Retirement and Disability Research Center
UNIVERSITY OF WISCONSIN–MADISON

Reem El Cheikh Taha
Auburn University

John M. Nunley
University of Wisconsin-La Crosse

R. Alan Seals, Jr.
Auburn University

Christopher John F. Cruz
Grand Valley State University

Automation and the Decline in Social Security Disability Insurance Applications

The research reported herein was performed pursuant to a grant from the US Social Security Administration (SSA) funded as part of the Retirement and Disability Consortium. The opinions and conclusions expressed are solely those of the author(s) and do not represent the opinions or policy of SSA or any agency of the Federal Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of the contents of this report. Reference herein to any specific commercial product, process or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply endorsement, recommendation or favoring by the United States Government or any agency thereof.

Abstract

SSDI application rates rose from the early 1990s until a sharp decline began in 2010. Various factors, such as improved economic conditions, aging Baby Boomers, better workforce support for disabled individuals, changes in claim processing, and lack of program awareness, have been suggested to explain this trend. This study examines the relationship between SSDI applications and the spread of automation technologies. Automation could influence SSDI applications by displacing workers and reducing wages or by making workplaces safer and less injury-prone, thus affecting the number of disability claims. Using confidential data on SSDI applications at the commuting-zone level, we estimate the effect of automation exposure on application rates across age and gender groups from 2005–2019. Our findings suggest that SSDI application rates for the 18–64 age group decline with greater automation exposure. This effect is more pronounced in the 35–54 and 55–64 age groups. While automation exposure has minimal impact on younger workers (18–34), the 35–54 age group experiences the largest negative effect, with smaller yet statistically significant effects for the 55–64 age group. These patterns hold across broadly (automation-intensive employment shares) and narrowly (industrial robots) defined measures of automation exposure.

Key words: Social Security Disability Insurance (SSDI), disability, industrial robots, automation

JEL Categories: J14, J23, J24, O33

1 Introduction

SSDI application rates generally followed an upward trend from the early 1990s until an unexpected and sharp reversal began in 2010.[1] Various hypotheses have been put forward to explain the abrupt change in applications, including improved economic conditions, the transition of Baby Boomers out of disability-prone age brackets, enhanced support for disabled individuals in the workforce, modifications in claim processing, and a lack of program awareness among eligible individuals.[2] This study explores the relationship between SSDI applications and the proliferation of robotics and other automation technologies.

Recent research shows that automation accounts for most of the change in the US wage structure since 1980 (Acemoglu and Restrepo, 2022b, 2024; Acemoglu et al., 2024). Automation could impact disability uptake by displacing workers and lowering wages or reducing injury-prone tasks (e.g., see Gihleb et al., 2022). The predicted impact on SSDI applications is therefore ambiguous, as automation may create slack conditions in the labor market, increasing uptake, or create safer work conditions, decreasing uptake. This paper presents estimates on the relationship between exposure to automation technologies and SSDI application rates.

Using confidential data on SSDI applications measured at the commuting-zone level, we estimate the effect of various measures of "automation" exposure on changes in per-capita SSDI applications by age group (18–64, 18–34, 35–54, and 55–64) and age group × sex. The extent of automation in the economy is measured via employment shares, which follows Autor and Dorn (2013), and industrial robot exposure, which is based on the measures used in Gihleb et al. (2022) and Acemoglu and Restrepo (2020). Long- and stacked-differences specifications are estimated using the The 2005–2019 to estimate the effects of automation exposure on changes in the SSDI application-to-population ratio.

The regression equation of interest is estimable by ordinary least squares (OLS), but the adoption of automation technologies can be influenced by other factors that affect firms' demands for labor within a commuting zone. If these factors were observable, we could account for them. However, many of these factors are likely unobserved and, as a result, an alternative estimation strategy is needed. To circumvent these empirical identification issues, the study employs an instrumental variables estimation strategy that follows Autor and Dorn (2013) for the employment-share measures and Gihleb et al. (2022) and Acemoglu and Restrepo (2020) for the measure of industrial robot exposure. For the former, we develop

[1] See `https://www.ssa.gov/oact/STATS/table6c7.html`.

[2] e.g., see `https://www.ssa.gov/policy/docs/briefing-papers/bp2019-01.html`.

a Bartik-type instrument from the 1990 routine employment share to predict automation-related employment shares in 2005.[3] For the latter, we leverage the fact that European countries led the US in terms of robot adoption and, therefore, industry-specific adoptions across France, Denmark, Finland, Italy, Germany, Norway, Spain, Sweden, and the United Kingdom are used as an instrument to predict US industrial robot exposure in 2005. For both the automation-related employment-share measures and the robot-exposure measure, the instruments are relevant in the first stage and contend that they affect SSDI outcomes via their impact on the suspected endogenous variable (i.e., the automation measures).

Application rates for the 18- to 64-year-olds decrease with greater automation exposure. These effects are not statistically significant with OLS but become significant with 2SLS. The estimates for 18–64 year-olds, however, mask heterogeneity in the effects of automation exposure across the 18–34, 35–54, and 55–64 age groups. The 35–54 age group shows the largest negative effect, while the 55–64 age group also shows significant but smaller negative effects.

The study's findings for the link between industrial robot exposure and SSDI application rates align with those from the analysis using the automation-related employment-share measures. One difference in the estimates of industrial robot exposure's impact and that of the employment-share measures is that the negative effects are statistically significant, regardless of whether OLS or 2SLS is used to estimate the parameters of interest. The similarities include null effects for the 18–34 age group and a negative and statistically significant effect for the 35–54 and 55–64 age groups. Moreover, industrial robot exposure has the largest effect on the 35–54 age group, which is also consistent with the findings for the employment-share measures.

2 Literature Review

2.1 Disability Applications

Since the 1970s, the SSDI program has undergone dramatic changes, mainly influenced by shifts in beneficiary demographics and a labor market with increasing female participation (Liebman, 2015). As baby boomers began to reach ages with higher SSDI risk, concerns arose about the program's solvency and policies (Autor and Duggan, 2006).

Maestas et al. (2021) find that the Great Recession led to nearly one million additional SSDI applications, of which 41.8 percent were awarded benefits, resulting in over 400,000

[3] See, for example, Goldsmith-Pinkham et al. (2020); Borusyak et al. (2022)

new beneficiaries who constituted 8.9 percent of all new beneficiaries during the recession. However, an unexpected large decrease in SSDI claiming began as the deleterious employment effects of the Great Recession began to wane. As of 2022, approximately 9 million individuals received $12.6 billion in SSDI payments and the program is expected to remain solvent through the long-run 75-year projection period (US Social Security Administration, 2023).

2.2 Automation and Tasks

2.2.1 General Automation Technologies

The increasing penetration of robots and technology in the labor market has raised concerns about the future of employment and wages (Johnson and Acemoglu, 2023). Acemoglu and Restrepo (2022b) reveal that automation-driven task displacement accounts for 50–70 percent of changes in the US wage structure over the past four decades, particularly affecting routine-task workers. Their study also shows that automation significantly reshapes wage structures, increasing the college premium while reducing real wages for less educated workers.

Autor and Dorn (2013) exploit differential exposure to routine tasks across local labor markets to identify the effects of automation technology adoption on labor-market outcomes. By combining past values of a commuting-zone-industry employment share with industry-level routine occupational share, the authors create a shift-share-type instrumental variable (e.g., see Goldsmith-Pinkham et al., 2020). They find that automation (particularly computers) increased wages for high-skilled workers and substituted for lower-skill workers in routine tasks. This also increased demand for service-sector jobs and reallocated lower-skilled labor to those jobs.

Bratsberg et al. (2022) study the Norwegian labor market over approximately the same sample period as this study. Individuals in occupations with higher routine task intensity (RTI) scores in 2003 were significantly less likely to remain employed and more likely to receive a disability pension or die by 2019. The study found that a standard-deviation increase in RTI score was associated with a 6.7 percent higher mortality rate for men and a 5.5 percent higher rate for women.

Acemoglu and Restrepo (2022a) show that aging populations promote firm demand for automation technologies, particularly as older workers are often less suited for physically demanding tasks. This scenario leads to a capital substitution effect, where labor-saving automation technologies become more prevalent because of the relative scarcity and higher wage demands of younger workers. Identifying an effect of automation on disability claiming requires taking account of these market forces, as aging populations will have both higher levels of disability and automation.

The literature on the health effects of automation risk has mixed results from labor markets around the world. Lordan and Stringer (2022) find small, negative effects of automation risk on the mental health of Australian workers. Blasco et al. (2024) also find negative effects on mental health for French workers, but the effect sizes are larger than those reported in Lordan and Stringer (2022). In Blasco et al. (2024), anxiety over the labor market is the primary mechanism through which the effect operates. Cheng et al. (2021) study a nationwide survey from Taiwan and find that people employed in jobs with a high likelihood of automation tended to experience lower levels of job control, greater job insecurity, and a higher prevalence of work-related injuries and illnesses. In contrast, those in positions with a low probability of automation faced greater psychological and physical demands, along with a higher incidence of burnout.

2.2.2 Industrial Robots

A growing body of literature examines the effects of industrial robot exposure on employment, wages, and worker health across different regions and market sectors. Acemoglu and Restrepo (2020) analyze the effect of the increase in industrial robot usage on US local labor markets. They show that robots may reduce employment and wages, and that the local-labor-market effects of robots can be estimated by regressing the changes in employment and wages on robot exposure. Dauth et al. (2021) estimate the effect of industrial robots on employment, wages, and the composition of jobs in German labor markets. They find that the adoption of industrial robots had no effect on total employment in local labor markets specializing in industries with high robot usage. Robot adoption led to job losses in manufacturing that were offset by gains in the business service sector. Acemoglu et al. (2020) study the firm-level implications of robot adoption in France and find that adopters experienced significant declines in labor shares and the share of production workers in employment and increases in value added and productivity and overall employment. However, the employment expansion comes at the expense of competitors, leading to a net negative effect on employment.

The literature on the health effects associated with industrial robot penetration is less mixed than for automation shares and automation risk. Gunadi and Ryu (2021) investigate the health impacts of increased industrial robot use in US cities. They find that a 10 percent increase in robots per 1,000 workers corresponds to a 10 percent decrease in poor health, work disability, and job quitting for health reasons among this group. This improvement is partly attributed to a shift away from physically demanding tasks. Gihleb et al. (2022) find that an increase of one standard deviation in industrial robot use reduces work-related injuries by approximately 1.2 per 100 full-time workers in the US and Germany. They report an annual injury cost reduction of $1.69 billion (2007 dollars) from 2005-2011, mainly due to decreased injury rates in manufacturing. However, US mortality and mental health data, shows a countervailing effect, as industrial robot penetration into a labor market significantly escalates substance abuse-related deaths and negatively impacts mental health. O'Brien et al. (2022) also found that robot penetration increases the rate of overdose deaths in the United States. However, Gihleb et al. (2022) find no consequential effects on suicide rates, suggesting that labor-market pressures and anxiety, exacerbated by robot penetration, are primarily responsible for the observed impacts.

3 Data

3.1 SSDI Applications

The dependent variables used in this study are population-weighted SSDI application counts across commuting zones and years by (a) age group (18–64, 18–34, 35–54, and 55–64) and (b) age group × sex. These data were provided by the SSA and are confidential.[4] In some instances, application counts from particular CZs are suppressed because of an insufficient number of applications in these locations. The missing information varies by (a) age group and (b) age group × sex. When sex is ignored in the application counts, the data cover 86 percent of CZs (643 out of 722) for the 18–64 age group;[5] 87 percent for the 18–34 age group, 95 percent for the 35–54, and 94 percent for the 55–64 age group. When examining application rates by sex and age group, the shares of CZs that are covered fall to 77, 78, 89, and 86 percent for the 18–64, 18–34, 35–54, and 55–64 age groups, respectively.[6] The CZs with missing SSDI application data tend to be located in the Mountain West and West North Central Midwest, and these locations have relatively small populations.

[4] All results involving the data provided by SSA must be approved by the SSA's Disclosure Review Board (DRB) for external presentation. The estimates reported herein have obtained the DRB's approval.

[5] Using 1990 commuting zones, the entire US has 741 commuting zones. Our analysis is restricted to the continental US, which results in the exclusion of 19 commuting zones located in Alaska and Hawaii.

[6] The 18- to 64-year-old outcome is created by summing the application counts for the 18–34, 35–54, and 55–64 age groups (and then dividing by the commuting zone's population). The SSA redacted data for commuting zones with applications below the threshold count required for external presentation, and the redactions vary by age group as well as age group × sex. Thus, when computing the measure for 18- to 64-year olds, the numbers of observations are less than those for the subgroups, as a commuting zone with a missing value for only one of the three age groups results in a missing value for the application rate for all age groups combined. The extent of redaction is greater for the 18–34 age group than the 35–54 and 55–64 age groups.

To adjust the raw application counts for population size, we use total population counts from the US Census's intercensal tables, which provide county-year population statistics by age and sex. Using the county-year population counts, the populations of counties within a CZ were then summed.[7] The outcome measure, termed the application-to-population ratio, is calculated as follows:

$$Apps^{*}_{a,s,c,t} = \frac{Apps_{a,s,c,t}}{Pop_{c,t}} \times 100. \tag{1}$$

The terms a, s, c, and t index age groups, sex, commuting zones, and years, respectively. The variable $Apps^{*}_{a,s,c,t}$ represents the SSDI application-to-population ratio for age group a and sex s in commuting zone c_d at year t; $Apps_{a,s,c,t}$ is the raw count of SSDI applications for age group a and sex s in commuting zone c at year t; and $Pop_{c,t}$ is the population count in commuting zone c_d at year t. In our econometric analysis, we estimate separate regressions for the 18–64, 18–34, 35–54, and 55–64 age groups with and without taking sex into account.

