Estimating the Impact of Artificial Intelligence on Jobs Within the Healthcare Industry

Sidharth Satya

University of California, Berkeley

Economics Undergraduate Honors Thesis

Advisor: Enrico Moretti

May 5th, 2021

Abstract

The rapid growth of the artificial intelligence field, as well as its ability to impact nearly every economic sector, underscores the importance of understanding its potential impact on labor. This paper examines how wages and employment levels for jobs in the healthcare industry are affected by artificial intelligence as well as the net effect of two direct impacts of AI through a difference-in-differences approach. The results show that wages and employment for Physicians and Surgeons increased after the introduction of IBM Watson for healthcare applications in 2013 while no significant effect was found for Secretaries and Administrative Assistants in the healthcare industry. Overall, the results suggest that when artificial intelligence augments labor on decision tasks, it has a net positive effect whereas the effect of automating prediction tasks remains ambiguous.

Acknowledgement

I would like to extend my deepest gratitude to my advisor, Professor Enrico Moretti for his support, guidance, and thoughtful suggestions throughout the completion of my thesis. I would also like to thank Professor Emily Tang for helping me formulate my research question and identify reliable readings and data sources on which to base my research.

1 Introduction

One of the topics that has captured much of the public's attention in recent years is the rise of artificial intelligence. Today, ordinary people interact with a variety of applications and machines that employ artificial intelligence capabilities. Whether it is voice recognition technologies such as Siri and Cortana, the powerful data-crunching algorithms backing social media giants such as Facebook and Twitter, or the self-driving capabilities of Tesla, artificial intelligence has had a large impact upon our daily lives. The opportunities of such a technology are seemingly endless, as companies rush to adopt more and more artificial intelligence applications in the hopes of automating as much of their businesses as possible. AI applications are being implemented in almost every single industry and every single job. For example, artificial intelligence has had a large impact on the financial industry and is even being used in applications focused on national defense (West & Allen, 2018). Investments in artificial intelligence are rising as well. The Brookings Institute estimates that in 2019 alone, AI-focused companies "attracted nearly $40 billion globally" (Arnold, 2020).

One of the industries that has been most impacted by advancements in artificial intelligence is the healthcare industry. The ability of artificial intelligence to detect anomalies in images and predict the existence of certain diseases has already shown promise. For example, algorithms have begun outperforming radiologists in detecting malignant tumors and the technology is showing even more promise in other areas of medical diagnosis (Yu et al., 2018). It has also become extremely prevalent in administrative applications in healthcare-related fields, due to the high overhead and time spent on regulatory and administrative tasks (Davenport & Kalakota, 2019). Despite the ethical and legal hurdles that artificial intelligence still has to overcome in order to fully integrate into the medical diagnosis field, it is clear that it has already shaped outcomes for jobs in the healthcare industry. Understanding how the technology has grown and impacted individuals within the healthcare industry can provide us with insight into the early effects of artificial intelligence for labor.

As artificial intelligence continues to grow as a field and as a major player in our daily lives, it is also being met with backlash. Due to the unpredictable nature of the state of artificial intelligence advancement, as well as its ability to perform some tasks even better than humans, there is much uncertainty surrounding AI and its potential impact on jobs (Schmelzer, 2019). For example, many call centers have been shut down during the Coronavirus Pandemic as the main labor force shifted towards AI-focused applications that can handle large amounts of callers at a fraction of the cost for employers (Samuels, 2020). The tradeoff between reducing costs with automation in exchange for human-led jobs reveals a potential downside to such unhindered technological advancement. This is not a new problem, either, as automation and its impact on labor has long been a large topic of research. However, artificial intelligence deviates from previous forms of automation, such as robotics, due to its predictive capabilities that make it possible to replace analytical, thinking tasks that were previously thought to be irreplaceable (Huang et al., 2019). Therefore, due to the role of artificial intelligence as a novel form of automation that already has a tremendous influence, it is extremely critical that we understand the economic impact that advances in AI can have upon wages and employment in order to ensure proper public policy and legislative action is taken to protect individuals' livelihoods.

Much of the current economic literature surrounding artificial intelligence is focused on understanding the potential impacts that it can have and coming up with new methods to empirically isolate its effect. Acemoglu and Restrepo (2018) develop a task-based model that accounts for AI's ability to affect both low-skill and high-skill labor -- something which did not exist in previous models due to the focus of previous automation on routine, manual tasks. Webb (2019) builds upon the task-based approach by scraping US patent texts and job task descriptions to develop "exposure scores" that measures how much a job is exposed to artificial intelligence. Agrawal et al. (2019) also adopt a task-based focus and identify four direct effects of automation via artificial intelligence on jobs. All three pieces of work identify ways in which artificial intelligence can have both a positive and negative effect on wages and employment and demonstrate how the net effect of automation via AI applications is still ambiguous.

In this research paper, I will build upon the conclusions formed in previous literature and examine the impact that artificial intelligence has had upon wages and employment for two types of jobs within the healthcare industry: Physicians and Surgeons and Secretaries and Administrative Assistants. This question is extremely relevant in the context of the current literature since the effect of commercial intelligence applications in specific industries is still relatively unknown and still a burgeoning field of research. Furthermore, while many of the findings in existing literature examine general trends in recent automation, specific case studies for different industries and jobs are scarce. Therefore, it is critical to understand the effect that artificial intelligence has already had on jobs in order to inform future research and public policy decision making. Due to the ability of artificial intelligence to support and increase labor productivity on decision making tasks, I hypothesize that it will have a positive effect on wages and employment for Physicians and Surgeons. Conversely, due to the routine nature of the tasks performed by Secretaries and Administrative Assistants, as well as AI's ability to automate such tasks, I hypothesize that their wages and employment level will be negatively affected. Furthermore, I hypothesize that jobs that are more human-focused and interpersonal, such as Nursing, Psychiatric, and Home Health Aides, will be unaffected by advancements in artificial intelligence.

For my research, I collected data from the Integrated Public Use Microdata Series (IPUMS USA) which provides samples over multiple years from the American Community Survey. Using data ranging from 2001 to 2017, I compiled data on yearly wages, employment, age, race, gender, and education levels at the individual-level. Through grouping the data into industry-occupation-year cells, I was able to generate observations for each industry and occupation over time. I also collected data from the OECD on that number of artificial intelligence related patents in the world, by year.

In order to test my hypotheses, I first designed a difference-in-differences approach, with Physicians and Surgeons and Secretaries and Administrative Assistants as two separate treatment groups and Nursing, Psychiatric, and Home Health Aides as my control group. I used the announcement of IBM Watson for commercial healthcare applications in 2013 as the treatment due to its role as a milestone in the development of commercial artificial intelligence applications. I then ran a series of regression with this design taking into account demographic controls such as race, gender, educational level, and age to test how wages and employment levels for the treatment groups were affected by the treatment.

I also ran a series of separate Ordinary Least Squares regressions with the outcome variables as wages and employment levels once again and the number of artificial intelligence patents each year as the explanatory variable for each occupation used in the difference-in-differences design. Once again, I included the same demographic controls used in the difference-in-differences design.

My results show that the treatment effect of the 2013 IBM Watson announcement for healthcare applications had a positive effect on wages and employment for Physicians and Surgeons. However, no significant treatment effect was observed for either wages nor employment for Secretaries and Administrative Assistants. Additionally, the results show that wages for Physicians and Surgeons are significantly associated with the number of artificial intelligence related patents while employment levels for Secretaries and Administrative Assistants are negatively associated with the number of patents. No significant effect of the number of AI-related patents on wages or employment was observed for the occupation that I used as the control group in the difference-in-differences regressions: Nursing, Psychiatric, and Home Health Aides.

The positive treatment effect on wages and employment for Physicians and Surgeons suggests that when automation in the form of artificial intelligence augments labor on decision tasks, the net effect is positive. However, the lack of a significant effect on wages and employment for Secretaries and Administrative Assistants suggest that the net effect of artificial intelligence when it automates prediction tasks remains ambiguous and requires further study. The lack of significant effects found for Nursing, Psychiatric, and Home Health Aides remains in line with my reasoning for including them as my control group.

The paper is developed in the following order. Section 2 performs a review of the existing literature surrounding the effect of artificial intelligence on jobs as well as how my research contributes to this field. Section 3 provides an in-depth overview of the data sources used as well as some descriptive statistics. Section 4 displays the regression models that I developed along with a justification for each of my choices and a discussion on potential shortcomings within my models. Section 5 presents the results of the regressions as well as an interpretation of the results. Finally, Section 6 summarizes the findings and connects them back to my research question as well as conclusions drawn in existing literature. The Appendix contains relevant figures and tables that are referenced throughout the paper.

2 Literature Review

The focus of this paper is centered around estimating the impact of artificial intelligence applications on wages and employment for different types of occupations within the healthcare industry. The approach taken to conduct this research is highly motivated by existing literature that is concerned with understanding the effect that artificial intelligence and automation may have on labor and inequality. Furthermore, it builds upon past research by providing a specific context for analysis as well as an econometric approach that is used to evaluate the conclusions formed by other authors.

2.1 Existing Literature and Current Findings

A key quality of artificial intelligence is that it is a form of automation. As a result, it is important to understand the impact that automation has on different types of jobs. A relevant and highly influential piece of work is that done by Acemoglu and Restrepo (2018). In this paper, the authors separate workers into two types: “low-skill” and “high-skill” workers. They then develop a task based model that compares the performance of low-skilled and high-skilled workers against machines on different types of tasks. The primary motivation for this design is that previous research surrounding automation has been concerned with the automation manual labor and routine tasks. However, recent advancements in artificial intelligence and “big data” applications have demonstrated the ability to perform better on tasks where human judgement has been thought to be irreplaceable. Using analysis from the model that they develop, the authors show that automation has a “displacement” effect, where it displaces the type of labor that it directly affects and depresses its wage. However, it also creates a positive “productivity” effect, which can increase wage and the price of all affected factors. Thus, the net impact of automation on any type of labor depends on which effect dominates the other. Furthermore, one of the most interesting conclusions of the paper is that while the effects of automation on wages are ambiguous, low-skill automation always increases wage inequality while the opposite is true for high-skill automation.

