. S. R., J. L. Moraga-González, and E. Motchenkova (2024). How do start-up acquisitions affect the direction of innovation? *The Journal of Industrial Economics 72*(1), 118–156.

Dinlersoz, E., C. Dogan, and N. Zolas (2024). Starting up AI. CES Discussion Papers No. 24-09, U.S. Census Bureau, Center for Economic Studies.

Du, W., S. L. Pan, N. Zhou, and T. Ouyang (2018). From a marketplace of electronics to a Digital Entrepreneurial Ecosystem (DEE): The emergence of a meta-organization in Zhongguancun, China. *Information Systems Journal 28*(6), 1158–1175.

Duan, Y., T. S. Hsieh, R. R. Wang, and Z. Wang (2020). Entrepreneurs' facial trustworthiness, gender, and crowdfunding success. *Journal of Corporate Finance 64*, 101693.

Dvouletỳ, O. and D. Nikulin (2023). Dependent self-employed individuals: are they different from paid employees? *Employee Relations: The International Journal 45*(3), 704–720.

Ehret, M. and J. Wirtz (2017). Unlocking value from machines: Business models and the industrial internet of things. *Journal of Marketing Management 33*(1-2), 111–130.

Elia, G., A. Margherita, and G. Passiante (2020). Digital entrepreneurship ecosystem: How digital technologies and collective intelligence are reshaping the entrepreneurial process. *Technological Forecasting and Social Change 150*, 119791.

Eloundou, T., S. Manning, P. Mishkin, and D. Rock (2023). GPTs are GPTs: An early look at the labor market impact potential of large language models. arXiv preprint arXiv:2303.10130.

Emmert-Streib, F. and M. Dehmer (2022). Taxonomy of machine learning paradigms: A data-centric perspective. *Wiley Interdisciplinary Reviews. Data Mining and Knowledge Discovery 12*(5), e1470–n/a. https://doi.org/10.1002/widm.1470.

European Digital SME Alliance (2021). DIGITAL SME position paper on the EU AI Act. https://www.digitalsme.eu/digital/uploads/DIGITAL-SME-Position-Paper-AI-Act-FINAL-DRAFT-1.pdf. Accessed on May 23, 2024.

Fairlie, R. W. (2013). Entrepreneurship, economic conditions, and the great recession. *Journal of Economics & Management Strategy 22*(2), 207–231.

Fairlie, R. W. and F. M. Fossen (2020). Defining opportunity versus necessity entrepreneurship: Two components of business creation. In S. W. Polachek and K. Tatsiramos (Eds.), *Change at Home, in the Labor Market and on the Job*, Volume 48 of *Research in Labor Economics*.

Fairlie, R. W. and J. Miranda (2017). Taking the leap: The determinants of entrepreneurs hiring their first employee. *Journal of Economics & Management Strategy 26*(1), 3–34.

Fan, Y. (2013). Ownership consolidation and product characteristics: A study of the US daily newspaper market. *The American Economic Review 103*(5), 1598–1628.

Fan, Y. and C. Yang (2020). Competition, product proliferation, and welfare: A study of the US smartphone market. *American Economic Eournal. Microeconomics 12*(2), 99–134.

Fedorets, A., S. Kirchner, J. Adriaans, and O. Giering (2022). Data on digital transformation in the German Socio-economic Panel. *Jahrbücher für Nationalökonomie und Statistik 242*(5-6), 691–705.

Felten, E., M. Raj, and R. Seamans (2021). Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses. *Strategic Management Journal 42*(12), 2195–2217.

Felten, E., M. Raj, and R. Seamans (2023). How will language modelers like ChatGPT affect occupations and industries? arXiv preprint arXiv:2303.01157.

Felten, E. W., M. Raj, and R. Seamans (2018). A method to link advances in artificial intelligence to occupational abilities. *AEA Papers and Proceedings 108*, 54–57.

Felten, E. W., M. Raj, and R. Seamans (2019). The occupational impact of artificial intelligence: Labor, skills, and polarization. *NYU Stern School of Business*.

Ferrati, F. and M. Muffatto (2021). Entrepreneurial finance: Emerging approaches using machine learning and big data. *Foundations and Trends® in Entrepreneurship 17*(3), 232–329.

Ferreira, J. J. M., E. Grigoroudis, E. G. Carayannis, and L. Dooley (2023). The role of entrepreneurial ecosystems in technological and social challenges. *R&D Management 53*(4), 563–583.

Fish, K. and P. Ruby (2009). An artificial intelligence foreign market screening method for small businesses. *International Journal of Entrepreneurship 13*, 65–.

Foss, N. J. and P. G. Klein (2012). *Organizing Entrepreneurial Judgment: A New Approach to the Firm*. Cambridge University Press.

Fossen, F. M. (2021). Self-employment over the business cycle in the USA: A decomposition. *Small Business Economics 57*(4), 1837–1855.

Fossen, F. M. and T. J. M. Buettner (2013). The returns to education for opportunity entrepreneurs, necessity entrepreneurs, and paid employees. *Economics of Education Review 37*, 66–84.

Fossen, F. M. and T. Martin (2018). Entrepreneurial dynamics over space and time. *Regional Science and Urban Economics 70*, 204–214.

Fossen, F. M., T. McLemore, and A. Sorgner (2022). The regional impacts of digitalization of work on entrepreneurship in the United States. In M. Keyhani, T. Kollmann, A. Ashjari, A. Sorgner, and C. E. Hull (Eds.), *Handbook of Digital Entrepreneurship*. Edward Elgar Publishing.

Fossen, F. M. and L. Neyse (2024). Entrepreneurship, management, and cognitive reflection: A preregistered replication study with extensions. *Entrepreneurship Theory and Practice 48*(4), 1082–1109.

Fossen, F. M., D. Samaan, and A. Sorgner (2022). How are patented AI, software and robot technologies related to wage changes in the United States? *Frontiers in Artificial Intelligence 5*, 869282.

Fossen, F. M. and A. Sorgner (2019). Mapping the future of occupations: Transformative and destructive effects of new digital technologies on jobs. *Foresight and STI Governance 13*(2 eng), 10–18.