Table 2 reports the averages (column 1) and standard deviations (column 2) of the SSDI application-to-population ratios for each age grouping. In addition, we decompose the standard deviation for each outcome variable into "within" and "between" components (column 3 and 4). The ratios for males and females combined are presented in Panel A, and those for males and females separately are shown in Panels B and C, respectively. Regardless of sex, the SSDI application-to-population ratios vary considerably across the age groups. For example, the ratio for the 35–54 age group is almost three times larger than that for the 18–34 age group and 1.5 times than that for the 55–64 age group. Moreover, the variation between commuting zones is greater than the within-commuting-zone-variation by a factor of 1.6 to 2.4. These patterns hold for males and females (see columns 3 and 4).

In Figure 2, we present the 2005 geographic variation (across the continental US) in the SSDI application-to-population ratio (Panel A) along with its change between 2005 and 2019 (Panel B) for the 18–64 age group. The application-to-population ratios tend to be highest in the south, but commuting zones in southern states, such as North Carolina, West Virginia, Tennessee, South Carolina, Arkansas, and Kentucky, experienced relatively larger reductions than commuting zones located in the Mountain West, West North Central, and Middle Atlantic regions. In total, about 75 percent of commuting zones across the United States experienced reductions in the application-to-population ratios.

Figures 3, 4, and 5 present maps analogous to those shown in Figure 2 separately

[7] The county-year population statistics were downloaded from the US Census's intercensal tables (2000–2009) and Vintage 2020 Population estimates (2010–2019).

for each age group: 18–34, 35–54, and 55–64. Across the three figures, there is a similar concentration in 2005 of higher application-to-population ratios in the southern states. However, the evolution of the application-to-population ratios between 2005 and 2019 varies considerably across the three age groups. For the 35–54 age group, the reduction in the ratio spans the vast majority of the US, with only around 3 percent of commuting zones experiencing upticks in their ratios. The share of commuting zones experiencing upticks in their application-to-population ratios is 54 percent for the 18–34 age group and 72 percent for the 55–64 age group. The patterns in Figures 3 and 5 indicate that it is the 35–54 age group (i.e., Figure 4) driving the patterns observed in Figure 2.

3.2 Measuring Automation

3.2.1 Automation-Related Employment Shares

We use three different employment-share variables to measure the extent of automation in the labor market. The measures are based on data from the Occupational Information Network (O*NET) and the American Community Survey (ACS). The computation of the three employment shares follow Autor and Dorn (2013), who connect the routine employment share to the hollowing out of middle-skill occupations in the US. Their measure identifies occupations in the top 1/3 of the 1950 routine task intensity distribution, and then employment in routine-intensive occupations as well as employment in general are totaled by commuting zone and year. We perform the same calculations but use variables from O*NET instead of the Department of Transportation (DOT) and rely on the 1990 task-intensity distributions instead of those from 1950.

The three employment-share variables are based on the routine task intensity measure from Deming (2017), which has two components. The first relates to the importance of performing repetitive tasks, and the second captures the extent to which a job is automated. The first employment share measure uses the composite measure, whereas the second and third employment-share variables use the individual components of the composite measure (i.e., repetition and automation). The three measures are referred to as the routine-, repetition-, and automation-intensive employment shares.

To perform the calculations, we first follow the approach commonly used in the literature and convert the measures based on the O*NET data, which are ordinal (typically 1-5 or 1-7 scales), to 0-10 scales (e.g., see Deming, 2017). This allows for identifying occupations in the top 1/3 of the variable's distribution. The occupation-identifier is then linked to the 2005-2019 ACSs via the 2018 Standard Occupation Classification (SOC) codes. For the occupations identified as being intensive in routine, repetition, and automation tasks, the number of workers employed in these occupations is counted as well as the number of workers employed overall in each CZ-year. The ratio of these employment counts yields the three employment-share variables used in our regression analysis. Formally, the employment shares are computed as follows:

$$Empsh_{c,t} = \left(\sum_{j=1}^{J} Emp_{c,j,t} \cdot 1 \left[TaskInt_{j,2005} > TaskInt^{p66}_{2005} \right] \right) \left(\sum_{j=1}^{J} Emp_{c,j,t} \right)^{-1} \quad (2)$$

The terms c, j, and t index commuting zones, occupations, and time periods, respectively. The variable $Empsh_{c_s,t}$ measures the share of workers employed in occupations in the top $1/3$ of the automation measure's task-intensity distribution in commuting zone c_s in year t; $Emp_{c_s,j,t}$ is employment in occupation j in commuting zone c_s at time t; and $1 \cdot [TaskInt_{j,2005} > TaskInt^{p66}_{2005}]$ is an indicator function that identifies occupations in the top $1/3$ of the task intensity distribution in 2005. The three variables that comprise $Empsh_{c_s,t}$ include the routine-intensive employment share ($RSH_{c_s,t}$), the repetition-intensive employment share ($RTSH_{c_s,t}$), and the automation-intensive employment share ($ASH_{c_s,t}$).[8]

Table 1 presents the top 10 occupations ranked for each of the automation-related employment shares. In some cases, there is overlap across the measures in terms of rankings. For example, the bookkeeping, accounting, and auditing clerks occupations are ranked #1, #1, and #8, respectively, across the routine, repetition, and automation intensity measures. However, travel agent and insurance claims occupations tend to be "automation intensive" and less "repetition intensive". By contrast, other occupations tend to be more "repetition intensive" than "automation intensive", such as tellers, pharmacists, and tax preparers.

In Figure 6, we present the initial routine-, repetition-, and automation-intensive employment shares across commuting zones for the year 2005 (Panels A, C, and E) as well as the change in each measure between 2005 and 2010 (Panels B, D, and F). In terms of the initial geographic distribution, no clear pattern emerges, as the heat map reveals both relatively high and relatively low employment shares in the same geographic regions.

[8] The computation of the employment shares requires two separate aggregations. The first is to the PUMA-year level, and the second is from the PUMA-year to the CZ-year level. Frequency weighting is applied in the first aggregation using the *perwt* variable from Ruggles et al. (2023). For the second aggregation, we incorporate the PUMA-commuting zone crosswalk from Autor and Dorn (2013), and sum the number of workers employed overall as well as those in occupations intensive in routine, repetitive, and general automation tasks across PUMAs within a CZ-year. Because there is uncertainty regarding the geographic location of survey respondents, we use the allocation factors from Autor and Dorn (2013), which are applied as importance weights in the aggregation from the PUMA-year to the CZ-year level.

The changes in these variables, however, vary considerably between 2005 and 2010, as the automation-intensive employment share becomes much more pervasive across commuting zones than the routine- and repetition-intensive employment shares. Similar to the initial geographic distributions, there is more of a pattern than the initial distribution, as Heartland disproportionately experiences increases between 2005 and 2010 (Panels B and D). For the automation-intensive employment share (Panel F), its increase is widespread, affecting each region across the US.

3.2.2 Industrial Robot Exposure

The International Federation of Robotics (IFR) has collected data on industrial robot usage across countries and industries since the early to mid-1990s. The IFR provides industry-year information on robot usage for various European countries dating back to its inception, but the inclusion of North America began in 2004. The IFR combined robot usage in the US, Canada, and Mexico into a single category from 2004 to 2010 before reporting separate statistics for each country beginning in 2010. Because our study period uses data reported before and after the reporting change, we use the North American statistics for the 2004–2010 period and then those for the United States thereafter. For the 2004–2010 period, the US accounts for over 90 percent of the North American market, making the use of North American robot usage a viable proxy (e.g., see Acemoglu and Restrepo, 2020). Measurement error is then inherent in the robot exposure measure used in our regression analysis, which results in attenuation bias. As such, an estimation strategy is needed to remove bias resulting from mismeasurement of the key explanatory variable. We return to this discussion in Section 4.3.

The IFR provides information on robot usage for nonmanufacturing industries (agriculture, forestry and fishing, mining, utilities, construction, education, research and development, and services) as well as disaggregated industries within the manufacturing sector (food and beverage, textiles, wood and furniture, paper and printing, plastics and chemicals, minerals, basic metals, metal products, industrial machinery, electronics, automotive, shipbuilding and aerospace, and "other" manufacturing). In total, there are 19 industries in the IFR classification.

The analysis includes two measures of industrial robot exposure. The first is from Gihleb et al. (2022) and the second follows Acemoglu and Restrepo (2020). The measure from Gihleb et al. (2022) requires two data sources: the robot usage data from IFR and industry-commuting employment statistics from the 1990 Census. Acemoglu and Restrepo (2020) rely on the same data sources but also incorporate an industry-level output growth measure from the EU KLEMS and the world supply of industrial robots from the IFR. The measure from Gihleb et al. (2022) is defined as follows:

$$RobotExp_{c,t}^{GGSW} = \sum_{i=1}^{I} l_{i,c,1990} \frac{M_{i,t}}{E_{i,1990}}. \tag{3}$$

The $GGSW$ superscript indicates that the measure is from Gihleb et al. (2022). The variable $M_{i,t}$ is the operation stock of autonomous robots in industry i in year t, and $E_{i,1990}$ is employment in industry i in 1990. The exposure measure is created by projecting the robots per worker to commuting zones via multiplication by the 1990 share of workers employed in industry i (i.e., $l_{i,c_d,1990}$).

The second measure, which is from Acemoglu and Restrepo (2020), is defined as follows:

$$RobotExp_{c,(t_0,t_1)}^{AR} = \sum_{i=1}^{I} l_{i,c,1990} \cdot APR_{i,(t_0,t_1)}, \tag{4}$$

in which

$$APR_{i,(t_0,t_1)} = \frac{M_{i,t_1} - M_{i,t_0}}{E_{i,1990}} - g_{i,(t_0,t_1)} \frac{M_{i,t_0}^{w}}{E_{i,1990}}. \tag{5}$$

The AR superscript indicates that the measure is from Acemoglu and Restrepo (2020). The variable $APR_{i,(t_0,t_1)}$ represents the adjusted robot-penetration ratio for industry i between time periods t_0 and t_1, which is projected to commuting zones via multiplication by the share of workers employed in industry i located in commuting zone c in 1990 (i.e., $l_{i,c,1990}$). In equation 5, the first term is the change in the operational stock of autonomous industrial robots in commuting zone c between time periods t_0 and t_1 relative to employment for industry i in 1990, and the second term is the growth rate of output for industry i between time periods t_0 and t_1 multiplied by the global stock of robots, M_{i,t_0}^{w}, for industry i at time t_0 relative to employment for industry i in 1990.

In Figure 7, we present the initial 2005 geographic distribution of robots per 1,000 workers (Panel A) and its changes between 2005 and 2010 (Panel B). Initially, robot usage is concentrated in the upper Midwest, but the usage in industrial productions expands in *all* areas. The expansion between 2005 and 2010 is greatest in the upper Midwest and surrounding areas.

4 Econometric Methodology

Our empirical strategy relies on both ordinary OLS and 2SLS estimation to compute the parameters of interest. In what follows, we outline our empirical approach for studying the relationship between the employment-share variables and SSDI applications (Section 4.2) as well as our strategy for empirically examining the link between exposure to industrial robots and SSDI applications (Section 4.3).

4.1 Dependent Variables

We transform SSDI outcomes presented in Table 2 into "long" or "stacked" differences for the purposes of our econometric analysis, which are based on differences in $Apps^*_{a,s,c,(t_0,t_1)}$ between a starting (t_0) and ending (t_1) period. We compute the following separately for each age-sex group:

$$Apps^*_{c,(t_0,t_1)} = Apps^*_{c,t_1} - Apps^*_{c,t_0}. \tag{6}$$

The starting and ending periods for our long and stacked differences specifications vary across the automation measures. When examining the employment-share measures and the robot-exposure measure from Gihleb et al. (2022), 2005 and 2019 are the starting and ending periods, respectively. For the stacked-differences specifications, there are two starting (s_1 and s_2) and ending (e_1 and e_2) periods. In our primary specification, $t_0^{s_1} = 2005$ and $t_0^{s_2} = 2010$ are the two starting periods and $t_1^{e_1} = 2010$ and $t_1^{e_2} = 2019$ are the two ending periods. The application of Acemoglu and Restrepo (2020)'s measure relies on 2004 and 2016 as the starting and ending years, respectively, for the long-differences specifications. For the stacked-differences specifications, the starting and ending years of the first "stack" are 2004 and 2007, respectively, and the starting and ending years of the second "stack" are 2013 and 2016, respectively.[9]

4.2 Automation-Related Employment Shares

The structural regression equation for the long-differences specifications is

[9]Our findings are insensitive to starting the analysis in 2004 versus 2005. Moreover, our findings are qualitatively robust to different starting and ending years across the two stacks. In the specification we present in this report, we use the early and late parts of the sample. The use of 2004-2007 and 2013-2016 also removes the influence of the Great Recession. We note that our results are even stronger if we examine, for example, 2004-2010 and 2010-2016 as the two stacked differences.

$$Apps^{*}_{c_s,(t_0,t_1)} = \alpha_0 + \alpha_1 Empsh_{c_s,t_0} + X^{'}_{c_s,t_{-1}}\Lambda + \phi_s + \epsilon_{c_s,(t_0,t_1)}. \tag{7}$$

The variable $Apps^{*}_{c_s,(t_0,t_1)}$ represents the dependent variables defined in equations (6) for commuting zone c_s (s indexes the state in which the commuting zone is located) between a starting period, t_0, and an ending period, t_1. The employment-share variables, which are defined in equation (2), are represented by the variable $Empsh_{c_s,t_0}$, which is for commuting zone c_s at the starting period, t_0. Separate models are estimated for the three employment-share variables: the routine-intensive employment share (RSH_{c_s,t_0}), the repetition-intensive employment share ($RTSH_{c_s,t_0}$), and the automation-intensive employment share (ASH_{c_s,t_0}).