A similar piece of research that builds upon the conclusions formed in Acemoglu and Restrepo (2018) is the work done by Michael Webb (2019). In his research, Webb compares job task descriptions, provided by O*NET, with patent descriptions to construct “exposure scores” that measure to what level different jobs, according to the *occ1990dd* classification, are exposed to automation. He separates automation into 3 parts: software, robotics, and artificial intelligence and develops separate exposure metrics for each of these categories. He then runs regressions with changes in wages and employment between 1980 and 2010 as the outcome variables as interest and exposure scores as the independent regressor for different occupations. The results show that low-skill occupations are most exposed to advances in robotics, middle-skill occupations are most exposed to advances in software, and high-skill occupations are most exposed to artificial intelligence. Interestingly, the results also show that individuals with higher levels of education and higher mean age are more exposed to artificial intelligence than robotics and software. Webb's research serves as a first step to estimating the impact of artificial intelligence on different types of jobs and supports the conclusions drawn in the work done by Acemoglu and Restrepo (2018).

Agrawal et al. (2019) also use a task-based approach to predict the impact that artificial intelligence may have upon jobs. The authors consider the predictive capabilities of artificial intelligence and aim to understand how AI applications could impact the decision making process. They identify four direct effects through which artificial intelligence could affect tasks: substituting capital for labor in prediction tasks, automating decision tasks, enhancing labor and increasing productivity on decision tasks, and creating new decision tasks. Real-world examples of artificial intelligence applications and advancements are identified to demonstrate how these forces could displace or enhance labor for specific jobs. Ultimately, the authors find that jobs are impacted by multiple effects and the net effect of artificial intelligence is ambiguous and varies across industries and occupations. Still, this work provides an important distinction between the impacts that artificial intelligence can have on jobs and provides a foundation that can be used to estimate the effect that artificial intelligence can have on wages and employment on labor.

Huang et al. (2019) is a work focused on understanding the impact that artificial intelligence has had on the types of tasks in the economy. The authors argue that a "Feeling Economy" is emerging, where mechanical and analytical tasks are performed by machines with AI capabilities. As a result, jobs are shifting to place more emphasis on empathetic and interpersonal tasks. The authors categorize the tasks provided by the O*NET database for jobs that exist between 2006 and 2016 into three categories: mechanical, thinking, and feeling. They then develop metrics that measure the relative importance of each type of task and calculate the change in these metrics between 2006 and 2016. They find that feeling tasks are becoming more important and that wages and employment for feeling tasks are growing. Furthermore, they find that industries that employ many jobs with feeling tasks will benefit the most from this emergence and that advances in artificial intelligence will aid this growth.

2.3 Contributions of Current Research to Existing Literature

Currently, existing research outlines potential impacts of artificial intelligence on tasks and finds general impacts on different types of jobs. However, a main limitation that still exists is that the general impact of artificial intelligence on wages and employment can be ambiguous due to its varying effects on different types of jobs and industries. My research aims to fill this gap in the literature by providing a specific case study that measures the changes in employment and wages for occupations within the healthcare industry. Furthermore, the results of my research can be used to evaluate whether the conclusions drawn by other authors hold in this context.

Both my approach and research design are heavily influenced by the work done by previous authors. I identify occupations within the healthcare industry that are heavily influenced by two of the four direct effects found in Agrawal et al. (2019): automation replacing labor on prediction tasks and automation augmenting labor on decision tasks. In order to identify these jobs, I employ a similar approach to that used in Hwang et al. (2019). I label the tasks provided by the O*NET database for each occupation and separate the jobs that I identify as only being affected by one of those two impacts and run separate regressions to compare outcomes for each occupation. In doing so, I am able to provide an example of how these separate impacts might affect the wages and employment for jobs within the healthcare industry and evaluate whether or not these estimated effects support the findings of Webb (2019) as well as the task-based model developed in Acemoglu and Restrepo (2018).

Overall, my research serves as a first step in using an econometric approach to understanding the impact of artificial intelligence on jobs within specific industries. This is something that is still lacking in the current literature due to the difficulty of finding reliable data as well as isolating the impact of artificial intelligence. My hope is that this research can be used as a foundation for future research that aims to measure and predict the effect of introducing different types of artificial intelligence applications in a commercial setting.

3 Data

3.1 Sources of Data

3.1.1 US Census

My primary source of data is the population survey data extracted from U.S. census records and the American Community Survey samples in the Integrated Public Use Microdata Series (IPUMS USA). I gathered samples from each year from 2000 to 2017, but I restricted my analysis to samples from the years 2005 to 2017. Furthermore, the sample is restricted to individuals aged 18-65 that are full-time, full-year employees that have been employed in the previous year. I define full-time employment as working more than 35 hours per week and full-year employment as working more than 40 weeks per year. Finally, I filtered out data points that had recorded statistics of wage and education equal to 0.

I clean the IPUMS data by collapsing the data into industry-occupation-year observations, using the *ind1990* and Standard Occupational Classifications (*occsoc)*. The full list of medical-related industries and occupational codes can be found in Appendix A. I construct the wage variable by converting the yearly income (*incwage)* for each observation into real (1999 dollars) using conversion rates provided by IPUMS[1]. Furthermore, I weight each wage

[1] The conversion rates to 1999 dollars can be found at https://usa.ipums.org/usa/cpi99.shtml

observation by the total number of individuals in each industry-occupation-year grouping using the IPUMS survey person weight (*perwt)* in order to create mean wage observations. I construct the employment variable by first restricting the observations with employment in the previous year. I then use the IPUMS survey person weight (*perwt*) for each observation. Since each observation in the IPUMS samples is representative of a larger group that *perwt* measures, by restricting the samples to observations with employment in the previous year, I am able to obtain total employment for each industry-occupation-year observation.

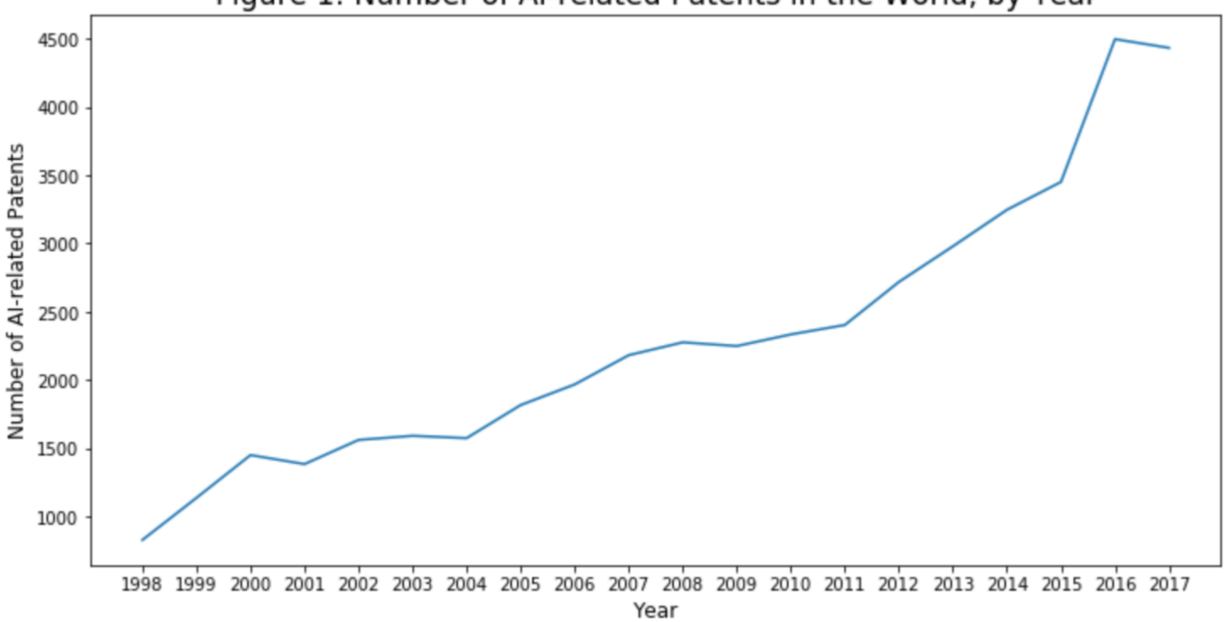
Additional variables that I extract from the IPUMS samples are *educ, sex, race,* and *age*. I use these variables to construct metrics regarding the fraction of individuals that have completed at least one year of college, the fraction of females, the fraction of white-identifying individuals, and the mean age for each industry-occupation-year grouping. For each of my regressions, I log-transform these variables.

One of the largest limitations regarding the IPUMS dataset that I had to consider before deciding on the design of my research was that the American Community Survey samples do not provide the ability to track individuals over time due to confidentiality concerns. As a result, it was not possible to get usable individual-level data over time. I therefore decided to aggregate all observations by industry-occupation-year cells as described above in order to track how outcomes for groups of individuals change over time.

3.1.2 OECD Data

I also extracted the number of patents related to artificial intelligence by country over time from the Organisation for Economic Cooperation and Development (OECD) statistical database. The data contains the number of patents related to artificial intelligence technologies for both OECD and non-OECD countries. I restrict the data to patents belonging to IP5 Patent Families, the priority date of the patent -- relating to a patent's earliest filing date -- and the inventor's country of residence. Finally, I collapse all observations into year cells which contain the total number of patents relating to artificial intelligence in the world by year. The data contains 1,000,000 observations for patents relating to different technologies and patent families by country. By grouping in the manner described above, I narrowed down the data to just 20 observations that hold the number of patents relating to AI in the world, by year. The graph below shows the trend in the number of artificial intelligence patents from 1998 to 2017. We can see a sharp increase in the number of patents starting in 2011, which is around the timeframe where I focus my analysis and treatment.