Fossen, F. M. and A. Sorgner (2021). Digitalization of work and entry into entrepreneurship. *Journal of Business Research 125*, 548–563.

Fossen, F. M. and A. Sorgner (2022). New digital technologies and heterogeneous wage and employment dynamics in the United States: Evidence from individual-level data. *Technological Forecasting and Social Change 175*, 121381.

Fountaine, T., B. McCarthy, and T. Saleh (2019). Building the AI-powered organization. *Harvard Business Review*, 63–73.

Franzoni, V., A. Milani, and D. Nardi (2019). Emotional machines: The next revolution. *Web Intelligence 17*(1), 1–7.

Frey, C. B. and M. A. Osborne (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change 114*, 254–280.

Fritsch, M., A. Sorgner, M. Wyrwich, and E. Zazdravnykh (2019). Historical shocks and persistence of economic activity: Evidence on self-employment from a unique natural experiment. *Regional Studies 53*(6), 790–802.

Fritsch, M. and M. Wyrwich (2023). Entrepreneurship in the long-run: Empirical evidence and historical mechanisms. *Foundations and Trends® in Entrepreneurship 19*(1), 1–125.

Furman, J. and R. Seamans (2019). AI and the economy. *Innovation Policy and the Economy 19*, 161–191.

Garbuio, M. and N. Lin (2019). AI as a growth engine for health care startups: Emerging business models. *California Management Review 61*(2), 59–83.

Garin, A., E. Jackson, and D. Koustas (2022). New gig work or changes in reporting? Understanding self-employment trends in tax data. *Understanding Self-Employment Trends in Tax Data (May 19, 2022). University of Chicago, Becker Friedman Institute for Economics Working Paper* (2022-67).

Garin, A., E. Jackson, D. K. Koustas, and A. Miller (2023). The evolution of platform gig work, 2012-2021. Technical report, National Bureau of Economic Research.

Gaspar, R., C. Pedro, P. Panagiotopoulos, and B. Seibt (2016). Beyond positive or negative: Qualitative sentiment analysis of social media reactions to unexpected stressful events. *Computers in Human Behavior 56*, 179–191.

Georgellis, Y. and H. J. Wall (2005). Gender differences in self-employment. *International Review of Applied Economics 19*(3), 321–342.

Georgieff, A. and R. Hyee (2022). Artificial intelligence and employment: New cross-country evidence. *Frontiers in Artificial Intelligence 5*, 832736.

Gerling, C., P. Meier, and C. Koehler (2022). AI meets digital: A critical review on artificial intelligence in digital entrepreneurship. In *ECIS 2022 Research Papers 21, European Conference on Information Systems*.

Ghose, A. and S. P. Han (2014). Estimating demand for mobile applications in the new economy. *Management Science 60*(6), 1470–1488.

Giering, O., A. Fedorets, J. Adriaans, and S. Kirchner (2021). Artificial intelligence in Germany: Employees often unaware they are working with AI-based systems. *DIW Weekly Report 11*(48), 369–375.

Gilbert, R., S. Stern, A. Jaffe, and J. Lerner (2006). Looking for Mr. Schumpeter: Where are we in the competition–innovation debate? *Innovation Policy and the Economy 6*, 159–215.

Giuggioli, G. and M. M. Pellegrini (2023). Artificial intelligence as an enabler for entrepreneurs: A systematic literature review and an agenda for future research. *International Journal of Entrepreneurial Behavior & Research 29*(4), 816–837.

Giuntella, O., J. König, and L. Stella (2023). Artificial intelligence and workers' well-being.

Glikson, E. and A. W. Woolley (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals 14*(2), 627–660.

Glupker, J., V. Nair, B. Richman, K. Riener, and A. Sharma (2019). Predicting investor success using graph theory and machine learning. *Journal of Investment Management 17*(1), 92–103.

Goebel, J., M. M. Grabka, S. Liebig, M. Kroh, D. Richter, C. Schröder, and J. Schupp (2019). The German Socio-economic Panel (SOEP). *Jahrbücher für Nationalökonomie und Statistik 239*(2), 345–360.

Goel, R. K. and M. A. Nelson (2024a). Awareness of artificial intelligence: Diffusion of AI versus ChatGPT information with implications for entrepreneurship. *The Journal of Technology Transfer*.

Goel, R. K. and M. A. Nelson (2024b). Do college anti-plagiarism/cheating policies have teeth in the age of AI? Exploratory evidence from the internet. *Managerial and Decision Economics 45*(4), 1739–2665.

Goel, R. K., J. W. Saunoris, and X. Zhang (2015). Innovation and underground entrepreneurship. *Journal of Technology Transfer 40*(5), 800–820.

Gofman, M. and Z. Jin (2024). Artificial intelligence, education, and entrepreneurship. *The Journal of Finance 79*(1), 631–667.

Goldfarb, A., B. Taska, and F. Teodoridis (2023). Could machine learning be a general purpose technology? A comparison of emerging technologies using data from online job postings. *Research Policy 52*(1), 104653.

Goldfarb, A. and C. Tucker (2019). Digital economics. *Journal of Economic Literature 57*(1), 3–43.

Graef, I. and J. Prüfer (2021). Governance of data sharing: A law & economics proposal. *Research Policy 50*(9), 104330.

Gries, T. and W. Naudé (2022). Modelling artificial intelligence in economics. *Journal for Labour Market Research 56*(1), 12.

Gupta, B. B., A. Gaurav, P. K. Panigrahi, and V. Arya (2023). Analysis of artificial intelligence-based technologies and approaches on sustainable entrepreneurship. *Technological Forecasting and Social Change 186*, 122152.

Haefner, N., J. Wincent, V. Parida, and O. Gassmann (2021). Artificial intelligence and innovation management: A review, framework, and research agenda. *Technological Forecasting and Social Change 162*, 120392.

Hall, J. V. and A. B. Krueger (2018). An analysis of the labor market for Uber's driver-partners in the United States. *ILR Review 71*(3), 705–732.