The vector of control variables, $X_{c_s,t_{-1}}$, is from Autor and Dorn (2013), whose study includes controls for the college-to-noncollege population, the ratio of immigrants to the noncollege population, the share of workers employed in manufacturing, the unemployment rate, the share of the population that is female and employed, the share of the population 65 or older, and the share of noncollege workers workers earning a real wage below the minimum wage that will prevail over the next decade (2000-2010). In addition, we include two additional control variables: the "China Shock" from Autor et al. (2013) and an SSDI processing efficiency variable from Kearney et al. (2021). Each of the variables in $X_{c_s,t_{-1}}$ is measured for commuting zone c_s in the year 2000, except the "China Shock" variable and the SSDI processing variable. The variable capturing rising import competition from China is the change in Chinese import exposure between 1990 and 2000, and the SSDI processing variable is measured in 2003 (the first year it is available). The inclusion of ϕ_s, which is a set of state dummy variables, means that the estimate for the parameter of interest, α_1, is identified based on temporal variation within states and across CZs. The error term, $\epsilon_{c_s,(t_0,t_1)}$, captures predictors of $Apps^{*}_{c_s,(t_0,t_1)}$ not held constant.

When estimating the stacked-differences specifications, equation (7) is altered as follows:

$$Apps^{*}_{c_s,(t_0^{s_i},t_1^{e_i})} = \beta_0 + \beta_1 Empsh_{c_s,t_0^{s_i}} + \beta_2 D_{(t_0^{s_i},t_1^{e_i})} + X^{'}_{c_s,t_{-1}^{s_i}}\Theta + \phi_s + \epsilon_{c_s,(t_0^{s_i},t_1^{e_i})}. \tag{8}$$

Equation (8) adds new superscripts to the time indices, t_0 and t_1, as well as an additional right-hand-side variable, $D_{(t_0^{s_i},t_1^{e_i})}$. The superscripts s_i and e_i index the different starting and ending years used in the stacked-differences specifications. The inclusion of $D_{(t_0^{s_i},t_1^{e_i})}$ accounts for differences in the application-to-population ratio across the difference periods. Thus, the stacked-differences estimates are based on within-period variation instead of variation over the entire time horizon, as is the case in the long-differences specifications.

We include two stacked differences, 2005–2010 and 2010–2019, in our empirical specifications. In effect, a panel of commuting zones with two time periods (one for each "stack") is formed for the dependent variables. Thus, each commuting zone has two observations. The first is for the 2005-2010 period and the second is for the 2010–2019 period. The variable $Empsh_{c_s,t_0^{2005}}$ is assigned to the first period (2005–2010) and $Empsh_{c_s,t_0^{2010}}$ is assigned to the second period (2010–2019). The variables in $X^{'}_{c_s,t^{s_i}_{-1}}$ are the same as those in equation (7), except the 1990 version of the control variables are linked to the first period (2005–2010) and the 2000 version of the control variables are linked to the second period (2010–2019). For the "China Shock," we use the change between two different points in time: the change in Chinese import penetration from 1990–2000 is linked to the 2005–2010 difference, and the analogous change between 2000 and 2007 is linked to the 2010–2019 difference.[10] The 2003 values for the SSDI processing time control variable are linked to the 2005–2010 difference period, and the 2008 values are linked to the 2010–2019 difference period.

When equations (7) and (8) are estimated via OLS, $\hat{\alpha}_1$ and $\hat{\beta}_1$ have causal interpretations if the following are true: (i) the automation proxies and population-adjusted SSDI applications are not jointly determined; (ii) after conditioning on the full set of control variables, the automation measures are uncorrelated with factors in the error term that also affect SSDI outcomes; and (iii) the automation-related variables are measured without error. Simultaneity bias is unlikely given that automation technologies generally affect economic outcomes with a delay. However, it is more likely that unmeasured variables in the error term or measurement error could bias our OLS estimates. Given these concerns, an alternative estimation strategy that circumvent these problems is needed.

We employ an instrumental variables (IV) estimation strategy to measure the causal effect of interest. For the three automation-related employment-share variables, we use an already-established instrument from Autor and Dorn (2013), who examine the effects of changes in the routine employment share on changes in service-sector employment via OLS and IV estimation. The instrument uses historical differences in industrial composition across commuting zones as a source of plausibly exogenous information with which to identify the causal effect of interest. We use the same instrument for $Empsh_{c_s,t_0}$ but make one change. The task-intensity distribution used in our cases is from 1990, whereas Autor and Dorn (2013) use the 1950 distribution.[11] The instrument has two components, both measured

[10]The data for Chinese import penetration into US commuting zones are from the replication package associated with Autor et al. (2013), which is available at the following link: `https://www.openicpsr.org/openicpsr/project/112670/version/V1/view`.

[11]Our use of 1990 instead of 1950 is due to the time periods of our study relative to Autor and Dorn (2013). Their study covers a long time span, 1950-2005. Our sample period spans 2005 through 2019.

15 years prior to the start of our sample period. The first is the industrial composition of commuting zones in 1990, and the second is the national structure of occupations across industries in 1990. The product of these separate components forms the instrument:

$$\widetilde{RSH}_{c_s,1990} = \sum_{i=1}^{I} \Big(E_{i,c_s,1990} \times R_{i,-c_s,1990}. \tag{9}$$

The variable $\widetilde{RSH}_{c_s,1990}$ is the routine employment share in commuting zone c in 1990, which is the instrument for $Empsh_{c_s,t_0}$; $E_{i,c_s,1990}$ is the employment share of industry i in commuting zone c_s in 1990; and $R_{i,-c_s,1990}$ is the routine occupation share among workers in industry i across all commuting zones except commuting zone c_s in 1990. When estimating the long-differences specifications, $\widetilde{RSH}_{c_s,1990}$ enters the first-stage regression equation as defined in equation 9, but $\widetilde{RSH}_{c_s,1990}$ and its interaction with $D_{(t_0^{s_i},t_1^{e_i})}$ comprise the instruments in the first-stage regression equation when estimating the stacked-differences specifications.

In the context of Autor and Dorn (2013), using equation (9) as instrument allows them to identify the quasi-permanent component of the routine employment share's impact on different outcomes, as the instrument would affect the long-run component but likely have no relationship with short-term fluctuations in outcomes. The same logic applies to our case. One would expect the routine employment share across occupations and industries in 1990 to be a powerful predictor of the routine-, repetition-, and automation-intensive employment shares in 2005. Indeed, this is shown to be the case in Section 5.1.

4.3 Industrial Robot Exposure

To estimate the relationship between industrial robot exposure and the change in the SSDI application-to-population ratio, both OLS and IV estimation are used. The structural regression equation for the long-differences specifications is

$$Apps^*_{c_d,(t_0,t_1)} = \gamma_0 + \gamma_1 RobotExp_{c_d,t_0} + X'_{c_d,t_{-1}}\Psi + \phi_d + \epsilon_{c_d,(t_0,t_1)}. \tag{10}$$

The subscript indexing the broader location of commuting zone c differs between the studies of Acemoglu and Restrepo (2020) from Autor and Dorn (2013). The former study identifies the effects via variation across commuting zones within US Census divisions, which is indexed in equation (10) with a d, rather than variation across commuting zones within states, which is indexed with an s. Each of the variables in equation (10) are defined above. The variables in $X_{c_d,t_{-1}}$ differ from those in equations (7) and (8). In particular, we hold constant the full set of control variables from Acemoglu and Restrepo (2020), which includes US Census division dummies (i.e., ϕ_d), the natural logarithm of the population, the unemployment rate, the shares of the population who are female, Asian, Black, Hispanic, White, over age 65, did not go to college, completed some college, graduated with a college degree or professional degree, completed a master's or doctorate degree, and employed in manufacturing in general as well as light manufacturing, and the share of females employed in manufacturing relative to total manufacturing employment. Each of these variables are measured in 1990, as in Acemoglu and Restrepo (2020). The "China Shock" and the SSDI processing time controls are held constant and are defined in Section 4.2.

We augment equation (10) and estimate stacked-differences specifications analogous to those described in Section 4.2. In particular, we add the variable $D_{(t_0^{s_i},t_1^{e_i})}$ to equation (10) and compute the stacked-differences estimates analogous to those described in equation (8) but with the right-hand-side variables shown in equation (10) held constant. When estimating the stacked-differences specifications, the 1990 version of the control variables held constant in equation (10) are linked to the first stack (either 2005-2010 or 2004-2007) and their 2000 versions are linked to the second stack (either 2010-2019 or 2013-2016).

We study the relationship between exposure to industrial robots via OLS and IV estimation. The same concerns of bias apply to robot exposure as applied to estimating the effects of the employment-share variables. As a result, we follow Acemoglu and Restrepo (2020) and Gihleb et al. (2022) and employ an IV estimation strategy that leverages the fact that European countries began adopting autonomous robots in production prior to their use in production throughout the US.

The instrument employed when using the robot-exposure measure from Gihleb et al. (2022) is defined as

$$RobotExp_{c_d,t_0}^{IV,GGSW} = \sum_{i=1}^{I} l_{i,c_d,1970} \frac{M_{i,t_0}^{EU}}{E_{i,1990}^{EU}}. \tag{11}$$

The variable $l_{i,c_d,1970}$ is the share of workers employed in commuting zone c_d in 1970; M_{i,t_0}^{EU} is the operation stock of autonomous industrial robots in industry i at time t_0 in nine European countries, which includes France, Denmark, Finland, Italy, Germany, Norway, Spain, Sweden, and the United Kingdom; and $E_{i,1990}^{EU}$ is employment in industry i across the aforementioned European countries in 1990. When estimating the stacked-differences specifications, M_{i,t_0}^{EU} is altered to $M_{i,t_0^{s_i}}$, which links the 2005 value to the starting period of the first difference (2005-2010) and the 2010 value to the starting period of the second difference (2010-2019).

When applying the measure from Acemoglu and Restrepo (2020), our instrument again relies on the previously described European countries. The measure is defined identically to the measure for the US, but for the 1994-2004 period. Formally, the instrument is

$$RobotExp^{IV,AR}_{c,(1994,2004)} = \sum_{i=1}^{I} \Big(l_{i,c,1990} \cdot APR^{EU}_{i,(1994,2004)}, \qquad (12)$$

When estimating the stacked-differences specifications, the instrument defined in equation (12) is included along with its interaction with $D_{(t_0^{s_i},t_1^{e_i})}$.

5 Results

We present the findings from our econometric analysis in two subsections. The first focuses on the automation-related employment shares (i.e.,routine, repetition, and automation), and industrial robot exposure is the subject of the second subsection.

5.1 Automation-Related Employment Shares and SSDI Applications

In Table 3, we present OLS and 2SLS estimates for the relationship between the three automation-related employment share measures and the change in the SSDI application-to-population ratio over the 2005-2019 period. The table includes eight columns, with the odd-numbered columns containing the OLS estimates and the even-numbered columns showing the 2SLS estimates. The table is separated into three panels, each of which focuses on a particular employment-share variable. The estimates for the 18–64 age group are shown in columns 1 and 2, the 18–34 age group in columns 3 and 4, the 35–54 age group in columns 5 and 6, and the 55–64 age groups in columns 7 and 8. In the final row of the table, we provide the unconditional average of the dependent variable employed in each specification.

From the table, we find little evidence of statistically significant relationships between the employment-share variables and the SSDI outcomes when using OLS estimation. In fact, we only find statistically significant effects for 35–54 age groups in Panels A and B. However, when estimating the parameters of interest with 2SLS, we find robust evidence of a statistically significant relationship between the employment-share variables and the SSDI outcomes, particularly for 35–54 and 55–64 age groups. In absolute value, the coefficients estimated via 2SLS are roughly twice as large as those estimated via OLS. The instrument is relevant in the first-stage regression, which is supported by the statistical significance of the coefficient on the instrument as well as the relatively large KP F-statistics.

In terms of interpretation, we use the standard deviations of employment-share variables to evaluate the magnitude of the effects shown in Table 3. For each employment-share measure, the standard deviation is approximately three. From column 2 in Panel A, the -0.0069 coefficient when multiplied by three becomes approximately -0.021. Thus, a standard-deviation increase in the routine-intensive employment share corresponds to approximately a -0.021 percentage-point reduction in the application-to-population ratio for the 18–64 age group. Given that the the application-to-population ratio fell by 0.083 percentage points over the 2005-2019 period, the 0.021 percentage-point reduction implied by the point estimate suggests that about 25 percent of the observed decline in the application-to-population ratio for 18- to 64-year-olds could be explained by increases in the routine-intensive employment share. The point estimate for the automation-intensive employment share is even larger (column 2, Panel C). The coefficient estimate when multiplied by 3 becomes -0.046, which explains 52 percent of the decline in the application-to-population ratio. The other point estimates in Table 3 can be evaluated by multiplying the coefficient estimate by 3 and then dividing by the unconditional change in the application-to-population ratio between 2005 and 2019 for each age group.

