Role of Job Characteristics and Employer Accommodation in Labor Supply and Disability Benefit Claiming Decisions in Later Life:

Analysis of Employee's Decisions to Continue Working after Disability Onset or Claim SSDI/SSI Benefits Before Normal Retirement Age

Ah-Reum Lee (Assistant Policy Researcher/PhD Fellow, RAND Corporation)

Table of Contents

List of Tables

List of Figures

1. Introduction

As individuals near retirement, the likelihood that they will experience an adverse health condition that affects their ability to work increases significantly (Wu and Hyde, 2019). Taking advantage of the large repository of longitudinal information from the Health and Retirement Survey (HRS)[1], Smith (2005) finds that about one in five respondents stated that they already had experienced a major health condition[2] onset by early 50s. Johnson, Mermin, & Murphy (2007) also report that one in four workers has experienced the onset of a work-limiting health condition[3] by their early 60s.

Consequently, older workers[4] who develop significant limitations in health or functioning in daily life face declines in income and consumption and an increased likelihood of poverty in the years prior to retirement. Dushi & Rupp (2013) examines three groups of adults aged 51–56 in 1992 with different disability experiences over 8 years using the HRS and finds that the newly disabled—people who started as nondisabled but suffered a disability shock since 1992 — experienced increased poverty rates and decreased median incomes[5]. Schimmel & Stapleton (2012) also study trajectories in earnings and income around onset of the work limitation using the HRS by comparing the household income and poverty status of the individuals who have a work-limiting conditions to those who do not. They find that mean earnings of the group that reported a new work limitation were on average 50% lower in the first period (2 years) after onset than they were among those in the group who did not experience onset. Furthermore, Meyer & Mok (2013) examine the lifetime prevalence of disability and how the disabled fare before and after the onset of disability using longitudinal data from the Panel Study of Income Dynamics (PSID) and find that ten years after disability onset, a person with a chronic and severe disability on average experiences a 79 percent decline in earnings, a 35 percent decline in after-tax income. In addition

[1] The HRS is a nationally representative biennial panel study of Americans over age 50 years. The HRS researchers began interviewing respondents in 1992, and its sample has been replenished with new cohorts of age-eligible respondents every 6 years, adding new cohorts of 51- through 56-year-olds in 1998 and again in 2004. It collects detailed information from age-eligible respondents and their spouses, including demographic, health, and functional and disability status measures, as well as information about income and wealth. The HRS survey is collected by the Institute of Survey Research at the University of Michigan, with funding from the National Institute on Aging and SSA. (See http://hrsonline.isr.umich)

[2] Major conditions are defined as cancer, heart condition, stroke, and diseases of the lung.

[3] Whether an impairment or health problem limits the kind or amount of paid work for the Respondent. Respondents are considered to have a work-limiting health condition if they answer "Yes" to the following question: "*Now I want to ask how your health affects paid work activities. Do you have any impairment or health problem that limits the kind or amount of paid work you can do*?"

[4] The United States government, through the Age Discrimination in Employment (ADEA), applies the term "older worker" to employees over the age of 40.

[5] Two years after the onset of a work-limiting health condition, earnings are about 50% lower and poverty rates are nearly double relative to the period before onset, with other sources of income offsetting relatively little of lost earnings.

to immediate effects on income, poverty, and consumption, exiting the labor force during their peak earnings years may mean lower wealth accumulation and has implications for accruing pension and Social Security wealth (Wu and Hyde, 2019).

Due to these financial consequences and the concerns about the sustainability of public programs to support these workers, a few studies have explored the possibility to delay disability and retirement of older workers by improving working conditions related to physical workload, job control, psychological job stress, which are increasingly identified as risk factors for disability and retirement. Blekesaune and Solem (2005) investigate the impact of working conditions on individual retirement for 19,114 Norwegian employees between the ages of 60 and 67 and 270 occupations, using the combination of survey data for estimates of job strains, income and social security data. They find that those who engage in hard physical work retire earlier than those with few physical strains in their jobs and typically with a disability pension. The authors also find that men in low-autonomy jobs retire earlier than do men with greater flexibility in deciding how their work should be conducted. Lahelma et al (2012) examine the contributions of work arrangements, physical working conditions and psychosocial working conditions to subsequent disability retirement using the data derived from the Helsinki Health Study cohort on employees of the City of Helsinki, Finland[6]. They also find that, among various working conditions, those that are physically demanding and those that imply low job control are potential risk factors for disability retirement. The authors suggest that improving the physical working environment and enhancing control over one's job is likely to help prevent early retirement due to disability. In the United States, Maestas et al (2017) studied the American Working Conditions Survey (AWCS)[7], a survey of individuals designed to collect detailed information on a broad range of working conditions in the American workplace, and find nearly three-fourths of Americans report either intense or repetitive physical exertion on the job at least one-quarter of the time. The same study also finds that more than one-half of Americans report exposure to unpleasant and potentially hazardous working conditions,[8] including heavy vibrations, loud noises, extreme temperatures, hazardous