Hall, J. V. and A. B. Krueger (2019). Reply to the comment by Berg and Johnston. *ILR Review 72*(1), 69–74.

Hannigan, T. R., A. R. Briggs, R. Valadao, M. D. L. Seidel, and P. D. Jennings (2022). A new tool for policymakers: Mapping cultural possibilities in an emerging AI entrepreneurial ecosystem. *Research Policy 51*, 104315.

Harris, E., J. Harris, and M. Beall (2024). Defense in depth: An action plan to increase the safety and security of advanced AI. Gladstone AI, available upon request at.

Harris, Z. S. (1954). Distributional structure. *Word (Worcester) 10*(2-3), 146–162.

Heath, R. (2024, February). Exclusive: States are introducing 50 AI-related bills per week. *Axios*.

Henley, A. (2021). The rise of self-employment in the UK: Entrepreneurial transmission or declining job quality? *Cambridge Journal of Economics 45*(3), 457–486.

High-level Expert Group on Artificial Intelligence (2019). Ethics guidelines for trustworthy AI. European Commission.

Himel, S. and R. Seamans (2017). Artificial intelligence, incentives to innovate, and competition policy. *Antitrust Chronicle 1*(3), 1–10.

Hoberg, G. and G. Phillips (2010). Product market synergies and competition in mergers and acquisitions: A text-based analysis. *Review of Financial Studies 23*(10), 3773–3811.

Hoffman, A. (2004). Reconsidering the role of the practical theorist: On (re)connecting theory to practice in organization theory. *Strategic Organization 2*(2), 213–222.

Hoffmann, J., S. Borgeaud, A. Mensch, E. Buchatskaya, T. Cai, E. Rutherford, and L. Sifre (2022). Training compute-optimal large language models. *arXiv preprint*.

Hoffreumon, C., C. Forman, and N. Van Zeebroeck (2024). Make or buy your artificial intelligence? Complementarities in technology sourcing. *Journal of Economics & Management Strategy 33*(2), 452–479.

Huang, M. H. and R. Rust (2018). Artificial intelligence in service. *Journal of Service Research 21*(2), 155–172.

Huang, M. H., R. Rust, and V. Maksimovic (2019). The feeling economy: Managing in the next generation of AI. *California Management Review 61*(4), 43–65.

Hurst, E. and B. W. Pugsley (2011). What do small businesses do? *Brookings Papers on Economic Activity 2011*(2), 73–118.

International Organization for Standardization (2022). Information technology - Artificial intelligence - Artificial intelligence concepts and terminology. ISO/IEC Standard No. 22989. https://www.iso.org/standard/74296.html.

Jacobides, M. G., S. Brusoni, and F. Candelon (2021). The evolutionary dynamics of the artificial intelligence ecosystem. *Strategy Science 6*(4), 412–435.

Jacquemin, A. and F. Janssen (2015). Studying regulation as a source of opportunity rather than as a constraint for entrepreneurs: Conceptual map and research propositions. *Environment and Planning C: Government and Policy 33*, 846–862.

Jia, J., G. Z. Jin, and L. Wagman (2018). The short-run effects of GDPR on technology venture investment. Technical report, NBER Working Paper 25248.

Jöhnk, J., M. Weißert, and K. Wyrtki (2021). Ready or not, AI comes—an interview study of organizational AI readiness factors. *Business and Information Systems Engineering 63*(1), 5–20.

Kang, C. and D. E. Sanger (2023, October). Biden issues executive order to create A.I. safeguards. The New York Times. https://www.nytimes.com/2023/10/30/us/politics/biden-ai-regulation.html.

Kazakova, S. S., A. A. Dunne, D. D. Bijwaard, J. Gossé, C. Hoffreumon, and N. van Zeebroeck (2020). European Enterprise Survey on the use of technologies based on artificial intelligence. Universite Libre de Bruxelles. https://EconPapers.repec.org/RePEc:ulb:ulbeco:2013/341443.

Keding, C. (2021). Understanding the interplay of artificial intelligence and strategic management: Four decades of research in review. *Management Review Quarterly 71*(1), 91–134.

Kerr, W. R. and F. Robert-Nicoud (2020). Tech clusters. *Journal of Economic Perspectives 34*(3), 50–76.

Keyhani, M., A. Ashjari, A. Sorgner, T. Kollmann, C. E. Hull, and Z. Jamshidi (2022). An introduction to digital entrepreneurship: Concepts and themes. In M. Keyhani, A. Ashjari, A. Sorgner, T. Kollmann, and C. E. Hull (Eds.), *Handbook of Digital Entrepreneurship*. Edward Elgar Publishing.

Kier, A. S. and J. S. McMullen (2018). Entrepreneurial imaginativeness in new venture ideation. *Academy of Management Journal 61*(6), 2265–2295.

Klapper, L. F. and I. Love (2011). Entrepreneurship and development: The role of information asymmetries. *The World Bank Economic Review 25*(3), 448–455.

Klein, E. H. (2023, July). A.I. could solve some of humanity's hardest problems. It already has. (no. 4). The Ezra Klein Show. The New York Times. [Audio podcast episode].

Knight, F. H. (1921). *Risk, Uncertainty and Profit*, Volume 31. Houghton Mifflin.

Kollmann, T., L. Kleine-Stegemann, C. de Cruppe, and C. Strauss (2022). Eras of digital entrepreneurship: Connecting the past, present, and future. In M. Keyhani, T. Kollmann, A. Ashjari, A. Sorgner, and C. E. Hull (Eds.), *Handbook of Digital Entrepreneurship*. Edward Elgar Publishing.

Lamine, W., A. Fayolle, S. Jack, and D. Audretsch (2023). Impact of digital technologies on entrepreneurship: Taking stock and looking forward. *Technovation 126*, 102823.

Lane, M., M. Williams, and S. Broecke (2023). The impact of AI on the workplace: Main findings from the OECD AI surveys of employers and workers. Technical report, OECD Social, Employment and Migration Working Papers No. 288.

Leendertse, J., M. Schrijvers, and E. Stam (2022). Measure twice, cut once: Entrepreneurial ecosystem metrics. *Research Policy 51*(9), 104336.