Figure 1: Number of AI-related Patents in the World, by Year



3.2 Summary Statistics of Key Variables

Appendix C contains tables that summarize key statistics of the data. We can see from Table C1 that Offices and Clinics of Physicians and Hospitals contain the majority of observations relating to Physicians and Surgeons, as well as Secretaries and Administrative Assistants. In contrast, Hospitals and Nursing and Personal Care Facilities contain a majority of the observations relating to Nursing, Psychiatric, and Home Health Aides. There are a total of 201,877 observations, which were then grouped into industry-occupation-year cells.

We can see in Table C2 that the average fraction of Physicians and Surgeons that have completed at least one year of college education is almost 1, the average fractions for white-identifying and male are over 0.5, and the average wage and employment is about $146,749 and 185,065, respectively. The average fraction of all Secretaries and Administrative Assistants that have completed one year of college education is about 0.53, the average fractions of white-identifying and female individuals are about 0.79 and 0.97, respectively, and the average wage and employment is about $25,438 and 95,433, respectively. Finally, the average fraction of all Nursing, Psychiatric, and Home Health Aides that have completed at least one year of college education is about 0.47, the average fraction that are white identifying is about 0.59, and the average fraction that are female is about 0.86. The average wage and employment level is $24,243 and 259,183, respectively. The means were calculated over all years of observations.

4 Description of Empirical Methods

4.1 Difference-in-Differences Regressions

The main focus of my empirical analysis was centered around a series of Difference-in-Differences regressions that I ran in order to understand how employment and wages for different jobs are affected by the introduction of artificial intelligence technologies. The estimated model for the Difference-in-Differences regressions are as follows:

$$Y_{iot} = \beta_0 + \beta_1 AI_o + \beta_2 AI_o Post_t + \beta_3 D_{iot} + \beta_4 \delta_t + \beta_5 C_i + \varepsilon_{iot}$$

Y_{iot} is the outcome variable of interest for occupation *o* in industry *i* in time *t*. The outcome variables that I measured are the mean of log wages and the mean of log employment for each occupation. $Post_t$ is a dummy variable equal to 1 if the current time period, *t,* is during or after the treatment has been applied. C_i is a vector related to industry-level fixed effects, δ_t is a vector relating to time-effects, and D_{iot} is a vector containing covariates such as the proportion of people that completed at least one year of college, the proportion of females, the proportion of white-identified people, and average age for each industry-occupation-year observation. Each of these covariates were log-transformed in the estimated regression. AI_o is a dummy variable that is equal to 1 if the occupation is part of the treatment group and ε_{iot} is the error term. The coefficient B_3 is the difference-in-difference estimator that relates to the treatment effect. An important note is that I omitted a separate indicator variable for the treatment time, $Post_t$, as is used in a simple two-period difference-in-differences model due to the inclusion of time-effects.

I restrict the analysis to full-time, full year employees in medical-related industries defined according to the IPUMS *ind1990* classification (Appendix A). The time-period for these regressions is limited to 2005 to 2017. The treatment in my regression analysis is the introduction of IBM Watson for medical applications and assistance with clinical procedures in 2013. I used the Nursing, Psychiatric, and Home Health Aides as the control group. I ran two separate difference-in-differences regressions for each outcome variable: one with Physicians & Surgeons as the treatment group and one with Secretaries & Administrative Assistants the treatment group. All occupations were restricted to those within healthcare-related industries. As stated above, the outcome variables that were measured are the mean of log wages and log employment for each group.

4.1.1 Justification of Treatment

One of the most important milestones in the recent history of artificial intelligence is the introduction of IBM Watson, a question-answering computer system that employed artificial intelligence techniques such as natural language processing (NLP) to handle query processing and information retrieval. Watson shocked the world in 2011 by participating in the famous game show, Jeopardy!, against human competitors and subsequently winning the first-place prize. IBM Watson was later transformed into a business-facing unit, whose data-crunching abilities and artificial intelligence capabilities could be used in all facets of business. Later in 2011, IBM announced several partnerships with medical companies to develop commercial applications to support clinical procedures. In February 2013, IBM and its partners, WellPoint and the Memorial Sloan-Kettering Center for Cancer Research, announced the first commercial application of Watson for utilization management decisions in lung cancer treatment (IBM, 2013).

Watson serves as an important point in the history of artificial intelligence not only for its remarkable performance in Jeopardy!, but also as a representative of one of the first use cases of artificial intelligence applications within the healthcare industry. This period marks a point in recent history when artificial intelligence applications were starting to be adapted for commercial use. Thus, I decided to use the announcement of the first commercial application of Watson for healthcare applications in 2013 as the treatment in the Difference-in-Differences regressions. An important point to note is that rather than trying to estimate the individual impact of the announcement of Watson for healthcare applications, I am using the 2013 announcement as a marker for when artificial intelligence applications started to affect jobs within the healthcare industry.

4.1.2 Justification of Treatment and Control Groups

In their research, Agrawal et al. (2019) highlight four different ways in which advancements in artificial intelligence can impact jobs. I look specifically at two of these potential impacts: Augmenting Labor on Decision Tasks and Automating Prediction Tasks. For Augmenting Labor on Decision Tasks, the authors assert that artificial intelligence can enhance labor when automating prediction tasks increases labor productivity. In contrast, for Automating Prediction Tasks, the authors state that capital will be substituted for labor in prediction tasks such as responding to emails or scheduling. Using these two impacts, I highlighted jobs within the Healthcare industry that have a disproportionate amount of tasks that would be affected by one of these impacts.

Occupations that fit labor augmentation of decision tasks are Physicians and Surgeons, as artificial intelligence applications can help them make more accurate diagnoses or assist with treatment. However, these applications are yet to fully automate these jobs since a human-led treatment or practice remains necessary for all legal procedures. The authors also highlight a couple jobs that have many prediction tasks that can be automated using artificial intelligence. One such job is Medical Secretaries, due to the variety of tasks that they perform relating to processing and retrieval of information, responding to emails and queries, and assisting physicians and surgeons. Therefore, I also used data on wages and employment for Secretaries and Administrative Assistants within healthcare-related industries.

In 2013, artificial intelligence applications demonstrated the ability to process and write data, perform scheduling tasks, and generate messages (Best, 2013). These capabilities are geared towards automating the tasks that Secretaries and Administrative Assistants perform, such as answering telephones and directing calls, maintaining medical records, and transmitting correspondence. While Physicians and Surgeons may interact with the technology, there are no commercial applications that can automate core tasks such as treating patients and performing surgery. We can see that 13 tasks out of 16 total tasks are susceptible to automation for prediction tasks for Secretaries and Administrative Assistants using artificial intelligence whereas only 19 tasks out of 251 total tasks can be automated for Physicians and Surgeons according to the 2019 O*NET Classifications (Appendix B, Table B3). Conversely, none of the tasks for Secretaries and Administrative Assistants are susceptible to augmenting labor on decision tasks while 81 out of 251 total tasks are susceptible for Physicians and Surgeons (Appendix B, Table B2). A full discussion on how I determine which tasks would be subject to each impact as well as descriptions of each task are included in Appendix B.

Due to the effect on tasks for the two separate occupations above, I classified both occupations as two separate treatment groups. Physicians and Surgeons represent a treatment group that is affected by artificial intelligence on decision tasks while Secretaries and Administrative Assistants represent a separate treatment group that is affected by artificial intelligence on prediction tasks that can be automated.

In order to select the control group, I repeat the steps above for the two treatment occupations and find an occupation in the healthcare industry that is not affected by advances in artificial intelligence. Nursing, Psychiatric, and Home Health Aides contain only 2 out of 65 tasks that are susceptible to automation on prediction tasks and no tasks that are subject to augmenting labor on decision tasks (Appendix B). Thus, I use Nursing, Psychiatric, and Home Health Aides as the control group for both regressions due to the minimal impact of artificial intelligence on the tasks for these occupations.

4.3 Assumptions and Potential Violations of the Difference-in-Differences Model

4.3.1 Parallel Trends Assumption

A major assumption for the difference-in-differences design is that in the absence of the treatment, the treatment and control group will display parallel trends over time. I account for this by observing that the change in the log of the outcome variable of interest for the treatment and control group displays parallel trends in the period before the treatment. Further discussion on the parallel trends between the treatment and control group is shown in the results section of this paper along with visualizations for each occupation's trends. Should the parallel trends assumption be violated however, the causal effect of the treatment cannot be determined since changes between the treatment and control group cannot be solely attributed to the treatment effect.

4.3.2 Omitted Variable Bias

Another major assumption that must hold in order for the causal treatment effect to be correctly estimated is that the error term in the regression is not correlated with the dependent variable, or outcome of interest. If this is violated, and if the error term is correlated with at least one of the independent covariates, then the estimates in the regression will be biased. For example, suppose the true causal model is

$$Y_{iot} = \beta_0 + \beta_1 X_{iot} + \beta_2 Z_{iot} + \varepsilon_{iot},$$

Where X_{iot} is a vector of all covariates included in the estimated difference-in-differences model as specified above, Z_{iot} is an omitted variable in the estimating model, and $Cov(X_{iot}, Z_{iot}) \neq 0$. Also, assume that the coefficient on Z_{iot} is not equal to 0, meaning that Z_{iot} is a determinant of Y_{iot}. Then, we can consider an auxiliary regression between the omitted variable and all other observed regressors:

$$Z_{iot} = \pi_0 + \pi_1 X_{iot} + \nu_{iot}$$

Substituting this into the true causal model uncovers the bias in estimates that is introduced via omitted variable bias.

$$Y_{iot} = \beta_0 + \beta_1 X_{iot} + \beta_2(\pi_0 + \pi_1 X_{iot} + \nu_{iot}) + \varepsilon_{iot}$$

$$\Rightarrow Y_{iot} = (\beta_0 + \beta_2\pi_0) + (\beta_1 + \beta_2\pi_1)X_{iot} + (\beta_2\nu_{iot} + \varepsilon_{iot})$$

$$\Rightarrow Y_{iot} = \beta_0^{OVB} + \beta_1^{OVB} X_{iot} + \eta_{iot}$$

As we can see above, by omitting a variable with predictive power from the regression, we obtain biased estimates. I attempt to account for this in my regression analysis by including covariates for education, race, sex, and age for each industry-occupation-year observation. Furthermore, I include industry fixed effects in order to account for time-invariant industry level characteristics. However, there may be other covariates that I have not accounted for that could affect the wages and employment for the occupations included in the regression. For example, the degree of tasks already automated by technologies not related to artificial intelligence -- such as robotics -- could contribute to the bias of the estimated effect. If this were the case, then I would expect that the degree of education required for jobs would increase in order to manage and operate these technologies. As a result, I would expect that the estimated treatment effect would be an overestimate since the changes in employment or wage for the treatment group cannot be contributed solely to advancements in artificial intelligence.