[6] Information on working conditions was obtained from the baseline surveys conducted in 2000, 2001 and 2002

[7] The AWCS was fielded on the RAND American Life Panel (ALP) in 2015. The ALP is a nationally representative (when weighted) sample of individuals residing in the United States who have agreed to participate in regular online surveys. Respondents who do not have a computer at home are provided both a computer and Internet access so that the panel is representative of all individuals in the country, not just Internet users. The survey instrument used by the AWCS was closely harmonized with the European Working Conditions Survey (EWCS), also fielded in 2015 across a representative sample of workers in 35 countries in Europe.

[8] Source: Physical Exposure Risks (2015 Working Conditions Survey / RAND American Life Panel)

contaminants, and verbal abuse[9], which disproportionally affect individuals without a college education. Lopez Garcia, et al (2022) examine the role of physical job requirements and hazardous working conditions on retirement and disability among older individuals in the US using the HRS linked with Occupational Requirements Survey (ORS) and find that 1 SD increase in the physical activity and physical work environment indices are associated with a 10-13 percentage points increase in the probability of being retired and a 3-5 percentage points increase in the probability of transitioning into retirement.

Improving working conditions and lessening physical job demands, however, may not automatically translate into all older individuals being able to work longer even if they are willing due to the factors such as individual's health and the employer characteristics (Lopez Garcia, Maestas, and Mullen 2019). The role played by employers in facilitating continued work through employer accommodation for individuals with new work-limiting conditions has been of particular interest in literature, because previous research has shown that many beneficiaries of disability benefits have substantial work capacity. For instance, Gruber and Kubik (1997) use differential decreases in state-level initial allowance rates for SSDI during the late 1970s that resulted from federal policy changes aimed at reducing program growth to identify the beneficiaries of disability benefits that have might have work capacity. They report that labor force participation among older men would have been 8.9 percentage points greater had they not been initially allowed benefits from Disability Insurance (DI). Chen and van der Klaauw (2008) also evaluate the work disincentive effects of the DI program during the 1990s using comparison group and regression-discontinuity methods with merged survey-administrative data. They find that during the 1990s the labor force participation rate of DI beneficiaries would have been at most 20 percentage points higher had none received benefits, implying substantial work capacity of DI program beneficiaries. Finally, Maestas et al (2013) find that the employment rate of beneficiaries on the margin of SSDI entry in 2005 and 2006 (23 percent of all applicants) would have been on average 28 percentage points higher two years later if they had never received SSDI benefits, showing that subset of SSDI beneficiaries either retains or recovers some degree of work capacity in the years immediately following their initial decision.[10]

[9] Are you exposed at work (at your MAIN JOB) to the following conditions at least 1/4 of the time or more: Heavy vibrations, noise, extreme temperature, breathing smoke/fumes/powder/dust/vapors, handling hazardous chemicals

[10] They make use of a unique workload management database called the Disability Operational Data Store (DIODS), which contains all initial medical determinations (that is, excluding technical denials) made between January 1, 2005, and December 31, 2006, linked to SSA's Master

Despite the potential work capacity of individuals with new work-limiting conditions, there has been limited research on the impact of employer accommodation on labor supply outcomes. Furthermore, the estimated impacts tend to be small and mixed depending on health conditions and intervention types. Stapleton et al. (2015) show a summary of 9 studies that show early intervention having positive impacts for musculoskeletal conditions and other low mortality conditions[11] and mental health conditions[12]. However, many of the cited studies do not provide evidence of impact (e.g., Linton et al. 2016), and others review evidence without regard to study quality and conclude with mixed findings (Dibben et al. 2012, p.26). In addition, three systematic analyses (Van Vilsteren et al. (2015), Palmer et al. (2011) Suijkerbuijk et al. (2017)), focus on the impact of workplace-based interventions[13] on health conditions, including musculoskeletal disorders and mental illnesses, and report mixed results on whether such interventions reduce time to first return to work and cumulative duration of absence from work.

