Quantifying the Impact of AI on Productivity and Labor Demand: Evidence from U.S. Census Microdata[1]

Dean Alderucci
Center for AI Analysis of Patents
Carnegie Mellon University

Sagar Baviskar
Carnegie Mellon University

Lee Branstetter
Carnegie Mellon University and NBER

Nathan Goldschlag
U.S. Census Bureau

Eduard Hovy
Language Technologies Institute, CMU
and University of Melbourne

Andrew Runge
Duolingo

Prasana Tambe
Wharton School of Management
University of Pennsylvania and NBER

Nikolas Zolas
U.S. Department of State

Draft

Abstract:

After decades of disappointment, artificial intelligence (AI) has entered a new era of rapidly advancing capabilities that are likely to raise productivity and reshape demand for labor within and across firms and industries. Accurately measuring these effects has been difficult due to a lack of detailed, firm-level data on AI innovation. We address that challenge by using a combination of machine learning algorithms to parse the text of U.S. patent grants and assess the degree to which they are AI-related. This approach indicates that AI-related invention is more pervasive than many previous analyses have suggested. We match our data on AI patenting to U.S. Census microdata collected on the innovating firms. We then perform an event study using these matched data to gauge the impact of these innovations on firm labor demand, labor productivity growth, and wage dispersion. We find that AI-related inventions are positively associated with growth in employment and increases in output per worker. In contrast, there is less evidence that AI invention is expanding income inequality. We also discuss ongoing efforts to measure the diffusion of AI technology to U.S. firms through the movement of workers trained at the technology frontier.

[1] Any opinions and conclusions expressed herein are those of the author(s) and do not necessarily represent the views of the U.S. Census Bureau, the National Bureau of Economic Research, or any research funder. All results have been reviewed to ensure that no confidential information is disclosed. We thank Andrew Toole and staff at USPTO for generously making their AI patent database available for comparison and also thank conference participants at the AEA meetings. The DRB codes for this project are: DRB-B0027-CED-20190205, CBDRB-FY19-414, CBDRB-FY20-105, CBDRB-FY22-182, CBDRB-FY22-CES007-004, and CBDRB-FY24-CES022-005. Branstetter and Hovy gratefully acknowledge funding from the National Science Foundation (Grant 2241237), the Heinz Endowments, and the CMU Digital Transformation and Innovation Center.

.

1. Introduction

After decades of disappointment, artificial intelligence (AI) has entered a new era of rapidly advancing capabilities (Agarwal et al., 2022; Baily et al., 2023). In the popular media, a vigorous debate is being waged between proponents of these new technologies, who believe they will bring a new era of rapid productivity growth and widespread prosperity, and skeptics, who fear an era of mass joblessness and wage stagnation for all but a small cognitive elite (Brynjolfsson and McAfee, 2014; Baily et al., 2023; Suskind, 2020).

Earlier industrial revolutions were characterized by significant and persistent increases in productivity growth that boosted living standards across the income distribution. Despite growing hype and concern over AI applications across the economy, aggregate productivity growth remains stuck at slow rates (Benzell et al., 2022). Will AI fail to live up to the enthusiasm of its advocates or are we merely in the early stages of an innovation and adoption process that will take years or decades to unfold? This paper seeks to address this question by examining the vanguard of firms that are already introducing AI-related innovations into the marketplace. If these early movers and innovators are already reaping significant productivity gains, then this augurs well for the ultimate positive impact of AI on the broader. economy. In our final section, we also describe ongoing work, in which we are beginning to examine the potential role played by Ph.D.-level academic experts in the development of AI-related new goods and services By linking these experts to the firms that employ them, we may obtain empirical leverage around the difficulty of measuring AI innovation that does not result in patents.

Our work complements many recent streams of research. One stream seeks to measure AI adoption and use through direct surveys of large, representative samples of U.S. firms (Zolas et al., 2019; McElheran et al., 2024). These extremely valuable efforts have shed useful light on AI adiption, but they have not yet demonstrated a strong relationship between AI adoption and faster productivity growth, and it will take time before these survey data acquire a sufficiently long time series dimension such that researchers can apply the usual econometric techniques for discerning plausibly causal effects from nonexperimental data. A second stream applies randomized controlled trials or quasi-experimental methods to measure the impact of AI on worker productivity in a particular work context (see Brynjolfsson, Li, and Raymond, 2023; Korinek, 2023; and Noy and Zhang, 2023). Some of these papers have found convincing evidence of a causal impact of AI adoption on productivity, but the results may not generalize from the particular work contexts in which these experiments or quasi-experiments were conducted. A third stream uses data on the recruitment of specialized labor to measure AI use and AI-related innovation (see Babina et al., 2022, 2024). Prominent papers in this stream have found positive effects of this investment on output and product innovation, but have failed to find robust evidence that investments in AI use led to increases in productivity growth. The productivity growth effects we document are potentially broader in scope than those found in the experimental literature and point to the optimistic possibility that AI could eventually lead to a significant and persistent acceleration in productivity growth across a broad range of industries.

Patent Data as an AI Innovation Indicator

Firms that succeed in using artificial intelligence to create new goods and services have an incentive to patent at least some of their inventions. If they fail to do so, other firms can copy their innovations without penalty or use patents to block the original innovator from applying their inventions in the marketplace. This means that firms throughout the world are filing thousands of AI-related patents with the U.S. Patent and Trademark Office (Cockburn et al., 2019; Webb et al., 2019; Giczy et al., 2021). By law, the vast majority of patent documents become public 18 months after filing, even if they are still being adjudicated by the patent office. Each patent application is supposed to provide sufficient detail such that the invention could be replicated by an individual who is proficient in the technology.

Patents are classified according to the technology they contain, and the U.S. Patent and Trademark Office has created a detailed taxonomy containing several hundred patent classes (and thousands of subclasses). However, if we only count patents in those classes and subclasses specifically and primarily associated with "artificial intelligence," we may vastly undercount the true scope and scale of AI-related invention. The reason a narrow focus is insufficient is precisely the reason that this emerging technology is so important – the applications, current and potential, of artificial intelligence, machine learning, "big data," and AI are so broad as to encompass virtually the entire economy. Similar machine learning algorithms can be used in combines, cars, jets, banks, insurance companies, and travel agencies, and the patents that apply them to these different domains could show up across a vast range of classes, including classes associated with the domain of application of the patent, such as patent classes traditionally associated with tractors and combines. A patent search procedure that examined only patents in the USTPO's designated patent class for artificial intelligence or focused only on a handful of key words will miss far more than it captures. Our results to date show this to be the case, and we demonstrate that in the next section.

Fortunately, developments in natural language processing permit a different approach. By training machine learning algorithms to parse the full text of patent documents, we can, in principle, capture nearly all AI-related inventions, regardless of the patent class to which they may be assigned by patent examiners. This gives us a broader, more complete data window through which to view the rise of AI-related invention. Our research team created an ensemble of algorithms that can quickly sort through millions of patents, assigning each to an appropriate AI-related bin, or not, as the text of the patent dictates. Our paper describes, in some detail, the creation and training of this algorithm, and efforts are currently underway to update our set of AI-related patents through patents granted by mid-2024.

Once we have correctly identified AI-related patents, we possess a highly granular map of AI invention that identifies the corporate owners of the patents, the geographic location of the inventors who created the new technology, and the time when the invention was originally conceived. We can thus trace AI-invention across time, geographic space, and industry space, and identify the firms that are most active in creating inventions in this new domain. We present some of the interesting findings one can infer from tabulations of the raw data in this map and quantify the impact of AI on American invention, labor demand, firm productivity, and other key variables.

Mapping to Census Data

In fulfillment of its institutional mission, the U.S. Census Bureau (hereafter Census) maintains detailed data on U.S. enterprises, including privately held enterprises that make few public disclosures about their business operations. Furthermore, the data collected by Census provides additional details that go beyond what publicly traded firms are required to disclose. For instance, Census gathers data not only on firms but also on the establishments – the individual business units – created by these firms, with identifiers that link them all to the parent firm; these data track mergers, acquisitions, and divestitures, ensuring that the mapping from parent firms to establishments remains current. Among other things, Census surveys establishments on their output (broken down by industry), material inputs, capital investment, and employment. Census has also invested in a detailed mapping that connects the patent owners (assignees) listed in USPTO patent data records to their own firm identification codes, adjusting for mergers and acquisitions This allows us to match data on the creation of AI technology by individual firms to the possible impact of that technology on their productivity growth. This, in turn, allows us to infer the impact of AI-invention on productivity. Our data, linking tens of thousands of AI-related patents to thousands of firms, are rich enough that we can explore the potential heterogeneity of this impact across time, industries, and firms of different types. Some observers worry that AI technologies will create a kind of winner-take-all industry dynamic, in which the benefits of the technology accrue to a small number of firms that implement it first. We can directly address this concern, establishing whether the productivity impact of AI invention is concentrated in a handful of leading firms or more broadly observed . In this way, we can determine to what extent AI is fulfilling the promise of enhanced productivity predicted by its proponents.

2. Artificial Intelligence as a General Purpose Technology

Leading consulting firms (McKinsey Global Institute, 2016), leading academics (Brynjolfsson and McAfee, 2014; Agarwal et al., 2021; Baily et al., 2023; Elandou et al., 2024), and leading CEOs have all claimed that industrial firms have now entered a new era, enabled by AI and related technologies, that can fundamentally transform business across the economy. If we translate the enthusiasm of these proponents into the language of economics, they are contending that rise of AI constitutes the emergence of a new "general purpose technology" (GPT). The economic literature on general purpose technologies (Bresnahan and Trajtenberg, 1995; Helpman, 1998) can be useful in helping us think about the impact of all this as a new technology of technological change. The confluence of AI-related technologies that has opened up these new opportunities are broadly applicable, potentially touching nearly every industry in the global economy. But effective application of this suite of new technologies may often require product and industry-specific knowledge.[2] So, every industry and firm needs to invest in new technology and new capabilities (Brynjolfsson, Rock, and Syverson, forthcoming). This suggests the possibility of inferring the impact of AI on the direction and pace of inventive activity in the U.S. by using patent data.

Precisely because the potential impact of AI is so broad, and the need for complementary innovation to adapt it to the vast array of contexts in which it can be applied is so great, it may take time for this impact to manifest itself in aggregate statistics. As David (1991) pointed out in his famous comparison of electricity and IT, the basic inventions necessary for the electrification of U.S. manufacturing were created decades before they were widely applied. Eventually, this process resulted in a significant and persistent surge in U.S. productivity growth. Of course, some firms and industries were in the vanguard of the process of electrification of manufacturing processes, achieving important gains years or even decades before other firms and industries.

The publication in 2019 of *The Economics of Artificial Intelligence: An Agenda* represented an important step forward in economists' investigation of this important technological shift. Several chapters in this volume take the view that AI is a general purpose technology, and explore that idea in a number of different ways. Agrawal, Gans, and Goldfarb (2019, 2021) advance the idea that AI algorithms are "prediction machines," enabling agents throughout the economy to better forecast outcomes, reducing uncertainty in a range of economic tasks. Agrawal, McHale, and Oettl (2019) suggest an important application of this idea, in terms of enabling firms to find productive new combinations of existing technologies. Cockburn et al. (2019) also view AI as a new technology for invention. However, a number of chapters in the volume point to the wide range of unanswered questions, and Raj and Seamans (2019), in particular, emphasize the importance of more firm-level data on AI invention, AI adoption, and its effects. Two recent nationwide surveys on technology adoption by U.S. firms from the Census, found that diffusion of AI and its various components (such as machine learning, machine vision, natural language processing, voice recognition and automated guided vehicles) is very low, with 3-6% of eligible firms adopting some form of AI, and with adoption skewed heavily towards the largest set of firms in the economy (Zolas et al. 2019; McElheran et al. 2024; Acemoglu et al. 2022). The limitations of the survey include the fact that the timing of adoptions are not identified. We seek to address this lack of firm-level data, by using patent to detect both evidence of AI inventions, as well as the timing of this invention and then linking data on patenting to the rich microdata maintained by Census on the innovating firms.

3. Using Machine Learning Techniques to Identify AI Patents

3.a Identifying AI Patents

Our goal is to construct a large dataset of AI-related and non AI-related patents in order to train a robust ensemble of machine learning models that can identify AI innovation across a wide range of invention domains. A natural approach to this task would be to leverage a schema such as the US Patent Classification (USPC), International Patent Classification (IPC) or Cooperative Patent

[2] Even the use of very general purpose technologies, like ChatGPT, requires some investment of time and thought on the part of the user. See the fascinating discussionin Korinek (2023).

Classification (CPC) systems, which have categories for different types of AI-related innovations. The World Intellectual Property Organization and other patent offices employ machine learning models to assign patent application to IPC classifications. (WIPO 2019) Other researchers have taken various machine learning approaches to emulating IPC classification. (Benites et. Al 2018; Grawe et. Al 2017). However, in practice this results in a few challenges. These classification systems use a deep, multi-label taxonomy in which a single patent can be assigned to multiple categories. The nebulous concept of AI is therefore captured using several non-overlapping classification labels. Unfortunately, this classification system is not perfect, with many inventions that describe intelligent technologies falling outside of the explicit AI categories in these systems, or assigned to categories in which there is a mix of both AI and non-AI inventions. Likewise, some inventions that leverage AI technologies as a component in a larger invention may not be assigned to an AI classification at all.

Much prior work in this area has also leveraged keyword-matching methods to identify patents that describe the use of particular algorithms, techniques, or areas of AI research. These methods achieve *high precision* in identifying AI patents, but suffer from *low recall* and may detect only a small fraction of the total patents leveraging AI technologies. For example, a component in an invention might be described as a model trained on a particular kind of data, without specifying the type of model in order to avoid restricting the scope of the patent. These inventions often clearly describe AI-related technologies, but do so without easily identifiable keywords, resulting in them being left out of these approaches. This is common in inventions in which the AI component is only a part of the greater invention, as opposed to the core of the proposed innovation.

Training a model using patents labeled automatically by either of these methods could therefore result in a model that is highly biased towards particular phrases, or towards certain subfields of AI innovation. We instead adopt a manual annotation approach, identifying patents that describe or utilize AI technologies with a focus on covering a wide range of different applications and domains.

The next several subsections of the paper provide some important details on our use of natural language processing (NLP) and machine learning (ML) techniques to build the patent dataset. It includes terms and descriptions that may be unfamiliar to readers who have not studied the language processing techniques now widely used in computer science. However, we include these details in the interests of providing transparency in terms of the techniques we have used and in the interests of facilitating comparisons with other efforts to identify AI-related invention using patent data. Economists and other social scientists less interested in these details may wish to proceed to the final portion of this section, *3.h*, found on page 10.

