Understanding the Geography of Job Automation Within the United States

Hunter Jones

December 8th

Abstract

This paper explains in brief the relationship innovations in artificial intelligence (AI) technologies and the labor market have and the differential effect it has depending on occupation and region. This paper then models automation risk in the US labor market at the county level, using a task-based classification method developed by Frey and Osborne (Frey & Osborne, 2017). This approach highlights the varying level of vulnerability to automation across the U.S. labor force. The analysis begins with a definition of AI and automation, explains the task classification framework, and provides background on regional impacts of automation on the labor market. This leads to a county-level risk assessment highlighting which counties are at highest risk of automation as well as provides regression models based on these findings demonstrating which individual and county attributes have the strongest correlation with automation risk. The paper provides evidence of the statement that education plays the largest role on automation impact which past findings also stated.

1 Introduction

With the growing popularity of tools like ChatGPT and the adoption of large language models and other artificial intelligence, or AI models, AI has become a frequent topic in the news. Technology has even been labeled as the critical driver of long-run growth with automation being a large factor in the magnitude of technological innovation (Crowley & Doran, 2023). Because of this, it is increasingly important to understand the relationship that AI has with the labor market. Understanding this correlation can be helpful for individuals trying to understand their resiliency against this technology as well as for government figures, especially with the differential effect that it will have on the labor market.

2 Background and Literature Review

2.1 Understanding AI And Automation

AI systems are computer programs which take in large quantities of data, apply statistical and computational techniques, and generate outputs based on the input data, allowing for the output of data-driven predictions (The White House, 2024). Generating predictions enables AI models to solve business problems like determining the expected number of sales a product might have, segmenting the market based on items that a consumer has purchased, and subsequently recommending new products for the consumer to purchase based on past purchases made by the consumer among many other tasks. AI predictions are also able to solve tasks as complex as classifying images as various objects based on the color of pixels and shape of lines within the picture as well as taking in an input text prompt from a user and generating a lengthy response using natural language.

In conjunction, automation is the process of replacing or complementing labor with capital through technological innovations in the labor market (Acemoglu & Restrepo, 2019). The ability for computers to make complex predictions allows for AI models to automate tasks which were typically seen as "too complex" for computers to solve. This process also is capable of increasing the productivity of previously automated technologies (The White House, 2024). The increase in automation driven by AI has the negative effect of displacing workers. However, not every effect caused by automation is negative for laborers. Automation also results in the positive effects of reinstating labor in newly created positions as well as increasing the productivity of automated positions. Reinstatement and productivity effects result in higher demand for labor in non-automated tasks (Acemoglu & Restrepo, 2019). Understanding the impact of automation, positive or negative, requires looking at displacement, reinstatement, and productivity effects the automated task has and whether it alters the share of tasks completed by labor relative to capital, called the task content (Acemoglu & Restrepo, 2019).

2.2 Past and Present Predictions About Automation

The idea that automation significantly impacts the labor market is not a new concept and existed far before the field of artificial intelligence existed. In 1930, Keynes overestimated the effects of automation on the labor market, predicting that labor would account for a smaller portion of task content and that workweeks would become 15 hours with workers completing simple tasks (The White House, 2024). These claims are a far cry from the labor market observable today and similar to many other claims by economists in the past who have predicted technology will substitute labor without looking at its capabilities of being a complementary force for laborers (The White House, 2024). It can be difficult when talking about current predictions of AI to not hear wild claims about it having a major productivity impact on the economy and subsequently, equally major impacts displacing the workforce. Current predictions by Goldman Sachs estimate a 7% increase in global GDP equating to 7 trillion USD as well as 1.5% productivity growth per year in the US over the next 10 years from the effects of AI (Acemoglu & Restrepo, 2018). At the same time, fears of job displacement driven by automation continue to rise. 72% of US adults are anxious about computers and automation accomplishing tasks typically completed by humans (Crowley & Doran, 2023). Furthermore, according to a 2017 Chicago Booth poll, 38% of US economists agreed that 'rising use of robots and artificial intelligence is likely to substantially increase the number of workers in advanced countries who are unemployed for long periods' with only 21% of economists disagreeing with the statement (Chicago Booth, 2017). However, given that economists have historically overestimated the effects of AI, it can be difficult to claim that this upcoming decade will be different.