4.3.3 Multiple Impacts for a Single Occupation

In order to create a valid difference-in-differences design I assume that the treatment and control group fall into only one of the two categories of impacts outlined in Agrawal, Gans, and Goldfarb (2019). To be precise, I assume that physicians and surgeons in the healthcare industry are only impacted by artificial intelligence as it pertains to augmenting labor on decision tasks and that secretaries and administrative assistants are only impacted by artificial intelligence as it pertains to automating prediction tasks. However, I account for this by choosing occupations that are disproportionately impacted by artificial intelligence in different ways. A large number of tasks for physicians and surgeons are not possible to be automated by the capabilities of IBM Watson, such as prescribing medication and in-person treatment. Therefore, I assume that Physicians and Surgeons will be largely only affected by artificial intelligence augmenting labor on decision tasks. In contrast, Secretaries and Administrative assistants perform many tasks that can be automated via artificial intelligence such as answering telephones, scheduling appointments, and transmitting correspondence. These tasks are largely affected by artificial intelligence automating prediction tasks, so I classify this occupation as the treatment group representing automation of prediction tasks.

4.3.4 Major Contributions to Commercial AI Prior to IBM Watson

In my research design, I make the strong assumption that there were no major contributions to the development of commercial artificial intelligence applications in the healthcare industry prior to the 2013 IBM Watson announcement. Since Watson serves as a major milestone for artificial intelligence as well as one of the first developed applications actually implemented within a commercial setting, I do not anticipate this to be a major issue. However, if this were the case, then the treatment period will have been incorrectly set in the difference-in-differences design and the causal effect of the treatment will not have been correctly estimated.

From the discussion above, it is clear that there are still some major limitations to the proposed difference-in-differences design, and I cannot claim causality in any of the methods that I have used. There remain factors that I was not able to include in the model which could contribute to a violation in the parallel trends assumption, omitted variable bias, and overall, an incorrect estimate of the treatment effect. My research focuses solely on determining whether there are significant relationships between the announcement of IBM Watson for healthcare applications in 2013 -- as a signal for one of the first uses of commercial artificial intelligence -- and the outcome variables of interest. Further research with a more rigorous experimental design is required to accurately estimate the treatment effect of artificial intelligence applications on jobs in the healthcare industry.

4.4 Ordinary Least Squares Regressions

In addition to the difference-in-differences regressions, I also ran a series of Ordinary Least Squares (OLS) regressions with the model below:

$$Y_{iot} = \beta_0 + \beta_1 NumPatents_t + \beta_2 D_{iot} + \beta_3 C_i + \varepsilon_{oit}$$

Y_{iot} is the outcome variable of interest for occupation *o* in industry *i* in time *t*. Similar to the Difference-in-Differences model, the outcome variables that I measured are the mean of log wages and the mean of log employment for each occupation. AI_{iot} is a dummy variable that is equal to 1 if the occupation is affected by automation of prediction tasks. $NumPatents_t$ is a measure of the total number of artificial intelligence patents in the world for each year. I divided the number of patents in each year by 1000 to obtain more observable coefficients. Further discussion is in the results section of this paper. C_i is a vector containing industry-fixed effects. Finally, D_{iot} is a vector of covariates controlling for the proportion of females, the proportion of white-identifying individuals, the mean age, and the proportion of individuals that completed at least one year of college education for each industry-occupation-year cell. Each of these covariates were log-transformed in the estimated regression. Time-fixed effects were omitted from this model due to issues with collinearity with the number of patents per year.

As above, I restrict the analysis to full-time, full year employees in medical-related industries defined according to the IPUMS *ind1990* classification (Appendix A). I run three separate regressions, one for each of the following occupations: Physicians and Surgeons, Secretaries and Administrative Assistants, and Nursing, Psychiatric, and Home Health Aides. I ultimately want to estimate the effect of the number of artificial patents on the outcomes for each occupation, separately.

4.5 Omitted Variable Bias in OLS Model

Just like the difference-in-differences design, there is the potential for omitted variable bias in the regression model described above -- which would result in biased estimates. For example, I am unable to account for advancements in technologies unrelated to artificial intelligence that may impact wages and employment for each of these occupations, such as computation. If advances in computational abilities of machines and computers used by these occupations contributed to increased wages and employment, I would expect an overestimate of the causal effect of the number of AI-related patents. If they instead had a negative effect, then I would expect an underestimate of the causal effect. Similar to the difference-in-differences design, I attempt to account for Omitted Variable Bias by including regressors for education, race, gender, and age, as well as industry fixed effects. However, by not including any covariate that is a determinant of the outcome variable and correlated with other regressors, I am exposing my model to omitted variable bias.

5 Results

In this section, I discuss the results of each of the regression that I ran as well as how to interpret them. In both the difference-in-differences regressions as well as the OLS regressions, all demographic control variables as well as the outcome variables were log-transformed. Thus, in order to interpret the coefficient on one of the control variables, I used the equation, $(1.1)^{\beta} \times 100$, where βis the coefficient of interest, to find the new percentage of the outcome variable. By subtracting this value by 1 for positive coefficients and subtracting it from 1 for negative coefficients, I am able to find the percent change in the outcome variable. For regressors that were not log-transformed, I used the equation $exp(\beta)$to find the percentage of the outcome variable, since all outcome variables were log-transformed. I then subtracted this value by 1 for positive coefficients and from 1 for negative coefficients to get the percent change in the outcome variable. An explanation of which regressors were log-transformed and which ones were not can be found in Section 4.

5.1 Difference-in-Differences For Physicians & Surgeons vs. Nursing, Psychiatric, & Home Health Aides

I first ran difference-in-differences regressions according to the model specified in Section 4.1.1 with Physicians and Surgeons as the treatment group and Nursing, Psychiatric, and Home Health Aides as the control group. The outcome variables of interest are *log(wage)* and *log(employment)* for each industry-occupation-year observation. As discussed in Section 4, this regression's aim is to estimate the treatment effect of artificial intelligence with regards to augmenting labor on decision tasks. The first year of treatment is 2013, indicating the introduction of IBM Watson for healthcare applications.

5.1.1 Parallel Trends

Figure 2: Plot of Log(Wage) over Time for Each Occupation

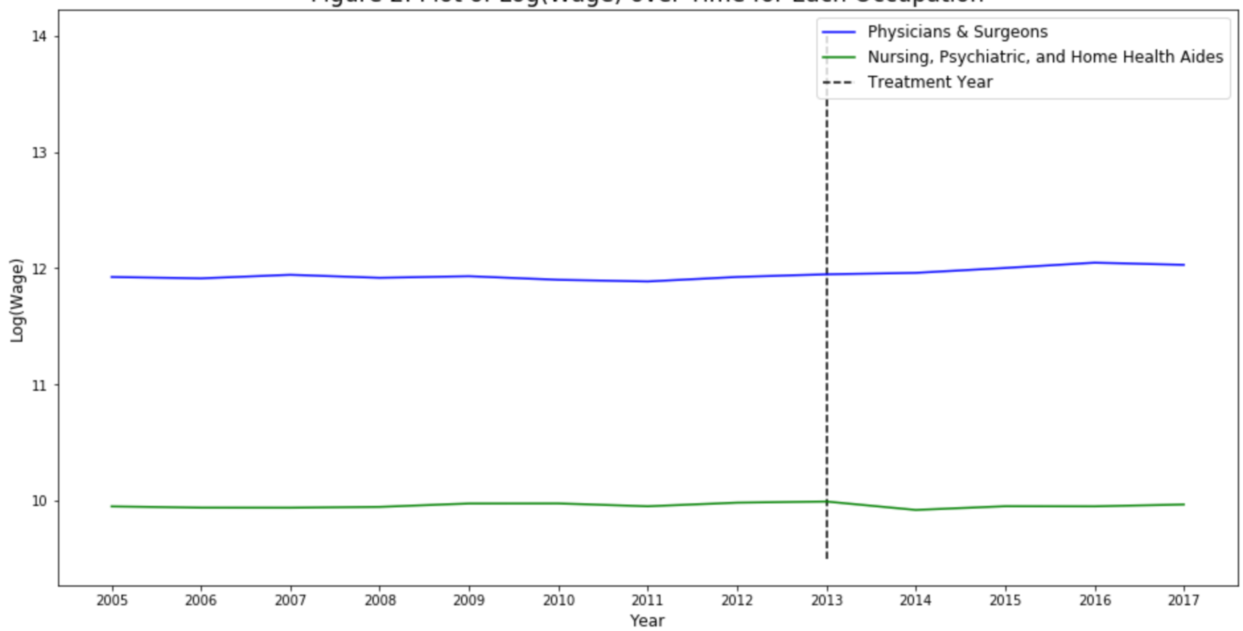
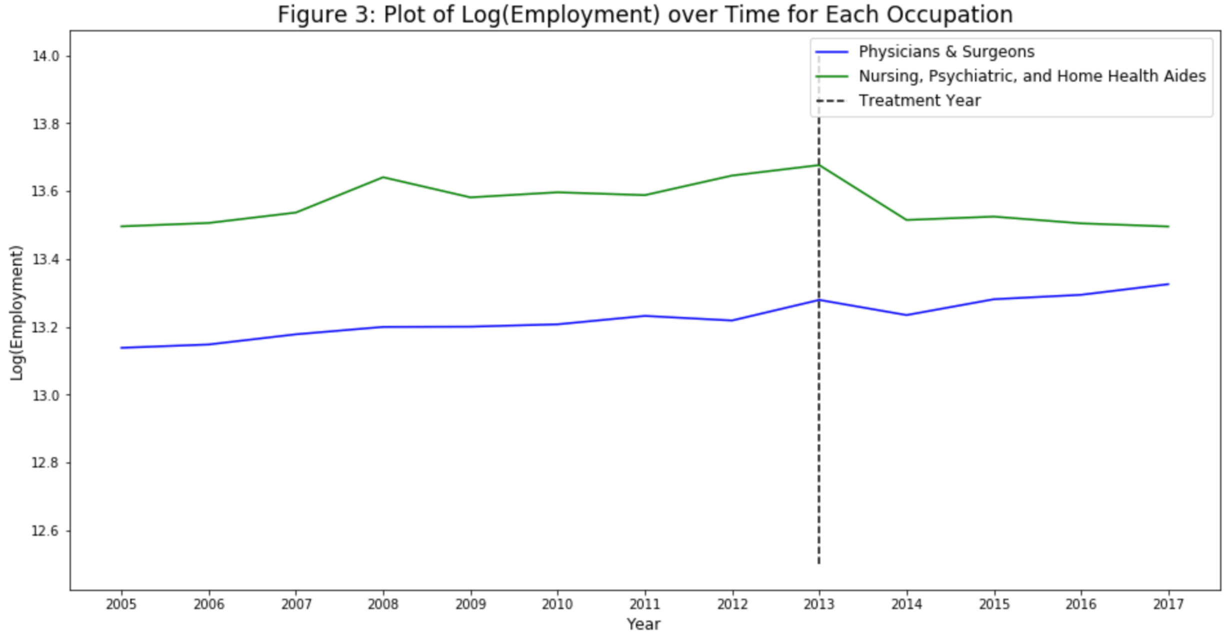