Lévesque, M., M. Obschonka, and S. Nambisan (2022). Pursuing impactful entrepreneurship research using artificial intelligence. *Entrepreneurship Theory and Practice 46*(4), 803–832.

Levine, R. and Y. Rubinstein (2017). Smart and illicit: Who becomes an entrepreneur and do they earn more? *Quarterly Journal of Economics 132*(2), 963–1018.

Lupp, D. (2023). Effectuation, causation, and machine learning in co-creating entrepreneurial opportunities. *Journal of Business Venturing Insights 19*, e00355.

Madridakis, S. (2017). The forthcoming Artificial Intelligence (AI) Revolution: Its impact on society and firms. *Futures 90*, 46–60.

Malecki, E. J. (2018). Entrepreneurship and entrepreneurial ecosystems. *Geography Compass 12*(3), e12359.

Mamedov, O., Y. Tumanyan, O. Ishchenko-Padukova, and I. Movchan (2018). Sustainable economic development and post-economy of AI. *Entrepreneurship and Sustainability Issues 6*(2), 1028–1040.

Mariani, M. M., I. Machado, and S. Nambisan (2023). Types of innovation and artificial intelligence: A systematic quantitative literature review and research agenda. *Journal of Business Research 155*, 113364.

Maslej, N., L. Fattorini, R. Perrault, V. Parli, A. Reuel, E. Brynjolfsson, J. Etchemendy, K. Ligett, T. Lyons, J. Manyika, J. C. Niebles, Y. Shoham, R. Wald, and J. Clark (2024). The AI index 2024 annual report. Technical report, AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA.

Matsumoto, A., U. Merlone, and F. Szidarovszky (2012). Some notes on applying the Herfindahl-Hirschman Index. *Applied Economics Letters 19*(2), 181–184.

Mazzeo, M. J., K. Seim, and M. Varela (2018). The welfare consequences of mergers with endogenous product choice. *The Journal of Industrial Economics 66*(4), 980–1016.

McElheran, K., J. F. Li, E. Brynjolfsson, Z. Kroff, E. Dinlersoz, L. Foster, and N. Zolas (2024). AI adoption in America: Who, what, and where. *Journal of Economics & Management Strategy 33*, 375–415.

McKenzie, D. and D. Sansone (2019). Predicting entrepreneurial success is hard: Evidence from a business plan competition in Nigeria. *Journal of Development Economics 141*, 102369.

McMullen, J. S. and D. A. Shepherd (2006). Entrepreneurial action and the role of uncertainty in the theory of the entrepreneur. *Academy of Management Review 31*(1), 132–152.

Méndez-Suárez, M., V. Simón-Moya, and J. Muñoz-de Prat (2023). Do current regulations prevent unethical AI practices? *Journal of Competitiveness 15*(3), 207–222.

Mickiewicz, T. and A. Rebmann (2020). Entrepreneurship as trust. *Foundations and Trends® in Entrepreneurship 16*(3), 244–309.

Millán, A., J. M. Millán, and L. Caçador-Rodrigues (2020). Disclosing ‘masked employees’ in Europe: job control, job demands and job outcomes of ‘dependent self-employed workers’. *Small Business Economics 55*, 461–474.

Montes, G. A. and B. Goertzel (2019). Distributed, decentralized, and democratized artificial intelligence. *Technological Forecasting and Social Change 141*, 354–358.

Muehlberger, U. (2007). Hierarchical forms of outsourcing and the creation of dependency. *Organization Studies 28*(5), 709–727.

Mueller, B. (2021). How much will the Artificial Intelligence Act cost Europe? Center for Data Innovation. https://www2.datainnovation.org/2021-aia-costs.pdf.

Nambisan, S. (2017). Digital entrepreneurship: Toward a digital technology perspective of entrepreneurship. *Entrepreneurship Theory and Practice 41*(6), 1029–1055.

Naudé, W. (2023). Destructive digital entrepreneurship. *IZA Discussion Paper* (16483). Institute for the Study of Labor.

Nickell, S. J. (1996). Competition and corporate performance. *The Journal of Political Economy 104*(4), 724–746.

Niebler, V., G. Pirina, M. Secchi, F. Tomassoni, et al. (2023). Towards 'bogus employment?'the contradictory outcomes of ride-hailing regulation in Berlin, Lisbon and Paris. *Cambridge Journal of Regions, Economy and Society*.

Nikolova, M. (2019). Switching to self-employment can be good for your health. *Journal of Business Venturing 34*(4), 664–691.

Nikolova, M., B. Nikolaev, and C. Boudreaux (2023). Being your own boss and bossing others: The moderating effect of managing others on work meaning and autonomy for the self-employed and employees. *Small Business Economics 60*(2), 463–483.

Nolan, M. (2023). Llama and ChatGPT are not open-source – Few ostensibly open-source LLMs live up to the openness claim. *IEEE Spectrum*. https://spectrum.ieee.org/open-source-llm-not-open.

Norval, C., H. Janssen, J. Cobbe, and J. Singh (2021). Data protection and tech startups: The need for attention, support, and scrutiny. *Policy & Internet 13*(2), 278–299.

Noy, S. and W. Zhang (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science 381*(6654), 187–192.

Obschonka, M. and D. B. Audretsch (2020). Artificial intelligence and big data in entrepreneurship: A new era has begun. *Small Business Economics 55*, 529–539.

Obschonka, M. and C. Fisch (2022). Artificial intelligence and entrepreneurship research. In *Oxford Research Encyclopedia of Business and Management*. Oxford University Press.

O'Donnell, P., M. Leger, C. O'Gorman, and E. Clinton (2024). Necessity entrepreneurship. *Academy of Management Annals 18*(1), 44–81.

OECD (2019). Recommendation of the council on artificial intelligence. OECD Legal Instruments.

OECD (2024). Explanatory memorandum on the updated OECD definition of an AI system. OECD Artificial Intelligence Papers, No. 8. https://www.oecd.org/science/explanatory-memorandum-on-the-updated-oecd-definition-of-an-ai-system-623da898-en.htm.