Quantifying the effects of AI and automation is even more difficult based on the speculative nature of the task. It is made harder based on only having access to currentday data in a rapidly evolving industry. According to a 2024 Council of Economic Advisors report, when looking at the 5-year moving average productivity growth over time, there appear to be large spikes when innovation like the .com boom occurred, reaching local maxima as high as 3.5% productivity gains per year before returning to approximately 2% per year, indicating the productivity gains do not have a substantial long-term effect (The White House, 2024). This leads to questions on whether productivity growth caused by automation will lead to a long-lasting increase or follow a similar trend as past innovations, with a sharp spike and a regression towards the 2% per year average productivity increase.

2.3 Classifying the Labor Market

Different tasks are more or less easily automated, making a model for the classification of tasks essential to understanding the risk of automation. Jobs can be classified as either routine or non-routine. Routine tasks are repetitive in nature and follow set rules and procedures, making it easier for the tasks to be translated into programs for a machine to follow (Autor, 2022). In contrast, non-routine tasks involve a level of complexity that is not typically feasible for a computer. The requirements for non-routine tasks can be viewed as a dynamic or variable set of procedures which are non-repetitive. Past findings have stated that routine jobs will more frequently substitute labor for capital. With the innovations in AI, this changes the capability of machines, giving computers the ability to take in current information about a task and create predictions based on the data as opposed to using rigid hard-coded instructions. This gives computers more flexibility in their potential to solve non-routine tasks (Frey & Osborne, 2017). As a result, many jobs which have been, in the past, resilient to automation are now facing the possibility of being displaced in the future.

Another way that jobs can be classified is based on whether the task is cognitive or manual. Cognitive tasks involve mental-labor and decision making skills whereas manual tasks focus on physical labor and tasks. Cognitive tasks have a higher emphasis placed on skilled labor (Frey & Osborne, 2017).

Combining whether a task is routine or non-routine and cognitive or manual creates a matrix and tasks can be labeled as cognitive routine, cognitive non-routine, manual routine, or manual non-routine (Frey & Osborne, 2017). Recent automation of non-routine tasks has been shown to have differing effects depending on if a position is cognitive or manual. Non-routine cognitive tasks often see a productivity improvement resulting in higher demand and increased wages whereas non-routine manual jobs are often displaced (The White House, 2024). This effect has resulted in task polarization, where the demand for high-paying non-routine cognitive positions has increased, while the demand for middle-paying non-routine manual jobs has decreased. As a result, many displaced workers have moved into low-paying positions, which, along with high-paying jobs, have seen an increase in demand (Autor, 2022).

Because of the higher level of education required for cognitive jobs, the task polarization results in a wage premium for attending college and an inequality in pay based on education (Autor, 2022). The increase in demand for educated or 'skilled' laborers is growing at a rate faster than the supply, resulting in even higher wage increases for educated workers (The White House, 2024). This education-based pay inequality is a leading cause of wage disparity over the past 40 years (Autor, 2022).

2.4 Regional Risk of Automation

The effects of automation, both positive and negative, are not limited to targeting specific occupations. There is also a regional disparity causing certain regions to be disproportionately impacted (The White House, 2024). Between the years of 1940 and 1980, less agglomerated areas experienced larger economic growth relative to their urban counterparts. However, starting in 1980 and continuing in subsequent years, one of the causes for this reversal was the automation of many positions commonly held in rural areas, including manufacturing jobs (Storper, 2018).

At the same time, there is also a push where more people are migrating to urban areas, and with the dense packing of many differing skills and professions, creating a breeding ground for technological innovation (Acemoglu & Restrepo, 2018). Urban areas are even further divided based on their size, with larger areas benefiting more from digital technologies. Areas with stronger digital infrastructure experienced increased productivity, innovation, and more job opportunities (Acemoglu & Restrepo, 2018). This disparity may lead some urban areas to decline while others may prosper from technological innovation.

Frey and Osborne estimate that 47% of US occupations are at 'high risk' of automation. This paper examines this data at a state and county level using more recent data on the US labor market as well as aims to tackle what factors cause certain regions to be at higher or lower risk of automation (Frey & Osborne, 2017).