Figure 3: Plot of Log(Employment) over Time for Each Occupation



Upon visual inspection of the trend of *log(wages)* for each occupation from 2005 to 2017 in Figure 2, we can see that the parallel trends assumption seems to hold. The two occupations seem to exhibit parallel trends in *log(wage)* and start to deviate slightly after the treatment year.

For the trend of *log(employment)* over time for both jobs in Figure 3, the parallel trends assumption also seems to hold for the pre-treatment period, although there seems to be some volatility in the trends for the pre-treatment period. Specifically, there is a slight increase in the *log(employment)* for Nursing, Psychiatric, and Home Health Aides in 2008. However, the average difference between the two trends seems to be roughly parallel in the pre-treatment period with deviations after the treatment year. In the post-treatment period, we observe a sharp dip in employment for Nursing, Psychiatric, and Home Health Aides which does not follow its slightly increasing pre-treatment period trend. This could be a violation of the parallel trends assumption, which may need to be inspected further to ensure correct estimates.

5.1.2 Empirical Results

Table 1: Difference-in-Differences: Physicians and Surgeons vs. Nursing, Psychiatric, and Home Health Aides

	Dependent variable:			
	log(wage)		*log(employment*	
	No Controls	With Controls	No Controls	With Controls
	(1)	(2)	(3)	(4)
Constant	10.308***	8.883***	11.809***	17.457**
	(0.049)	(0.843)	(0.910)	(7.013)
AI	1.776***	1.366***	−1.284**	5.576***
	(0.030)	(0.120)	(0.569)	(1.002)
AI : Post	0.061	0.081*	0.143	−1.254***
	(0.049)	(0.046)	(0.917)	(0.381)
Frac.College		−0.036		−9.916***
		(0.100)		(0.834)
Frac.White		0.645***		−1.772
		(0.169)		(1.403)
Frac.Female		−0.259***		−2.637***
		(0.089)		(0.742)
Age		0.425*		−3.987**
		(0.225)		(1.875)
Industry FE?	Yes	Yes	Yes	Yes
Year FE?	Yes	Yes	Yes	Yes
Observations	78	78	78	78
R^2	0.990	0.993	0.343	0.910
Adjusted R^2	0.987	0.991	0.170	0.878
Note:			*p<0.1; **p<0.05; ***p<0.01	

Table 1 displays the results for the difference-in-difference regressions with *log(wage)* as the outcome variable of interest for regressions (1) and (2) and *log(employment)* as the dependent variable for regressions (3) and (4). Regressions (1) and (3) are run without demographic controls while (2) and (4) are run with the controls. The estimated treatment effect for each regression is represented by the coefficient for *AI:Post*.

In regression (2), we observe a significant positive coefficient on the estimated treatment effect at the 90% significance level. This estimated treatment effect is 0.081, meaning that the impact of the introduction of commercial artificial intelligence in the healthcare industry led to an increase of about 8.44% in the difference of wages for Physicians and Surgeons compared to Nursing, Psychiatric, and Home Health Aides. We also observe a highly significant positive coefficient for the fraction of white-identifying individuals, 0.645, and a highly significant negative coefficient for the fraction of females, -0.259. There is also a significant positive coefficient for age of 0.425 and no significant effect on the fraction of individuals that completed at least one year of college. These results suggest that we expect about a 6.34% increase in wages for both groups if there is a 10% increase in the fraction of white-identifying individuals, a 2.44% decrease in wages for a 10% increase in the fraction of females and about a 4.13% increase in wages for a 10% increase in the mean age of both groups. The R^2 and Adjusted R^2 for regression (2) are both above 0.987, which means that the regression explains at least 98.7% of the variability in the data.

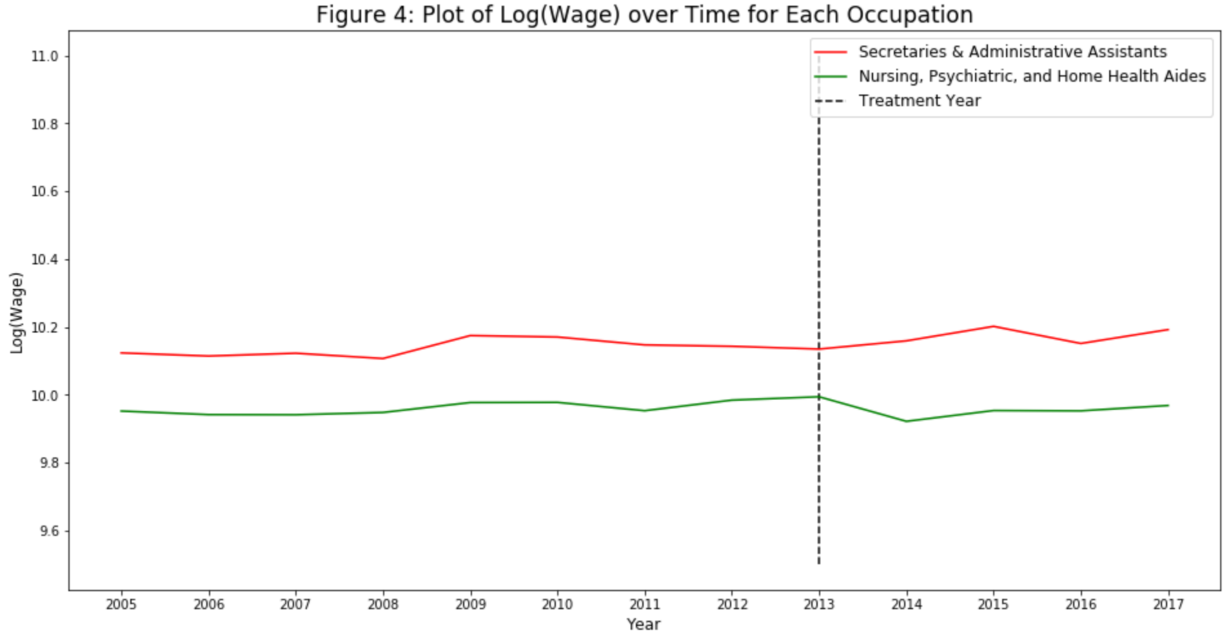
In regression (4), we observe a significant negative coefficient of -1.254 on the estimated treatment effect at the 95% significance level. This suggests that the impact of the introduction of commercial artificial intelligence in the healthcare industry led to a decrease of about 71.46% in the difference of employment for Physicians and Surgeons compared to Nursing, Psychiatric, and Home Health Aides. We also observe a highly significant negative coefficient of -9.916 for the fraction of individuals that completed at least one year of college, a significant negative coefficient of -2.637 for the fraction of females, and a significant negative coefficient of -3.987 on the age variable. These results suggest that we expect about a 61.14% decrease in the wages for both groups for a 10% increase in the fraction of individuals that completed at least one year of college, a 22.22% decrease in the wages for both groups for a 10% increase in the fraction of females, and a 31.61% decrease in wages for a 10% increase in the mean age over both groups. The R^2 and Adjusted R^2 for regression (2) are both above 0.878, which means that the regression explains at least 87.8% of the variability in the data.