In Table 4, we repeat the analysis shown in Table 3 separately for males (odd-numbered columns) and females (even-numbered columns). OLS and 2SLS estimates are presented for each age group (columns 1 and 2, 3 and 4, 5 and 6, and 7 and 8) and employment-share measure (Panels A, B, and C). Again, we find that 2SLS estimates are considerably larger than those produced when using OLS estimation. The estimates for males and females are, in large part, consistent with each other. One deviation is for the 55–64 age group. We find strong evidence of a negative effect of each employment-share variable among females, but find no evidence of a statistical link when focusing on the male application-to-population ratio.

In Table 5, we present stacked-difference estimates. The table layout is identical to that used in Table 3. Using the stacked-difference specifications, we tend to find statistically significant effects using both OLS and 2SLS. The only exception is the 18–34 age group for whom we find limited evidence of a statistically significant link between the automation-related employment shares and the SSDI outcomes. Similar to the results in Table 3, the 2SLS estimates are at least twice as large as those based on OLS estimation. For the most part, the stacked-difference estimates are consistent with those based on the long-differences specifications. Using the standard deviation of the employment-share variables (≈ 3)

Table 6 examines separately the effects of the employment-share variables on males and females using the stacked-differences estimation strategy. Although there are exception, the OLS estimates tend to be statistically insignificant in the majority of cases, but 2SLS estimates are, for the most part, statistically significant. Relative to the long-difference estimates (See Table 4), the stacked-differences estimates suggest that the 35–54 age group is driving the aggregate estimates in columns 1 and 2, as we find little evidence of a statistical relationship for the 18–34 age group and relatively small effects for the 55–64 age group.

5.2 Exposure to Industrial Robots and SSDI Applications

We now examine how a specific technology that tends to be purely labor saving affects SSDI outcomes. In Table 7, we present overall estimates for each age group using the measure proposed by Gihleb et al. (2022). The table is organized into two panels: Panel A presents the long-differences estimates, and Panel B presents the stacked-difference estimates. OLS and 2SLS estimates are presented in each column, and the columns give estimates for males (odd numbered) and females (even numbered) in particular age groups: columns 1 and 2 (18–64), columns 3 and 4 (18–34), columns 5 and 6 (35–54), and columns 7 and 8 (55–64). We note the similarities in the coefficient estimates between the sexes as well as the stability of the estimates when using OLS and 2SLS estimation. In general, the long- and stacked-differences estimates are consistent with each other, but the stacked-differences estimates tend to be smaller than those estimated via the long-differences specifications.

Using a one-unit change for evaluating the coefficient estimates is problematic due to the uncommonness of values of 1 or more in the data. Therefore, we, again, use the variables' standard deviations to assess the effect sizes, which are around 0.33 for the long-differences specifications and 0.72 for the stacked-differences specifications. When we use the values to evaluate the coefficients, for example, in column 1 from Table 7, we find similarly sized effects. For example, multiplying the -0.0355 coefficient (the 2SLS estimate in Panel A) by 0.33 yields a -0.0112 percentage-point change in the application-to-population ratio. Likewise, the product of 0.72 and the -0.0138 coefficient (the 2SLS estimate in Panel B) is a -0.0099 percentage-point change in the SSDI outcome. These estimates imply that–when compared relative to the sample mean for the dependent variable (last row in Table 7), a standard-deviation increase in robots per 1000 workers explains about 12 percent of the decline in the SSDI application-to-population ratio between 2005 and 2019.

We repeat the analysis from Table 7 separately for males and females in Table 8. The estimates for males and females within each age group are similar when estimating via OLS and 2SLS. The OLS and 2SLS estimates, likewise, are similar and statistically different from zero at conventional levels. One discrepancy between the long- and stacked-differences estimates is the statistically significant negative effects shown in Panel A and null effects shown in Panel B.

Lastly, in Table 9, we use the robot-exposure measure advanced by Acemoglu and Restrepo (2020). The table presents estimates for males and females combined (Panel A), males only (Panel B), and females only (Panel C). In columns 1-4, we present the long-differences estimates, and columns 5–8 show the stacked-difference estimates. The vast majority of the estimated coefficients presented in Table 9 are statistically different from zero at conventional levels. We rely on the standard deviations of the robot-exposure measures to interpret the magnitudes of the effects. For the long-differences estimates, the standard deviation of the robot-exposure measure is 1.75, and it is 1.00 for the stacked-differences specifications. Thus, if we multiply the the -0.0058 coefficient in column 1 (2SLS estimate from Panel A) by 1.75, the estimates imply a reduction in the application-to-population -0.01 percentage points. The implied impact is smaller for the stacked-differences specifications, as the point estimate (i.e.,the 2SLS estimate from column 5 from Panel A)–after multiplying by 1–is about half the size of the long-differences estimates. Thus, the robot-exposure measure captures about 12 percent of the observed decline in the overall SSDI application-to-population ratio when using the long-differences estimates and about 6 percent when using the stacked-differences specification.

6 Discussion

While the literature focuses more on the negative effects of automation on lower-skilled-workers labor market outcomes, this paper indicates a robust, negative effect of automation on SSDI claiming. Insofar as SSDI claiming rates proxy for the health of the workforce, our results are congruent with much of the literature on automation and worker health. Although further inquiry into the mechanisms underpinning the negative effect on SSDI applications is needed, it appears that the replacement of more dangerous and injury-prone tasks with machine labor (e.g., see Gihleb et al., 2022) has dominated the negative psychological effects of job destruction (e.g., see O'Brien et al., 2022). More research on different worker populations (e.g., race/ethnicity) is warranted to better identify the relationship between automation-type-technologies and workers' health outcomes.[12]

[12]Further analysis of SSDI claiming rates for racial/ethnic sub-populations may require aggregation to larger geographic areas, as many commuting zone observations by race were censored due to privacy concerns.

7 Conclusion

We study the effects of both broad and narrow measures of automation technology on disability claiming in the United States. We find that the broader measure of automation, which uses employment shares of routine tasks, is negatively related to SSDI claiming. However, the negative relationship is not as robust as we find when we examine the effect of industrial robot penetration on SSDI applications. Automation accounts for about a third of the drop in SSDI application-to-population ratio.

There are several obstacles to establishing a credible causal relationship between automation and SSDI claiming. First, omitted-variable bias is a concern. Automation technology is likely correlated with unobserved variables that also affect SSDI applications. Second, our measures of automation are imperfect, as we follow the literature and project national industry and occupational statistics related to automation onto regional labor markets called commuting zones. Third, the passage of time is required, as the effect of any automation technology requires time to spread through the economy. Lastly, the combination of the first three estimation issues could be compounded by large movements in aggregate economic activity, such as the Great Recession, which is in the middle of our sample period.[13]

We attempt to address each of these concerns. With respect to timing issues, we use long- and stacked-differences models to both allow automation time to diffuse through the labor market and minimize the impact of short-run disruptions in aggregate economic activity on estimates. To identify the causal relationship between automation and SSDI claiming, we implement a shift-share-type estimation strategy (Goldsmith-Pinkham et al., 2020; Borusyak et al., 2022). This strategy is similar to that of prominent studies in the economics automation literature, primarily the work of Autor and Dorn (2013) and Acemoglu and Restrepo (2020). Recent research on technological change, within a task framework of production (e.g., see Acemoglu et al., 2024), shows that automation shifts tasks away from certain groups of workers, which also affects the wage distribution. The confluence of these two labor-market changes makes it more difficult to identify a specific mechanism that would affect SSDI claiming–that is, does automation affect SSDI by replacing dangerous tasks and/or by changing the wage structure? By using two measures of automation, one broad and one specific to industrial robots, we hoped to gain insight into the underlying mechanisms driving the reduction in SSDI claiming. In future research, we plan to investigate the extent to which automation has replaced human labor in more dangerous and/or physical tasks versus the destruction of jobs that might incentivize SSDI applications.

[13] We do not consider the COVID-19 period in our data.

References

Acemoglu, D., F. Kong, and P. Restrepo (2024). Tasks At Work: Comparative Advantage, Technology and Labor Demand. Publisher: National Bureau of Economic Research.

Acemoglu, D., C. Lelarge, and P. Restrepo (2020, May). Competing with Robots: Firm-Level Evidence from France. *AEA Papers and Proceedings 110*, 383–388.

Acemoglu, D. and P. Restrepo (2020, June). Robots and Jobs: Evidence from US Labor Markets. *Journal of Political Economy 128*(6), 2188–2244.

Acemoglu, D. and P. Restrepo (2022a, January). Demographics and Automation. *The Review of Economic Studies 89*(1), 1–44.

Acemoglu, D. and P. Restrepo (2022b). Tasks, automation, and the rise in us wage inequality. *Econometrica 90*(5), 1973–2016. Publisher: Wiley Online Library.

Acemoglu, D. and P. Restrepo (2024). Automation and Rent Dissipation: Implications for Wages, Inequality, and Productivity. Technical report, National Bureau of Economic Research.

Autor, D. H. and D. Dorn (2013). The growth of low-skill service jobs and the polarization of the US labor market. *American economic review 103*(5), 1553–1597. Publisher: American Economic Association.

Autor, D. H., D. Dorn, and G. H. Hanson (2013, October). The China Syndrome: Local Labor Market Effects of Import Competition in the United States. *American Economic Review 103*(6), 2121–2168.

Autor, D. H. and M. G. Duggan (2006). The growth in the social security disability rolls: a fiscal crisis unfolding. *Journal of Economic perspectives 20*(3), 71–96. Publisher: American Economic Association.

Blasco, S., J. Rochut, and B. Rouland (2024, January). Displaced or depressed? Working in automatable jobs and mental health. *Industrial Relations: A Journal of Economy and Society*, irel.12356.

Borusyak, K., P. Hull, and X. Jaravel (2022, January). Quasi-Experimental Shift-Share Research Designs. *The Review of Economic Studies 89*(1), 181–213.

Bratsberg, B., O. Rogeberg, and V. Skirbekk (2022). Technology-induced job loss risk, disability and all-cause mortality in Norway. *Occupational and Environmental Medicine 79*(1), 32–37. Publisher: BMJ Publishing Group Ltd.

Cheng, W., L. Pien, and Y. Cheng (2021, February). Occupation-level automation probability is associated with psychosocial work conditions and workers' health: A multilevel study. *American Journal of Industrial Medicine 64*(2), 108–117.

Dauth, W., S. Findeisen, J. Suedekum, and N. Woessner (2021, December). The Adjustment of Labor Markets to Robots. *Journal of the European Economic Association 19*(6), 3104–3153.

Deming, D. J. (2017). The growing importance of social skills in the labor market. *The Quarterly Journal of Economics 132*(4), 1593–1640. Publisher: Oxford University Press.

Gihleb, R., O. Giuntella, L. Stella, and T. Wang (2022). Industrial robots, workers' safety, and health. *Labour Economics 78*, 102205. Publisher: Elsevier.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2020). Bartik instruments: What, when, why, and how. *American Economic Review 110*(8), 2586–2624. Publisher: American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Gunadi, C. and H. Ryu (2021). Does the rise of robotic technology make people healthier? *Health Economics 30*(9), 2047–2062. Publisher: Wiley Online Library.

Johnson, S. and D. Acemoglu (2023). *Power and Progress: Our Thousand-Year Struggle Over Technology and Prosperity*. Hachette UK.

Kearney, M. S., B. M. Price, and R. Wilson (2021). Disability Insurance in the Great Recession: Ease of Access, Program Enrollment, and Local Hysteresis. Volume 111, pp. 486–490. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Liebman, J. B. (2015). Understanding the increase in disability insurance benefit receipt in the United States. *Journal of Economic Perspectives 29*(2), 123–150. Publisher: American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203-2418.

Lordan, G. and E.-J. Stringer (2022). People versus machines: The impact of being in an automatable job on Australian worker's mental health and life satisfaction. *Economics & Human Biology 46*, 101144. Publisher: Elsevier.

Maestas, N., K. J. Mullen, and A. Strand (2021, July). The effect of economic conditions on the disability insurance program: Evidence from the great recession. *Journal of Public Economics 199*, 104410.

O'Brien, R., E. F. Bair, and A. S. Venkataramani (2022). Death by robots? Automation and working-age mortality in the United States. *Demography 59*(2), 607–628. Publisher: Duke University Press.

Ruggles, S., S. Flood, M. Sobek, D. Backman, A. Chen, G. Cooper, S. Richards, R. Rogers, and M. Schouwiler (2023). IPUMS USA: Version 14.0.

US Social Security Administration (2023). Annual Statistical Report on the Social Security Disability Insurance Program, 2022.