3 Data Description and Methodology

3.1 Occupation Classification and Automation

The O*NET Program is a database sponsored by the US Department of Labor/Employment Training Administration that includes Standard Occupational Classification (SOC) codes for various occupations (O*NET). This SOC code is a 6 digit number with the first 2 digits representing the major group, the 3rd digit representing the minor group, the 4th and 5th digits representing the broad occupation category, and the 6th digit representing the detailed occupation classification. For example, the SOC code 11-2021 has the major group 11-XXXX representing management occupations, the minor group 11-20XX representing Advertising, Marketing, Promotions, Public Relations, and Sales Managers, the broad occupation 11-202X representing Marketing and Sales Managers, and lastly the detailed occupation 11-2021 representing Marketing Managers. O*NET also details the anatomy of each occupation, encompassing the knowledge, skills, and abilities required.

Frey and Osborne utilized the O*NET variables from an O*NET survey in which respondents were asked to rate various computerization bottlenecks for each occupation. They selected 9 variables of interest: finger dexterity, manual dexterity, cramped work space/ awkward positions, originality, fine arts, social perceptiveness, negotiation, persuasion, and assisting and caring for others. Frey and Osborne then asked a group of machine learning researchers to hand-label 70 occupations as 1 if the position is automatable and 0 if not based on the question: "Can the tasks of this job be sufficiently specified, conditional on the availability of big data, to be performed by state of the art computer-controlled equipment. (Frey & Osborne, 2017)" Using the automatable variable as the target variable and the 9 properties as features, Frey and Osborne were able to develop a classification algorithm which outputs a value between 0 and 1, where 0 represents very low risk of automation and 1 represents a very high risk of automation. They then classified 702 of the 2010 occupation SOC codes using their algorithm. The number ranging from 0 to 1 representing the risk of automation is then used within this paper to quantify the risk of automation.

3.2 Combining Automation Risk with Geographic Data

IPUMS USA provides U.S. census microdata, which includes household and personal information such as income, occupation, ethnicity and racial demographic, and education (Ruggles et al, 2024). For this paper, the 2021 American Community Survey (ACS) was used, containing a 1Key variables used in this analysis include a person's Federal Information Processing Standards (FIPS) code. The FIPS code is a 5-digit number with the first 2 digits representing the state and the last 3 digits represent the county). FIPS codes were collected at the county-level, however, the Census only includes county-level data for counties with populations over 100,000. Counties below this threshold were combined into an aggregate-state location including all individuals from counties with a population less than 100,000. Additionally demographic information including age, sex, race, education level, and poverty status was included.

For this paper, the Frey and Osborne automation risk levels were joined with the 2021 ACS data based on the SOC code. This integration allowed analysis on automation risk at the county level as well as insights into demographic-specific patterns related to automation risk. Because the 2021 ACS sample used the 2018 version of the SOC codes while Frey and Osborne's analysis was based on the 2010 version, a crosswalk provided by O*NET, which maps 2010 codes to their 2018 counterparts, was utilized. In cases where a 2010 SOC code corresponded to multiple 2018 SOC codes, the automation risk was calculated as the average of the automation levels for those 2018 codes. Another shortcoming of this combination was that Frey and Osborne primarily classified detailed occupations, while the US Census data included occupations at varying levels of detail. Census data for detailed occupations were used as is, while for broad occupations, the automation risk was estimated by averaging the automation risks of all corresponding detailed occupations classified by Frey and Osborne. Because of the large range in automation risk based on the aggregation of all occupations within one minor group, occupations missing their minor group in the ACS 2021 survey were discarded.

The data was then aggregated to obtain the proportion of individuals that Frey and Osborne calculate as 'high' automation risk for each county. The Census only provides the county if the individual is in a county with a population greater than 100,000. Because of this, not all counties could be represented and instead the automation risk for the aggregate-state was calculated in these circumstances. This state-wide value typically represented rural areas, while the counties with specific counts were usually larger cities. By looking at the data in this manner, it allows for the analysis of automation risk associated with each state and county, providing valuable insights into how demographic factors and geographic locations influence automation vulnerability.

In joining the data, there was some information that had to be modified or discarded. For example, the occupation FIPS code was the primary variable of interest, however if there is county-level data on a respondent's residency but not occupation, then the residency FIPS code was used instead. Another issue is that IPUMS USA includes FIPS codes for counties that are part of multiple states (e.g., the code 61 represents Maine-New Hampshire-Vermont). Respondents with these FIPS codes were excluded, eliminating less than 1% of the sample. After filtering for individuals who had a job and corresponding SOC code, whose SOC code was classified by Frey and Osborne, and whose FIPS location was not aggregated, the final sample size from the 2021 ACS was reduced from 15,721,123 to 5,751,689 individuals utilized for analysis. This reduction was primarily based on removing individuals without an occupation or with an occupation that was not measurable using data from Frey and Osborne. This also includes analysis of 478 individual counties and aggregate states.