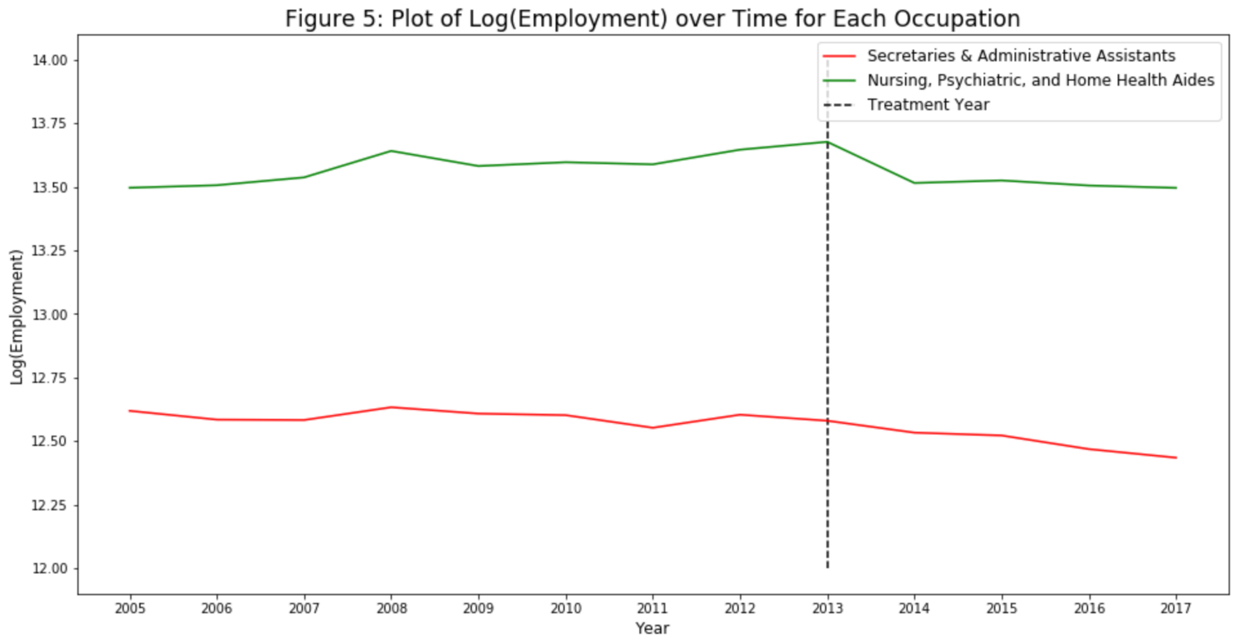
5.2 Difference-in-Differences For Secretaries & Administrative Assistants vs. Nursing, Psychiatric, & Home Health Aides

I next ran difference-in-differences regressions with Secretaries and Administrative Assistants as the treatment group and Nursing, Psychiatric, and Home Health Aides as the control group. The outcome variables of interest are *log(wage)* and *log(employment)* for each industry-occupation-year observation. As discussed in Section 4, this regression's aim is to estimate the treatment effect of artificial intelligence with regards to automating prediction tasks. The first year of treatment is 2013, indicating the introduction of IBM Watson for healthcare applications.

5.2.1 Parallel Trends

Figure 4: Plot of Log(Wage) over Time for Each Occupation





Upon visual inspection of the trend of *log(wages)* for each occupation from 2005 to 2017 in Figure 4, we can see that the parallel trends assumption seems to hold. The two occupations seem to exhibit parallel trends in *log(wage)* and start to deviate slightly after the treatment year, with *log(wage)* for Secretaries and Administrative Assistants seemingly increasing and decreasing slightly for Nursing, Psychiatric, and Home Health Aides. Similar to the discussion in Section 5.1.1, looking at the trend of *log(employment)* over time for both jobs in Figure 5, the parallel trends assumption also seems to hold for the pre-treatment period, although there seems to be some volatility in the trends for the pre-treatment period. Specifically, there is a slight increase in the *log(employment)* for Nursing, Psychiatric, and Home Health Aides in 2008. However, the average difference between the two trends seems to be roughly parallel in the pre-treatment period with deviations after the treatment year. In the post-treatment period, we observe a sharp dip in employment for Nursing, Psychiatric, and Home Health Aides which does not follow its slightly increasing pre-treatment period trend. This could be a violation of the parallel trends assumption, which may need to be inspected further to ensure correct estimates.

5.2.2 Empirical Results

Table 2: Difference-in-Differences: Secretaries and Administrative Assistants vs. Nursing, Psychiatric, and Home Health Aides

	Dependent variable			
	log(wage)		*log(employment)*	
	No Controls	With Controls	No Controls	With Controls
	(1)	(2)	(3)	(4)
Constant	10.160***	5.422**	11.098***	40.268***
	(0.056)	(2.277)	(0.485)	(12.800)
AI	0.076**	−0.363***	−0.820***	2.674***
	(0.035)	(0.080)	(0.303)	(0.452)
AI : Post	−0.018	−0.033	−0.063	0.029
	(0.056)	(0.037)	(0.489)	(0.207)
Frac.College		0.400***		−3.868***
		(0.107)		(0.603)
Frac.White		0.588***		−5.914***
		(0.195)		(1.095)
Frac.Female		0.255		0.288
		(0.234)		(1.313)
Age		1.452**		−9.446***
		(0.617)		(3.467)
Industry FE?	Yes	Yes	Yes	Yes
Year FE?	Yes	Yes	Yes	Yes
Observations	78	78	78	78
R^2	0.511	0.850	0.434	0.927
Adjusted R^2	0.382	0.798	0.285	0.901
Note:			*p<0.1; **p<0.05; ***p<0.01	

Table 2 displays the results for the difference-in-difference regressions with *log(wage)* as the outcome variable of interest for regressions (1) and (2) and *log(employment)* as the dependent variable for regressions (3) and (4). The estimated treatment effect for each regression is represented by the coefficient for *AI:Post*.

We do not observe a statistically significant coefficient on any of the estimated treatment effects for the regressions in Table 2. As a result, we cannot attribute any causality to the treatment effect estimates between Secretaries and Administrative Assistants and Nursing, Psychiatric, and Home Health Aides.

Despite the lack of significant estimates for the treatment effect for the above regressions, we observe significant coefficients for some of the covariates included as controls. In regression (2), we observe a coefficient of 0.400 for the fraction of individuals that completed at least one year of college and a coefficient of 0.588 for the fraction of white-identifying individuals. Both of these coefficients are highly statistically significant at the 99% significance level. These results indicate that we expect a 3.89% increase in wages for both groups for a 10% increase in the fraction of individuals that completed at least one year of college and a 5.76% increase in wages for both groups for a 10% increase in the fraction of individuals that are white-identifying. We also observe a positive coefficient of 1.452 on age that is significant at the 95% confidence level. Thus, we expect a 14.84% increase in wages for both groups for a 10% increase in the mean age.The R^2 and Adjusted $R^2$for regression (2) are both above 0.798, which means that the regression explains at least 79.8% of the variability in the data.

In regression (4), we observe a coefficient of -3.868 on the variable denoting the fraction of individuals that completed at least one year of college and a coefficient of -5.914 for the fraction of individuals that are white-identifying. Additionally, we also observe a coefficient of -9.446 on age. All three of these estimates are highly significant, at the 99% significance level. Similar to above, these results suggest that we should expect a 30.83% decrease in employment for both groups for a 10% increase in the fraction of individuals that completed at least one year of college, a 43.09% decrease in employment for both groups for a 10% increase in the fraction of individuals that are white-identifying, and a 59.36% decrease in employment for both groups for a 10% increase in the mean age. The R^2 and Adjusted R^2 for regression (2) are both above 0.901, which means that the regression explains at least 90.1% of the variability in the data.

5.3 Difference-in-Differences: Interpretation of Results

The results from the difference-in-differences regressions are very interesting. We can see from the results in Section 5.1.2 the estimated treatment effect is negative for wages and positive for employment. Upon visual inspection of the graphs in Section 5.1.1, this can be interpreted as a positive effect on wages and employment for Physicians and Surgeons, since the gap for wages increased and the gap for employment between the treatment and control group decreased. In the case of employment, the control group, Nursing, Psychiatric, and Home Health Aides, actually had a higher total employment than Physicians and Surgeons. Thus, since the gap between the treatment and control group decreased, this would suggest that the introduction of artificial intelligence in commercial healthcare applications had a positive effect on employment for Physicians and Surgeons.

The results from section 5.2.2, however, show that we fail to reject the null hypothesis that the estimated treatment effects for Secretaries and Administrative Assistants when compared to the control group are different from 0. Thus, it would seem that the introduction of artificial intelligence applications in healthcare in 2013 had no significant effect on the wages and employment levels for Secretaries and Administrative Assistants in the healthcare industry.

One potential cause for concern with these regressions are the particularly large negative coefficients on the control variables in the regressions related to employment. Upon closer inspection of the data, we can see that, within the control group, the mean proportion of individuals that attended college is about 0.467, the mean proportion of white-identifying individuals is about 0.591, the mean proportion of females is about 0.864, and the mean age is

40.821 (Appendix C). In comparison Secretaries and Administrative Assistants have a larger proportion of individuals that have completed one year of college, a larger proportion of white individuals, a larger proportion of females, and a larger mean age (Appendix C). Furthermore, Physicians and Surgeons have a much larger proportion of individuals that completed one year of college, a larger proportion of white-individuals, a smaller proportion of females, and a larger mean age (Appendix C). Thus, one possible explanation for these high estimates is that since the employment of Nursing, Psychiatric, and Health Aides is higher than that of either treatment group, the regression estimates lower levels of employment for variables where the treatment group has a higher average proportion or average number than the control group.

Another possible explanation for these large estimates is that the parallel trends assumption is violated, which would lead to incorrect estimates. The high estimate on the intercept term in regression (4) of Section 5.2.2 implies that this may be the case. This is something that must be investigated further in order to ensure that this research design yields reliable results.

Overall, these results suggest that the impact of artificial intelligence on jobs where labor can be augmented on decision making tasks, such as Physicians, and Surgeons, is positive for wages and employment. However, we cannot conclude any significant effect on employment and wages for jobs where tasks can be automated via prediction, such as Secretaries and Administrative Assistants.

5.4 OLS Regressions

As described in Section 4.4, I ran 3 separate Ordinary Least Squares regressions for each outcome variable: *log(wage)* and *log(employment)*. I separated the data for each occupation, Physicians and Surgeons, Secretaries, and Nursing, Psychiatric, and Home Health Aides in order to understand how the number of patents relating to artificial intelligence technologies impacts their wages and employment. The full specification of the regression model can be found in Section 4.4.