Otis, N., R. Clarke, S. Delecourt, B. Holtz, and R. Koning (2024). The uneven impact of generative AI on entrepreneurial performance. http://dx.doi.org/10.2139/ssrn.4671369.

Özgul, P., M. C. Fregin, M. Stops, S. Janssen, and M. Levels (2024). High-skilled human workers in non-routine jobs are susceptible to AI automation but wage benefits differ between occupations. *arXiv preprint arXiv:2404.06472*.

Packard, M. D., B. B. Clark, and P. G. Klein (2017). Uncertainty types and transitions in the entrepreneurial process. *Organization Science 28*(5), 840–856.

Paolillo, A., F. Colella, N. Nosengo, F. Schiano, W. Stewart, D. Zambrano, I. Chappuis, R. Lalive, and D. Floreano (2022). How to compete with robots by assessing job automation risks and resilient alternatives. *Science Robotics 7*, eabg5561.

Parent-Rocheleau, X. and S. K. Parker (2022). Algorithms as work designers: How algorithmic management influences the design of jobs. *Human Resource Management Review 32*(3), 100838.

Parker, S. C. (2009). *The Economics of Entrepreneurship*. Cambridge University Press.

Phillips, G. M. and A. Zhdanov (2013). R&D and the incentives from merger and acquisition activity. *Review of Financial Studies 26*(1), 34–78.

Raab, M., S. Schlauderer, S. Overhage, and T. Friedrich (2020). More than a feeling: Investigating the contagious effect of facial emotional expressions on investment decisions in reward-based crowdfunding. *Decision Support Systems 135*, 113326.

Ramesh, G. S., T. V. R. Kanth, and D. Vasumaathi (2018). Analysis of customer data using hybridized machine learning technique along with data exploration methods. *International Journal of Engineering and Technology 7*(4), 4388–4392.

Rammer, C., G. P. Fernández, and D. Czarnitzki (2022). Artificial intelligence and industrial innovation: Evidence from German firm-level data. *Research Policy 51*(7), 104555.

Ransbotham, S., P. Gerbert, M. Reeves, D. Kiron, and M. Spira (2018). Artificial intelligence in business gets real. Accessed on May 23, 2024.

Ransbotham, S., S. Khodabandeh, D. Kiron, F. Candelon, M. Chu, and B. LaFountain (2020). Expanding AI's impact with organizational learning. *MIT Sloan Management Review*.

Rasmusen, E. (1988). Entry for buyout. *The Journal of Industrial Economics 36*(3), 281–299.

Reimers, N. and I. Gurevych (2019a). Sentence-BERT: Sentence embeddings using siamese BERT-networks. *arXiv:1908.10084v1*.

Reimers, N. and I. Gurevych (2019b). Sentence-BERT: Sentence embeddings using siamese BERT-networks. *arXiv preprint arXiv:1908.10084*.

Richter, D. and J. Schupp (2015). The SOEP Innovation Sample (SOEP-IS). *Journal of Contextual Economics–Schmollers Jahrbuch 135*(3), 389–400.

Ritsilä, J. and H. Tervo (2002). Effects of unemployment on new firm formation: Micro-level panel data evidence from Finland. *Small Business Economics 19*, 31–40.

Robertson, K. (2024, April). 8 daily newspapers sue OpenAI and Microsoft over A.I. *New York Times*. April 30, 2024, https://www.nytimes.com/2024/04/30/business/media/newspapers-sued-microsoft-openai.html.

Roma, P. and D. Ragaglia (2016). Revenue models, in-app purchase, and the app performance: Evidence from Apple's App Store and Google Play. *Electronic Commerce Research and Applications 17*, 173–190.

Ruggles, S., S. Flood, M. Sobek, D. Brockman, G. Cooper, S. Richards, and M. Schouweiler (2023). Ipums USA: Version 13.0 [dataset].

Saheb, T. and T. Saheb (2023). Topical review of artificial intelligence national policies: A mixed method analysis. *Technology in Society 74*, 102316.

Samoili, S., M. López Cobo, E. Gómez, G. De Prato, F. Martínez-Plumed, and B. Delipetrev (2021). AI Watch. Defining artificial intelligence 2.0. EUR 30873 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-76-42648-6, doi:10.2760/019901, JRC126426.

Scherpenzeel, A. (2011). Data collection in a probability-based internet panel: How the LISS panel was built and how it can be used. *Bulletin of Sociological Methodology 109*(1), 56–61.

Schiavone, F., M. C. Pietronudo, A. Sabetta, and F. Bernhard (2022). Designing AI implications in the venture creation process. *International Journal of Entrepreneurial Behavior & Research 29*(4), 838–859.

Schwab, A. and Z. Zhang (2019). A new methodological frontier in entrepreneurship research: Big data studies. *Entrepreneurship Theory and Practice 43*(5), 843–854.

Serrano, C. (2024). AI can level global playing field. Knowledge@HEC, https://www.hec.edu/en/ai-can-level-global-playing-field. Accessed 5/24/2024.

Shane, S. (2000). Prior knowledge and the discovery of entrepreneurial opportunities. *Organization Science 11*(4), 448–469.

Shepherd, D. A. and A. Majchrzak (2022). Machines augmenting entrepreneurs: Opportunities (and threats) at the nexus of artificial intelligence and entrepreneurship. *Journal of Business Venturing 37*(4), 106227.

Shepherd, D. A., K. Wennberg, R. Suddaby, and J. Wiklund (2019). What are we explaining? A review and agenda on initiating, engaging, performing, and contextualizing entrepreneurship. *Journal of Management 45*(1), 159–196.

Short, C. E. and J. C. Short (2023). The artificially intelligent entrepreneur: ChatGPT, prompt engineering, and entrepreneurial rhetoric creation. *Journal of Business Venturing Insights 19*, e00388.

Shrestha, Y. R., S. M. Ben-Menahem, and G. von Krogh (2019). Organizational decision-making structures in the age of artificial intelligence. *California Management Review 61*(4), 66–83.