3.3 Constructing Regression Models

The first model utilized performs analysis at an individual level without being aggregated into a county. The equation for Model 1 is as follows: $Y_{high\ automation\ probability} = \beta_0 + \beta_{1\ ethnicity/race} + \beta_{2\ sex} + \beta_{3\ age} + \beta_{4\ education} + \beta_{5\ poverty} + \beta_{6\ log\ earnings} + \beta_{7\ city} + \beta_{8\ region} + \epsilon$ and uses the statistical person weights provided by IPUMS. This model uses all census individuals who were capable of being estimated by Frey and Osborne's model with the dependent variable being the binary value classifying an occupation as high automation or not and then variables for the individuals race and ethnicity, sex, binned age, binned education level, poverty status, log earnings, whether or not the person lives in a city, and the region the individual lived in. This model allows for analysis of which attributes about a person have the greatest impact on automation. Each dependent variable, outside of log earnings used, is a binary variable. The race/ethnicity variable included separate categories for Non-Hispanic White, Non- Hispanic Black, Non-Hispanic Asian, Hispanic individualis, and then individuals who are part of multiple groups or other groups not listed. This model also segmented an individual's age into groups for those who are 16-24, 25-35, 36-50, 51-64, and 65+, creating a binary variable for each segment. Similarly it segmented education based on if a person: did not receive a high school diploma, graduated high school or received a GED, completed some college, obtained a bachelor's degree, and if an individual obtained a masters or higher level graduate degree. The regions used are: northeast, midwest, west, and south following IPUMS classification methodology. The model also uses the natural log value of an individual's earnings from their occupation. Lastly the city variable is a binary variable depending on if IPUMS classifies the individual as within a city.

Another model was constructed, making use of the county and aggregate state data. The equation for Model 2 is: $Y_{high\ automation\ probability} = \beta_0 + \beta_{1\ region} + \beta_{2\ income} + \beta_{3\ estimated\ population} + \epsilon$. Data at the county level, outside of the region and population variable, uses proportional information, with each variable representing the proportion of the county which is part of the binary group. For example, a value of .3 for the dependent variable indicates that 30% of the population of the specific county or aggregate state is at high risk of automation. Similarly, the income variable is broken down into 5 quintiles with each quintile variable representing the proportion of the county within each quintile based on their total income. For example, if a county has a value of .12 for the lowest income quintile, this indicates that only 12% of the county's population was in the bottom 20% of total income values for survey respondents. The region variable instead is a binary variable where each county will have a 0 for each region besides the region they are part of. The regions are the same breakdown used within Model 1. The estimated population is then calculated based on the number of individuals within the ACS 2021 sample multiplied by 100, since it is a 1% sample. Because individuals in aggregate states are from counties where there is a population of less than 100,000 but the aggregate states often include many counties, the population for aggregate states was cutoff at a max of 100,000.

The final model examined within this paper is another model at the county level, however it included additional information about individuals within the county. The equation for Model 3 is: $Y_{high\ automation\ probability} = \beta_0 + \beta_{1\ region} + \beta_{2\ income} + \beta_{3\ estimated\ population} + \beta_{4\ ethnicity/race} + \beta_{5\ age} + \beta_{6\ education} + \epsilon$. This model includes ethnicity/race, age, and education using the same methodology of breaking down each factor from Model 1 as well as the proportional representation used within Model 2 for automation probability and income for the additional variables. This model allows for a wider ranged analysis of counties than was possible in Model 2 to identify which factors are significant in determining a county's risk of automation.

4 Findings

4.1 Developing a Mapping of Automation Risk

Frey and Osborne identified that 47% of the US workforce is categorized as being at 'high risk' of automation. Frey and Osborne defined high risk as a value of .7 or greater predicted from their model. These positions face the potential of being automated within the next 10 or 20 years (Frey & Osborne, 2017). When analyzing the 478 counties and aggregate states included in the IPUMS USA data, it can be observed that each county and aggregate state has a percentage of the population at high risk varying between roughly 30% and 54%. Figure 1 demonstrates locations estimated with a lower proportion of high risk individuals tend to be metropolitan areas. This highlights a pattern where areas surrounding major metropolitan locations are predominantly shaded 'blue,' while most rural and less-populated regions are shaded 'red.' It can also be observed that counties along the West Coast, in the Southwest, and parts of the Great Lakes region have the highest percentage of population at high risk of automation, while many of the East Coast regions appear to have a majority of the counties with the lowest risk of automation.