Tables

Table 1: Top 10 Detailed Occupations for the Automation-Related Employment Shares

Ranking	Routine-Intensive (1)	Repetition-Intensive (2)	Automation-Intensive (3)
#1	Bookkeeping, Accounting, and Auditing Clerks	Bookkeeping, Accounting, and Auditing Clerks	Extruding, Forming, Pressing, and Compacting Machine Setters, Operators, and Tenders
#2	Insurance Claims and Policy Processing Clerks	Gambling Cage Workers	Travel Agents
#3	Eligibility Interviewers, Government Programs	Radiation Therapists	Library Technicians
#4	Medical Records Specialists	Reservation and Transportation Ticket Agents and Travel Clerks	Insurance Claims and Policy Processing Clerks
#5	Health Information Technologists and Medical Registrars	Eligibility Interviewers, Government Programs	Accountants and Auditors
#6	Tire Builders	Payroll and Timekeeping Clerks	Budget Analysts
#7	Loan Interviewers and Clerks	Brokerage Clerks	Sawing Machine Setters, Operators, and Tenders, Wood
#8	Budget Analysts	Tax Preparers	Bookkeeping, Accounting, and Auditing Clerks
#9	Atmospheric and Space Scientists	Tellers	Financial and Investment Analysis
#10	Library Technicians	Pharmacists	Financial Risk Specialists

Notes: The table presents the top 10 occupations for the routine-, repetition-, and automation-intensive employment shares generally defined in equation (2).

Table 2: Summary Statistics for SSDI Applications by Age Group

		Standard Deviation		
	Average	Overall	Between	Within
	(1)	(2)	(3)	(4)
Panel A: Males and Females				
18-64 Year Olds	0.6098	0.2089	0.1921	0.0837
18-34 Year Olds	0.1035	0.0424	0.0395	0.0215
35-54 Year Olds	0.2988	0.1227	0.1118	0.0551
55-64 Year Olds	0.1971	0.0594	0.0565	0.0273
Panel B: Males				
18-64 Year Olds	0.3229	0.1158	0.1038	0.0530
18-34 Year Olds	0.0523	0.0233	0.0214	0.0130
35-54 Year Olds	0.1531	0.0663	0.0581	0.0333
55-64 Year Olds	0.1106	0.0343	0.0313	0.0168
Panel C: Females				
18-64 Year Olds	0.3003	0.0992	0.0931	0.0373
18-34 Year Olds	0.0538	0.0212	0.0205	0.0105
35-54 Year Olds	0.1521	0.0591	0.0547	0.0256
55-64 Year Olds	0.0885	0.0275	0.0254	0.0138

Notes: The table presents sample means, standard deviations, and the within and between components of the standard deviations for the four dependent variables. The sample sizes vary in Panel A: 9,755 observations for the 18-64 year olds; 9,846 observations for the 18-34 year olds; 10,685 observations for the 35-54 year olds; and 10,586 observations for 55-64 year olds. The sample sizes vary across the age groups due to the suppression of data from commuting zones with application counts below the threshold for external presentation determined by SSA's Disclosure Review Board.

Table 3: Routine, Repetition, and Automation Intensive Employment Shares and the SSDI Application-to-Population Ratio: OLS and 2SLS Estimates, Long Differences, 2005-2019

	Age Group							
	18-64		18-34		35-54		55-64	
	OLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Routine-Intensive Employment Share							
$RSH_{c,2005}$	-0.0018	-0.0069^{***}	0.0004	0.0008	-0.0021^{**}	-0.0055^{***}	-0.0001	-0.0020^{***}
	(0.0013)	(0.0026)	(0.0004)	(0.0007)	(0.0009)	(0.0018)	(0.0004)	(0.0008)
First Stage Coeff.	–	0.6016^{***}	–	0.6016^{***}	–	0.6062^{***}	–	0.6069^{***}
KP F-Statistic	–	65.1875	–	65.2252	–	67.0934	–	66.9569
	Panel B: Repetition-Intensive Employment Share							
$RTSH_{c,2005}$	-0.0009	-0.0077^{***}	0.0006	0.0009	-0.0020^{**}	-0.0061^{***}	0.0005	-0.0023^{***}
	(0.0012)	(0.0029)	(0.0004)	(0.0008)	(0.0009)	(0.0020)	(0.0004)	(0.0008)
First Stage Coeff.	–	0.5432^{***}	–	0.5432^{***}	–	0.5459^{***}	–	0.5464^{***}
KP F-Statistic	–	76.3211	–	76.3632	–	79.6606	–	79.3729
	Panel C: Automation-Intensive Employment Share							
$ASH_{c,2005}$	-0.0017	-0.0142^{***}	0.0002	0.0016	-0.0012	-0.0116^{***}	-0.0006	-0.0044^{***}
	(0.0015)	(0.0052)	(0.0004)	(0.0016)	(0.0009)	(0.0039)	(0.0005)	(0.0015)
First Stage Coeff.	–	0.2938^{***}	–	0.2938^{***}	–	0.2853^{***}	–	0.2847^{***}
KP F-Statistic	–	45.4623	–	45.4796	–	45.7129	–	44.8765
N	623	623	626	626	687	687	679	679
$\overline{Apps^{*}_{c,(2005,2019)}}$	-0.0826	-0.0826	-0.0038	-0.0038	-0.0931	-0.0931	0.0151	0.0151

Notes: The table presents OLS and 2SLS long-differences estimates based on equation (7). For each age group, we present the OLS and 2SLS estimates side-by-side separately for 18-64, 18-34, 35-54, and 55-64 year-olds. For the specifications estimated via 2SLS, we report the coefficient on the instrument (defined in equation 9) and the Kleibergen-Paap Wald *rk* F-statistic. The sample sizes, N, vary across the age groups due to the suppression of data from commuting zones with application counts below the threshold for external presentation determined by SSA's Disclosure Review Board. Each specification includes the control variables held constant in Autor and Dorn (2013)'s study, which include the college-to-noncollege population, the ratio of immigrants to the noncollege population, the share of workers employed in manufacturing, the unemployment rate, the share of the population who is female and employed, the share of the population 65 or older, and the share of noncollege workers workers earning a real wage below the minimum wage that will prevail over the next decade (2000-2010). Each of these variables is measured in 2000. We also hold constant the "China Shock" from (Autor et al., 2013), which is measured as the change in Chinese import exposure between 1990 and 2007 and the SSDI processing efficiency measure from Kearney et al. (2021) for the year 2003. We report standard errors clustered at the state level in parentheses. *, **, and *** indicate statistical significance at the 10, 5 and 1 percent levels, respectively.

Table 4: Routine, Repetition, and Automation Intensive Employment Shares and Per-capita SSDI Applications: Males and Females Separately, OLS and 2SLS Estimates, Long Differences, 2005-2019

	Age Group							
	18-64		18-34		35-54		55-64	
	Male	Female	Male	Female	Male	Female	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Routine-Intensive Employment Share							
OLS	−0.0010	−0.0007	0.0003*	0.0001	−0.0013**	−0.0008**	−0.0001	−0.0001
	(0.0008)	(0.0007)	(0.0002)	(0.0002)	(0.0006)	(0.0004)	(0.0002)	(0.0003)
2SLS	−0.0030**	−0.0039***	0.0005	0.0002	−0.0029***	−0.0026***	−0.0005	−0.0016***
	(0.0014)	(0.0015)	(0.0004)	(0.0005)	(0.0010)	(0.0009)	(0.0004)	(0.0005)
First Stage Coeff.	0.6015***	0.6015***	0.6016***	0.6016***	0.6030***	0.6030***	0.6014***	0.6014***
KP F-Statistic	63.3168	63.3168	63.3654	63.3654	65.5689	65.5689	65.5885	65.5885
	Panel B: Repetition-Intensive Employment Share							
OLS	−0.0006	−0.0002	0.0004**	0.0002	−0.0013**	−0.0007*	0.0003	0.0002
	(0.0008)	(0.0007)	(0.0002)	(0.0003)	(0.0006)	(0.0004)	(0.0002)	(0.0002)
2SLS	−0.0033**	−0.0043***	0.0005	0.0003	−0.0033***	−0.0029***	−0.0006	−0.0017***
	(0.0016)	(0.0016)	(0.0004)	(0.0005)	(0.0011)	(0.0010)	(0.0004)	(0.0006)
First Stage Coeff.	0.5417***	0.5417***	0.5418***	0.5418***	0.5442***	0.5442***	0.5431***	0.5431***
KP F-Statistic	73.1754	73.1754	73.2304	73.2304	77.2897	77.2897	77.1331	77.1331
	Panel C: Automation-Intensive Employment Share							
OLS	−0.0003	−0.0014*	0.0002	−0.0001	−0.0005	−0.0007	0.0000	−0.0006**
	(0.0008)	(0.0008)	(0.0002)	(0.0002)	(0.0005)	(0.0005)	(0.0003)	(0.0003)
2SLS	−0.0060**	−0.0078***	0.0010	0.0005	−0.0061***	−0.0054***	−0.0011	−0.0032***
	(0.0029)	(0.0026)	(0.0009)	(0.0009)	(0.0023)	(0.0017)	(0.0008)	(0.0009)
First Stage Coeff.	0.3014***	0.3014***	0.3014***	0.3014***	0.2903***	0.2903***	0.2918***	0.2918***
KP F-Statistic	45.6387	45.6387	45.6574	45.6574	45.5121	45.5121	46.9524	46.9524
N	559	559	562	562	639	639	621	621
$\overline{Apps^{*}_{s,c,(2005,2019)}}$	−0.0664	−0.0029	−0.0593	−0.0012	−0.0207	−0.0016	−0.0353	0.0175

Notes: The table presents OLS and 2SLS long-differences estimates based on equation (7) separately for males (odd-numbered columns) and females (even-numbered columns). The sample sizes, N, vary across the age groups due to the suppression of data from commuting zones with application counts below the threshold for external presentation determined by SSA's Disclosure Review Board. Each specification includes the control variables held constant in Autor and Dorn (2013)'s study, which include the college-to-noncollege population, the ratio of immigrants to the noncollege population, the share of workers employed in manufacturing, the unemployment rate, the share of the population who is female and employed, the share of the population 65 or older, and the share of noncollege workers workers earning a real wage below the minimum wage that will prevail over the next decade (2000-2010). Each of these variables is measured in 2000. We also hold constant the "China Shock" from (Autor et al., 2013), which is measured as the change in Chinese import exposure between 1990 and 2007 and the SSDI processing efficiency measure from Kearney et al. (2021) for the year 2003. We report standard errors clustered at the state level in parentheses. *, **, and *** indicate statistical significance at the 10, 5 and 1 percent levels, respectively.

Table 5: Routine, Repetition, and Automation Intensive Employment Shares and Per-capita SSDI Applications: OLS and 2SLS Estimates, Stacked Differences, 2005-2010 and 2010-2019

	Age Group							
	18-64		18-34		35-54		55-64	
	OLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Routine-Intensive Employment Share								
$RSH_{c,2005}$	-0.0035^{**}	-0.0083^{*}	-0.0002	-0.0011	-0.0026^{**}	-0.0055^{**}	-0.0008^{**}	-0.0017^{**}
	(0.0016)	(0.0044)	(0.0004)	(0.0010)	(0.0010)	(0.0028)	(0.0004)	(0.0008)
First Stage Estimates								
Main Effect Coeff.	–	0.5804^{***}	–	0.5804^{***}	–	0.5845^{***}	–	0.5840^{***}
Coeff. on Interaction Term	–	-0.1124^{**}	–	-0.1124^{**}	–	-0.1094^{**}	–	-0.1093^{**}
KP F-Statistic	–	21.9941	–	21.9968	–	22.5637	–	22.4685
Panel B: Repetition-Intensive Employment Share								
$RTSH_{c,2005}$	-0.0041^{**}	-0.0099^{*}	-0.0004	-0.0021^{*}	-0.0030^{***}	-0.0062^{*}	-0.0008^{**}	-0.0016^{*}
	(0.0016)	(0.0052)	(0.0004)	(0.0013)	(0.0010)	(0.0032)	(0.0004)	(0.0009)
First Stage Estimates								
Main Effect Coeff.	–	0.5424^{***}	–	0.5424^{***}	–	0.5453^{***}	–	0.5449^{***}
Coeff. on Interaction Term	–	-0.1812^{***}	–	-0.1812^{***}	–	-0.1770^{***}	–	-0.1771^{***}
KP F-Statistic	–	30.5157	–	30.5187	–	31.6373	–	31.4715
Panel C: Automation-Intensive Employment Share								
$ASH_{c,2005}$	0.0003	-0.0153^{**}	0.0004	0.0003	0.0007	-0.0120^{**}	-0.0008^{*}	-0.0043^{***}
	(0.0015)	(0.0076)	(0.0004)	(0.0017)	(0.0009)	(0.0053)	(0.0005)	(0.0015)
First Stage Estimates								
Main Effect Coeff.	–	0.2343^{***}	–	0.2343^{***}	–	0.2295^{***}	–	0.2304^{***}
Coeff. on Interaction Term	–	0.0181	–	0.0181	–	0.0174	–	0.0171
KP F-Statistic	–	15.1533	–	15.1554	–	14.8478	–	14.9314
N	1,266	1,266	1,268	1,268	1,388	1,388	1,361	1,361
$\overline{Apps^{*}_{c,(2005,2019)}}$	-0.0826	-0.0826	-0.0038	-0.0038	-0.0931	-0.0931	0.0151	0.0151

Notes: The table presents OLS and 2SLS stacked-differences estimates based on equation (8). For each age grouping, we present the OLS and 2SLS estimates side-by-side separately for 18-64, 18-34, 35-54, and 55-64 year-olds. For the specifications estimated via 2SLS, we report the coefficient on the instrument (defined in equation 9) and the Kleibergen-Paap Wald rk F-statistic. The sample sizes, N, vary across the age groups due to the suppression of data from commuting zones with application counts below the threshold for external presentation determined by SSA's Disclosure Review Board. Each specification includes the control variables held constant in Autor and Dorn (2013)'s study, which include the college-to-noncollege population, the ratio of immigrants to the noncollege population, the share of workers employed in manufacturing, the unemployment rate, the share of the population who is female and employed, the share of the population 65 or older, and the share of noncollege workers workers earning a real wage below the minimum wage that will prevail over the next decade. We link the 1990 version of each of these variables to the starting year of the first difference in the stack (i.e. 2005), and the 2000 version of these variables is linked to the starting period of the second difference in the stack (i.e. 2010). For the share of workers earning below the minimum wage that will prevail in the decade, the relevant decade for the starting period of the first difference is 1990-2000 and the decade relevant to the starting period of the second difference is 2000-2010. We hold constant the "China Shock" from (Autor et al., 2013), which is measured as the change in Chinese import exposure between two points in time. We link the change between 1990 and 2000 to the starting period of the first difference and the change between 2000 and 2007 to the starting year of the second difference. We report standard errors clustered at the state level in parentheses. Lastly, we hold constant the SSDI processing efficiency measure. The 2003 value is linked to the first difference in the stack, and the 2008 value is linked to the second difference in the stack. *, **, and *** indicate statistical significance at the 10, 5 and 1 percent levels, respectively.