3.b Dataset Construction

We construct our dataset using a semi-supervised iterative approach, in which we train an ensemble of machine learning models to identify AI patents and then apply this ensemble to the full USPTO patent corpus from 1990-2018. We then manually review selections at different decision thresholds, which can help us quickly find and validate high confidence AI and non-AI patents for our dataset, as well as find and more carefully review patents that the model classified with lower confidence.

We start with an initial set of 330 patents that had been manually identified as AI-related under a previous effort led by Dean Alderucci, a member of our research team with years of experience as a patent lawyer. For our initial assignment model, we augment this dataset with patents from class 706, the artificial intelligence category in the U.S. patent classification system to serve as the positive class, and select the negative class from all other categories in the system, excluding a few other AI-related categories. In this initial step, we train a simple decision tree model on a tf.idf weighted bag of words representation of the claims sections of the patent and extract the term features used by the model. We retrieve a ranked list of patents from a Lucene-based text index using these keywords and then label 50-100 documents at several points through the ranked list using the labeling process described in the previous section. This allows us to quickly find 1) AI patents that do not have these keywords and thus appear lower in the list and 2) non-AI patents that ranked highly in the list, therefore helping us expand our AI and non-AI sets beyond what keyword matching might have identified.

At the end of this first phase, we have a set of 1,200 evenly split AI and non-AI patents, which we now use to train a Support Vector Machine (SVM) model (Cortes and Vapnik, 1995). We applied the trained model to the full artificial intelligence class of the U.S. patent classification system and ranked them by how likely they were to be AI patents according to the model. We vet the highest confidence decisions, using a quicker process to verify the model's predictions, and more carefully review 350 low confidence predictions generated by the model. After this second phase, we have a set of 2,000 patents, evenly split between AI and non-AI which we can use to run a broader set of experiments.

3.c Classification

We then train and apply several alternative classifiers to label each patent as "AI" or "non-AI". We compare the performance of these modules by measuring their F1 scores on a held-out test set of patents. The classifiers implement two types of machine learning: statistical and neural.

We split the set of 2,000 patents (1000 each AI and non-AI) into three subsets: assigning 10% for validation, 10% for evaluation, and 80% for training, maintaining even numbers of AI and non-AI patents in all sets.

3.d Use of Statistical Models

Using the text from each patent's Abstract, Claims, and Description fields, we represent the patent as a bag-of-words vector with tf.idf weight scores. We train the following statistical machine learning models using the Python scikit-learn toolkit: 1) Naïve Bayes, 2) logistic regression, 3) random forest (Ho, 1995), and 4) linear support vector machine. We tune the hyperparameters for each model using three-fold cross-validation (in which 1/3 of the data is held out as a validation set while the model is trained on the remaining two thirds), and average performance scores across the three trained models to determine the best parameters.

3.e Use of Neural Models

We implement three neural classification models in addition to the above models. As before, we use the text of the Abstract, Claims, and Description fields only. We use the fastText toolkit built by Facebook (https://en.wikipedia.org/wiki/FastText) to build skipgram word embeddings (Mikolov et al. 2013) on the text of all patents from 1990–2018. (Doing so provides word embeddings tailored specifically to the language of patents.) These embeddings are the first layer in all our neural models that form vectors for use in later layers. In each neural model, the embeddings forming the input text are converted via several layers in the neural network to a final AI or Non-AI decision. We apply our three neural transformation architectures separately to each of the three patent sections, and at the end concatenate the final representations for the three sections together to obtain a final representation of the document, which is then used by the final classifier to predict whether the document is AI or not.

Convolutional neural network (CNN): A convolutional neural network, also called a time-delay neural network (Waibel et al. 1989) applies a sliding window left-to-right over the text to extract local patterns that may be useful for classification. After traversing the whole sequence of text, we pool the extracted features using a max pooling layer to obtain a single vector, which is then used for the final classification decision. We use multiple convolutional layers concurrently with different sliding window sizes, or kernel widths, in order to extract patterns of varying lengths, and concatenate the final representations together for each section of the patent.

Recurrent neural network (RNN): An RNN model (Rumelhart et al. 1988) applies a function sequentially to a series of input word embeddings, and includes a mechanism to allow the model to retain information encountered early in the sequence for later application. The function outputs a vector after reading each input word, and after processing the whole sequence we again use a max pooling function to combine all the outputs into a single vector representation. We apply a single recurrent layer in both directions over the text using a Gated Recurrent Unit (GRU) (Cho et al. 2014). We concatenate the final outputs in both directions to create the final vector representation for the classification decision.

Hierarchical attention network (HAN) (Yang et al. 2016): The model proceeds 'top-down', breaking the patent down into sentences, and subsequently words. For each sentence, a recurrent layer is passed over its word embeddings and the outputs for the words are combined using an attention layer. This network layer allows the model to dynamically weight each of the outputs from the recurrent function, essentially allowing the model to decide how important each word in the input is. The weighted outputs are then summed to create a single representation for the sentence, which is passed to the final classification decision.

We tune the hyperparameters of our neural models using ten-fold cross-validation and generate predictions by taking the average of the predictions for the ten trained models.

3.f Model Performance and Comparison

Table 1 Statistical and neural model performance on AI patent classification

Model	Test Micro F1
Naïve Bayes	.765

Logistic Regression	.885
Random Forest	.910
Linear SVM	.890
CNN	.901
RNN	.905
HAN	.911

The evaluation performance of each model is shown in Table 1, using the F1 score, which weights Precision (accuracy) and Recall (coverage) equally. Average model performance is fairly strong on the evaluation set, ranging from .81 to .91 F1 across the various different model architectures.

However, the predicted likelihood that any given document is AI can widely vary across the different models, so we use these differences to identify areas where our models may be overfitting to the data due to the data size. For each pair of models, we identify the 25–30 documents with the largest differences in model prediction scores. These indicate documents that may be very challenging for one of the models and therefore might most benefit from further manual review. We collect a set of 328 patents and manually review them, creating a small "challenge" set which will help indicate how well any model is generalizing to a set of challenging sub-areas. While in some cases we find examples that indicate strong overfitting to particular terms, such as "training" or "network", many documents in the challenge set come from conceptually similar technology areas that would be difficult to classify as AI or Non-AI even for humans. These include particular algorithm formulations that resemble rule-based AI systems, patents that contain boilerplate language about how an invention may incorporate statistical modeling components, and certain image processing techniques commonly used as inputs to more advanced AI systems that in themselves might be considered not-AI.

3.g Model Ensembling

It is fairly common to obtain better and more stable classification performance by merging (ensembling) the results of classifiers built on different principles. Using the above challenge set, we combine the various model architectures to create a more robust model ensemble. For each model pair from the challenge set selection step, we assign a portion of the labeled documents for that pair to either a new training, validation, or test set. We further supplement these new sets with documents from the original training and evaluation sets. For the evaluation set, we use the entire original evaluation set, creating a new evaluation set with 298 documents. For the validation set, we use 50 each of AI and non-AI labeled documents from the original training set. Finally we add the rest of the original training set (1700 documents) to the challenge set documents to construct the ensemble training set.

As inputs to the ensemble model, we use both the tf.idf bag-of-words representation (used by the statistical models) and the AI likelihood scores produced by each trained model. In addition we use one feature for each statistical model and eleven features for each neural model (namely the predictions from each of the ten cross-fold models plus their average). For the ensemble models, we experiment with a random forest, a support vector machine using an RBF kernel, and logistic regression, tuning the parameters for the models on the validation set.

We report micro F1 statistics on the validation and test sets in Table 2. We show results on our challenge-augmented datasets using the model ensemble features as well as just using the tf.idf representation of the patents. The ensemble features produce large performance increases: 0.7–6.0 F1 scores on the validation set and 4.3–8.0 F1 scores on the test set. The more-varied representations learned by the different models therefore prove to be quite valuable in discerning between the more challenging documents in these sets.

Table 2: Validation and Tests of Various ML models

Model	Features	Validation Micro F1	Test Micro F1
Logistic Regression	TFIDF	.887	.792
Logistic Regression	TFIDF+Ensemble	.908	.872
Random Forest	TFIDF	.901	.839
Random Forest	TFIDF+Ensemble	.908	.886
SVM	TFIDF	.873	.819
SVM	TFIDF+Ensemble	.937	.862

Using the random forest model, we label all patents granted by the USPTO between 1990–2018 and define two sets based on prediction score thresholds at 0.7 and 0.95. These contain 146,952 and 52,896 patents respectively, and are listed below. A full representation of the machine learning process to identify the AI patents is given in Figure 1.

3.h Comparison with Other Studies of AI-Related Patenting

A rapidly growing number of other studies also use patent data to identify AI-related inventions. One early study that influenced our own work is Cockburn et al. (2019). These authors select patents from two U.S. Patent Classes: 706, which is the designated patent class for artificial intelligence patents, and 901, which is the designated patent class for robotics. In addition, the authors search patent titles for a small number of specific keywords, such as “neural networks,” and include patents containing these keywords regardless of which Combining these two approaches, the authors find 13,615 unique “AI” patents granted between 1990 and 2014. Webb et al. (2019) take a broadly similar approach with a focus on narrower categories of advanced technology patenting, identifying over 2,000 patents related to “machine learning” and over 4,000 patents related to “neural networks.” In contrast, our methods identify between 52,896 and 146,952 AI-related patents granted between 1990 and 2018, depending on the stringency of our definition. The larger numbers we identify reflect the differences between our empirical approach and that taken by these other studies. First, we examine the core text of the patent (including the patent claims and the patent description), not just the title and abstract. Second, our ensemble of machine learning algorithms allows us to detect AI-related patents even if they do not explicitly include the small number of keywords used by, for instance, Cockburn et al. (2019). Third, we do not limit our purview to a handful of patent classes – and one of the most important findings of this paper is the (very) large number of AI-related inventions that are classified into patent classes associated with the domain of application of the invention rather thant the patent class traditionally associated with the "upstream" AI concepts.

On the other hand, we *exclude*, by design, a number of the patents Cockburn et al. (2019) include in their designated AI patent set, and this difference is easiest to describe in the context of the "robotics" patents from class 901. Patents in this category can include "hardware" patents that represent important advances in the arms, sensors, and other components of a robot "body," but have little to do with the algorithmic advances or software innovations that have equipped recent generations of robots with more artificial intelligence, and it is the latter categories of invention on which we seek to focus. We do not exclude all patents with a hardware component, but our methods are designed to exclude those that are not somehow also related to the ability of a robot or other AI-enabled system to learn, think, and respond intelligently to its environment.

Our focus is therefore quite different from a number of other closely related patent studies, which include patent classes or keywords that give significant weight to hardware innovations in robotics, computing machinery, semiconductors, and related domains, including Keisner et al. (2015), De Prato et al. (2018), and Van Roy et al. (2019). The latter study identifies roughly 155,000 patents around the world that the authors link to AI, which would appear roughly the same as the number of AI patents we identify using our less stringent definition of AI-relatedness. However, a significant number of the patents identified by Van Roy et al. (2019) are "pure hardware" patents of the kind we seek to exclude. These authors use an extensive set of keywords to indentify AI patents, but that list explicitly includes keywords related to hardware inventions describing robot bodies but not brains. Van Roy et al. (2019) also look far beyond the USPTO, identifying "AI-related" patents with the requisite keywords anywhere in the world. While it is certainly useful to look at AI-patenting outside the United States, a large fraction of the total AI patents Van Roy et al. identify are granted by the China's patent office to entities which patent their AI-related inventions in China and nowhere else. This is problematic, because China's patent system is, to put it mildly, still a work in progress, and there is strong reason to believe that China's patent office grants patents to inventions whose limited novelty might not qualify for patent protection elsewhere (Branstetter et al., 2018). Well-publicized efforts by the Chinese government to promote AI and generous subsidies for domestic patenting have opened the gates to a flood of domestic patenting of questionable quality. Van Roy et al.'s list of top AI patenting entities includes organizations like China's State Grid Company, which is well known to China energy experts for its monopolistic control over much of the nation's power grid, but is not globally recognized as a leader in AI. Van Roy et al.'s data suggests that Japan is close follower behind China in terms of AI patenting, but this almost surely reflects high levels of Japanese patenting in hardware categories of the kinds discussed above. Many AI industry insiders view Japanese firms as seriously lagging behind American counterparts in AI capabilities.

To the best of our knowledge, the only other papers explicitly using machine learning techniques to identify AI-related patenting are Mann and Puttman (2018) and Giczy et al. (2021). The former looks at a much broader category of invention than just AI-related patenting and seek to measure all innovations related to automation, and purely mechanical devices like an automatic taco machine and a hair dye applicator qualify as automation patents, even though they would not appear to qualify as AI patents. In the latter study, the USPTO authors identified eight AI component technologies (e.g. natural language processing, machine learning, as well as hardware applications) and then trained a machine learning model for each of these eight components to generate the Artificial Intelligence Patent Database (AIPD). We incorporate these sets of patents (with 95% confidence) into our analysis and explicitly compare and contrast the firm-level outcomes between the two patent sets. The AIPD (95% confidence) consists of approximately 290,000 patents (compared to the 52,896 in our 95% confidence set), with approximately 46,000 patents overlapping across both datasets. Despite the overlap, the results from our more narrowly-defined set of patents differ substantially from that of the USPTO, and we believe our approach offers significant advantages.

4. Mapping AI Invention in Geographic Space and Time

Our methodology yields a large number of patents originating from many different technological fields. We highlight some of the interesting patterns in the data before we begin our econometric analysis of the matched patent-census data set.[3]

We start by noting the counts of AI patents over our time period, which currently spans through 2018. Figure 2 plots the growth of AI patents found in the set of USPTO data. The blue bar represent the counts with 70% confidence, while the orange bar represents the counts with 95% confidence. For the purposes of our analysis, we will predominantly focus on the latter (95% confidence). The number of AI-related innovations increased dramatically between 2000 and 2018. In 2000, there were 539 AI-related patents granted (95% confidence) and this number increased to more than 6,300 in 2018.

We can utilize the geographic data for the assignees or inventors of these patents to chart the origination of these patents. A number of well-known experts have raised the concern that the U.S. is now lagging behind other countries, especially China, in terms of research investments in AI (Lee, 2018). Figure 3 assigns a geographic location to each patent, based on the location of its inventors. Since we are using patents granted by the USPTO, these data will have a well-documented U.S. bias, since inventors tend to take out patents in their home market first and patent selectively abroad. However, even allowing for this bias, the overwhelming dominance of U.S. inventors in the AI space is striking. Conversely, the tiny numbers of AI patents ascribed to mainland Chinese inventors is equally striking. Our data identify more AI patents created by Taiwanese inventors than on the Chinese mainland. Given the enormous size of the U.S. economy – still larger than China's at market exchange rates by a conventional measure – and the highly developed nature of the AI economy within the U.S., if Chinese inventors have valuable new technology that they eventually wish to deploy abroad, they would seem to be running a nontrivial risk in not patenting that technology in the U.S. These data do not support the notion that the U.S. is falling behind in AI invention.,

[3] We are currently in the process of updating our data on AI-related patents through the most recent cohort of U.S. patent grants.