Furthermore, Figure 2 shows that the distribution of counties and aggregate states appears to be normally distributed, with a mean and median at approximately 44%. Both Figure 2 and Figure 3 outline one location with an extreme outlier with only 24% of its population classified as being at high risk of automation. This location is Arlington, Virginia, located near Washington, D.C. which would, as expected, have a unique labor market contributing to its resiliency against automation. The distribution of the proportion of residents at high risk of automation has no upper-bound outliers, suggesting that no counties or aggregate states exhibit exceptionally high vulnerability. Based on the boxplot Arlington, Virginia is not the only lower-bound outlier, however, only 3 of the counties were identified as lower-bound outliers and, even among these outliers, Arlington is substantially more resilient to automation.

However, even with minimal outliers, automation risk still spans a considerable range. This ranges from approximately 30% to 54% of individuals in various counties being classified as high risk. This highlights the regional disparity of automation risk and the importance of further analysis on automation risk based on region.

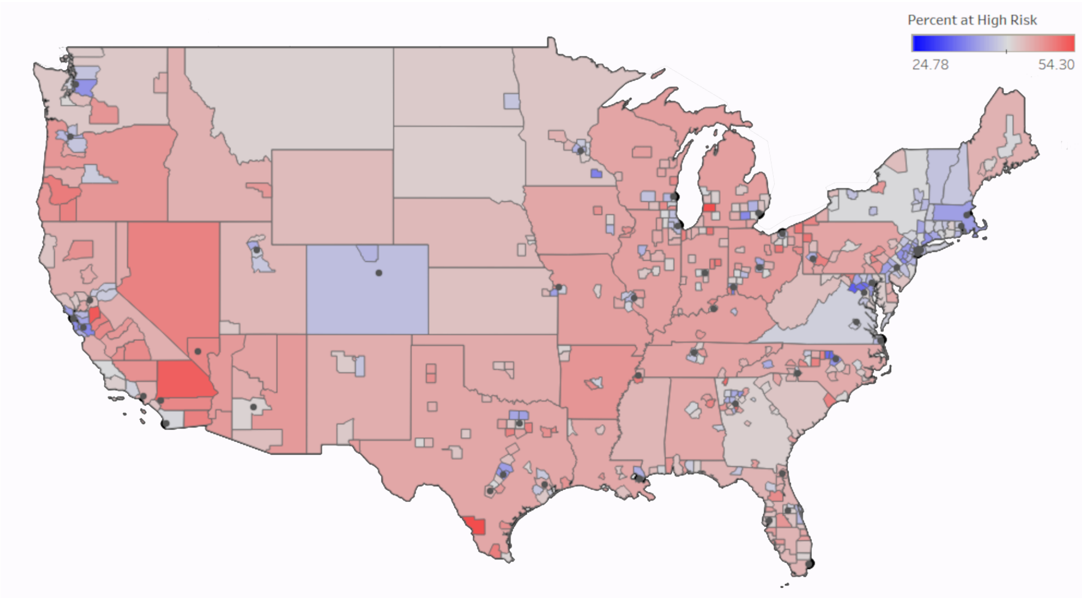

Figure 1: Metropolitan areas with a lower proportion of individuals at high risk of automation. Black dots indicate metropolitan areas with populations over 1,000,000.

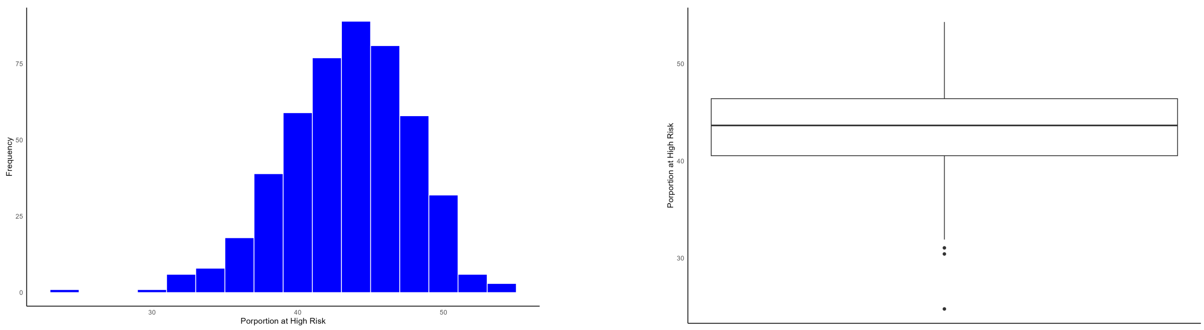

Figure 2: Fig 2 & 3: Normally distributed data with three lower-bound outliers: one extreme and two slightly beyond the threshold of a standard value.