In this research, I aimed to study the major predictors of disability onset for older workers in the United States and the role of various employer accommodations in retaining newly disabled workers in the workforce using a nationally representative data. I used machine learning (ML)[14] as a primary tool to identify major predictors of disability onset for older workers due to its ability to predict unforeseen events in the future[15] and generalizability achieved through cross-validation (Allen (1974), Stone (1974), Geisser (1975))[16], and propensity score matching (PSM)[17] to measure the impact of employer accommodations on newly disabled workers.

Beneficiary Record (MBR) and administrative annual earnings records between 1995 and 2009 from SSA's Detailed Earnings Record (DER) in order to identify which applicants ultimately received SSDI benefits.

[11] Williams et al. 2007, Waddell et al. 2008, Van Oostrom et al. 2009, Wickizer et al. 2011, Dibben et al. 2012, Hoefsmit et al. 2012, Bevan 2015, Linton et al. 2016, Richmond et al. 2015

[12] Killackey et al. 2008, Dibben et al. 2012, Hoefsmit et al. 2012, Rost et al. 2004

[13] Interventions included prescribed exercises, behavioral change techniques, workplace modifications, provision of extra services, vocational rehabilitation interventions

[14] Machine Learning (ML) is an analytical approach in which users can build statistical models that 'learn' from data to make accurate predictions and decisions, and the focus is generally on enhancing prediction rather than explaining why a phenomenon happens (Mitchell 1997).

[15] Models are evaluated only on how well they predict on data they were not created from, not how they explain.

[16] Cross-validation is a resampling procedure used to evaluate machine learning models on a limited data sample. The procedure has a single parameter called k that refers to the number of groups that a given data sample is to be split into. As such, the procedure is often called k-fold cross-validation. When a specific value for k is chosen, it may be used in place of k in the reference to the model, such as k=10 becoming 10-fold cross-validation. Cross-validation is primarily used in applied machine learning to estimate the skill of a machine learning model on unseen data. That is, to use a limited sample in order to estimate how the model is expected to perform in general when used to make predictions on data not used during the training of the model. Summarize the skill of the model using the sample of model evaluation scores (Allen (1974), Stone (1974), Geisser (1975)).

[17] Propensity score matching estimates each individual's propensity to receive a binary treatment (in this case, employer accommodation) as a function of observable characteristics and matches individuals with similar propensities. Estimated propensity scores are used to reweight the distribution of covariates X in the control group to match the distribution of X observed in the treated group. Essentially, this method places more weight on individuals who are not accommodated but are similar to individuals who are accommodated based on observable characteristics and less weight on unaccommodated individuals that are less similar, so that these two groups of individuals are more comparable (Rosenbaum, 1983)

The research questions were:

1. Are there significant differences in job characteristics, including occupational job requirements, working conditions and workplace-related benefits, as well as worker characteristics, by disability and accommodation status of workers near retirement in the United States? To answer this question, I conducted descriptive analysis of job and worker characteristics of workers near retirement by their employment, disability and employer accommodation status and test statistical difference across these groups.
2. What are the major predictors of disability onset before normal retirement ages for workers near retirement? To answer this question, I trained and compared different machine learning models for binary classification which conduct binary classification tasks (e.g., predict whether an individual will be disabled or not), including random forests, gradient boosted trees and ℓ_1- penalized logistic regression.
3. What is the impact of different types of employer accommodation on labor supply and disability claiming decisions of workers who experience disability onset before normal retirement ages? To answer this question, I used propensity score matching, which estimates each individual's propensity to receive a binary treatment (i.e., employer accommodation) as a function of observable characteristics, matches individuals with similar propensities, and measures the impact of the treatment on those individuals.

The goal of this study was to provide the policymakers and the academic community with a better understanding of the role of job characteristics in predicting disability onset and how employer accommodations could predict labor force exit, and explore new methodologies such as machine learning to improve the predictive ability for out-of-sample data[18] in the future. However, this research did not extend to investigating which firm and job characteristics, worker characteristics and their disability types explain the variation in provision of disability accommodation at the local level (e.g., variations within state or cities), as well as heterogeneity in accommodation rates by industry, which could be further investigated using a rich administrative database in the future. This research will inform the design of DOL's policies on providing reasonable job accommodations for older workers with disabilities and preventing the

[18] Out-of-sample is data that was unseen and used only to produce the prediction/forecast on it.

early exit of labor market before the normal retirement age[19], which will help minimize fiscal burden associated with the prevalence of disability retirement and improve retirement security of older workers.