5.4.1 Wages

Table 3: Wage OLS for Each Occupation Against Number of AI Related Patents by Year

	Dependent variable: log(wage)		
	Physicians and Surgeons	Secretaries	Health Aides
	(1)	(2)	(3)
Constant	14.503***	10.245***	0.348
	(3.064)	(1.538)	(2.683)
Num.Patents	0.077***	0.0003	0.015
	(0.026)	(0.012)	(0.016)
Frac.College	−0.073	0.184*	−0.030
	(0.956)	(0.094)	(0.110)
Frac.White	0.737***	−0.178	−0.385*
	(0.211)	(0.234)	(0.221)
Frac.Female	−0.142	−0.113	−0.565
	(0.188)	(0.390)	(0.462)
Age	−0.669	−0.006	2.639***
	(0.744)	(0.403)	(0.714)
Industry FE?	Yes	Yes	Yes
Observations	39	39	39
R^2	0.788	0.255	0.938
Adjusted R^2	0.741	0.087	0.924
			*p<0.1; **p<0.05; ***p<0.01

Note: Secretaries is shorthand for Secretaries and Administrative Assistants. Health Aides is shorthand for Nursing, Psychiatric, and Home Health Aides.

Table 3 displays the regression results for each regression that I run with *log(wage)* as the outcome variable of interest. Regression (1) was restricted to Physicians and Surgeons, (2) was restricted to Secretaries and Administrative Assistants, and (3) was restricted to Nursing, Psychiatric, and Home Health Aides.

From the results of regression (1), we can observe a coefficient of 0.077 that is statistically significant at the 95% confidence level. Since the number of patents was divided by 1000 in the model, we can interpret this coefficient as about an 8% increase in the wages for Physicians and Surgeons for an increase of 1000 new patents relating to artificial intelligence technologies. We can also observe a coefficient of 0.737 that is highly significant at the 99% confidence level for the proportion of individuals that are white-identifying among each industry-occupation-year cell. We can interpret this coefficient as about a 7.28% increase in wages for Physicians and Surgeons for 10% increase in the fraction of individuals that are white-identifying. The R^2 and Adjusted R^2 for regression (1) are both above 0.741, which means that the regression explains about 74.1% of the variability in the data.

From the results of regression (2), we do not observe a significant coefficient on the number of patents. However, we do observe a coefficient of 0.184 that is significant at the 90% confidence level for the fraction of individuals that completed at least one year of college among Secretaries and Administrative Assistants. Thus, we would expect a 1.77% increase in wages for Secretaries and Administrative Assistants for a 10% increase in the fraction of individuals that completed one year of college. The R^2 and Adjusted R^2 for regression (2) are both above 0.087, which means that the regression explains at least 8.7% of the variability in the data. This value is extremely low, and suggests that the coefficients may not have been estimated correctly. A further discussion on this low R^2 is included in Section 5.5.

From the results of regression (3), we again do not observe a significant coefficient on the number of patents. As a result, we cannot determine a causal effect on wages that is not 0. We do, however, observe significant coefficients of -0.385 and 2.639 for the fraction of white-identifying individuals and the mean age among Nursing, Psychiatric, and Home Health

Aides, respectively. Thus, we would expect a 3.6% decrease in wages for a 10% increase in the fraction of white identifying individuals and a 28.6% increase in wages for a 10% increase in the mean age for Nursing, Psychiatric, and Home Health Aides. The R^2 and Adjusted R^2 for regression (2) are both above 0.924, which means that the regression explains at least 92.4% of the variability in the data.

5.4.2 Employment

Table 4: Employment OLS for Each Occupation Against Number of AI Related Patents by Year

	Dependent variable: log(employment		
	Physicians and Surgeons	Secretaries	Health Aides
	(1)	(2)	(3)
Constant	6.340	15.245***	7.932*
	(7.031)	(2.984)	(3.909)
Num.Patents	0.084	−0.065***	−0.022
	(0.060)	(0.023)	(0.024)
Frac.College	3.256	0.166	0.119
	(2.194)	(0.183)	(0.160)
Frac.White	0.088	0.030	0.531
	(0.485)	(0.455)	(0.322)
Frac.Female	0.181	−0.890	−1.576**
	(0.432)	(0.757)	(0.674)
Age	1.563	−0.954	0.618
	(1.708)	(0.783)	(1.041)
Industry FE?	Yes	Yes	Yes
Observations	39	39	39
R^2	0.992	0.995	0.997
Adjusted R^2	0.991	0.994	0.996
			*p<0.1; **p<0.05; ***p<0.01

Note: Secretaries is shorthand for Secretaries and Administrative Assistants. Health Aides is shorthand for Nursing, Psychiatric, and Home Health Aides.

Table 4 displays the regression results for each regression that I run with *log(employment)* as the outcome variable of interest. Regression (1) was restricted to Physicians and Surgeons, (2) was restricted to Secretaries and Administrative Assistants, and (3) was restricted to Nursing, Psychiatric, and Home Health Aides.

From the results of regression (1), we do not observe any significant coefficients at the 90%, 95%, or 99% confidence level. The R^2 and Adjusted R^2 for regression (1) are both above 0.991, which means that the regression explains at least 99.1% of the variability in the data.

From the results of regression (2), we observe a coefficient of -0.065 on the number of patents that is highly significant at the 99% confidence level. Since the number of patents is divided by 1000 in the regression model, we would expect a decrease of 6.29% in employment for Secretaries and Administrative Assistants for an increase of 1000 new AI-related patents. We do not observe a significant coefficient on any of the other controls used in the regression. The R^2 and Adjusted R^2 for regression (2) are both above 0.994 which means that the regression explains at least 99.4% of the variability in the data.

Finally, from the results of regression (3), we again do not observe a significant coefficient on the number of patents. We do observe, however, a coefficient of -1.576 that is highly significant at the 99% confidence level for the fraction of females among Nursing, Psychiatric, and Home Health Aides. Thus, we would expect a decrease of about 13.9% in employment for a 10% increase in the fraction of females among Health Aides. The R^2 and Adjusted R^2 for regression (2) are both above 0.996 which means that the regression explains at least 99.6% of the variability in the data.

5.5 OLS: Interpretation of Results

The regressions in Sections 5.4.1 and 5.4.2 yield some intriguing results. The regressions suggest that the number of artificial intelligence related patents has a positive effect on the wages for Physicians and Surgeons and no significant effect on the wages of the other two occupations.

Additionally, the number of patents seems to have a negative effect on the employment for Secretaries and Administrative Assistants but no significant effect on either of the two occupations.

One potential cause for concern in these results is the very low R^2 and Adjusted R^2 values for regression (2) in Section 5.4.1. Although the estimated coefficient on the number of artificial intelligence related patents is significant, the model does not do a good job of explaining most of the variability within the data for Secretaries and Administrative Assistants. This suggests that the model may incorrectly estimate the coefficients in the model or that the regressors used are not sufficient to explain the trend in wages for Secretaries and Administrative Assistants. This could mean that there is omitted variable bias that must be accounted for.

Overall, however, these results suggest that the number of artificial intelligence related patents has a positive effect on wages for jobs where labor can be augmented on decision making tasks and a negative effect on employment for jobs where tasks can be automated via prediction. It is noteworthy that no significant estimate was found on wages or employment for Nursing, Psychiatric, and Home Health Aides, which is consistent with the justification that I use for including it as the control group in the difference-in-differences regressions in Section 4.1.2.

5.6 Multicollinearity

In each of the regressions I ran, the absence of multicollinearity is an important assumption to ensure correct estimates and standard errors. I ran the variance inflation factor command (VIF) in R in order to check for collinearity between variables. I did not observe a VIF above 5 for any of the treatment effect variables in my difference-in-differences regressions or on the variable relating to the number of patents in my OLS regressions. I did, however, observe a VIF greater than 10 for some of the control variables in my regressions. While this may raise cause for concern, only the control variables may be collinear whereas the variables of interest in each of the regressions had low VIF estimates. Thus, I elected to not remove any of the variables from my regressions since the performance of the variables interest would not be affected.

6 Conclusion

In this paper, I study the impact of artificial intelligence on two jobs within the healthcare industry: Physicians and Surgeons and Secretaries and Administrative Assistants. By estimating the effect of artificial intelligence on the wages and employment for these occupations, I fill in gaps in the current literature by providing a industry-specific case study that builds upon the models developed in past work. This research is both important and relevant due to the rapid rise of artificial intelligence applications in recent years as well as the pervasiveness that AI already holds in our daily lives. As such, it is critical to develop research that examines the current impacts of artificial intelligence advancements in order to inform future public policy decision making.

The results of this study show that the treatment effect of the 2013 announcement of IBM Watson for healthcare, which serves as an important milestone in the development of commercial AI applications, had a positive effect on the wages and employment levels of Physicians and Surgeons in the healthcare industry. Meanwhile, no significant effect was found on wages and employment levels for Secretaries and Administrative Assistants within the healthcare industry. Further regression results also show that the number of artificial intelligence related patents had a positive effect on wages for Physicians and Surgeons and negative effect on employment for Secretaries and Administrative Assistants. Furthermore, the number of artificial intelligence patents had no observable effect on wages nor employment for Nursing, Psychiatric, and Home

Health Aides, which further justifies its use as the control group in the difference-in-differences design.

The results of this study connect back to existing literature in multiple ways. Firstly, I identified Physicians and Surgeons as an occupation that has a disproportionate amount of tasks that are impacted by artificial intelligence augmenting labor on decision tasks and Administrative Assistants as an occupation that has a disproportionate amount of tasks that are impacted by artificial intelligence automating prediction tasks. These identifications are drawn from the work done in Agrawal et al. (2019). Thus, the results show that the ability of artificial intelligence to augment labor on decision tasks may have a net positive effect on wages and employment whereas the ability to automate prediction tasks does not have a clear effect on wages and employment and remains ambiguous. Additionally, the results seem to partially corroborate the results of Webb (2019), where the author concludes that high-skill jobs are more exposed to artificial intelligence. Physicians and Surgeons can definitely be defined as a high-skill occupation due to the educational requirements as well as the specialization needed in each role. The positive significant effect on both wages and employment for this occupation suggest that not only are high skill jobs exposed to advancements in artificial intelligence but also that this increased exposure may have a positive effect. Furthermore, the tasks for Physicians and Surgeons are highly interpersonal and treatment-focused. Thus, the positive effect on wages and employment support the conclusions drawn in Huang et al. (2019) . Finally, by adopting the task-based approach presented in Acemoglu and Restrepo (2018), we can observe that the “productivity” effect seems to dominate the “displacement” effect for Physicians and Surgeons while the dominating effect remains ambiguous for Secretaries and Administrative Assistants.