4.2 Analysis of Individual-Level Regression Model

When looking at Table 1, which includes individual-level regression results, each variable had a p-value which was below the .001 threshold meaning there is a less than 0.1% chance these relationships occurred by random chance. This can partially be explained based on the very large sample size allowing for precise estimates. Despite the high probability values, the R2 value is relatively low. This R2 value is approximately 11%, which indicates that although this model highlights key variables in quantifying the risk of automation, it has a poor predictive power at estimating if an individual's position is at risk of automation. However, this is a nonissue because the purpose of this model is to identify factors which have the greatest impact on an individual's risk of automation.

Among the factors observed within the model, the variable which has the strongest magnitude appears to be education status. This aligns with past findings with the wage premium of skilled labor and automation's differential effect based on education. Individuals who did not receive a high school degree have an average magnitude this is .34 higher than those with a bachelor's degree when determining the likelihood of being at high risk of automation. This is the strongest magnitude within the model. Even individuals with a high school degree and college experience that is less than a bachelor's degree had a higher risk of automation with a magnitude of .24. Similarly there is a .11 magnitude in the inverse for individuals with a graduate's degree compared to a bachelor's degree. This suggests that graduate education often leads to jobs with greater protection from automation. This further highlights the wage premium on skilled labor.

Another key variable in determining an individual's likelihood of being at high risk of automation is age. Individuals within the 16-24 category were identified as being the move susceptible to having their position be automated with individuals in the 65+ category being the second most likely. When compared to individuals within the 36-50 category, 16-24 year old laborers have a magnitude of .13. This aligns with trends where routine entry-level positions, which are commonly held by younger workers, are easily automated and replaced. For example, self-checkout systems replaced cashiers at many grocery stores and fast food locations. The group that is least likely to be automated is the 36-50 age group. This creates a situation where the youngest and oldest laborers have the highest threat of automation while middle-age laborers have the highest level of safety creating a parabolic model.

While every variable in the model is below the .001 significance level, the magnitudes of the remaining variables remain close to 0 indicating a lesser impact on an individual's resilience to automation. This includes the findings relating to ethnicity/race, poverty, city status, and earnings. This is interesting because although, being in a city is associated with a negative magnitude which approaches zero, insignificant impact living in a city has on job resiliency. While this lack of magnitude could be based on the sensitivity IPUMS uses to label a location as a city, this suggests that unlike prior knowledge, cities are not more resilient to automation.

4.3 Extending Regression to County Level Risk Assessment

Table 2 has the results for county-level data. This table shows that all income quintiles are statistically significant at the 0.05 level. This shows a strong likelihood of a relationship between county income and automation risk with the largest effect size being .056 difference between the highest income quintile and the average quintile. This creates a trend where counties with larger proportions of high income individuals have strong protection against automation. There is also an effect of the inverse where counties with high proportions of low income individuals are at greater risk of automation, however the magnitude is much smaller. This indicates that the few wealthy counties are much safer while most other counties are at a similar level of risk. This appears consistent with visual findings where there were a few lower-bound outliers, but no upper-bound outliers.

Table 2 also shows that northeast counties are the least at risk of automation, as shown in figure 1. Midwest counties and western counties were both shown to be at more risk of automation than southern counties. However, the difference between southern and western was the only comparison with a sizable coefficient. This comparison was associated with an effect size of .012. The difference between northeast and southern counties was not even being labeled as statistically significant.

Finally, the method of calculating estimated population used in the model found the log-population to be statistically insignificant.

The county-level data was further extended with demographic-specific features about individuals within the counties in Model 3. This includes findings from proportional variables for race/ethnicity, age, and education. When utilizing this less constrained model, the only variables which are statistically significant are: Non-Hispanic Black in comparison with Non-Hispanic White individuals, the 65+ age category when compared to 36-50 individuals, all education variables being compared to a bachelor's degree, the two highest earnings quintiles, west and midwest regions when compared to southern regions, and lastly and interestingly the estimated population variable.