Table 6: Routine, Repetition, and Automation Intensive Employment Shares and Per-capita SSDI Applications: Males and Females Separately, OLS and 2SLS Estimates, Stacked Differences, 2005-2010 and 2010-2019

	Age Group							
	18-64		18-34		35-54		55-64	
	Male	Female	Male	Female	Male	Female	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Routine-Intensive Employment Share								
OLS	−0.0025**	−0.0010	−0.0002	−0.0000	−0.0017***	−0.0009**	−0.0006**	−0.0002
	(0.0010)	(0.0007)	(0.0002)	(0.0002)	(0.0006)	(0.0004)	(0.0003)	(0.0002)
2SLS	−0.0051*	−0.0032*	−0.0008	−0.0004	−0.0032**	−0.0023*	−0.0011**	−0.0006*
	(0.0026)	(0.0019)	(0.0006)	(0.0005)	(0.0016)	(0.0012)	(0.0005)	(0.0003)
First Stage Estimates								
Main Effect Coeff.	0.5765***	0.5765***	0.5765***	0.5765***	0.5817***	0.5817***	0.5795***	0.5795***
Coeff. on Interaction Term	−0.1125**	−0.1125**	−0.1125**	−0.1125**	−0.1116**	−0.1116**	−0.1100**	−0.1100**
KP F-Statistic	21.1402	21.1402	21.1437	21.1437	22.1461	22.1461	22.0142	22.0142
Panel B: Repetition-Intensive Employment Share								
OLS	−0.0031***	−0.0009	−0.0003	−0.0000	−0.0020***	−0.0009**	−0.0008***	−0.0000
	(0.0009)	(0.0007)	(0.0002)	(0.0002)	(0.0006)	(0.0004)	(0.0002)	(0.0002)
2SLS	−0.0069**	−0.0033	−0.0016**	−0.0007	−0.0037**	−0.0024*	−0.0014**	−0.0002
	(0.0032)	(0.0022)	(0.0008)	(0.0006)	(0.0018)	(0.0014)	(0.0007)	(0.0004)
First Stage Estimates								
Main Effect Coeff.	0.5384***	0.5384***	0.5384***	0.5384***	0.5433***	0.5433***	0.5417***	0.5417***
Coeff. on Interaction Term	−0.1818***	−0.1818***	−0.1818***	−0.1818***	−0.1799***	−0.1799***	−0.1788***	−0.1788***
KP F-Statistic	29.2231	29.2231	29.2270	29.2270	30.8400	30.8400	30.7101	30.7101
Panel C: Automation-Intensive Employment Share								
OLS	0.0002	0.0001	0.0002	0.0002	0.0003	0.0003	−0.0003	−0.0005**
	(0.0010)	(0.0006)	(0.0002)	(0.0002)	(0.0006)	(0.0004)	(0.0003)	(0.0002)
2SLS	−0.0070	−0.0074**	0.0004	−0.0001	−0.0063**	−0.0053**	−0.0017*	−0.0025***
	(0.0044)	(0.0032)	(0.0009)	(0.0008)	(0.0031)	(0.0022)	(0.0010)	(0.0007)
First Stage Estimates								
Main Effect Coeff.	0.2420***	0.2420***	0.2420***	0.2420***	0.2312***	0.2312***	0.2343***	0.2343***
Coeff. on Interaction Term	0.0167	0.0167	0.0167	0.0167	0.0189	0.0189	0.0164	0.0164
KP F-Statistic	14.8690	14.8690	14.8708	14.8708	14.9907	14.9907	15.1521	15.1521
N	1,137	1,137	1,139	1,139	1,301	1,301	1,256	1,256
$\overline{Apps^*_{s,c,(2005,2019)}}$	−0.0664	−0.0029	−0.0593	−0.0012	−0.0207	−0.0016	−0.0353	0.0175

Notes: The table presents OLS and 2SLS stacked-differences estimates based on equation (8) separately for males (odd-numbered columns) and females (even-numbered columns). For the 2SLS estimates, we report the coefficient on the instrument (defined in equation 9), its interaction with the difference-period indicator (see equation (7), and the Kleibergen-Paap Wald rk F-statistic. The sample sizes, N, vary across the age groups due to the suppression of data from commuting zones with application counts below the threshold for external presentation determined by SSA's Disclosure Review Board. Each specification includes the control variables held constant in Autor and Dorn (2013)'s study, which include the college-to-noncollege population, the ratio of immigrants to the noncollege population, the share of workers employed in manufacturing, the unemployment rate, the share of the population who is female and employed, the share of the population 65 or older, and the share of noncollege workers workers earning a real wage below the minimum wage that will prevail over the next decade. We link the 1990 version of each of these variables to the starting year of the first difference in the stack (i.e. 2005), and the 2000 version of these variables is linked to the starting period of the second difference in the stack (i.e. 2010). For the share of workers earning below the minimum wage that will prevail in the decade, the relevant decade for the starting period of the first difference is 1990-2000 and the decade relevant to the starting period of the second difference is 2000-2010. We hold constant the "China Shock" from (Autor et al., 2013), which is measured as the change in Chinese import exposure between two points in time. We link the change between 1990 and 2000 to the starting period of the first difference and the change between 2000 and 2007 to the starting year of the second difference. We report standard errors clustered at the state level in parentheses. Lastly, we hold constant the SSDI processing efficiency measure. The 2003 value is linked to the first difference in the stack, and the 2008 value is linked to the second difference in the stack. *, **, and *** indicate statistical significance at the 10, 5 and 1 percent levels, respectively.

Table 7: Exposure to Industrial Robots and Per-capita SSDI Applications: OLS and 2SLS Estimates, Long Differences (2005-2019) and Stacked Differences (2005-2010, 2010-2019)

	Age Group			
	18-64	18-34	35-54	55-64
	(1)	(2)	(3)	(4)
	Panel A: Long Differences Estimates			
OLS	-0.0389^{***}	-0.0085^{***}	-0.0215^{***}	-0.0089^{**}
	(0.0080)	(0.0016)	(0.0040)	(0.0038)
2SLS	-0.0355^{***}	-0.0066^{***}	-0.0224^{***}	-0.0065
	(0.0092)	(0.0015)	(0.0038)	(0.0051)
First Stage Estimates				
First Stage Coeff.	0.1776^{***}	0.1776^{***}	0.1775^{***}	0.1775^{***}
KP F-Statistic	$1,247.8025$	$1,247.7057$	$1,236.5783$	$1,237.8741$
N	623	628	689	681
	Panel B: Stacked Differences Estimates			
OLS	-0.0169^{***}	-0.0005	-0.0122^{***}	-0.0042^{**}
	(0.0034)	(0.0010)	(0.0020)	(0.0019)
2SLS	-0.0138^{***}	0.0010	-0.0109^{***}	-0.0039
	(0.0050)	(0.0007)	(0.0026)	(0.0025)
First Stage Estimates				
Main Effect	0.1626^{***}	0.1626^{***}	0.1625^{***}	0.1626^{***}
Coeff. on Interaction	0.2850^{***}	0.2850^{***}	0.2850^{***}	0.2850^{***}
KP F-Statistic	$4,626.7748$	$4,626.7956$	$4,617.6992$	$4,629.3804$
N	1,266	1,268	1,388	1,361
$\overline{Apps^{*}_{c,(2005,2019)}}$	-0.0826	-0.0038	-0.0931	0.0151

Notes: The table presents OLS and 2SLS long-differences estimates based on equation (10) in Panel A, and OLS and 2SLS stacked-differences estimates based on equation (8 are presented in Panel B. For each age grouping, we present the OLS and 2SLS estimates side-by-side separately for 18-64, 18-34, 35-54, and 55-64 year-olds. For the specifications estimates via 2SLS, we report the coefficient on the instrument (defined in equation (11)) and the Kleibergen-Paap Wald rk F-statistic. The sample sizes, N, vary across the age groups due to the suppression of data from commuting zones with application counts below the threshold for external presentation determined by SSA's Disclosure Review Board. Each specification includes the control variables held constant in Acemoglu and Restrepo (2020)'s study, which includes US Census division dummies (i.e. ϕ_d), the natural logarithm of the population, the "China Shock" from Autor et al. (2013), the shares of the population who are female, Asian, Black, Hispanic, White, over age 65, did not go to college, completed some college, graduated with a college degree or professional degree, completed a masters or doctorate degree, and employed in manufacturing in general as well as light manufacturing, and the share of females employed in manufacturing relative to total manufacturing employment. When estimating the long-differences specifications, the 1990 version of these variables is used. By contrast, when estimating the stacked-differences specifications, we link the 1990 and 2000 versions of these variables to the first and second periods in the stack, respectively. We also hold constant the "China Shock" from (Autor et al., 2013) and the SSDI processing efficiency variable from Kearney et al. (2021). See the notes from Tables 3 and 5 for details on the Chinese import penetration and SSDI processing efficiency controls. We account for these variables in the same way that we do when estimating the models focused on the employment-share automation measures. We report standard errors clustered at the state level in parentheses. *, **, and *** indicate statistical significance at the 10, 5 and 1 percent levels, respectively.

Table 8: Exposure to Industrial Robots and Per-capita SSDI Applications:
Males and Females Separately, OLS and 2SLS Estimates, Long Differences (2005-2019) and Stacked Differences (2005-2010, 2010-2019)

	Age Group							
	18-64		18-34		35-54		55-64	
	Male	Female	Male	Female	Male	Female	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Long Differences Estimates							
OLS	-0.0213^{***}	-0.0180^{***}	-0.0044^{***}	-0.0044^{***}	-0.0118^{***}	-0.0098^{***}	-0.0049^{*}	-0.0040^{***}
	(0.0059)	(0.0032)	(0.0010)	(0.0009)	(0.0027)	(0.0018)	(0.0028)	(0.0013)
2SLS	-0.0213^{***}	-0.0146^{***}	-0.0039^{***}	-0.0029^{***}	-0.0134^{***}	-0.0093^{***}	-0.0037	-0.0027
	(0.0063)	(0.0035)	(0.0011)	(0.0007)	(0.0023)	(0.0019)	(0.0036)	(0.0017)
First Stage Estimates								
First Stage Coeff.	0.1776^{***}	0.1776^{***}	0.1776^{***}	0.1776^{***}	0.1776^{***}	0.1776^{***}	0.1776^{***}	0.1776^{***}
KP F-Statistic	1,243.1879	1,243.1879	1,243.1220	1,243.1220	1,245.8722	1,245.8722	1,244.1930	1,244.1930
N	561	561	564	564	641	641	623	623
	Panel B: Stacked Differences Estimates							
OLS	-0.0085^{***}	-0.0082^{***}	0.0001	-0.0006	-0.0065^{***}	-0.0057^{***}	-0.0023^{*}	-0.0019^{**}
	(0.0023)	(0.0013)	(0.0005)	(0.0006)	(0.0012)	(0.0008)	(0.0012)	(0.0008)
2SLS	-0.0068^{**}	-0.0067^{***}	0.0009^{*}	0.0002	-0.0059^{***}	-0.0050^{***}	-0.0020	-0.0019^{*}
	(0.0033)	(0.0019)	(0.0005)	(0.0004)	(0.0016)	(0.0011)	(0.0016)	(0.0010)
First Stage Estimates								
Main Effect	0.1625^{***}	0.1625^{***}	0.1625^{***}	0.1625^{***}	0.1626^{***}	0.1626^{***}	0.1626^{***}	0.1626^{***}
Coeff. on Interaction	0.2851^{***}	0.2851^{***}	0.2851^{***}	0.2851^{***}	0.2850^{***}	0.2850^{***}	0.2850^{***}	0.2850^{***}
KP F-Statistic	4,634.0076	4,634.0076	4,634.0831	4,634.0831	4,633.5938	4,633.5938	4,624.9458	4,624.9458
N	1,137	1,137	1,139	1,139	1,301	1,301	1,256	1,256
$\overline{Apps^{*}_{s,c,(2005,2019)}}$	-0.0664	-0.0029	-0.0593	-0.0012	-0.0207	-0.0016	-0.0353	0.0175