Song, A. K. (2019). The digital entrepreneurial ecosystem – a critique and reconfiguration. *Small Business Economics 53*, 569–590.

Sorgner, A. (2023). COVID-19 and entrepreneurship. In K. F. Zimmermann (Ed.), *Handbook of Labor, Human Resources and Population Economics*. https://doi.org/10.1007/978-3-319-57365-6˙321-1.

Sorgner, A. and M. Wyrwich (2022). Calling Baumol: What telephones can tell us about the allocation of entrepreneurial talent in the face of radical institutional changes. *Journal of Business Venturing 37*(5), 106246.

Spigel, B. (2017). The relational organization of entrepreneurial ecosystems. *Entrepreneurship Theory and Practice 41*(1), 49–72.

Stam, E. (2015). Entrepreneurial ecosystems and regional policy: A sympathetic critique. *European Planning Studies 23*(9), 1759–1769.

Stam, E. and B. Spigel (2018). Entrepreneurial ecosystems. In *The SAGE Handbook of Small Business and Entrepreneurship*, pp. 407–422.

Stix, C. (2020). A survey of the European Union's artificial intelligence ecosystem. https://doi.org/10.48550/arXiv.2101.02039.

Storey, D. J. (1991). The birth of new firms–does unemployment matter? A review of the evidence. *Small Business Economics 3*, 167–178.

Stucke, M. E. and A. P. Grunes (2016). *Introduction: Big Data and Competition Policy*. Oxford University Press.

Subramaniam, M., B. Iyer, and V. Venkatraman (2019). Competing in digital ecosystems. *Business Horizons 62*(1), 83–94.

Sussan, F. and Z. J. Acs (2017). The digital entrepreneurial ecosystem. *Small Business Economics 49*, 55–73.

Sutskever, I., O. Vinyals, and Q. V. Le (2014). Sequence to sequence learning with neural networks. *arXiv:1409.3215v3*.

Sweeting, A. (2010). The effects of mergers on product positioning: Evidence from the music radio industry. *The Rand Journal of Economics 41*(2), 372–397.

Tangredi, S. J. and G. e. Galdorisi (2021). *AI at War: How Big Data, Artificial Intelligence, and Machine Learning Are Changing Naval Warfare.* Naval Institute Press.

Thurik, A. R., M. A. Carree, A. Van Stel, and D. B. Audretsch (2008). Does self-employment reduce unemployment? *Journal of Business Venturing 23*(6), 673–686.

Timoteo, M., B. Verri, and Y. Wang (2021). Ethics guidelines for artificial intelligence: Comparing the European and Chinese approaches. *China and WTO Review 2*, 305–330.

Tomy, S. and E. Pardede (2018). From uncertainties to successful start ups: A data analytic approach to predict success in technological entrepreneurship. *Sustainability 10*(3), 602.

Torres, P. and P. Godinho (2022). Levels of necessity of entrepreneurial ecosystems elements. *Small Business Economics 59*, 29–45.

Torrisi, S., A. Gambardella, P. Giuri, D. Harhoff, K. Hoisl, and M. Mariani (2016). Used, blocking and sleeping patents: Empirical evidence from a large-scale inventor survey. *Research Policy 45*(7), 1374–1385.

Townsend, D. M. and R. A. Hunt (2019). Entrepreneurial action, creativity, & judgment in the age of artificial intelligence. *Journal of Business Venturing Insights 11*, e00126.

Townsend, D. M., R. A. Hunt, J. Rady, P. Manocha, and J. H. Jin (2023). Are the futures computable? Knightian uncertainty and artificial intelligence. *Academy of Management Review*. https://doi.org/10.5465/amr.2022.0237.

University of Essex (2023). Understanding society: Innovation panel, waves 1-15, 2008-2022. [data collection]. 12th Edition. UK Data Service, Institute for Social and Economic Research. SN: 6849, http://doi.org/10.5255/UKDA-SN-6849-15.

U.S. Department of Justice and the Federal Trade Commission (2023). Merger guidelines. https://www.ftc.gov/system/files/ftc˙gov/pdf/2023˙merger˙guidelines˙final˙12.18.2023.pdf.

Van der Zwan, P., R. Thurik, I. Verheul, and J. Hessels (2016). Factors influencing the entrepreneurial engagement of opportunity and necessity entrepreneurs. *Eurasian Business Review 6*, 273–295.

Van Gelderen, M., J. Wiklund, and J. S. McMullen (2021). Entrepreneurship in the future: A delphi study of ETP and JBV editorial board members. *Entrepreneurship Theory and Practice 45*(5), 1239–1275.

Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin (2023). Attention is all you need. *arXiv:1706.03762v7*.

Venancio, A., W. Picoto, and I. Pinto (2023). Time-to-unicorn and digital entrepreneurial ecosystems. *Technological Forecasting & Social Change 190*, 122425.

Wang, W., W. Chen, K. Zhu, and H. Wang (2020). Emphasizing the entrepreneur or the idea? The impact of text content emphasis on investment decisions in crowdfunding. *Decision Support Systems 136*, 113341.

Wang, W., H. Zheng, and Y. J. Wu (2020). Prediction of fundraising outcomes for crowdfunding projects based on deep learning: A multimodel comparative study. *Soft Computing 24*(11), 8323–8341.

Wang, Y., L. Q. Yue, N. Rajagopalan, and B. Wu (2024). The entry-deterring effects of synergies in complementor acquisitions: Evidence from Apple's digital platform market, the iOS app store. *Strategic Management Journal 45*(13), 2791–2817.

Webb, M. (2020). The impact of artificial intelligence on the labor market. Working Paper, Stanford University. https://www.michaelwebb.co/webb˙ai.pdf.

Weber, M., M. Beutter, J. Weking, M. Böhm, and H. Krcmar (2022). AI startup business models: Key characteristics and directions for entrepreneurship research. *Business & Information Systems Engineering 64*(1), 91–109.

Wilkinson, G. (2018). General data protection regulation: No silver bullet for small and medium-sized enterprises. *Journal of Payments Strategy & Systems 12*(2), 139–149.