It is important to note that there are shortcomings to my research design and therefore I cannot attribute causality to my results. Omitted variable bias is a large concern in the models that I developed and it is very likely that there are other variables that are determinants of the outcome variables of interest that I did not include. Furthermore, there are potential violations of the parallel trends assumption for the difference-in-differences regressions regarding employment levels that must be taken into account, along with collinearity issues found with the demographic control variables that I used. Further research is needed with a more rigorous and controlled design in order to reliably estimate the causal effect of artificial intelligence on wages and employment for these occupations.

Overall, this research uncovers key insights on the impact of artificial intelligence on wages and employment levels. It is clear that artificial intelligence has and will continue to have a significant impact on labor as more applications are developed for commercial use. Therefore, it is more important than ever that research as well as public policy keeps pace with the rapid development of artificial intelligence in order to ensure that jobs and individuals are not negatively affected by unbridled technological adoption and advancement.

7 Appendix

Appendix A

Table A1: Chosen Occupations and Their Corresponding IPUMS *OCCSOC* Code

Occupation	**IPUMS USA *OCCSOC* Code**
Physicians & Surgeons	290611
Secretaries & Administrative Assistants	436010
Nursing, Psychiatric, and Home Health Aides	311010

Table A2: Chosen Medical Industries and Their Corresponding IPUMS *IND1990* Code

Industry	IPUMS USA *IND1990* Code
Offices and Clinics of Physicians	812
Hospitals	831
Nursing and Personal Care Facilities	832

I selected the industries displayed in Table A2 based on those that contained the most number of observations for each selected occupation out of all medical-related industries.

Appendix B

In order to determine the tasks for each occupation, I used crosswalk between the 2010 O*NET SOC Codes to the 2019 O*NET SOC Codes. I then looked at current task descriptions for jobs in the O*NET 2019 classification and labeled each job with a Yes or No for each of two categories: Augmenting Labor on Decision Tasks, and Automating Prediction Tasks. For example, the first few tasks for Secretaries and Administrative Assistants is shown here:

Occupation	O*NET 2019 Code	Description	Automated via Prediction?	Augment Labor on Decision Tasks?
Anasthesiologists	29-1211	Monitor patient before, during, and after anesthesia and counteract adverse reactions or complications.	No	No
Anasthesiologists	29-1211	Record type and amount of anesthesia and patient condition throughout procedure.	Yes	No
Anasthesiologists	29-1211	Provide and maintain life support and airway management and help prepare patients for emergency surgery.	No	No
Anasthesiologists	29-1211	Administer anesthetic or sedation during medical procedures, using local, intravenous, spinal, or caudal methods.	No	No
Anasthesiologists	29-1211	Examine patient, obtain medical history, and use diagnostic tests to determine risk during surgical, obstetrical, and other medical procedures.	No	Yes

My methodology for labeling each task for each of these categories is as follows. For the category of Automated via Prediction, I looked at tasks involved with routine or manual labor, such as answering telephones, recording information, or transmitting correspondence. For these types of tasks, I labeled the task as "Yes" for this category and "No" otherwise. For the category of Augmenting Labor on Decision Tasks, I looked at decision tasks where the individual would decide on a course of action, such as what treatment to administer. The tasks identified involved those with words such as "diagnose" and "prescribe" in their descriptions. These tasks were labeled with a "Yes" for this category and "No" otherwise.

The full list of each labeled task for each occupation is too large to include within this paper. Instead, below are individual links to Google Sheets containing tables that match the same formatting as the example above:

Table B1: Spreadsheet Links of Labeled Tasks for Each Occupation

Occupation	**Spreadsheet Link**
Physicians & Surgeons	https://docs.google.com/spreadsheets/d/15Z2y_0OOplQSaa6uEk7fIOguRYBy1845Z6KmZYqpvCU/edit?usp=sharing
Secretaries & Administrative Assistants (In Medical Industries)	https://docs.google.com/spreadsheets/d/16j-Bk3r8v-99Paw7s5astl90ff8x_9D-OFAEW1CtgLU/edit?usp=sharing
Nursing, Psychiatric, and Home Health Aides	https://docs.google.com/spreadsheets/d/1ffaz-55jhx4jRmLLUehby6m8j2nqr_-W3sEP1Y_j2pI/edit?usp=sharing

Table B2: Count of Tasks Affected by Augmenting Labor on Decision Tasks by Occupation

	Physicians and Surgeons	**Secretaries**	**Health Aides**
Augment Labor on Decision Tasks?			
No	170	16	65
Yes	81	0	0

Table B3: Count of Tasks Affected by Automating Prediction Tasks by Occupation

Automated via Prediction?	Physicians and Surgeons	Secretaries	Health Aides
No	232	3	63
Yes	19	13	2

Appendix C

Table C1: Number of Observations and Means, split by Industry and Occupation

Industry	Occupation	Num. Observations	Mean Frac. College	Mean Age	Mean Frac. White	Mean Frac. Female
Offices and clinics of physicians	**Physicians & Surgeons**	33007	0.997001	48.528130	0.771400	0.272850
	Nursing, Psychiatric, and Home Health Aides	2907	0.581406	39.881321	0.733219	0.895050
	Secretaries & Administrative Assistants	11303	0.535266	44.295497	0.859418	0.972670
Hospitals	**Physicians & Surgeons**	40990	0.996168	41.881313	0.674343	0.375245
	Nursing, Psychiatric, and Home Health Aides	37618	0.501144	41.770163	0.529979	0.787264
	Secretaries & Administrative Assistants	24342	0.566259	46.329513	0.723596	0.961713
Nursing and personal care facilities	**Physicians & Surgeons**	178	0.993043	48.994382	0.733241	0.427207
	Nursing, Psychiatric, and Home Health Aides	49146	0.323127	40.862512	0.513446	0.908935
	Secretaries & Administrative Assistants	2386	0.508329	46.483655	0.802800	0.967025

Table C2: Means for Each Variable of Interest, Split by Occupation

	Physicians & Surgeons	Secretaries & Administrative Assistants	Nursing, Psychiatric, and Home Health Aides
Frac. College	0.994412	0.537526	0.467264
Frac. White	0.727349	0.795275	0.591146
Frac. Female	0.353524	0.966970	0.864478
Real Wage	146749.647481	25438.125379	24243.412502
Employment	185065.153846	95433.615385	259183.538462

Table C3: Standard Deviations for Each Variable of Interest, Split by Occupation

	Physicians & Surgeons	Secretaries & Administrative Assistants	Nursing, Psychiatric, and Home Health Aides
Frac. College	0.020426	0.051286	0.120127
Frac. White	0.073906	0.059985	0.104669
Frac. Female	0.081651	0.019081	0.057320
Real Wage	30509.734387	1036.171233	5195.029594
Employment	135743.500582	68493.601639	172881.349548

References

Agrawal, A., Gans, J. S., & Goldfarb, A. (2019). "Artificial Intelligence: The Ambiguous Labor Market Impact of Automating Prediction." Journal of Economic Perspectives, 33 (2): 31-50. https://www.aeaweb.org/articles?id=10.1257/jep.33.2.31.

Arnold, Z. (2020, September 29). "What investment trends reveal about the global AI landscape". Brookings Institution. https://www.brookings.edu/techstream/what-investment-trends-reveal-about-the-global-ai-landscape/.

Best, J. (2013, September 9). "IBM Watson: The inside story of how the Jeopardy-winning supercomputer was born, and what it wants to do next". TechRepublic. https://www.techrepublic.com/article/ibm-watson-the-inside-story-of-how-the-jeopardy-winning-supercomputer-was-born-and-what-it-wants-to-do-next/.

Daron, A. & Restrepo, P. (2018). "Low-Skill and High-Skill Automation," Journal of Human Capital, vol 12(2), pages 204-232, 2018. http://doi.org/10.1086/697242.

Davenport, T. & Kalakota, R. (2019). "The potential for artificial intelligence in healthcare". Future Healthcare Journal, 6 (2): 94–98. doi: 10.7861/futurehosp.6-2-94.

Huang, M., Rust, R., & Maksimovic, V. (2019). "The Feeling Economy: Managing in the Next Generation of Artificial Intelligence (AI)." California Management Review. https://doi.org/10.1177/0008125619863436.

IBM. (2013, February 8). "IBM Watson Hard At Work: New Breakthroughs Transform Quality Care for Patients". IBM. https://www-03.ibm.com/press/us/en/pressrelease/40335.wss.

OECD. (2020). Patents by Technology [dataset]. OECD.Stat, Accessed on March 25, 2021. https://stats.oecd.org/index.aspx?r=462108&erroCode=403&lastaction=login_submit

Ruggles, S., Flood, S., Foster, S., Goeken, R., Pacas, Jose., Schouweiler, M., & Sobek, M. IPUMS USA: Version 11.0 [dataset]. Minneapolis, MN: IPUMS, 2021. https://doi.org/10.18128/D010.V11.0.

Schmelzer, R. (2019, October 31). "Should We Be Afraid of AI?". Forbes https://www.forbes.com/sites/cognitiveworld/2019/10/31/should-we-be-afraid-of-ai/?sh=6259024331df.

Semuels, A. (2020, August 6). "Millions of Americans Have Lost Jobs in the Pandemic—And Robots and AI Are Replacing Them Faster Than Ever." Time. https://time.com/5876604/machines-jobs-coronavirus/.

Webb, M. (2019). "The Impact of Artificial Intelligence on the Labor Market." SSRN, http://dx.doi.org/10.2139/ssrn.3482150.

West, D. & Allen, J. (2018, April 24). "How artificial intelligence is transforming the world". Brookings Institution. https://www.brookings.edu/research/how-artificial-intelligence-is-transforming-the-world/.

Yu, K., Beam, A. L., Kohane, I. S. (2018). "Artificial intelligence in healthcare". Nature Biomedical Engineering, 2, 719–731. https://doi.org/10.1038/s41551-018-0305-z.