In addition to the country codes of the assignees, we can utilize the geographic data of the inventors to plot where in the U.S. the AI-related innovations are taking place. This picture is rendered in Figure 4. Perhaps unsurprisingly, we find that AI patenting activity is concentrated in the high-tech areas of Silicon Valley, Seattle, Austin, and New York.

Finally, we can plot the technology classes that have seen the most activity in terms of AI. Figure 5 plots the 25 most common USPC codes found on the patents of AI innovations. The most common USPC codes consist of 382, Image Analysis, and Data Processing (USPC 702 – 709). However, it is clear from this graph that AI patents appear to be widely distributed across a very large number of patent classes, with patents applying AI to particular domains showing up in the classes associated with those domains of application. This is, perhaps, what we would expect if AI truly is a general purpose technology, and these findings lend support to our methodology for the identification of the patents.

In order to gain a better understanding of how AI is impacting the larger economy, it is necessary to link the AI patents and accompanying assignees with data on firm performance. We can do this through an existing mapping between U.S. patents granted to U.S. firms and detailed firm-level data housed at the U.S. Census Bureau. The steps for creating this linkage is described below.

5. AI Invention at the Firm Level

5.a Linking Patents to Firms Using the Patent to Census Crosswalk

Once the patents have been identified, we can link the U.S. assignees of these innovations to firm-level microdata using an extended USPTO Patent to Census Crosswalk found in Fort et al. (2020) and first generated by Graham et al. (2018). This crosswalk builds upon previous efforts by Kerr & Fu (2008) and Balasubramanian & Sivadasan (2011) that have linked the NBER patent database to Census data. In the Graham et al. (2018) approach, the authors bring in the full USPTO database from PatentsView and incorporate a triangulation approach that combines fuzzy name and address matching of the assignee with the firm name and address found in the Census Business Register (BR), and inventor links in the Longitudinal Employer Household Dynamics (LEHD) data, that match employees with their employers. The extension of this approach occurs in the years prior to widespread availability of the LEHD and in the most recent years. The resulting crosswalk improves upon previous efforts to link patent data with Census data that relied solely on the assignee matches. The inventor links are used to disambiguate many-to-one firm-level matches and thus provide a cleaner and more accurate linkage then previous efforts.

The result of this crosswalk is a patent-to-Census firm identifier for all U.S. patents granted between 2000 and 2019. Using this crosswalk, we are able to link approximately 85% of AI-related innovations. These firms are then linked to the available sets of Census microdata described below. The resulting set of firms and their AI innovations form the basis of the analysis.

In addition to the firm-level data across various Census datasets, we also incorporate worker-level data from the LEHD to construct measures of earnings ratios for employees at the 90th, 50th, and 10th percentiles of their firms' wage distributions. The LEHD provides us with quarterly earnings data compiled from state unemployment insurance (UI) wage records and the Quarterly Census of Employment and Wages (QCEW) data (see Abowd et al. (2009) for notes on data construction and source files)). The coverage of this data is broad (roughly 98% of private sector employers submit wage records), but state level participation has varied over time, ranging from approximately 20 states at the start of our analysis (1997) to 49 states by the later years of our sample.

The firm links from the USPTO match are able to identify the set of workers and their corresponding earnings in each quarter. These earnings consist of traditional hourly or salary earnings, along with bonuses and other potentially irregular large payments (including income derived from the exercise of stock options by employees at high-tech AI-inventing firms). Since these irregular payments can have a significant impact on the measured wage distribtuion, we winsorize the earnings distribution at 99 percent of the state-year-quarter distribution (approximately $125,000 per quarter on average). We also only keep the "full-quarter" earnings of workers (defined as being employed in the previous quarter and the following quarter) and, to construct the annual measures of earnings ratio for each firm, we first construct the earnings ratio and then take the four quarter average.

5.b The Characteristics of AI-Inventing Firms

Before delving into our analysis, it will be helpful to review the types of firms that are innovating in AI and how they might differ from the typical manufacturing firm, or even typical innovating firm. We categorize each firm as follows and compare their 2018 firm characteristics. For firms with at least 1 patent in their portfolio, these firms are classified as patenting firms, while firms with at least 1 AI patent (95% confidence from our classifier) are categorized as an AI firm. Table 3 compares their baseline characteristics in 2018.

Table 3: Summary Statistics by Firm Type, 2018

Variable	**ALL US FIRMS**	**PATENTING FIRMS**	**FIRMS WITH AI PATENTS**
Mean Employment	16.67	154.1	657.3
Mean Age	14.42	21.95	22.81
Mean Payroll per Employee	40.19	67.4	103.4
Mean Revenue per Employee	207.1	297.0	356.1
% Multi-unit	3.21%	21.7%	44.4%
% Multinational	1.24%	17.4%	32.7%
Observations	6,000,000	90,000	2,600

The first thing to notice is how much larger firms with AI patents are relative to their counterparts. Firms with patents are on average nearly 10x larger than the average U.S. firm, while firms with at least 1 AI patent areon average nearly 40x times larger. Innovating firms are also significantly older on average, more capitalized (pay higher earnings per employee), and significantly more likely to be multi-unit and multinational.These findings correspond with national indicators on AI adoption found in Zolas et al. (2021) and McElheran et al. (2022), who find that adoption of AI is skewed towards very large and older firms.

When we limit the sample to manufacturing firms, similar scale effects appear and are even more stark, as the size difference between the average manufacturing firm and manufacturing firms with an AI patent are 75x larger. We see similar trends with regards to multi-unit and multinational status, as well as capitalization and productivity per worker.

Table 4: Summary Statistics by Manufacturing Firm Type, 2017[4]

Variable	ALL US MANU FIRMS	PATENTING MANU FIRMS	MANU FIRMS WITH AI PATENT
Mean Employment	69.8	517.8	5,253
Mean Age	19.62	27.34	32.52
Mean Payroll per Employee	41.79	62.48	93.72
% Multi-unit	8.0%	33.1%	79.1%
% Multinational	9.3%	40.5%	79.5%
Manufacturing Statistics			
Production Worker Share	71.7%	65.7%	59.8%
Mean Capital Stock	6,529	57,400	875,900
Mean VA per Employee	123.2	173.0	298.5
Mean Capital per Employee	35.92	72.41	87.30
Mean TFP	2.081	1.980	2.097
Mean Sales per Employee	228.6	332.6	607.8
Observations	236,000	20,500	450

Looking at the distribution of these same statistics, the differences are just as stark. Figure 6 plots the Kernel density for size and earnings across the 3 firm types in 2018 and then looks at the distribution of the different manufacturing characteristics, namely: value-added per worker, production worker share, total factor productivity (TFP) and capital-labor ratio. We can see that the distribution of employment across the firm types systematically differ as the majority of U.S. firms tend to be relatively small. Patenting firms have a more normalized distribution, with a slightly fatter tail, while firms with AI patents are primarily concentrated in large firms (firms with 1000+ employees). In the distribution of payroll per employee, we find that the distribution across all firms is normal, with lowest variability and highest means in firms with AI patents. This same pattern persists in one of the measures of productivity for manufacturing firms (value-added per employee), but with more similarity across the firm types with regards to TFP. Our capital-labor ratios also show that firms with AI patents tend to have higher capital-labor ratios.

[4] We focus on 2017 for manufacturing firms as that was the most recent Economic Census year at the time the original analyses was undertaken, ensuring adequate coverage of the key manufacturing variables, such as capital stock, value-added, TFP and production worker share. Note that the data are collected at the establishment-level and aggregated to the firm-level. TFP is calculated as the employment-weighted average TFP of each manufacturing establishment within a firm.

Our measure of production worker share is intended to capture the labor demand across the different firm types, with production workers typically being classified as lower skilled. Our density plots reveal consistent patterns as AI patenting firms tend to have the lowest labor demand for production workers, relative to patenting and all manufacturing firms.

To summarize, firms with AI patents differ across a number of important dimensions from the typical U.S. firm and from innovating firms (defined as firms with patents). They are on average much larger, older, better capitalized, more productive and significantly more likely to be multinational and multi-unit. They also employ fewer lower production workers on average. The next section describes how we estimate the impact of AI on both productivity and labor demand.

6. Estimating the Impact of AI on Productivity and Labor Demand

The next section describes our methodology for assessing the impact of AI innovations on productivity and labor demand. This section is descriptive as the decision to innovate in artificial intelligence is likely to be endogenous with other firm decisions that could potentially impact the outcomes that we are measuring. In the previous section, we demonstrated that the firms that invest in AI-related innovations are categorically different from the vast majority of firms in our data. These differences are unlikely to result entirely from the decision to innovate in AI, but are due to a combination of several factors, many of which are related. Thus, while our language may, in places, suggest a causal relationship between AI invention and other variables of interest, we are not, at this point, making any strong claims about causality.

We first take a standard approach, using the advent of AI as our treatment and employing simple linear regression models to measure the impact of this treatment on four separate outcomes: employment, revenue per employee, value-added per employee and production worker share. The sample of firms used in these regressions is the entire set of manufacturing firms in the U.S. between 1997 and 2018. In these regressions, we make no effort to "match" our AI firms with firms that are very similar in observable characteristics, but do not create AI patents.

Our second approach attempts to partially control for endogeneity and confounding factors that we cannot measure in the data by matching each firm in our set of AI-inventing firms with a closely-related counterpart that has not generated AI patents. We then measure the before and after effects of the AI patent. In both approaches, we find that AI innovations are positively and significantly associated with higher employment, more revenue per employee, greater value-added per employee and fewer production workers (lower-skilled labor). We also find that the strength of this treatment grows in the years following the initial innovation, with employment being approximately 25% higher and revenue being 40% higher five years after treatment.

6.a Revenue per Worker and Value Added

If AI is raising firm productivity, than it should lead not only to more patents but also higher levels of revenue, by increasing product quality, and thus product demand, or it should lead to lower production costs. Our approach to the measurement of these effects can be motivated by a standard Cobb-Douglas production function, in which counts of AI patents or a dummy variable equal to 1 when AI patenting begins are introduced as a separate regressor. Thus, suppressing time subscripts, output can be described as:

$$Q_i = K_i^{\alpha} L_i^{\beta} A_i^{\varphi} e^{\varepsilon_i}$$

taking the logs of both sides and normalizing output by employment gives us

$$q_i - l_i = \alpha k_i + \varphi a_i + \varepsilon_i$$

Here q is output, k is capital, l is labor input, and a is the firm-level measure of AI innovation. We allow for the existence of individual effects which are potentially correlated with right hand side regressors, such that

$$\varepsilon_{it} = \lambda_{it} + u_i$$

The traditional procedure is to use a "within" panel estimator to eliminate the individual effect, which is what we do. The coefficient of interest, φ, picks up the effect of changes in firm i's own AI intensity on its productivity or outcome variable. Our primary outcome measures will be revenue per employee and value-added per employee. Our firm controls include the firm's capital stock, multinational status, age, as well as individual and yearly fixed effects.

Our primary variable of interest is an indicator variable (1/0) showing whether the firm has obtained at least one AI patent or not. When incorporating firm fixed effects, our identification hinges on firms which transition into AI innovation over our time period, 1997-2018. The results from our initial specification applied to manufacturing firms can be found in Table 5 below

Table 5: Impact of AI Innovations on Firm Productivity, 1997-2018 (manufacturing only)

	(1)	(2)	(3)	(4)	(5)	(6)
	Ln Total Value of Shipments per Employee		Ln Value Added Per Employee		Ln Total Factor Productivity (TFP)	
AI Treatment (1/0)	0.272*** (0.0322)		0.225*** (0.0337)		0.0830*** (0.0225)	
IHS AI Patents		0.148*** (0.0161)		0.104*** (0.0166)		0.0565*** (0.0119)
Ln Capital Stock	0.310*** (0.00195)	0.310*** (0.00195)	0.301*** (0.00203)	0.301*** (0.00203)	-0.0633*** (0.000888)	-0.0633*** (0.000888)
Age Bins	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,124,000	1,124,000	1,124,000	1,124,000	1,124,000	1,124,000
R-squared	0.674	0.674	0.921	0.921	0.706	0.706

Robust Standard Errors clustered at the 4-digit NAICS industry level. *, ** and *** denotes significance at the 5%, 1% and 0.1% respectively. Each regression includes a constant, age bins, multi-unit and multinational indicator controls, which are not displayed here. Note that our Multi-Unit regressor drops out from the within-firm specification as the firm-identifier for multi-unit status does not change.

Taking our standard linear approach, we find that the presence of AI patents among manufacturing firms has a positive and significant effect on both revenue per employee and value-added per employee. According to the coefficient values from our fixed effects regression, an AI patent is associated with a 27.2% increase in total value of sales (revenue) per employee, a 22.5% increase in value-added per employee, and an 8% increase in TFP. When we convert our treatment into a continuous measure using the inverse hyperbolic sine transformation of the cumulative total number of AI patents held by the firm (i.e., the AI patent "stock"), the measured impact of AI patents on productivity remains positive and significant.

It is also worth comparing the treatment coefficients across the different types of AI assignments. In the next exercise, we analyze the treatment effect for firms that have AI patents with 75% confidence, firms with AI patents at 50% confidence, firms with the USPTO's AIPD patent designation (95% confidence), firms with AI patents that were identified by both our algorithm and USPTO's algorithm and firms with AI patents identified by USPTO's algorithm only.

Table 6: Impact of different classifiersfor AI innovations on labor productivity and TFP, 1997-2018 (manufacturing only)

	(1) Ln Total Value of Shipments per Employee	(2) Ln Value Added Per Employee	(3) Ln Total Factor Productivity (TFP)
Baseline AI Treatment (1/0)	0.272*** (0.0322)	0.225*** (0.0337)	0.0830*** (0.0225)
AI Treatment (75% Confidence)	0.255*** (0.0253)	0.199*** (0.0267)	0.0589*** (0.0166)
AI Treatment (50% Confidence)	0.232*** (0.0219)	0.178*** (0.0234)	0.0421** (0.0147)
AIPD Treatment (95% Confidence)	0.221*** (0.0186)	0.179*** (0.0204)	0.0291* (0.0123)
Baseline & AIPD Treatment	0.271*** (0.0328)	0.222*** (0.0344)	0.0802*** (0.0228)
AIPD Only Treatment	0.117*** (0.0186)	0.0933*** (0.0201)	0.000596 (0.0117)

Robust Standard Errors clustered at the 4-digit NAICS industry level. *, ** and *** denotes significance at the 5%, 1% and 0.1% respectively. Each regression is identical to the specification in Table 5, but with a different binary treatment effect. AIPD denotes patents indentified as AI-related by the USPTO as described in Giczy et al., 2021.