Wilson, H. J., P. Daugherty, and N. Bianzino (2017). The jobs that artificial intelligence will create. *MIT Sloan Management Review 58*(4), 14.

Winkler, C., B. Hammoda, E. Noyes, and M. Van Gelderen (2023). Entrepreneurship education at the dawn of generative artificial intelligence. *Entrepreneurship Education and Pedagogy 6*(4), 579–589.

Wollmann, T. G. (2018). Trucks without bailouts: Equilibrium product characteristics for commercial vehicles. *The American Economic Review 108*(6), 1364–1406.

Wurth, B., E. Stam, and B. Spigel (2022). Toward an entrepreneurial ecosystem research program. *Entrepreneurship Theory and Practice 46*(3), 729–778.

Wurth, B., E. Stam, and B. Spigel (2023). Entrepreneurial ecosystem mechanisms. *Foundations and Trends® in Entrepreneurship 19*(3), 224–339.

Yao, S., D. Yu, J. Zhao, I. Shafran, T. Griffiths, Y. Cao, and K. Narasimhan (2024). Tree of thoughts: Deliberate problem solving with large language models. *Advances in Neural Information Processing Systems 36*.

Yuan, H., R. Y. K. Lau, and W. Xu (2016). The determinants of crowdfunding success: A semantic text analytics approach. *Decision Support Systems 91*, 67–76.

Zellweger, T. and T. Zenger (2023). Entrepreneurs as scientists: A pragmatist approach to producing value out of uncertainty. *Academy of Management Review 48*(3), 379–408.

Appendix

Appendix A: Do apps play follow the leader

Download cutoffs

We filter the dataset by the number of downloads to minimize the inclusion of apps in our models that have no business ambitions. Filtering out apps under 5k downloads was arbitrarily selected. In Table 15, we run the baseline regression from equation (4) with different download thresholds up to a million or more downloads required. The results are robust to changes in download threshold with the tested variables increasing significance from the 5% level, to the 1% level when the filter is set to a million or more downloads.

Table 15: Quadratic regression varying download thresholds, outcome: similarity

	$> 5k$	$> 10k$	$> 50k$	$> 100k$	$> 500k$	$> 1M$
HHI/10k	-0.274**	-0.271**	-0.246**	-0.237**	-0.212**	-0.213***
	(0.122)	(0.119)	(0.106)	(0.098)	(0.081)	(0.075)
HHI/10k squared	0.717**	0.705**	0.630**	0.610**	0.540**	0.544***
	(0.315)	(0.309)	(0.276)	(0.255)	(0.208)	(0.193)
Constant	0.594***	0.595***	0.598***	0.599***	0.602***	0.604***
	(0.006)	(0.006)	(0.005)	(0.005)	(0.004)	(0.004)
R-squared	0.0220	0.0220	0.0192	0.0181	0.0155	0.0162

425,817 observations; Google Playstore 2019
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
HHI is divided by 10k to scale the results

Control variables

All of our variables are endogenous at some level, including similarity and HHI. We break the control variables into tiers, adding a single tier at a time, to observe tradeoffs and sensitivity in the model. Anything that is an inherent feature of the app, such as whether users interact, is minimally endogenous. Strategic decisions by the app developers are endogenous. We break the control variables into three groups which are designed characteristics (minimally endogenous), decided about the app (minimally endogenous), and strategy (expected endogeneity). Table 16 shows that the estimated U-shape is robust. In general, apps become less similar to other apps in their categories over time, larger apps are more similar to other apps in their categories, 4 star and 3 star apps are more similar to other apps in their categories than 1 star apps, apps containing ads or digital purchases are more similar to apps in their categories, apps that are not free are more similar to other apps in their categories, apps rated mature, containing gambling, violence, or drug references are more similar to other apps in their categories, while apps with adult language or suggestive themes are less similar. There is some evidence, lacking robustness, that apps that share location or receive the editor's choice distinction are similar to other apps in their categories.

Table 16: Quadratic regression with control variables, outcome: similarity

	(1) Base	(2) Designed Characteristics	(3) External Decisions	(4) Strategy
HHI/10k	-0.274(0.122)**	-0.262(0.125)**	-0.258(0.124)**	-0.243(0.125)*
HHI/10k squared	0.717(0.315)**	0.675(0.324)**	0.666(0.321)**	0.643(0.322)*
creation year		-0.000883(0.000376)**	-0.000799(0.000370)**	-0.000880(0.000312)***
users interact		0.00156(0.00219)	0.00149(0.00215)	0.000936(0.002106)
location sharing		0.00307(0.00212)	0.00308(0.00207)	0.00318(0.00165)*
info sharing		0.000376(0.00189)	0.000335(0.001836)	-0.000127(0.001809)
size (gigabytes)		0.103(0.044)**	0.102(0.044)**	0.0618(0.0347)*
5 star rating			0.00268(0.00533)	0.000182(0.005152)
4 star rating			0.00739(0.00278)**	0.00459(0.00263)*
3 star rating			0.00564(0.00210)***	0.00399(0.00202)*
2 star rating			0.00244(0.00199)	0.00219(0.00199)
not rated			0.00412(0.00310)	0.00217(0.00287)
editor's choice			0.00451(0.00240)*	0.00101(0.00190)
ads				0.00456(0.00233)*
digital purchases				0.00701(0.00232)***
not free				0.00786(0.00175)***
mature				0.00539(0.00316)*
gambling				0.0225(0.0045)***
violence				0.00962(0.00395)**
adult language				-0.00595(0.00226)**
suggestive themes				-0.00720(0.00328)**
drug references				0.00130(0.00192)*
Constant	0.594(0.006)***	0.596(0.006)***	0.590(0.006)***	0.587(0.006)***
R-squared	0.0220	0.0266	0.0281	0.0394

425,817 observations; Google Playstore 2019
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
HHI is divided by 10k to scale the results

Text similarity example

To check and demonstrate the effectiveness of our procedure to detect similarity among app descriptions, we select an app and run a search for the 4 nearest neighbors based on similarity scores. Our results show the strength of the embedding approach and that copying occurs among developers.