Notes: Separately for males (odd-numbered columns) and females (even-numbered columns), the table presents OLS and 2SLS long-differences estimates based on equation (10) in Panel A, and OLS and 2SLS stacked-differences estimates described in Section 4.3 in Panel B. For each age grouping, we present the OLS and 2SLS estimates side-by-side separately for 18-64, 18-34, 35-54, and 55-64 year-olds. For the specifications estimated via 2SLS, we report the coefficient on the instrument (defined in equation (11) and the Kleibergen-Paap Wald rk F-statistic in Panel A and also report the coefficient on the interaction effect between the instrument and the difference-period indicator variable. he sample sizes, N, vary across the age groups due to the suppression of data from commuting zones with application counts below the threshold for external presentation determined by SSA's Disclosure Review Board. Each specification includes the control variables held constant in Acemoglu and Restrepo (2020)'s study, which includes US Census division dummies (i.e. ϕ_d), the natural logarithm of the population, the "China Shock" from Autor et al. (2013), the shares of the population who are female, Asian, Black, Hispanic, White, over age 65, did not go to college, completed some college, graduated with a college degree or professional degree, completed a masters or doctorate degree, and employed in manufacturing in general as well as light manufacturing, and the share of females employed in manufacturing relative to total manufacturing employment. When estimating the long-differences specifications, the 1990 version of these variables is used. By contrast, when estimating the stacked-differences specifications, we link the 1990 and 2000 versions of these variables to the first and second periods in the stack, respectively. We also hold constant the "China Shock" from (Autor et al., 2013) and the SSDI processing efficiency variable from Kearney et al. (2021). See the notes from Tables 3 and 5 for details on the Chinese import penetration and SSDI processing efficiency controls. We account for these variables in the same way that we do when estimating the models focused on the employment-share automation measures. We report standard errors clustered at the state level in parentheses. *, **, and *** indicate statistical significance at the 10, 5 and 1 percent levels, respectively.

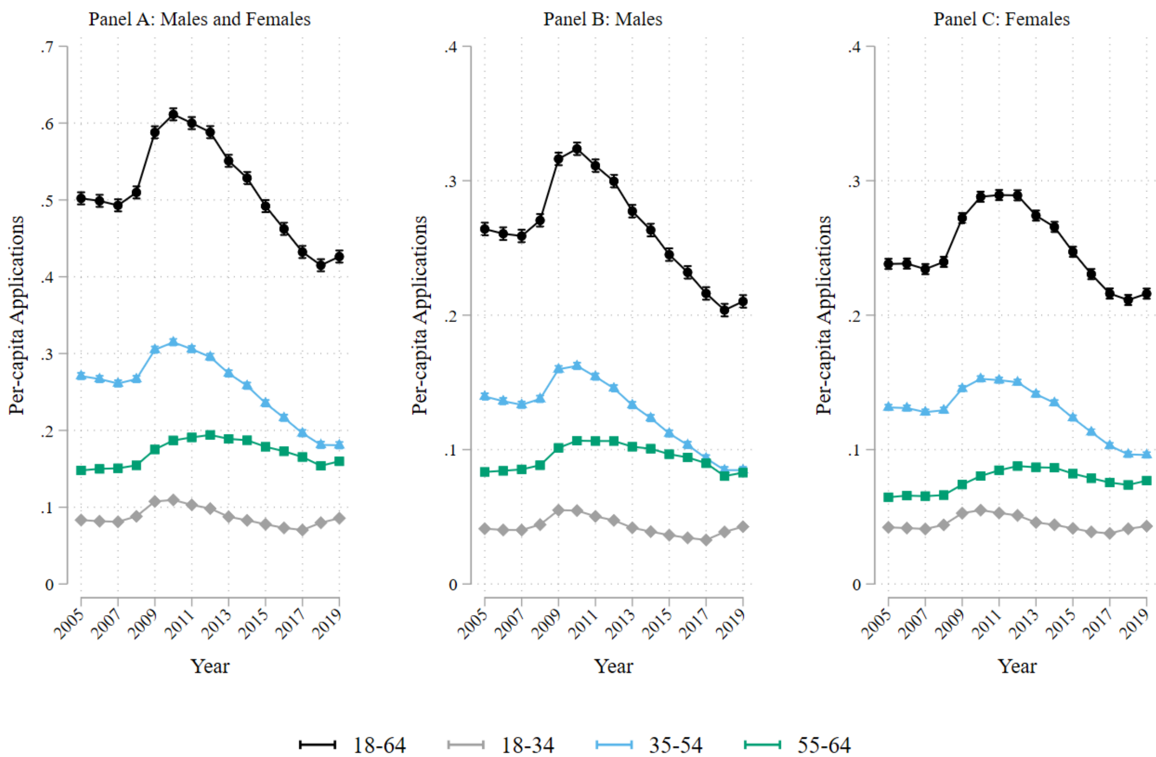
Table 9: Exposure to Industrial Robots and Per-capita SSDI Applications, OLS and 2SLS Estimates, Long Differences (2004-2016) and Stacked Differences (2004-2007, 2013-2016)

	Long Differences, 2004-2016				Stacked Differences, 2004-2007 and 2013-2016			
	18-64	18-34	35-54	55-64	18-64	18-34	35-54	55-64
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Males and Females								
OLS	-0.0051^{***}	-0.0015^{***}	-0.0020^{**}	-0.0016^{**}	-0.0045^{***}	-0.0008^{***}	-0.0023^{***}	-0.0013^{*}
	(0.0017)	(0.0005)	(0.0009)	(0.0006)	(0.0013)	(0.0003)	(0.0007)	(0.0007)
2SLS	-0.0058^{***}	-0.0014^{***}	-0.0027^{***}	-0.0016^{***}	-0.0046^{***}	-0.0006^{**}	-0.0022^{***}	-0.0017^{***}
	(0.0016)	(0.0005)	(0.0008)	(0.0006)	(0.0015)	(0.0003)	(0.0008)	(0.0006)
First Stage Coeff.	1.3422^{***}	1.3422^{***}	1.3431^{***}	1.3429^{***}	0.6476^{***}	0.6476^{***}	0.6481^{***}	0.6480^{***}
Coeff. on Interaction	–	–	–	–	0.0974^{***}	0.0974^{***}	0.0974^{***}	0.0974^{***}
KP F-Statistic	2, 293.2306	2, 293.7997	2, 329.0867	2, 323.7103	4, 271.1866	4, 271.8920	4, 351.5636	4, 318.4567
N	625	627	692	685	1,246	1,248	1,381	1,362
Panel B: Males								
OLS	-0.0030^{***}	-0.0010^{***}	-0.0011^{*}	-0.0008^{*}	-0.0014	-0.0003^{*}	-0.0012^{**}	0.0000
	(0.0011)	(0.0002)	(0.0006)	(0.0004)	(0.0010)	(0.0002)	(0.0005)	(0.0005)
2SLS	-0.0037^{***}	-0.0011^{***}	-0.0016^{***}	-0.0009^{**}	-0.0016	-0.0003^{*}	-0.0012^{**}	-0.0002
	(0.0009)	(0.0002)	(0.0005)	(0.0004)	(0.0010)	(0.0001)	(0.0005)	(0.0005)
First Stage Coeff.	1.3408^{***}	1.3408^{***}	1.3427^{***}	1.3423^{***}	0.6472^{***}	0.6472^{***}	0.6479^{***}	0.6477^{***}
Coeff. on Interaction	–	–	–	–	0.0967^{***}	0.0967^{***}	0.0975^{***}	0.0973^{***}
KP F-Statistic	2, 240.2580	2, 240.9812	2, 306.9613	2, 295.4873	4, 217.8029	4, 218.6427	4, 292.4579	4, 258.9084
N	557	559	644	630	1,119	1,121	1,293	1,254
Panel C: Females								
OLS	-0.0022^{***}	-0.0005^{*}	-0.0009^{**}	-0.0008^{***}	-0.0029^{***}	-0.0005^{**}	-0.0011^{***}	-0.0013^{***}
	(0.0007)	(0.0003)	(0.0004)	(0.0002)	(0.0005)	(0.0002)	(0.0003)	(0.0003)
2SLS	-0.0022^{***}	-0.0003	-0.0011^{***}	-0.0008^{***}	-0.0029^{***}	-0.0003^{*}	-0.0010^{**}	-0.0016^{***}
	(0.0007)	(0.0003)	(0.0004)	(0.0002)	(0.0006)	(0.0002)	(0.0004)	(0.0002)
First Stage Coeff.	1.3408^{***}	1.3408^{***}	1.3427^{***}	1.3423^{***}	0.6472^{***}	0.6472^{***}	0.6479^{***}	0.6477^{***}
Coeff. on Interaction	–	–	–	–	0.0967^{***}	0.0967^{***}	0.0975^{***}	0.0973^{***}
KP F-Statistic	2, 240.2580	2, 240.9812	2, 306.9613	2, 295.4873	4, 217.8029	4, 218.6427	4, 292.4579	4, 258.9084
N	557	559	644	630	1,119	1,121	1,293	1,254

Notes: The table presents OLS and 2SLS long-differences estimates based on equation (10) in columns 1-4, and OLS and 2SLS stacked-differences estimates described in Section 4.3 are presented in columns 5-8. Estimates are presented for males and females combined in Panel A; males only in Panel B; and females only in Panel C. For the specifications estimated via 2SLS, we report the coefficient on the instrument (defined in equation (12) and the Kleibergen-Paap Wald rk F-statistic in Panel A and also report the coefficient on the interaction effect between the instrument and the difference-period indicator variable. he sample sizes, N, vary across the age groups due to the suppression of data from commuting zones with application counts below the threshold for external presentation determined by SSA's Disclosure Review Board. Each specification includes the control variables held constant in Acemoglu and Restrepo (2020)'s study, which includes US Census division dummies (i.e. ϕ_d), the natural logarithm of the population, the "China Shock" from Autor et al. (2013), the shares of the population who are female, Asian, Black, Hispanic, White, over age 65, did not go to college, completed some college, graduated with a college degree or professional degree, completed a masters or doctorate degree, and employed in manufacturing in general as well as light manufacturing, and the share of females employed in manufacturing relative to total manufacturing employment. When estimating the long-differences specifications, the 1990 version of these variables is used. By contrast, when estimating the stacked-differences specifications, we link the 1990 and 2000 versions of these variables to the first and second periods in the stack, respectively. We also hold constant the "China Shock" from (Autor et al., 2013) and the SSDI processing efficiency variable from Kearney et al. (2021). See the notes from Tables 3 and 5 for details on the Chinese import penetration and SSDI processing efficiency controls. We account for these variables in the same way that we do when estimating the models focused on the employment-share automation measures. We report standard errors clustered at the state level in parentheses. *, **, and *** indicate statistical significance at the 10, 5 and 1 percent levels, respectively.

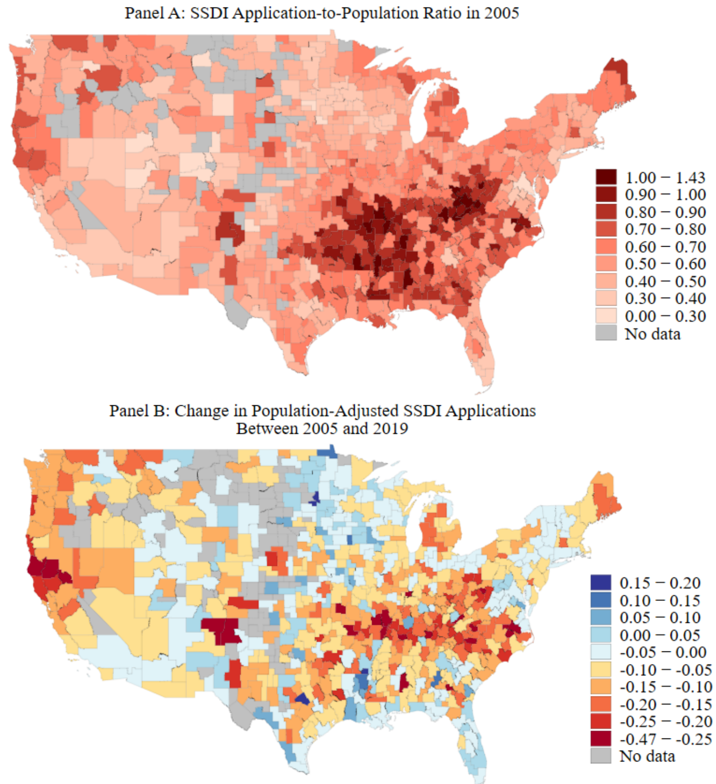
Figures

Figure 1: SSDI Application-to-Population Ratio, 2005-2019



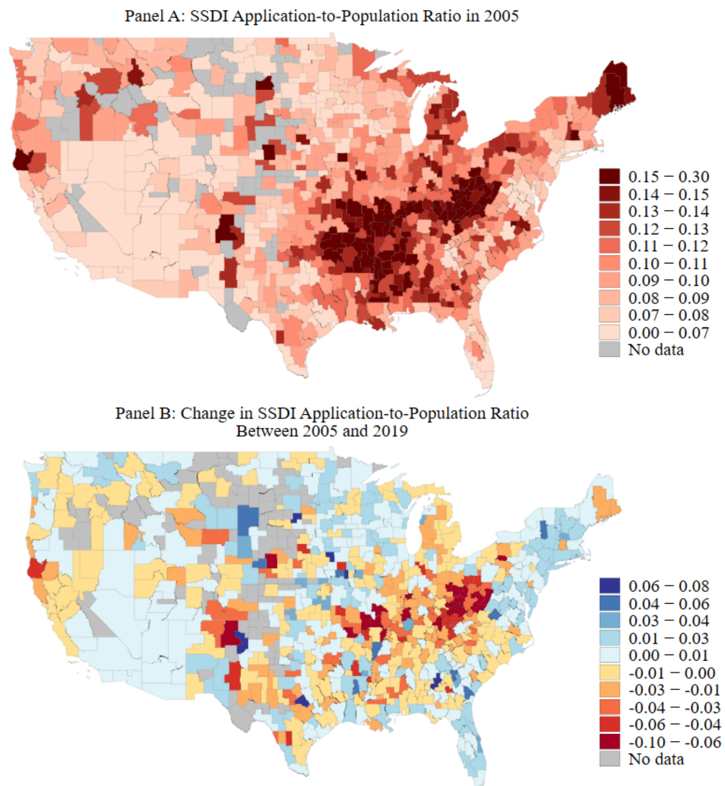
Notes: For the 2005-2019 period, the figure plots predicted values from a regression of SSDI application-to-population ratio, as defined in equation 1, in the commuting zone on a constant, year dummies, and state dummies. The regression estimates are weighted by the commuting zone's share of the US population in 2000. The year-by-year estimates are presented separately in Panels A (males and females combined), B (males only), and C (females only) for the 18–64, 18–34, 35–54, and 55–64 age groups.