Across all of the various outcome measures, our baseline AI classifier has the strongest measured association with manufacturing firm productivity. As the classifier broadens to include patents where we are less confident in their connection to AI, we see a declining measured association with sales per employee, value-added per employee, and total factor productivity. We see a similar decline in the magnitude when using patents in the USPTO's AI patent database that are not also included in our own.

The next section describes our method for estimating the changes to labor demand.

6.b Labor Demand

Matching firm-level data on AI patenting to firm-level data on employment allows us to make potentially significant contributions to the current understanding of the impact of AI on the labor market. Due to the limited availability of firm-level data, some prior research has used textual descriptions of the task content of various occupations and information on the emerging capabilities of AI algorithms to “predict” the possible impact of AI on employment (e.g., Brynjolfsson, Mitchell, and Rock, 2018). We find this line of research useful, but because it is not based on actual observations of real firms altering their labor demand as they create and deploy AI technology within the firm, this line of research is necessarily speculative. In contrast, our data match allows us explore how leading firms adjust their employment as they create AI technology.[5]

A large literature in labor economics documents the impact of earlier generations of information technology on the relative demand for skilled labor (Autor et al., 1998; Autor et al., 2003). As earlier waves of automation and computerization advanced, much evidence suggests that demand for the most skilled workers increased but demand for the less skilled workers decreased, accounting for a significant degree of the rise in income inequality that has characterized U.S. labor markets since the 1970s.[6] Many observers worry that AI will continue, and perhaps even exacerbate these longstanding trends.

As we noted in our introduction, the richness of our data enables us to probe for the existence of these effects at the level of the firm in a number of different ways. One empirically feasible approach is to estimate an equation along the lines of Berman, Bound, and Griliches (1994), who derived an equation explaining the nonproduction worker share of total employment in manufacturing industries as a function of relative wages, capital intensity, and a series of additional variables proxying for skill-biased technological change, including industry-level measures of R&D and computer investment. Following their basic logic, though not their exact specfication, we can use firm-level data from Census to estimate the following equation:

$$dln(S_i) = \beta_0 + \beta_1 dln(K_i) + \beta_2 dln(A_i) + \varepsilon_i$$

where S measures changes in the nonproduction worker *share* of total firm *i* employment over some period of time, modeled as a function of changes in the log of the capital stock of firm *i*, K_i

[55] The work of McElheran et al. (2024) and Babina et al. (2022, 2024) are examples of studies that use firm-level data on AI adoption or use to measure their impact on firm-level outcomes, including employment. Our work complements this recent research.

6 The evolution has been complicated by employment “polarization,” with American job creation in recent decades concentrated in high-skill intensive jobs and low-skill intensive jobs. A hollowing out of middle-skill, middle-income jobs has led to wage declines for workers in these categories – unable to compete for high-skill intensive jobs they have fallen down the skill ladder. See Autor and Dorn (2013) for a recent explanation.

and a measure of changes in the AI-intensity of firm innovation, A_{it}, over the same time period.[7] In our case, we will use both the dummy variable indicating the inception of AI-related patenting in firm *i* and the inverse hyberbolic sine transformation of the firm's AI patent stock, rather than a log transformation, as a way of contending with the well known "zero problem." Our current specification lacks the relevant wage data to control directly for changes in production worker and nonproduction worker wages, so we incorporate individual effects into our model in attempt to control for these within-firm changes. In all cases we take natural logs of the key variables.

Table 7: Impact of AI Innovations on Labor Demand, 1997-2018 (manufacturing only)

	(1)	(2)	(3)	(4)
	Ln Production Worker Share			
AI Treatment (1/0)	-0.114***	0.00796		
	(0.0164)	(0.0174)		
IHS AI Patents			-0.0445***	0.00313
			(0.00744)	(0.00960)
Ln Capital Stock	0.00643***	0.00990***	0.00641***	0.00990***
	(0.000415)	(0.000620)	(0.000414)	(0.000620)
Age Bins	Yes	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes
Observations	1,124,000	1,124,000	1,124,000	1,124,000
R-squared	0.833	0.949	0.833	0.949

Robust Standard Errors clustered at the 4-digit NAICS industry level. *, ** and *** denotes significance at the 5%, 1% and 0.1% respectively. Each regression includes a constant, multi-unit status and multinational status, which are not displayed here

Our results using this approach suggest that transition into AI invention has little to no impact on the firm's own production worker share when controlling for firm fixed effects. Likewise, inclusion of the firm's (transformed) patent stock has no measured impact on production worker share once firm effects are taken into account.

6.c Earnings

Although within-firm changes in the share of production workers resulting from transition into AI invention or expansion of the AI patent stock appear to be negligible, there may changes in the earnings distributions of workers within these firms. Three decades of research in labor economics document the dispiriting reality that less educated Americans have faced relatively weak demand for their services, stagnant or declining wages, and an increasingly polarized job market (Autor, Katz, and Kearney, 2006; Autor and Dorn, 2013). While a number of other economic forces, including globalization, declining union density, and falling minimum wages,

[7] Berman, Bound, and Griliches (1994) use a measure of capital intensity (K/Y) where we use a measure of capital stock. However, we do not find robust effects of AI invention on production worker share and do not expect that substiting capital intensity for capital stock would change these results. In the Census data, capital stock is only computed accurately for manufacturing firms, so inclusion of capital-stock relates variables effectively limits purview of the study to the manufacturing sector.

have also contributed to rising income inequality over the past four decades, economists generally agree that skill-biased technological change may be the single most important cause (Autor, 2014). If AI invention really does transform the economy to the degree that proponents expect, then it may significantly accelerate and exacerbate the kind of skill-biased technological change already documented by labor economists. We examine whether this is the case by linking our firms – both the AI-inventing firms and their same-industry peer firms – to the Longitudinal Employer Household Dynamics (LEHD) database widely used by labor economists. We are able to conduct this match for firms and establishments in 49 out of 50 states. In principle, this allows to probe the impact of transition into AI invention on the distribution of employee earnings within firms over time.

Our analysis on the changes to the earnings distribution follows the general methodology in Autor, Katz and Kearny (2008), who look at changes to the 90-10, 90-50 and 50-10 earnings ratios. We do something similar for our set of firms and estimate an equation along the lines of:

$$dER_i = \beta_0 + \beta_1 dln(L_i) + \beta_1 dln(K_i) + \beta_2 dln(A_i) + \varepsilon_i$$

where ER reflects either the 90-10, 90-50 or 50-10 earnings ratios of firm *i* over some period of time, modeled as a function of changes in the employment *L*, the log of the capital stock of firm i, K_{it}, and a measure of changes in the AI-intensity of firm innovation, A_{it}, over the same time period. Again, we use both the dummy variable indicating inception of AI-related patenting and, alternatively, the IHS-transformed stock of AI-related patents as alternative measures of AI invention. We incorporate individual firm fixed effects into our model. In all cases, we take natural logs of the key variables.

Table 8: Impact of AI Innovations on 90-10, 90-50 and 50-10 Earnings Ratio, 1997-2018 (manufacturing only)

	(1)	(2)	(3)	(4)	(5)	(6)
	90-10 Earnings Ratio		90-50 Earnings Ratio		50-10 Earnings Ratio	
AI Treatment (1/0)	0.0420* (0.0200)		0.0189* (0.00846)		0.0244 (0.0150)	
IHS AI Patents		0.0265*** (0.00759)		0.00758 (0.00394)		0.0188*** (0.00566)
Ln Employment	-0.110*** (0.00264)	-0.110*** (0.00264)	-0.0670*** (0.00163)	-0.0670*** (0.00163)	-0.0420*** (0.00161)	-0.0420*** (0.00161)
Ln Capital Stock	-0.0175*** (0.00118)	-0.0175*** (0.00118)	-0.00656*** (0.000701)	-0.00656*** (0.000701)	-0.0108*** (0.000843)	-0.0108*** (0.000843)
Age Bins	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,012,000	1,012,000	1,012,000	1,012,000	1,012,000	1,012,000
R-squared	0.713	0.713	0.729	0.729	0.624	0.624

Robust Standard Errors clustered at the 4-digit NAICS industry level. *, ** and *** denotes significance at the 5%, 1% and 0.1% respectively. Each regression includes a constant, which is not displayed here.

As Table 8 highlights, the within-firm effect of taking out an AI patent leads to a positive and significant rise in the top earnings decile. An AI patent is associated with a 4.2% change in the

90-10 earnings ratio and a 1.9% rise in the 90-50 ratio and 2.4% rise in the 50-10 ratio (albeit insignificant). These effects remain positive and significant when we include a continuous measure of AI patents. This is suggestive that the adoption of AI technology is leading to further earnings inequality, but the considerable differences between AI-inventing firms and the rest of the manufacturing sample suggest caution in drawing this inference.

Continuing with our comparison of different AI-related patent classifiers, Table 9 provides the AI Treatment effect for a broader set of classifiers, including the one provided by USPTO.

Table 9: Impact of different classifiersfor AI innovations on Earnings Ratios, 1997-2018 (manufacturing only)

	(1) 90-10 Earnings Ratio	(2) 90-50 Earnings Ratio	(3) 50-10 Earnings Ratio
Baseline AI Treatment (1/0)	0.0420* (0.0200)	0.0189* (0.00846)	0.0244 (0.0150)
AI Treatment (75% Confidence)	0.0521*** (0.0156)	0.0202** (0.00730)	0.0330** (0.0115)
AI Treatment (50% Confidence)	0.0393** (0.0136)	0.0200** (0.00642)	0.0213* (0.0100)
AIPD Treatment (95% Confidence)	0.0515*** (0.0118)	0.0205*** (0.00583)	0.0320*** (0.00872)
Baseline & AIPD Treatment	0.0445* (0.0200)	0.0198* (0.00850)	0.0262 (0.0147)
AIPD Only Treatment	0.0329** (0.0102)	0.0124* (0.00528)	0.0210** (0.00763)

Robust Standard Errors clustered at the 4-digit NAICS industry level. *, ** and *** denotes significance at the 5%, 1% and 0.1% respectively. Each regression is identical to the specification in Table 8, but with a different binary treatment effect.

Table 9 shows that using a different classifier has a mostly similar effect on measured earnings inequality across all firms. In many of these cases, the magnitudes are a bit higher than the baseline classifier, but not significantly different when we account for the standard errors.

To summarize, our preliminary results using standard linear approaches to measuring the impact of AI show that innovating in AI is positively and significantly associated with higher sales per employee, increased value added per worker and TFP after controlling for firm fixed effects. On the other hand, we do not see a within-firm change in the production worker share for firms who innovate in AI within the time horizon, but do see a slight rise in income for the 90th percentile workers relative to the median and bottom decile worker. This rise is highest among the 90th percentile workers, suggesting increased demand for the highest-skilled workers.

While this section attempts to control for these across-firm differences by limiting our analysis to within-firm changes, in the next section, we attempt to do a better job by performing an event study that pairs our AI-inventing firms with a comparable set of similar control firms.

7. Estimating the Impact of AI on Productivity and Labor

Demand with an Event Study Approach

Our first pass of the data found positive and significant within-firm changes to sales and value-added resulting from innovations in AI. This section attempts to better control for some of the endogeneity described earlier, as well as look at firm behavior and outcomes before and after the AI innovation. We do this by conducting an event study analysis (e.g. a difference-in-differences specification with a group-specific time trend) centered around the timing of the first AI-related patent. Our identification relies on matching each firm with at least one AI-related patent as closely as possible with a similar same-industry counterpart which does not obtain an AI-related patent. We can accomplish this using the full richness of the Census data.

7.a Exact Matching of Firms with AI patents

Our event study begins by attempting to identify an exact match of the firm with an AI-related patent. Our matching criteria are based on firm size, firm industry, multi-unit status, and firm age and are centered around the timing of the first AI patent application (e.g. if a firm took out its first AI patent in 2001, we would attempt to identify a matching non-AI patenting firm in 2001). We group the AI firms into 10 employment bins based on the decile within a 4-digit NAICS-year. We then attempt to match AI firms with their precise non-AI counterpart by age, multi-unit status, primary industry (largest 4-digit NAICS employment within a firm), primary state (largest headquarter by employment, or largest state by employment if no headquarter), and employment decile. For unmatched and matched firms with multiple controls, we look at the closest employment counterpart by measuring the Euclidean distance of employment differences in the current and 3 years prior. Once this is complete, we are able to identify unique matches for 87% of the firms with an AI patent. Unmatched firms have no clear counterpart by industry-age-state and multi-unit status.

7.b Event Study Plots

Before looking at the regression results, it will help to look at the event study plots centered around the time of the first AI innovation. Our figures below are centered 2 years prior to the first AI innovation and track firm performance in the 5 years following the AI innovation. We start with employment. Figure 7 looks at the relative employment growth in the pre and post AI patent application date for the AI firms and their matched counterpart. We set our relative employment to 1 at the time of the AI patent application.

Our exact match of the firm types should show that the pre-period trends follow relatively closely, which is confirmed in the plot depicted in Figure 7. In the years following the AI patent application however, employment growth for AI patenting firms deviates relative to the employment growth of non-AI patenting firms. In fact, we find that employment growth is almost 50% higher in the five years following the first AI patent application.

We construct a similar event study plot, charting the revenue per employee for AI patenting firms and non-AI patenting firms. Figure 8 shows the results. While the pre-trends show some differences prior to the AI patent, these differences show that the non-AI patent holders had slightly larger revenue prior to the AI patent, with similar patterns as found in Figure 7 once the

AI patent is applied for. AI firms have 15-20% higher revenues per employee than their closely matched counterparts.

7.c Event Study Specification

We now move to a more formal specification that includes firm controls and assesses whether the deviation in employment and revenue growth persists after controls are introduced. Our event study looks at the difference-in-differences of the firm outcomes and includes a group specific time trend. Our specification is as follows:

$$y_{it} = \alpha + \beta_1 AI_{it}(1|0) + \beta_2 TIME + \beta_3 AI_{it} \times TIME + X_{it} + \varepsilon$$

with firm controls for earnings, age, multi-unit status and multinational status. The β_1 coefficient indicates the aggregate effects of having an AI patent on the outcome variable, while β_2 measures aggregate time trends. The β_3 coefficient indicates whether the impact of AI patents is changing over time on the outcome variable.

In our empirical analyses, we split our sample between the full matched dataset and a manufacturing-only dataset with different outcome variables for each. The full matched dataset will include employment and revenue per employee as the primary outcome variables, while the manufacturing-only sample will examine value-added per employee and production worker share.[8] Table 10 provides the first set of results.