The ethnicity/race variable shows counties with a higher population of Non-Hispanic Black individuals are disproportionately more likely to be at high risk of automation, however all other race/ethnicity variables were not found to be statistically significant. Similarly counties with a higher population of 65+ individuals were found to be less resilient to automation. The findings between these 2 race/ethnic and age demographics highlight demographic groups that are particularly vulnerable at the county level.

The income variable's significance for only high income counties provides further evidence that the wealthier counties are resilient against automation, whereas lower income counties do not appear to be significantly less resilient than middle-income counties.

Interestingly, counties in the midwest and west appear to have a disadvantage compared to southern counties with both regions having positive but small coefficients. Even more interesting is the findings that estimated populations are now statistically significant indicating that information within ethnicity/race, age, and education is explained by the estimated population. Furthermore, the coefficient is positive and close to 0 which indicates that counties with higher populations also have a higher proportion of high risk occupations. This issue may stem from the methodology the log estimated population was calculated.

The findings presented within these 3 models identify key features influencing automation susceptibility at an individual and county level. At an individual level all features were found to be statistically significant with education being associated with the largest coefficient, followed by age. At the county level, education remains statistically significant in all groups however only select groups in the race/ethnicity (Non-Hispanic Black), age (65+), income (top 2 quintiles), and region (all but northeast), as well as estimated population appear significant. These findings illustrate the different demographic and regional attributes, highlighting the complexity of the situation.

5 Conclusion

In conclusion, it is difficult to understand the impact that automation will have on the economy. It is especially the case given the nature of economists to overestimate the harmful effects that automation will have on the labor market. However, with recent innovation in the fields of AI and machine learning technology, many non-manual positions, which are too complex for hard-coded computer program solutions, now face a risk of automation. The impact of this automation threat does not affect regions equally with different counties within the US being at higher or lower risk of automation. Different attributes of an individual or county have differing effects on automation with education, income, and age being identified as key factors. Understanding the regional impact of automation is important at an individual level in understanding ways to protect yourself from automation, as well as at a governmental level in creating policies which operate as a safeguard against the harmful effects brought on by automation.

6 References

Acemoglu, Daron, and Pascual Restrepo. 2019. "Automation and New Tasks: How Technology Displaces and Reinstates Labor." Journal of Economic Perspectives, 33 (The White House, 2024): 3–30.

The White House. (2024). An economic framework for understanding artificial intelligence. In Economic report of the president (Chapter 7). Executive Office of the President.

Acemoglu, D. (2024). The simple macroeconomics of AI. Preprint. https://doi.org/10.3386/w32487

Acemoglu, Daron, and Pascual Restrepo. 2018. "The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment." American Economic Review, 108 (Crowley & Doran, 2023): 1488–1542.

Chicago Booth. (2017, September). Robots and artificial intelligence. Chicago Booth. https://www.kentclarkcenter.org/surveys/robots-and-artificial-intelligence-2/

Crowley, F., Doran, J. The geography of job automation in Ireland: what urban areas are most at risk?. Ann Reg Sci 71, 727–745 (2023). https://doi.org/10.1007/s00168-022-01180-4

Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? Technological Forecasting and Social Change, 114, 254–280. https://doi.org/10.1016/j.techfore.2016.08.019

Autor, D. (2022). The labor market impacts of technological change: From unbridled enthusiasm to qualified optimism to vast uncertainty (Working Paper No. 30074). National Bureau of Economic Research. https://doi.org/10.3386/w30074

Michael Storper, Separate Worlds? Explaining the current wave of regional economic polarization, Journal of Economic Geography, Volume 18, Issue 2, March 2018, Pages 247–270, https://doi.org/10.1093/jeg/lby011

ONET Resource Center. (n.d.). About ONET. O*NET Resource Center. Retrieved October 2021, from https://www.onetcenter.org/overview.html

Steven Ruggles, Sarah Flood, Matthew Sobek, Daniel Backman, Annie Chen, Grace Cooper, Stephanie Richards, Renae Rodgers, and Megan Schouweiler. IPUMS USA: Version 15.0 [dataset]. Minneapolis, MN: IPUMS, 2024. https://doi.org/10.18128/D010.V15.0