Figure 2: Geographic Variation in Population-Adjusted SSDI Applications, 18–64 Year Olds



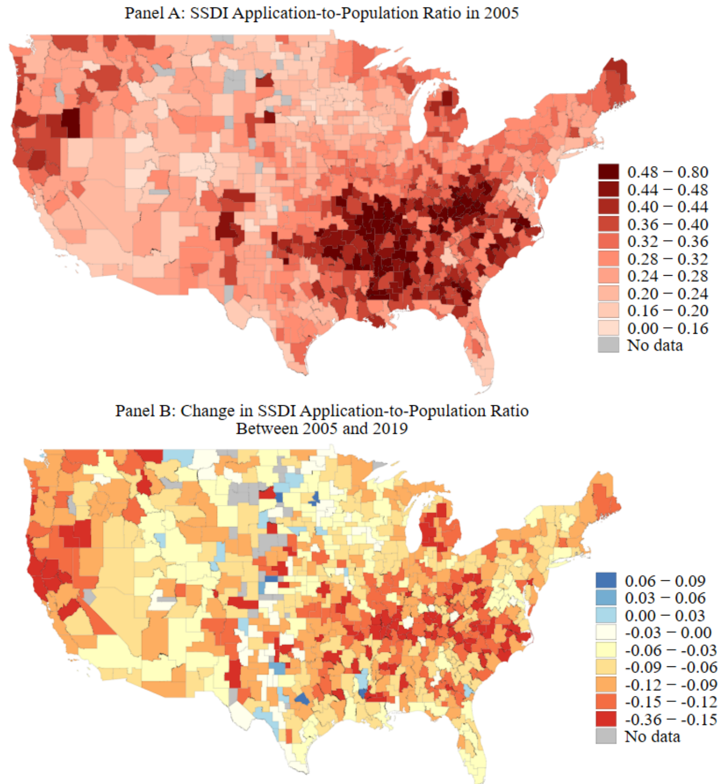
Notes: For 18–64 year-olds, the figure presents heat maps of the SSDI application-to-population ratio, as defined in equation 1, in 2005 (Panel A) as well as its change between 2005 and 2019 (Panel B). In Panel A, lighter red colors indicate a lower application-to-population ratios, and darker red colors indicate a greater prevalence of per-capita applications. Commuting zones for which insufficient applications were submitted to allow for external presentation are shown in gray. In Panel B, we use a diverging color scheme, in which light to dark blue indicate positive changes in application rates, and yellow to red indicate negative changes in application rates. The positive changes are greater when the color is a darker blue. Likewise, negative changes are greater when the color is a darker red.

Figure 3: Geographic Variation in Population-Adjusted SSDI Applications, 18–34 Year Olds



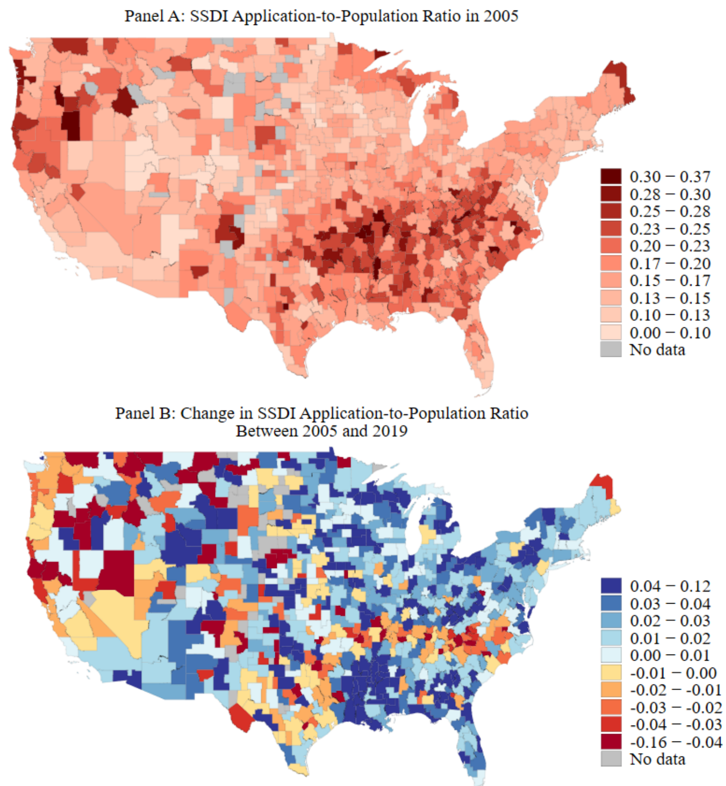
Notes: For 18–34 year-olds, the figure presents heat maps of the SSDI application-to-population ratio, as defined in equation 1, in 2005 (Panel A) as well as its change between 2005 and 2019 (Panel B). In Panel A, lighter red colors indicate a lower application-to-population ratios, and darker red colors indicate a greater prevalence of per-capita applications. Commuting zones for which insufficient applications were submitted to allow for external presentation are shown in gray. In Panel B, we use a diverging color scheme, in which light to dark blue indicate positive changes in application rates, and yellow to red indicate negative changes in application rates. The positive changes are greater when the color is a darker blue. Likewise, negative changes are greater when the color is a darker red.

Figure 4: Geographic Variation in Population-Adjusted SSDI Applications, 35–54 Year Olds



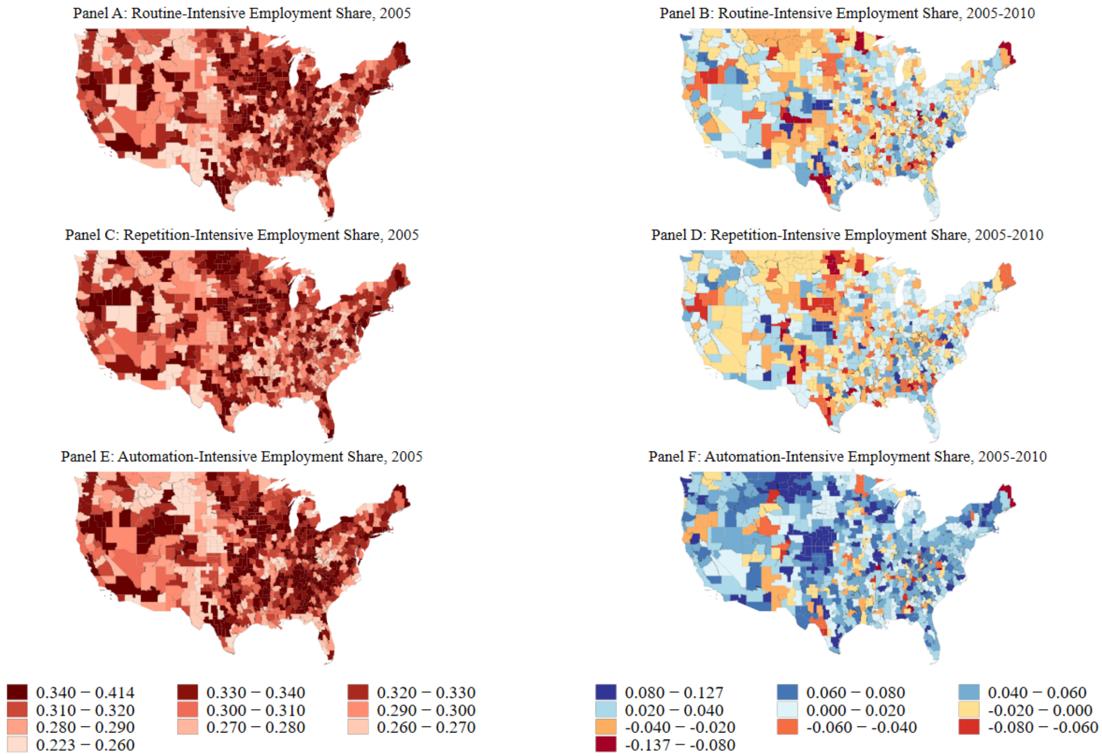
Notes: For 35–54 year-olds, the figure presents heat maps of the SSDI application-to-population ratio, as defined in equation 1, in 2005 (Panel A) as well as its change between 2005 and 2019 (Panel B). In Panel A, lighter red colors indicate a lower application-to-population ratios, and darker red colors indicate a greater prevalence of per-capita applications. Commuting zones for which insufficient applications were submitted to allow for external presentation are shown in gray. In Panel B, we use a diverging color scheme, in which light to dark blue indicate positive changes in application rates, and yellow to red indicate negative changes in application rates. The positive changes are greater when the color is a darker blue. Likewise, negative changes are greater when the color is a darker red.

Figure 5: Geographic Variation in Population-Adjusted SSDI Applications, 55–64 Year Olds



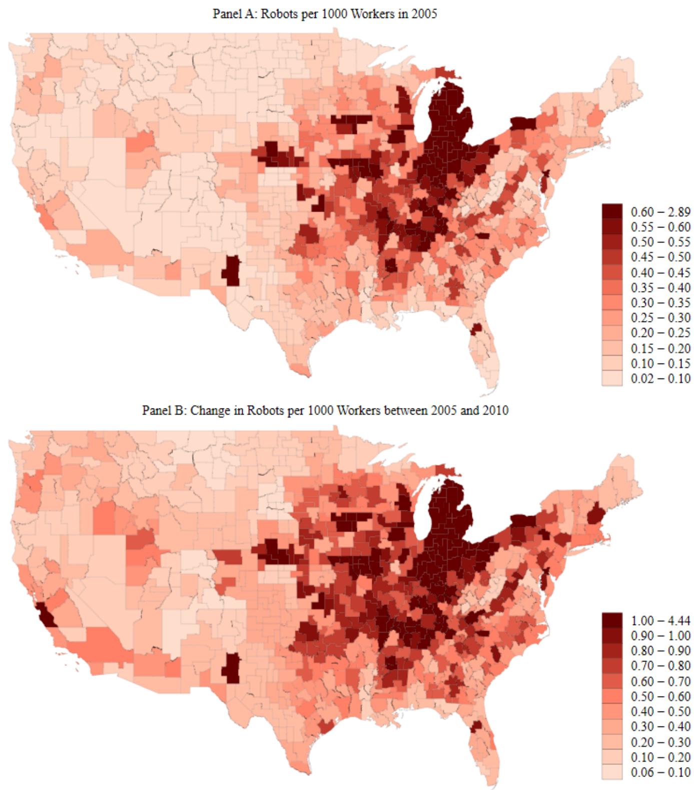
Notes: For 55–64 year-olds, the figure presents heat maps of the SSDI application-to-population ratio, as defined in equation 1, in 2005 (Panel A) as well as its change between 2005 and 2019 (Panel B). In Panel A, lighter red colors indicate a lower application-to-population ratios, and darker red colors indicate a greater prevalence of per-capita applications. Commuting zones for which insufficient applications were submitted to allow for external presentation are shown in gray. In Panel B, we use a diverging color scheme, in which light to dark blue indicate positive changes in application rates, and yellow to red indicate negative changes in application rates. The positive changes are greater when the color is a darker blue. Likewise, negative changes are greater when the color is a darker red.

Figure 6: Geographic Variation in the Routine, Repetition, and Automation Intensive Employment Shares



Notes: The figure presents heat maps for the automation measures described in Section 3.2.1 for 2005 (Panels A, C, and E) as well as the changes in the measures between 2005 and 2010 (Panels B, D, and F). Panels A and B focus on the routine-intensive employment share; Panels C and D on the repetition-intensive employment share; and Panels E and F on the automation-intensive employment share.

Figure 7: Geographic Variation in Exposure to Industrial Robots



Notes: The figure presents heat maps for the automation measures described in Section 3.2.2 for 2005 (Panel A) as well as the changes in the measure between 2005 and 2010 (Panel B).