Table 10: Impact of AI Innovations on Employment and Revenue, 1997-2018 (matched only)

	(1)	(2)	(3)	(4)
	Ln Employment		Ln Revenue per Employee	
AI Treatment (1/0)	Dropped	Dropped	Dropped	Dropped
Post AI Year	0.0206 (0.0144)		-0.124*** (0.0192)	
AI Treatment x Post AI Year	0.138*** (0.0194)		0.164*** (0.0263)	
AI Treatment x Year = -2		0.0229 (0.0224)		-0.162*** (0.0346)
AI Treatment x Year = -1		0.0320 (0.0167)		-0.128*** (0.0291)
AI Treatment x Year = 0	Dropped	Dropped	Dropped	Dropped
AI Treatment x Year = +1		0.157*** (0.0149)		-0.0146 (0.0259)
AI Treatment x Year = +2		0.224*** (0.0189)		0.0249 (0.0300)
AI Treatment x Year = +3		0.233*** (0.0217)		0.0701* (0.0335)
AI Treatment x Year = +4		0.245*** (0.0256)		0.0307 (0.0373)
AI Treatment x Year = +5		0.274*** (0.0275)		0.0439 (0.0375)

[8] Value-added and production worker numbers are not available for the full sample.

Age Bins	Yes	Yes	Yes	Yes
Industry-Year Fixed Effects	Yes	Yes	Yes	Yes
Firm Fixed Effects	Yes	Yes	Yes	Yes
Observations	36,000	36,000	36,000	36,000
R-squared	0.975	0.976	0.787	0.787

Robust Standard Errors clustered at the 4-digit NAICS industry level. *, ** and *** denotes significance at the 5%, 1% and 0.1% respectively. Each regression includes a constant, which is not displayed here.

Relative to the untreated and pre-AI patenting time period, we find that the impact of AI patents is positive on both employment and revenue per employee in the post-AI period. Having an AI patent is associated with 13.8% higher employment and 16.4% higher revenue per employee, with the impact for each growing over time, at least for employment. After 5 years from the AI patent application, the treated firms have 27.4% higher employment.

We can similarly compare the impact of each different classifier, performing a similar matching exercise for each type of treated firm. The effects can be found in Table 11.

Table 11: Impact of different AI classifiers on Employment and Revenue, 1997 – 2018 (Event Study)

	(1) Ln Employment	(2) Ln Revenue per Employee
Baseline AI Treatment (1/0)	0.138***	0.164***
	(0.0194)	(0.0263)
AI Treatment (75% Confidence)	0.122***	0.132***
	(0.0141)	(0.0185)
AI Treatment (50% Confidence)	0.128***	0.120***
	(0.0121)	(0.0163)
AIPD Treatment (95% Confidence)	0.129***	0.131***
	(0.00966)	(0.0131)
Baseline & AIPD Treatment	0.118***	0.162***
	(0.0243)	(0.0346)
AIPD Only Treatment	0.0972**	0.104*
	(0.0357)	(0.0458)

Robust Standard Errors clustered at the 4-digit NAICS industry level. *, ** and *** denotes significance at the 5%, 1% and 0.1% respectively. Each regression is identical to the specification in Table 10, but with a different binary treatment effect.

The effects are strongest for our baseline classifier on both employment and revenue per employee.

7.d Earnings Ratios (Event Study)

We perform a similar estimation as in Table 8 looking at the earnings ratio for the matched set of firms. To frame our analysis, we start by plotting the change in the earnings ratio centered around the time of the first AI patent, as in the previous event study plots in Figures 7 and 8. This can be found in Figure 10. As Figure 10 shows, we see almost no difference between the treated and untreated set of firms in terms of widening wage gaps. While the 90-10 and 90-50 both rise slightly in the post-AI period, there is little separation across the firm types.

We formally control for other firm characteristics and estimate the potential impact of the treatment, as in Tables 8 and 10 and include firm fixed effects. Table 12 provides the set of results.

Table 12: Impact of AI Innovations on 90-10, 90-50 and 50-10 Earnings Ratio, 1997-2019 (full matched set of firms)

	(1)	(2)	(3)	(4)	(5)	(6)
	90-10 Earnings Ratio		90-50 Earnings Ratio		50-10 Earnings Ratio	
AI Firm (1/0)	Dropped	Dropped	Dropped	Dropped	Dropped	Dropped
Post AI (1/0)	-0.0170		-0.00563		-0.0101	
	(0.0159)		(0.00825)		(0.0117)	
AI x Post (1/0)	-0.00464		-0.0135		0.00920	
	(0.0211)		(0.0109)		(0.0149)	
AI x Year = -2		-0.00633		-0.00442		0.00107
		(0.0274)		(0.0149)		(0.0193)
AI x Year = -1		0.00795		0.00647		-0.00394
		(0.0216)		(0.0114)		(0.0160)
AI x Year = 0	Dropped	Dropped	Dropped	Dropped	Dropped	Dropped
AI x Year = +1		-0.000185		0.00146		-0.00348
		(0.0181)		(0.00978)		(0.0134)
AI x Year = +2		0.00475		-0.0143		0.0161
		(0.0223)		(0.0123)		(0.0161)
AI x Year = +3		-0.0249		-0.0236		-0.00268
		(0.0242)		(0.0137)		(0.0171)
AI x Year = +4		-0.00885		-0.0306*		0.0177
		(0.0267)		(0.0148)		(0.0187)
AI x Year = +5		0.000276		-0.0217		0.0249
		(0.0290)		(0.0160)		(0.0201)
ln (Emp)	0.104***	0.105***	0.0422***	0.0432***	0.0646***	0.0646***
	(0.0130)	(0.0131)	(0.00743)	(0.00746)	(0.00784)	(0.00787)
Age Bins	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	33,500	33,500	33,500	33,500	33,500	33,500
R-squared	0.727	0.727	0.728	0.729	0.679	0.679

Robust Standard Errors clustered at the 4-digit NAICS industry level. *, ** and *** denotes significance at the 5%, 1% and 0.1% respectively. Each regression includes a constant, which is not displayed here.

The post-treatment effect shows an insignificant effect for each of the earnings ratio, indicating that once we include a proper control sample, the widening wage inequality seen in Table 8 mostly disappears. This seems to be indicative that the attributes of the firm that are similarly associated with AI patents, such as being in active in certain sectors and being of a certain size, are the main contributors to widening earnings inequality and that the earnings gap for firms with AI patents does not change much relative to similar firms who may not possess those AI patents.

To conclude, the event-study matched firm analyses did not yield the same results as the full sample panel regressions. In general, the strong positive and significant conditional correlations between the inception of AI patenting and labor productivity detected in the full sample panel regressions remain robust and relatively large in magnitude when we shift to our event study analysis. However, the evidence that AI is conditionally correlated with rising wage dispersion within the firm which we detected in the full sample regressions appears to dissipate into insignificance when undertake event study analyses. If we regard the event study approach as more robust to omitted variable bias and other inference challenges, then we might conclude that AI invention appears to be raising firm labor productivity but, so far, at least, not expanding income inequality within the AI-inventing firms.

While the propensity-score matched firms do a reasonable job of identifying a set of control firms, there are still flaws with this approach as we were unable to include innovation characteristics in generating our matches. Likewise, our event-study specification is intended to identify treatment impacts that may occur several years after the initial treatment, but we neglect to include an intensity measure for the treated, and have a limited post-period at our disposal due to the recentness of the patent data. Therefore, while these statistical associations are suggestive, it would be premature to view them as providing strong evidence of a *causal* relationship between AI innovation and productivity. Stronger conclusions about causality will require additional data and further analyses.

8. Next Steps: Tracking AI-related Innovation through the Movement of "Frontier" AI Talent

As we have noted in this paper, we are currently engaged in an effort to update our database of AI-related patents, which will then allow us to update our empirical analysis of their impact on labor productivity and income inequality. However, not all innovations are patented, and some industries investing heavily in AI to generate new products and services hardly patent at all. How can we move beyond the limitations of patent data to examine AI-related innovation in these contexts? Here, we describe ongoing research efforts to address this issue.

To do so, we are currently using publication data from Elsevier to identify the top academic scientists working in domains related to AI and the graduate students whom they supervise and with whom they coauthor. We are using a combination of publication data, website data, and data from professional career profile services and online resumes to track the movement of these students of scientific thought leaders across geographic space, organizational boundaries, and time. We can also use publication data to track the direct interaction between top academic scientists and the companies they work with when that interaction results in a publication. Once we can link the star scientists (which the research team refers to in their internal dialog as the "immortals") and their students to the firms with which (and for whom) they have worked, and trace these linkages over time, we can leverage our access to U.S. Census microdata, obtained through our ongoing collaboration with Census microeconomists, to ask whether these linkages have provided the receiving firms with a statistically discernable advantage over their same industry peers who lack them in terms of output, employment, or, most importantly, productivity.

Table 13, reproduced below, provides an illustrative breakdown of the identified star scientists ("immortals"), students, direct collaborations, and papers.

Table 13: Counts of star scientists, student advisees, and direct collaborations between corporate labs and star scientists (April 2024) by AI subdomain

No.	Data point \ Domain	NLP	ML	Robotics	Agents	HCI	Speech	KR	CV	IR
1.	# of Immortals	208	320	160	29	17	140	4	79	459
2.	# of journals/conference venues	722	79	17	58	18	12	32	80	146
3.	# of papers associated with immortals	50,414	74,165	56,476	9,362	4,451	36,396	873	33,509	89,490
4.	# of children of immortals identified/information tagged (so far)	1272	2938	342	297	86	126	No NA based immortal	Yet to be found	Yet to be tagged
5.	Balance # of children of immortals (yet to be tagged)	124	420	5622	101	19	251	-	-	6102
6.	# of direct Immortal-Corporate collaborative papers	8,553	16,026	4,582	665	987	6,859	30	4,301	12,047

This line of inquiry is related to the work of Babina et al. (2022, 2024), but focuses on the role played by elite scientists, who may be disproportionately important in the defining the technology frontier, and their doctoral students, who may be disproportionately important in bringing this frontier technology into industrial practice (Agarwal and Henderson, 2000; Zucker and Darby, 1998). We can imagine that any firm seeking to apply frontier AI to the substantive reengineering of its current products and services or the creation of new products and services needs to create within itself a "pyramid" of AI talent graphically depicted in the left portion of Figure 9 (Arora et al., 2013; Branstetter et al., 2018). At the lower ranks of the pyramid, the firm could productively employ programmers with "self-taught" AI skills who use standard AI tools and techniques. At the middle levels of the talent pyramid, the firm might need professionals with bachelors or masters degrees that include specialized AI training, but these professionals need not have trained at elite universities. However, at the very apex of the pyramid, a firm seeking to out-engineer its rivals may seek to acquire "software architects" who have been trained up to the technology frontier by elite academic scientists based at the top universities.[9] The data and approach taken by Babina et al. (2022, 2024) use data on the entire pyramid; our approach focuses on the star scientists and their students who could constitute the disproportionately important part of the apex of that pyramid. The role played by these individuals is related to that of the "architects" described in the theoretical work of Benzell et al. (2022). As in that paper, we consider the idea that the limited supply of these software architects could constitute an important constraint on the ability of firms to fully leverage frontier AI technologies.

[9] We acknowledge that some frontier work is done by elite AI scientists at leading tech companies, rather than universities, but these companies do not play the same teaching/training role as universities. Hence, they do not exert the same influence through their trainees.

To fix ideas further, imagine that our data sources identify CMU doctoral recipient "Dr. Who" as one of the Ph.D. advisees of an elite academic scientist. Dr. Who's subsequent movement to Google DeepMind could further augment the intellectual resources of this impressive corporate research operation. It is possible that Dr. Who begins to specialize at Google DeepMind in the application of advanced AI algorithms to medical imaging. Then, he carries this skill to Siemens Healthineers and from there to diagnostic imaging start-up Arterys. By following star scientists' students like Dr. Who from firm to firm, we could trace out their differential impact, if any, on the enhancement of host firms' output, employment, and productivity. The hypothesis that these movements predict success can be tested using access to Census data on the hiring firms and their same industry peers who have hired fewer or no advanced AI experts.

Table 14 The Impact of Collaboration with Elite AI Scientists

	Employment	**Payroll per Employee**	**Revenue**	**Revenue per Employee**	**AI Patents**
Coauthored Publications	+***	+***	+***	+***	+***
Industry-Year FE Firm FE Publication>0 Firms	Y N N	Y N N	Y N N	Y N N	Y N N
Cumulative Coauthored Publications	+***	+***	+***	-	+***
Industry-Year FE Firm FE Publication>0 Firms	N Y N	N Y N	N Y N	N Y N	N Y N
Cumulative Coauthored Publications	-	+***	+	-	+***
Industry-Year FE Firm FE Publication>0 Firms	N N Y	N N Y	N N Y	N N Y	N N Y

DRB disclosure code associated with these results is CBDRB-FY24-CES022-005. This table illustrates the sign and significance of regression coefficients obtained from a panel data regression of the designated dependent variable on various measures of direct collaboration between elite AI scientists and U.S. companies, as evidenced by the appearance of coauthored scientific publications or conference presentations in journals or venues tracked by the authors which feature the elite scientist and co-authors affiliated with the firm as named authors. The magnitudes of the coefficients are suppressed pending additional disclosure review.

Results in Table 14 present preliminary evidence on the conditional correlations between one of our measures of interaction between elite AI scientists and firms and firm-level outcomes, including employment, revenue, pay per employee, revenue per employee, and the number of AI patents. We consider both a "flow" measure of the number of co-authored publications that appear in a particular year and a cumulative "stock" of the number of co-authored publications that exists at a point of time. We also use a range of fixed effects (industry-year, firm, etc.) to mitigate (but not eliminate) some of the more obvious examples of potential omitted variable bias. While we are not yet permitted to disclose the regression coefficients associated with these variables, we have been permitted to share the signs and significance levels of those coefficients in Table 14.

Prior to any effot to match these data to Census data, we could already see that direct scientific collaborations with elite AI scientists were quite narrowly concentrated in a small group of firms. This implies that the cross-sectional dimension of the matched data set is limited, and this, in turn, limits the strength of statistical correlations we might expect to find. With these caveats in mind, results in Table 14 show some robust relationships between our measures of direct collaboration and other firm-level outcome variables of interest. The statistical correlations that appear to be the most robust are those associated with pay per employee (i.e., average worker income) and AI patents. There do not appear to be strong and robust direct impacts on labor productivity, but other results in this paper suggest an impact of AI patents on labor productivity, and AI patents appear to be positively associated with directl collaboration. Ongoing work will seek to further clarify to the nature of these relationshps and also probe the impact of the movement of the advisees of these scientists into firms.

9. (Preliminary) Conclusions

Significant breakthroughs in AI and related technologies have led some economists to predict dramatic effects on firm productivity and labor demand. However, empirical assessment of the actual impact has been limited by the scarcity of firm-level data on AI innovation and its economic effects. This paper advances our understanding of the economic impact of AI by using patent data to capture AI-related innovation. Previous analyses by economists have generally sought to measure AI-related innovation by focusing on a small number of keywords and patent classes (Giczy et al., 2021, is an important exception). Instead, we use a suite of machine learning algorithms to parse the entire text of patent documents; our more comprehensive search yields a much larger count of AI-related patents than have been identified in some earlier studies. The sharp rise in patenting and its wide distribution across patent classes and firms through 2018 are consistent with the characterization in the literature of AI as a general purpose technology.