Reference App

Flight Simulator: Fly Plane 3D. Developed by i6 Games in 2014. 50M downloads.

Flight Simulator: Fly Plane 3D is an awesome new 3D Airplane Simulator game, become the pilot and fly your commercial jet to the destination. Guide and steer your plane through all of the waypoints to ensure you head to the correct destination, go through all of the waypoints and land at your destination airport within the time limit to earn yourself more pilot stripes. When arriving at your destination zone, slow the plane down and prepare for landing, be careful not to crash! Guide your plane towards the runway and park within the marked zone to complete the level. While parking your plane, be very careful to avoid the buses, helicopters etc parked along the run way.

Nearest Neighbor

Fly Flight Landing Simulator. Developed by RG Games in 2017. 10k downloads.

Fly Flight Landing Simulator: Fly Plane 3D is an awesome new 3D Airplane Simulator game, become the pilot and fly your commercial jet to the destination. Guide and steer your plane through all of the waypoints to ensure you head to the correct destination, go through all of the waypoints and land at your destination airport within the time limit to earn yourself more pilot stripes. When arriving at your destination zone, slow the plane down and prepare for landing, be careful not to crash! Guide your plane towards the runway and park within the marked zone to complete the level. While parking your plane, be very careful to avoid the buses, helicopters etc parked along the run way.

2nd Nearest Neighbor

Flight Simulator Fly plane. Developed by GR Mobile Games in 2016. 100k downloads.

Flight Simulator Fly plane is an awesome new 3D Airplane Simulator game, become the pilot and fly your commercial jet to the destination. Guide and steer your plane through all of the waypoints to ensure you head to the correct destination, go through all of the waypoints and land at your destination airport within the time limit to earn yourself more pilot stripes. When arriving at your destination zone, slow the plane down and prepare for landing, be careful not to crash! Guide your plane towards the runway and park within the marked zone to complete the level. While parking your plane, be very careful to avoid the buses, helicopters etc parked along the run way.

3rd Nearest Neighbor

Airplane Flight Simulator 3D. Developed by i6 Games in 2014. 1M downloads.

Airplane Flight Simulator 3D is an awesome 3D Air Plane Sim flying game. Start off with learning how to fly a plane, you simply take off from airport island and flying the plane to another Airport Runway. You must then park the airplane on the runway. In this aircraft flying simulation game, complete the levels within 50% of the given time and collect all the rings to earn all 3 badges; so racing through the levels can get you all of the badges. As you develop your skills and progress through the game, you have the opportunity to unlock bigger and faster planes. Be very careful, there's a lot of passengers relying on your pilot skills.

4th Nearest Neighbor

Taxi & Bus Driver 3D. Developed by Han's Games in 2015. 10k downloads.

With its smooth full hd graphics and small size, Taxi & Bus Driver 3D has never been this much fun to pick up passengers and drift at the same time. Show your excellent driving skills and pick up your passengers and reach the destination on time while drifting. You must also improve your driving skills in challenging conditions with the active traffic system and take care to obey traffic rules.

Revisiting the Uber effect

Berger et al. (2018) comparison

Table 17: Comparison of results with Berger et al. (2018)

	This paper 2009-2022 (1)	This paper 2009-2015 (2)	Berger et al. (2018) 2009-2015 (3)
		Panel A All drivers	
$Uber_{it}$(=1) on earnings	-0.053	-0.067	-0.140**
	(0.057)	(0.073)	(0.059)
$Uber_{it}$(=1) on wage	-0.041	-0.068	-0.185***
	(0.061)	(0.079)	(0.067)
$Uber_{it}$(=1) on hrly earnings	-0.042	-0.041	-0.116*
	(0.034)	(0.043)	(0.059)
$Uber_{it}$(=1) on hrly wage	-0.045	-0.048	-0.162**
	(0.037)	(0.050)	(0.065)
$Uber_{it}$(=1) on labor supply	-0.030	0.023	0.004
	(0.043)	(0.058)	(0.081)
		Panel B Wage-employed drivers	
$Uber_{it}$(=1) on earnings	-0.079	-0.125	-0.183***
	(0.067)	(0.089)	(0.065)
$Uber_{it}$(=1) on wage	-0.051	-0.094	-0.155**
	(0.069)	(0.093)	(0.064)
$Uber_{it}$(=1) on hrly earnings	-0.085*	-0.101**	-0.125*
	(0.045)	(0.049)	(0.068)
$Uber_{it}$(=1) on hrly wage	-0.058	-0.070	-0.097
	(0.049)	(0.057)	(0.071)
$Uber_{it}$(=1) on labor supply	-0.077*	-0.038	-0.076
	(0.046)	(0.069)	(0.098)

Notes: Columns (2) and (3) should be identical. All values are controlled for MSA and year FE, additional MSA controls to include mean ln earnings, the share of workers with a college degree, the female share of the labor force, ln MSA population, the unemployment rate, and the share of the labor force that falls in four age categories (16–25, 26–39, 40–54, and 55), and for MSA-specific time trends. Robust standard errors reported in parentheses. Statistical significance is denoted by:*** p<0.01, ** p<0.05, * p<0.10.

Appendix C: Artificial intelligence and entrepreneurship

Glossary

ABS Annual Business Survey AI Artificial Intelligence

AIOE Artificial Intelligence Occupational Exposure

BTOS Business Trends and Outlook Survey

CCPA California Consumer Protection Act

CP Computerization Probability

DEE Digital Entrepreneurial Ecosystem

EE Entrepreneurial Ecosystem

EIDES European Index of Digital Entrepreneurship Systems

GDPR General Data Protection Regulation

GPT Generative Pre-trained Transformer

ICT Information, Communication, and Technology

IoT Internet of Things

IT Information Technology

LBD Longitudinal Business Database

LLM Large Language Model

MIP Mannheim Innovation Panel

ML Machine Learning

OECD Organisation for Economic Co-operation and Development

SML Suitability for Machine Learning

SOEP German Socio-economic Panel

SOEP-IS German Socio-economic Panel – Innovation Survey