7 Model Output

Table 1: Individual Level Regression

Variable	**Coefficient**	**Std. Error**	**P-Value**
Intercept	0.3404	0.0010	< 0.001 ***
Ethnicity/Race: Black (Non-Hispanic)	0.0508	0.0006	< 0.001 ***
Ethnicity/Race: Asian (Non-Hispanic)	0.0340	0.0008	< 0.001 ***
Ethnicity/Race: Hispanic	0.0541	0.0006	< 0.001 ***
Ethnicity/Race: Multiple	0.0318	0.0007	< 0.001 ***
Ethnicity/Race: Other	0.0339	0.0022	< 0.001 ***
White (Non-Hispanic)	–	–	–
Sex: Female	-0.0061	0.0004	< 0.001 ***
Sex: Male	–	–	–
Age: 16-24	0.1279	0.0006	< 0.001 ***
Age: 25-35	0.0189	0.0005	< 0.001 ***
Age: 51-64	0.0272	0.0005	< 0.001 ***
Age: 65+	0.0848	0.0009	< 0.001 ***
Age: 36-50	–	–	–
Education: Less than High School	0.3407	0.0008	< 0.001 ***
Education: High School Diploma	0.2737	0.0006	< 0.001 ***
Education: Some College	0.1733	0.0005	< 0.001 ***
Education: Graduate Degree	-0.1190	0.0007	< 0.001 ***
Education: Bachelor's Degree	–	–	–
Poverty Level	0.0601	0.0009	< 0.001 ***
Log Earnings (ln)	-0.0084	0.0001	< 0.001 ***
City	-0.0044	0.0006	< 0.001 ***

Significance codes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

R^2: 0.1178 **Residual Std. Error:** 2.192 **Adjusted R-Squared:** 0.1178

Table 2: County Level Regression

Variable	**Coefficient**	**Std. Error**	**P-Value**
Intercept	0.4357	0.0158	<0.001 ***
Region: Northeast	-0.0026	0.0034	0.446
Region: Midwest	0.0086	0.0030	0.004 **
Region: West	0.0123	0.0034	<0.001 ***
Region: South	–	–	–
Log Population (ln)	0.0005	0.0012	0.668
Income: Very Low	0.0171	0.0036	<0.001 ***
Income: Low	0.0134	0.0036	<0.001 ***
Income: High	-0.0188	0.0037	<0.001 ***
Income: Very High	-0.0565	0.0039	<0.001 ***
Income: Medium	–	–	–

Significance codes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

R^2: 0.5275 **Residual Std. Error:** 0.0263 **Adjusted R-squared:** 0.5203

Table 3: County Level Regression w/ Demographic Information

Variable	**Coefficient**	**Std. Error**	**P-Value**
Intercept	0.2675	0.0368	$<$0.001 ***
Ethnicity/Race: Black (Non-Hispanic)	0.0356	0.0110	0.001 **
Ethnicity/Race: Asian (Non-Hispanic)	0.0068	0.0182	0.709
Ethnicty/Race: Hispanic	-0.0105	0.0161	0.515
Ethnicity/Race: Multiple	0.0400	0.0312	0.200
Ethnicity/Race: Other	-0.0213	0.0337	0.528
Ethnicity/Race: White (Non-Hispanic)	–	–	–
Age: 16-24	0.0695	0.0445	0.119
Age: 25-35	-0.0146	0.0530	0.784
Age: 51-64	0.0272	0.0553	0.623
Age: 65+	0.1683	0.0532	0.002 **
Education: Pre-HS	0.3061	0.0526	$<$ 0.001 ***
Education: High School	0.2386	0.0299	$<$ 0.001 ***
Education: Some College	0.1070	0.0333	0.001 **
Education: Graduate	-0.3303	0.0492	$<$ 0.001 ***
Education: Bachelor's Degree	–	–	–
Income: Very Low	0.0037	0.0023	0.113
Income: Low	0.0030	0.0021	0.158
Income: High	-0.0065	0.0022	0.003 **
Income: Very High	-0.0111	0.0029	$<$ 0.001 ***
Income: Medium	–	–	–
Region: Northeast	0.0037	0.0025	0.143
Region: Midwest	0.0092	0.0021	$<$ 0.001 ***
Region: West	0.0101	0.0028	$<$ 0.001 ***
Region: South	–	–	–
Log Population (ln)	0.0045	0.0008	$<$ 0.001 ***

Significance codes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

R^2: 0.8533 **Residual Std. Error:** 0.0148 **Adjusted R-Squared:** 0.8472