Because we know the identity of patent assignees, we can match the patent data to extensive confidential microdata on these firms collected by the U.S. Census Bureau. These rich data are available for publicly traded and privately held firms, and they allow us to take a first look at the impact of AI innovation on the productivity and labor demand of the innovating firms. The results of such analyses may be useful for researchers and policymakers, because the impact observed on the limited number of leading firms in this space could be indicative of the larger impact we will see as this technology spreads across firms and industries. Such analyses also provide a useful counterpoint to recent scholarship that seeks to predict labor market impacts based on textual analysis of occupations, tasks, and emerging AI capabilities.

We have expanded our analyses of the impact of growing AI invention on labor demand by matching our firm-level data to the Longitudinal Employer-Household Dynamics (LEHD) data set. This enables us to quantify the impact of AI innovation on the income distribution of AI-inventing firms.

Our current explorations are preliminary, and it would be premature to interpret our statistical results as strong evidence of causal impacts. Nevertheless, our results suggest that the inception of AI innovation is associated with economically and statistically significant increases in revenue per employee, suggesting strong productivity effects. While there is some limited evidence of a coincident increase in income inequality within AI-inventing firms, these results are not robust to the use of our most demanding empirical approaches.

Efforts to refine our estimates of the impact of AI led us to conduct event-study style analyses on matched samples of AI firms and similar firms in the same industries that did not transition into AI patenting over our sample period. Our regression results on this matched sample continue to suggest substantial labor productivity effects that may strengthen over time. There is no indication, in these more demanding specifications, of a similar increase in within-firm wage dispersion. Despite the limitations of our current analyses, preliminary results suggest that this is promising line of research, which we will continue to explore.

Not all AI innovations are patented, but significant efforts to develop new AI-enabled goods and services or to re-engineer existing goods and services to leverage AI may require the employment of experts trained up to the technological frontier. This paper describes our effort to identify elite AI scientists and track their direct collaboration with firms, as evidenced by co-authored publications. It also describes our ongoing efforts to track the movement of the students of these elite scientist into firms, and the methods we hope to use trace out the impact of these hires on the output, employment, and productivity of the hiring firms. These efforts may also yield interesting and useful results.

References

Acemoglu, Daron, Gary Anderson, David Beede, Catherine Buffington, Eric Childress, Emin Dinlersoz, Lucia Foster, Nathan Goldschlag, John Haltiwanger, Zachary Kroff, Pascual Restrepo and Nikolas Zolas, (2022). "Automation and the Workforce: A Firm-Level View from the 2019 Annual Business Survey". *Mimeo*.

Ajay, Agraway, Joshua Gans, and Avi Goldfard, (2021). *Power and Prediction: The Disruptive Economics of Artificial Intelligence*, Harvard Business Review Press.

Agrawal, Ajay, Joshua Gans, and Avi Goldfarb, (Eds.), 2019, *The Economics of Artificial Intelligence: An Agenda*, University of Chicago Press and NBER.

Agrawal, Ajay, Joshua Gans, and Avi Goldfarb, "Prediction, judgment, and complexity: a theory of decisionmaking and artificial intelligence," in Agrawal, Gans, and Goldbarb, (Eds.), (2019), *The Economics of Artificial Intelligence: An Agenda*, University of Chicago Press and NBER.

Arora, Ashish, Lee G. Branstetter, and Matej Drev. "Going soft: How the rise of software-based innovation led to the decline of Japan's IT industry and the resurgence of Silicon Valley." *Review of Economics and Statistics* 95, no. 3 (2013): 757-775.

Autor, David, Frank Levy, and Richard Murnane. "The skill content of recent technological change: an empirical exploration." *The Quarterly Journal of Economics* 118, no. 4 (2003): 1279-1311.

Autor, David, Lawrence Katz, and Alan Krueger. "Computing inequality: have computers changed the labor market?" *The Quarterly Journal of Economics* 133, no. 4 (1998): 1169-1213.

Autor, David and David Dorn. "The growth of low-skill service jobs and the polarization of the U.S. labor market." *The American Economic Review* 103, no. 5 (2013): 1553-1597.

Babina, Tania, Anastassia Fedyk, Alex He, and James Dodson. (2024). "Artificial Intellgience, Firm Growth, and Product Innovation," Journal of Financial Economics.

Babina, Tania, Anastassia Fedyk, Alex He, and James Dodson, forthcoming, "Firm Investments in Artificial Intelligence Technologies and Changes in Workforce Composition," NBER Conference Voume on Technology, Productivity and Economic Growth.

Bail, Martin N, Erik Brynjolfsson, and Anton Korinek, "Machines of Mind: The Case for an AI-Powered Productivity Boom," Brookings, May 2023.

Balasubramanian, Natarajan, and Jagadeesh Sivadasan. "What happens when firms patent? New evidence from US economic census data." *The Review of Economics and Statistics* 93, no. 1 (2011): 126-146.

Benzell, Seth, Erik Brynjolfsson, and Guillame Saint-Jacques. (2022). "Digital Abundance Meets Scarce Architects: Implications for Wages, Interest Rates, and Growth," Working Paper.

Berman, Eli, John Bound, and Zvi Griliches. "Changes in the demand for skilled labor within US manufacturing: evidence from the annual survey of manufactures." *The Quarterly Journal of Economics* 109, no. 2 (1994): 367-397.

Branstetter, Lee G., Matej Drev, and Namho Kwon. "Get with the program: Software-driven innovation in traditional manufacturing." *Management Science* 65, no. 2 (2018): 541-558.

Bresnahan, Timothy F., and Manuel Trajtenberg. "General purpose technologies 'Engines of growth'?." *Journal of econometrics* 65, no. 1 (1995): 83-108.
Erik Brynjolfsson, Danielle Li, Lindsey Raymond, "Generative AI at Work" NBER WP 31161, April 2023
Brynjolfsson, Erik, and Andrew McAfee. *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. WW Norton & Company, 2014.
David Byrne, John Fernald, and Marshall Reinsdorf. ""Does the United States Have a Productivity Slowdown or a Measurement Problem?" *Brookings Papers on Economic Activity* Vol. 47 (2016) 109-182.
Brynjolfsson, Erik, Tom Mitchell, and Daniel Rock, "What can machines learn and what does it mean for occupations and the economy." *The American Economic Review Papers and Proceedings* 108 (2018): 43-47.
Brynjolfsson, Eric, Daniel Rock, and Chad Syverson, "Artificial intelligence and the modern productivity paradox: a clash of expectations and statistics," in Agrawal, Gans, and Goldbarb, (Eds.), (2019), *The Economics of Artificial Intelligence: An Agenda*, University of Chicago Press and NBER.
Cockburn, Iain, Rebecca Henderson, and Scott Stern. "The impact of artificial intelligence on innovation," in Agrawal, Gans, and Goldbarb, (Eds.), (2019), *The Economics of Artificial Intelligence: An Agenda*, University of Chicago Press and NBER.
Cortes, C., & Vapnik, V. (1995). Support-Vector Networks. *Machine Learning, 20*, 273-297.
Cho, K., Van Merriënboer, B., Bahdanau, D., & Bengio, Y. (2014). On the properties of neural machine translation: Encoder-decoder approaches. *arXiv preprint arXiv:1409.1259*.
David, P. A. "The Computer and the Dynamo The modern productivity paradox in a not-too -distant mirror," in *Technology and Productivity: The Challenge for Economic Policy*, Paris: OECD, (1991).
Elandou, Tyna, Sam Manning, Pamela Mishkin, and Daniel Rock. (2023). GPTs are GPTs: An Early Look at the Labor Market Impact of Large Language Models. Working Paper.
Fort, T.C., Keller, W., Schott, P.K., Yeaple, S. and Zolas, N., (2020). Colocation of Production and Innovation: Evidence from the United States.*Mimeo.*
Giczy, A.V., Pairolero, N.A. and Toole, A.A., (2021). "Identifying artificial intelligence (AI) invention: A novel AI patent dataset." *The Journal of Technology Transfer*, pp.1-30.
Graham, Stuart JH, Cheryl Grim, Tariqul Islam, Alan C. Marco, and Javier Miranda. "Business dynamics of innovating firms: Linking US patents with administrative data on workers and firms." *Journal of Economics & Management Strategy* 27, no. 3 (2018): 372-402.
Helpman, Elhanan, ed. *General purpose technologies and economic growth*. MIT press, (1998).
Ho, Tin Kam. (1995). Random decision forests. In Third International Conference on Document Analysis and Recognition, ICDAR 1995, August 14 - 15, 1995, Montreal, Canada. Volume I, pages 278–282.
Jaffe, Adam B., and Manuel Trajtenberg. *Patents, citations, and innovations: A window on the knowledge economy*. MIT press, (2002).
Kerr, William R., and Shihe Fu. "The survey of industrial R&D—patent database link project." *The Journal of Technology Transfer* 33, no. 2 (2008): 173-186.
Korinek, Anton, (2023), "Language Models and Cognitive Automation for Economic Research," NBER Working Paper 30957.

Mann, Katja and Puttmann, Lukas, (2018). Benign Effects of Automation: New Evidence from Patent Texts. Working paper, Copenhagen Business School.

McElheran, Kristina, J. Frank Li, Erik Brynjolfsson, Zachary Kroff, Emin Dinlersoz, Lucia Foster and Nikolas Zolas, (2024). "AI Adoption in America: Who, What and Where.*" Journal of Economics and Management Strategy*, January.

Mikolov, Thomas, Sutskever, Ilya, Chen, Kai, Corrado, Gregory S., and Dean, Jeffrey. (2013). Distributed representations of words and phrases and their compositionality. In Advances in Neural Information Processing Systems 26: 27th Annual Conference on Neural Information Processing Systems 2013. Proceedings of a meeting held December 5-8, 2013,Lake Tahoe, Nevada, United States., pages 3111–3119National Academies of Sciences, Engineering, and Medicine. *Information Technology and the US Workforce: Where Are We and Where Do We Go from Here?*. National Academies Press, 2017.

Noy, Shaken and Whitney Zhang, (2023). "Expermental Evidence on the Productivity Effects of Generative Artificial Intelligence." Science 381 (6654), pp. 187-192.

Raj, Manav and Robert Seamans. "Artificial intelligence, labor, productivity, and the need for firm-level data," in in Agrawal, Gans, and Goldbarb, (Eds.), (2019), *The Economics of Artificial Intelligence: An Agenda*, University of Chicago Press and NBER.

Rumelhart, David E., Hinton, Geoffrey E. and Williams, Ronald J. (1988). Neurocomputing: Foundations of research. Learning Representations by Back-propagating Errors, pages 696–699. MITPress, Cambridge, MA, USA

Suskind, Daniel. (2020). *A World Without Work.* Allen Lane: London.

Van Roy, Vincent, Vertesy, Daniel and Damioli, Giacomo. (2019). AI and Robotics Innovation: a Sectoral and Geographic Mapping using Patent Data. GLO Discussion Paper No. 433.

Waibel, A.H., Hanazawa, T., Hinton, G.E., Shikano, K., & Lang, K.J. (1989). Phoneme recognition using time-delay neural networks. *IEEE Trans. Acoustics, Speech, and Signal Processing, 37*, 328-339.

Webb, Matthew, Nick Short, Nicholas Bloom, and Joshua Lerner. (2019) "Some facts on high-tech patenting." NBER working paper no. 24793.

Yang, Z., Yang, D., Dyer, C., He, X., Smola, A., & Hovy, E. (2016, June). Hierarchical attention networks for document classification. In *Proceedings of the 2016 conference of the North American chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 1480–1489).

Zolas, Nikolas., Zachary Kroff, Erik Brynjolfsson, Kristina McElheran, David Beede, Catherine Buffington, Nathan Goldschlag, Lucia Foster, and Emin Dinlersoz, (2021). "Advanced technologies adoption and use by US firms: Evidence from the annual business survey". NBER working paper no. 28290.

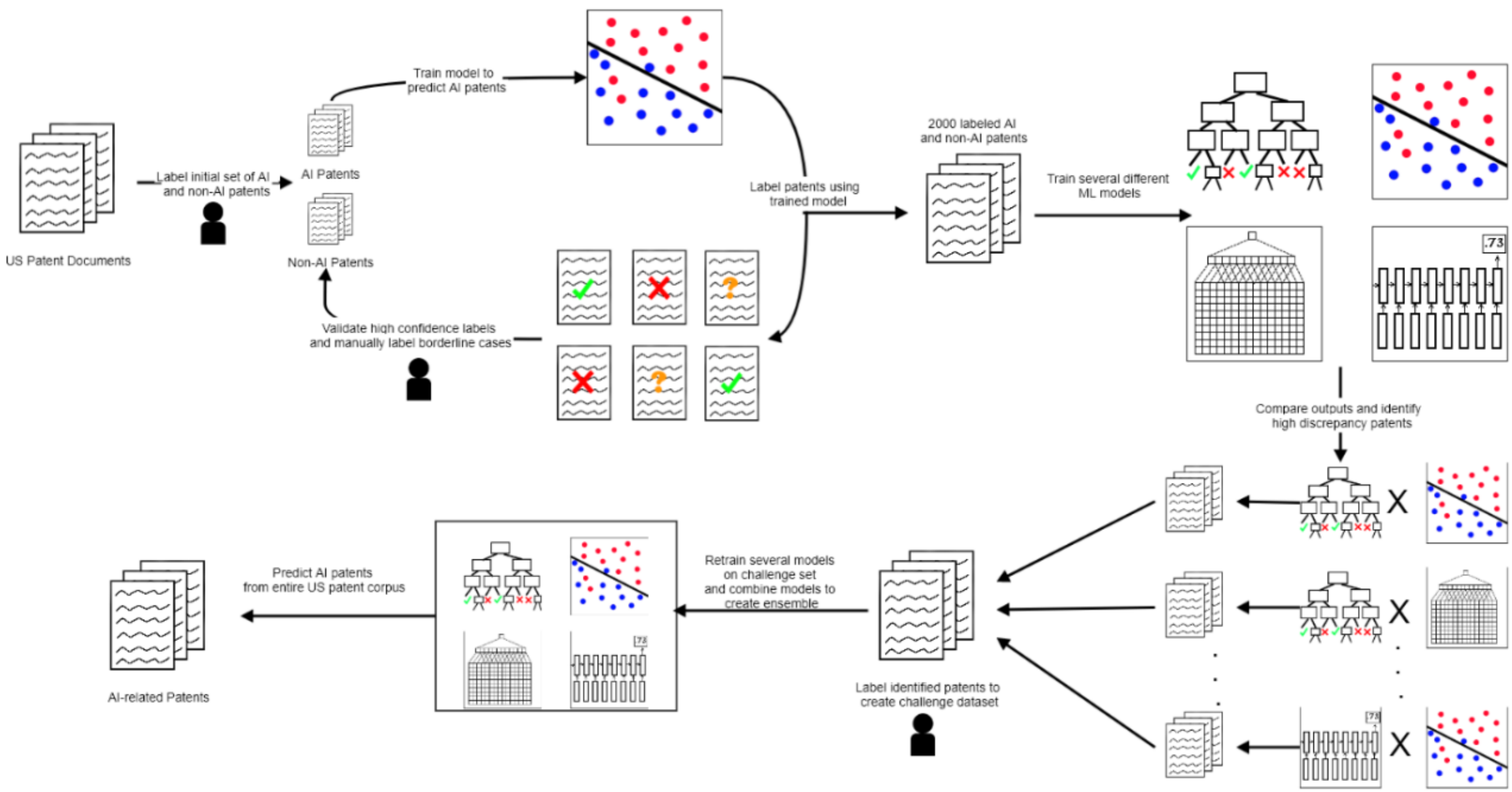


Figure 1. AI Patent Identification Algorithm

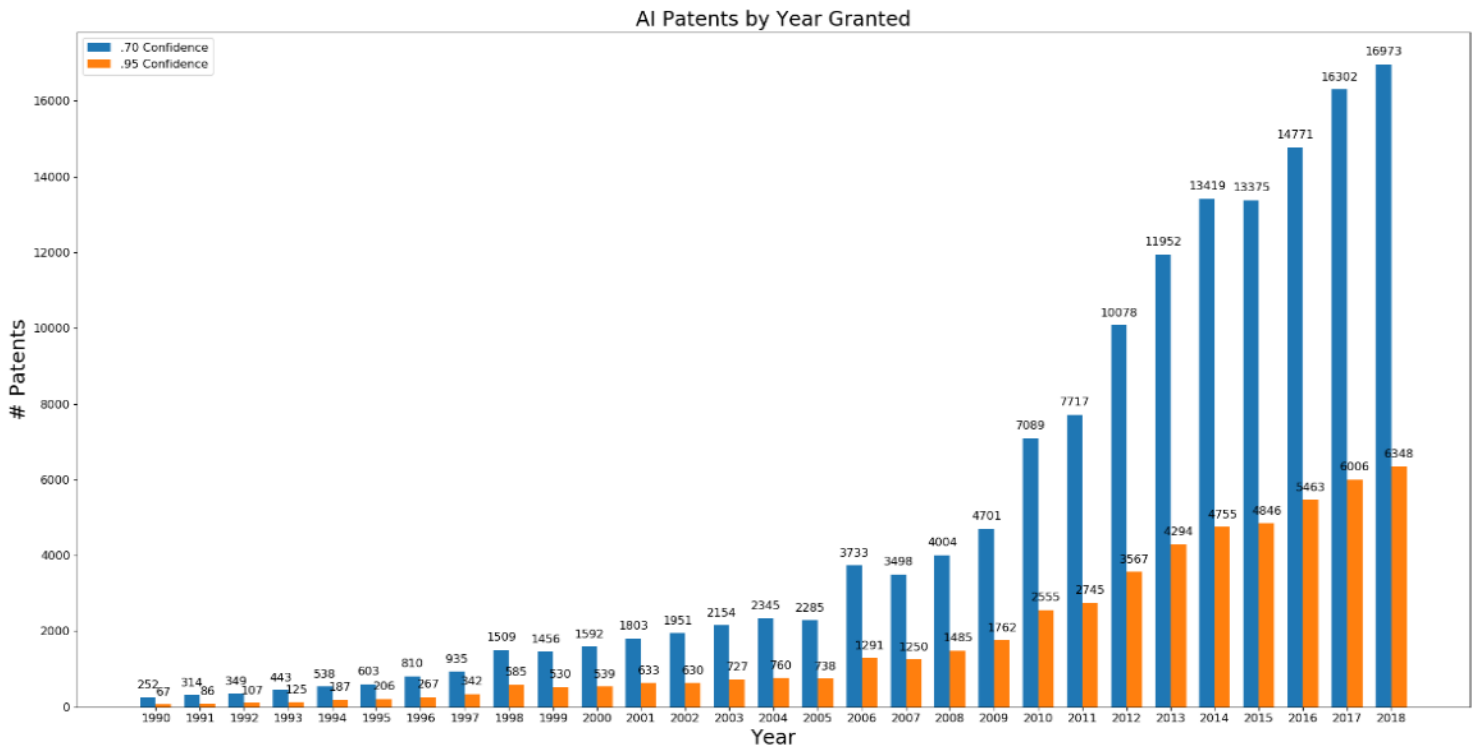


Figure 2. AI Patents by Grant Year

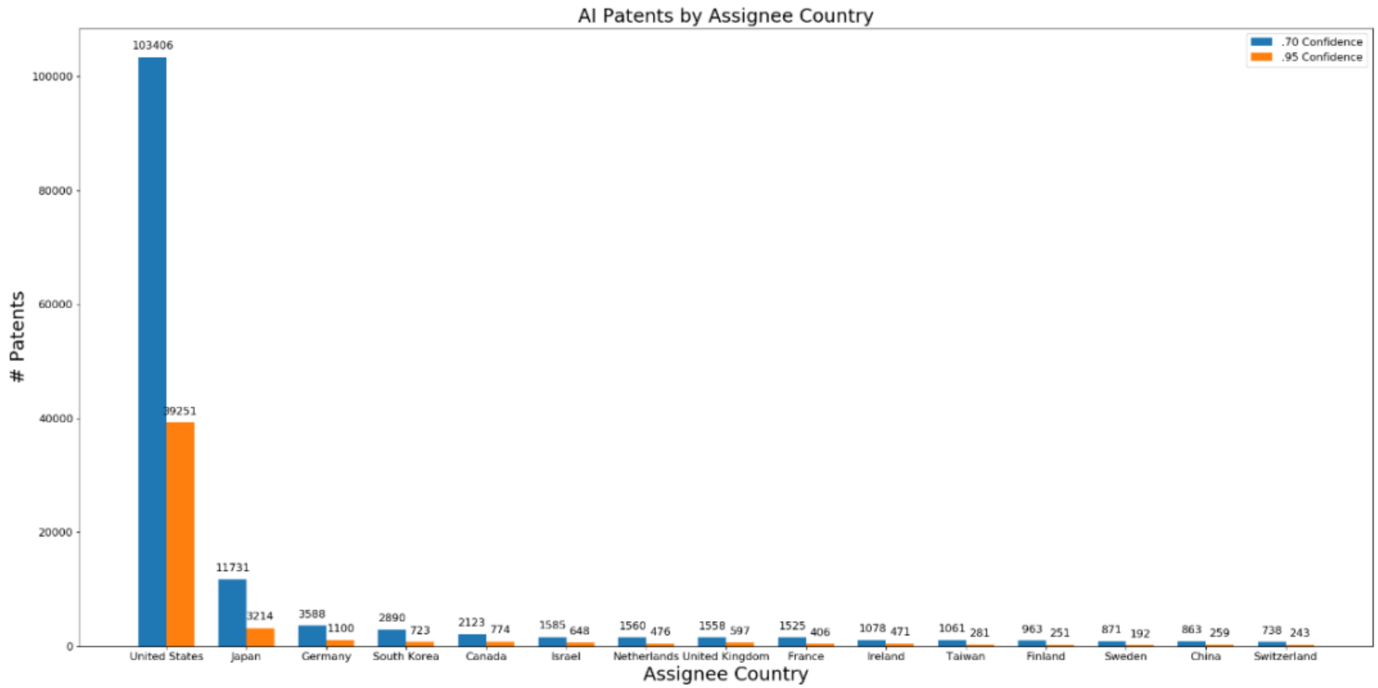


Figure 3. AI Patents by Country

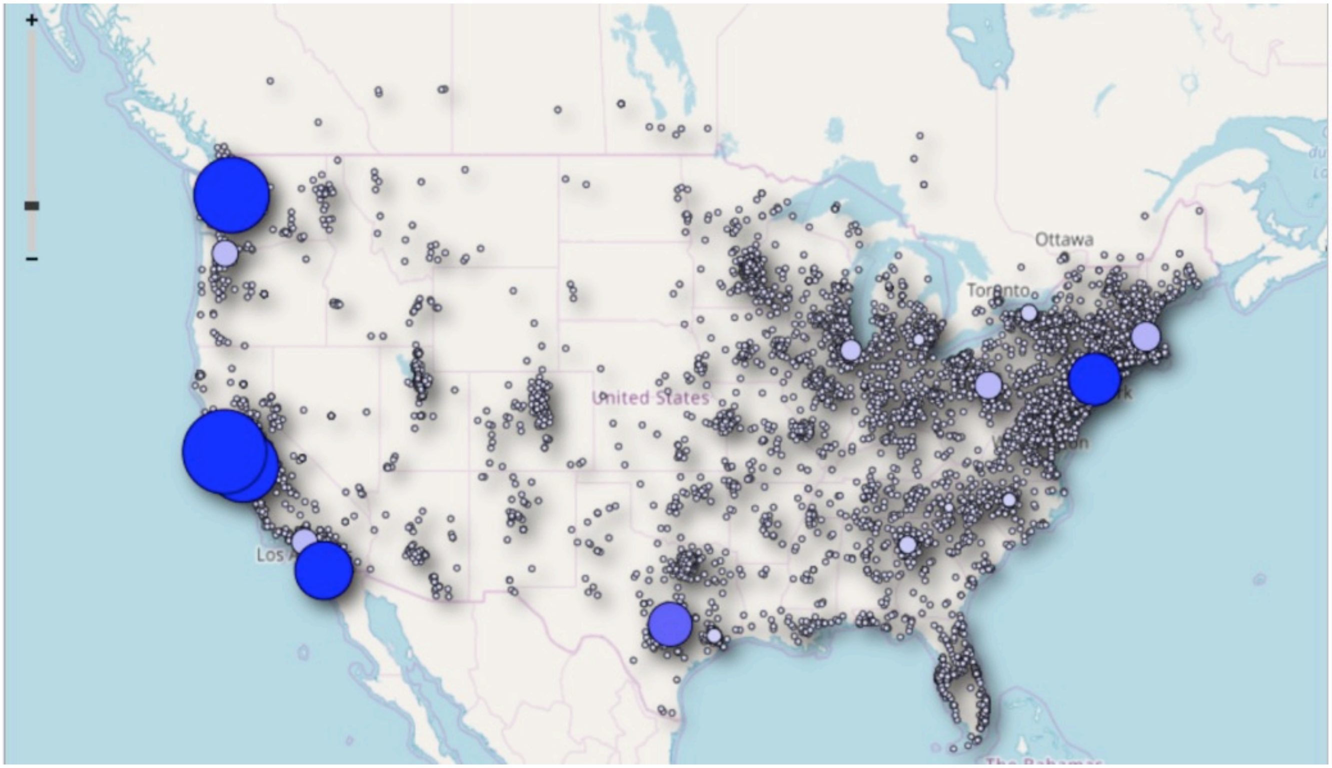


Figure 4 Inventor Heat Map of AI Patents in U.S

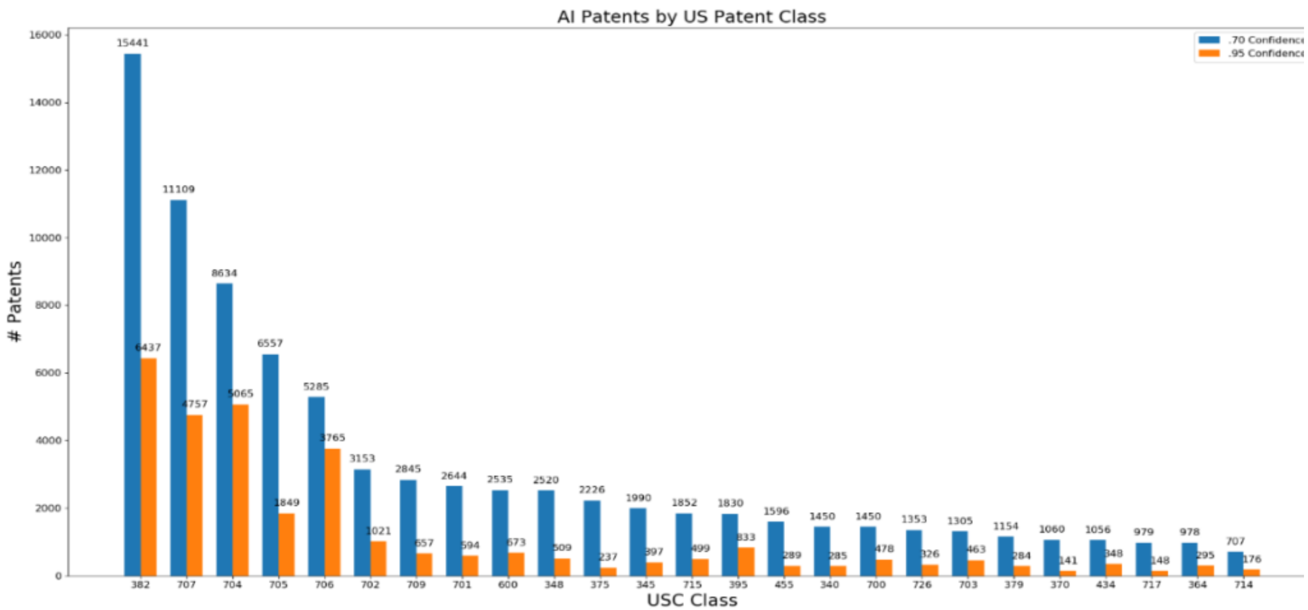
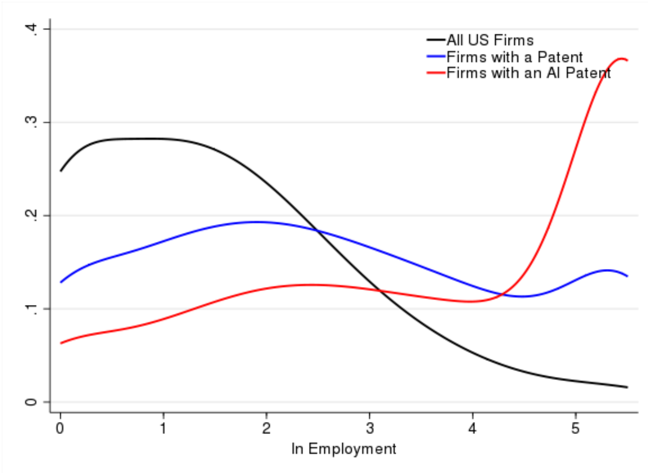
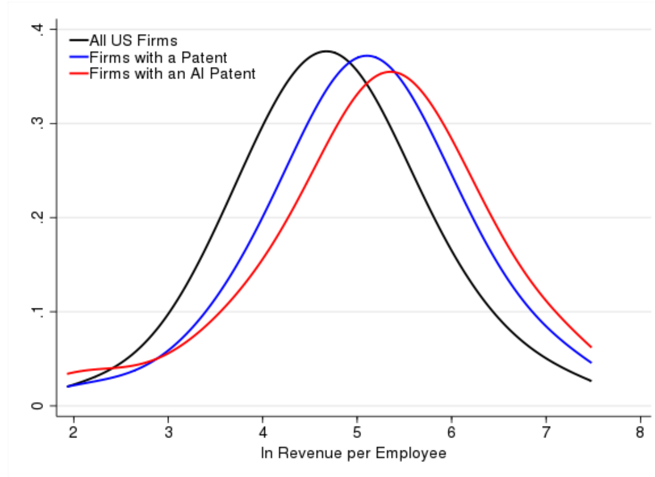


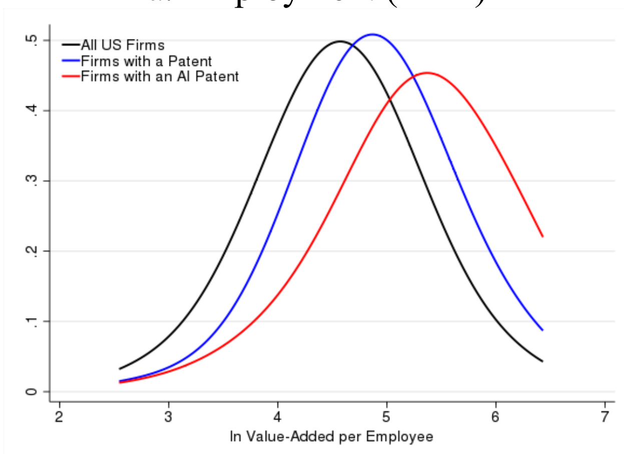
Figure 5. AI Patents by USPC Class



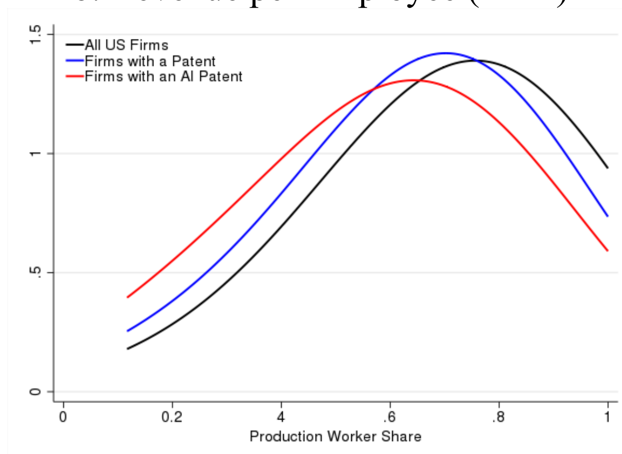
a. Employment (KDE)



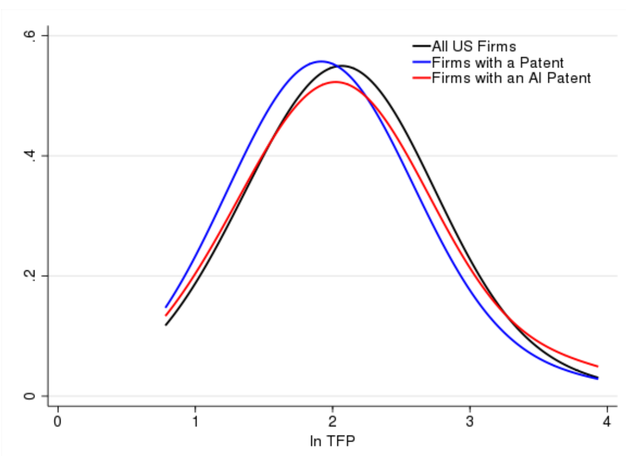
b. Revenue per Employee (KDE)



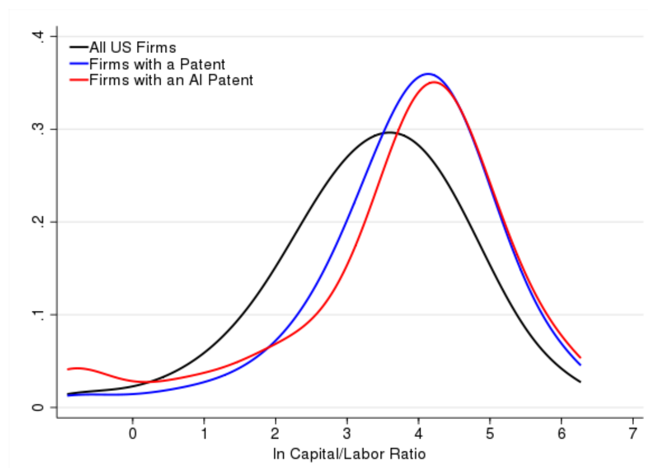
c. Value-Added per Employee (KDE)



d. Production Worker Share (KDE)



e. Total Factor Productivity (KDE)



f. Capital/Labor Ratio (KDE)

Figure 6: Kernel Density Plots of Firm Characteristics by Patenting Type

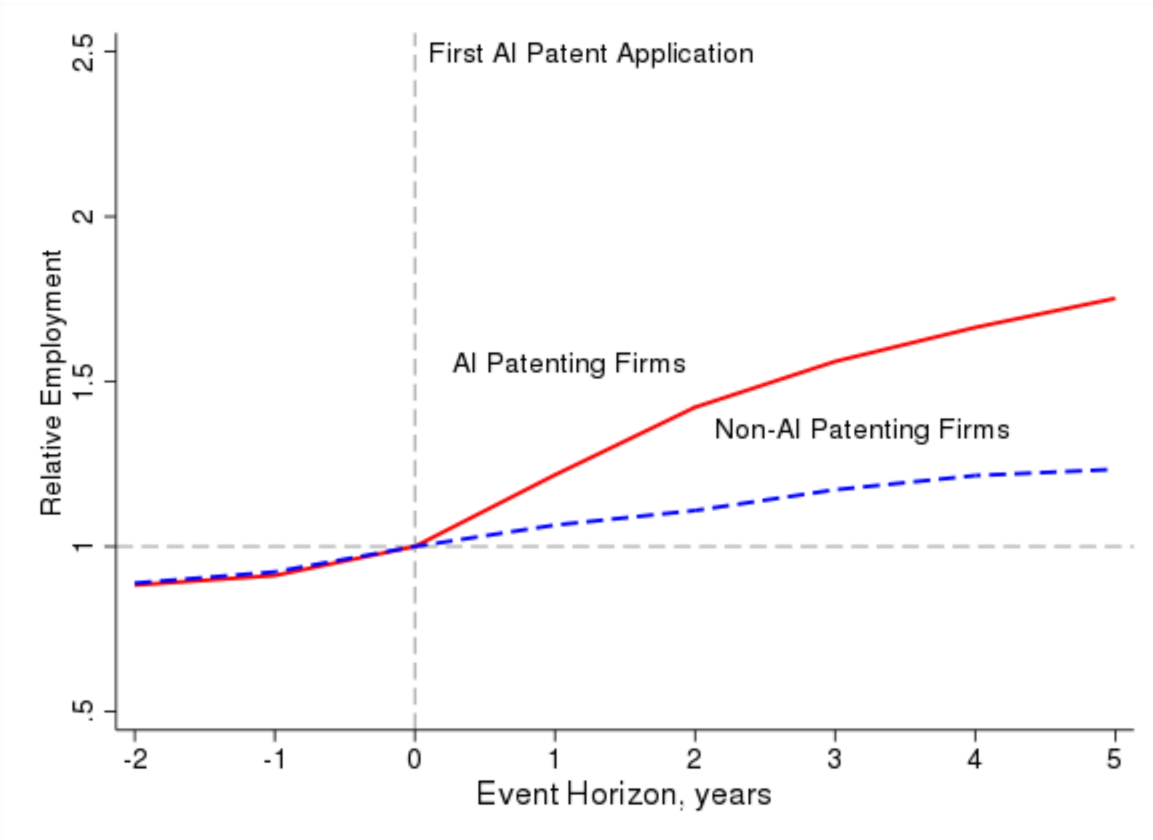


Figure 7: Event Study Plot on Relative Employment between AI Patenting and non-AI Patenting Firms

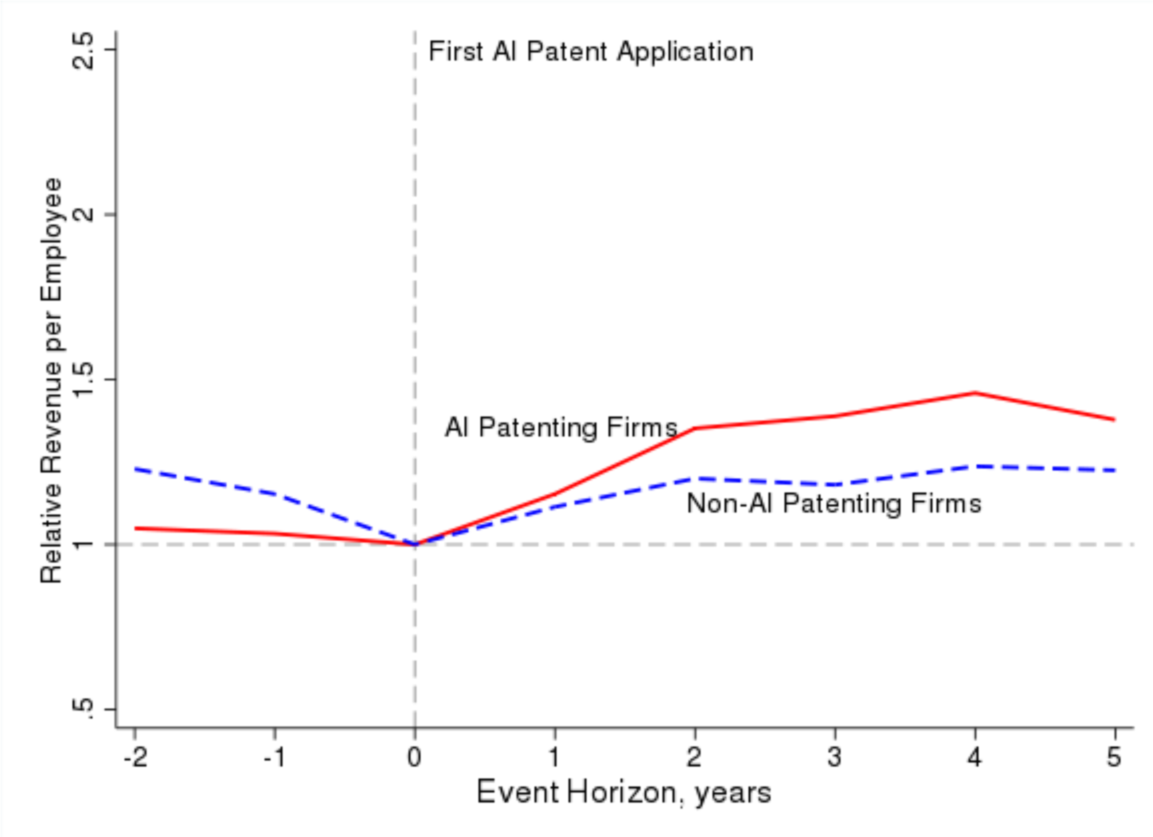


Figure 8: Event Study Plot on Relative Revenue per Employee between AI Patenting and non-AI Patenting Firms

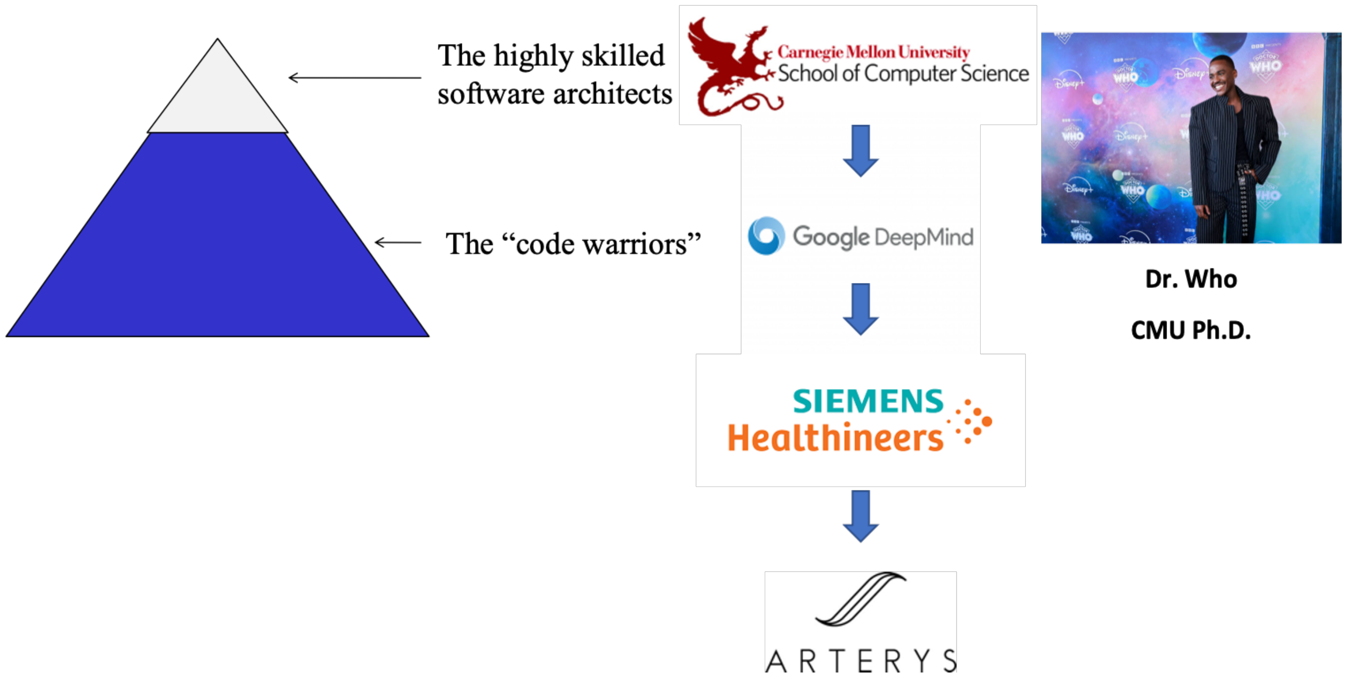


Figure 9. Tracing the Impact of AI Software "Architects": A Suggestive Illustration