2. Conceptual and Operational Definitions

This section provides conceptual/operational definitions of key terms used in this paper:

- *Disability*: Disability is defined as a work-limiting health condition. I consider individuals to be disabled if they answer "Yes" to the following question in the HRS: "Do you have any impairment or health problem that limits the kind or amount of paid work you can do?"
- *Disability onset:* I identify disability onset as the first wave when the individuals who are not disabled when they enter the panel report disability in the HRS.
- *Newly disabled workers*: I identify newly disabled workers as workers who are not disabled when they enter the panel but who subsequently report a work disability that began when they were employed in the HRS.
- *Labor force exit*: Individuals are considered to have exit the labor force if their reported labor force status is "fully retired" or "Not in the labor force" in the HRS.
- *Employer accommodations*: As defined in the Americans with Disabilities Act of 1990, "an accommodation is any change in the work environment or in the way things are customarily done that enables an individual with a disability to enjoy equal employment opportunities." It includes modification of job requirements and work schedules, or provision of assistive equipment; it does not extend to other interventions thought to promote return-to-work such as coordination of medical care, career counseling, vocational rehabilitation, or education and re-training.
- *Job characteristics*[20]: Every occupation requires a different mix of knowledge, skills, and abilities, and is performed using a variety of activities and tasks. Job characteristics are defined as these distinguishing characteristics of an occupation. Job characteristics considered in this study include:
    - *Physical demand:* Physical activities required to perform tasks in a job

[19] Age 66 for respondents who were born in 1943-1954. For more details, see "Social Security Administration: Full Retirement and Age 62 Benefit By Year Of Birth" (https://www.ssa.gov/benefits/retirement/planner/agereduction.html)
[20] Source: https://www.bls.gov/opub/hom/ors/concepts.htm

- *Environmental conditions:* The various tangible or concrete hazards or difficulties that are in proximity to the location where jobs' critical tasks are performed
- *Cognitive and mental requirements:* The qualifications that workers need to use judgment, make decisions, interact with others, and adapt to changes in a job
- *Work activities*: Work activities that are common across a very large number of occupations
- *Work context*: Physical and social factors that influence the nature of work
- *Basic skills*: Developed capacities that facilitate learning or the more rapid acquisition of knowledge
- *Abilities:* Enduring attributes of the individual that influence performance of a job

3. Data

I used three data sources for this research. The first was the Occupational Requirements Survey (ORS), collected by the Bureau of Labor Statistics under the Department of Labor (BLS). The ORS is a survey of establishments in private industry and state and local government. The first wave of ORS, collected over a three-year period between 2015 and 2018, provides information on the physical demands, environmental conditions, and education, training and experience that are required in each job as reported by Bureau of Labor Statistics (BLS) field economists who are extensively trained and given detailed instructions on data collection techniques.[21] The second wave, planned for collection over five years from September 2018 to July 2023, provides the same information with additional data on reported cognitive and mental job requirements. For this research, I used the preliminary second wave data (reference year 2021 complete dataset) collected through July 2021[22] to include cognitive and mental requirements to the analysis. Specifically, the

---

[21] The Occupational Requirements Survey (ORS) is a survey of establishments in private industry and state and local government conducted by the Bureau of Labor Statistics (BLS). Private industry and state and local government establishments in the 50 states and the District of Columbia are eligible for selection. Major exclusions from the survey are workers in federal and quasi-federal agencies (examples include the military, postal service, and Federal Reserve), establishments in the agriculture, forestry, fishing, and hunting industry sector, workers employed by private households, contractors (onsite workers at the surveyed establishment who are paid by another party are not included in data collection from the surveyed establishment), the self-employed, volunteers, unpaid workers, individuals receiving long-term disability compensation, and those working overseas. Individuals who set their own pay, such as business owners, and family members who are paid token wages are also excluded. Employees in sampled jobs must receive market-based payments, such as salary, commission, or hourly wages, from the establishment for services performed in the labor market and the establishment must pay the employer's portion of Medicare taxes on the worker's wages. The ORS publishes information about job requirements, including physical demands; environmental conditions; education, training, and experience; as well as cognitive and mental requirements.

- Key measures: Cognitive and mental requirements, Education, training, and experience, Environmental conditions, Physical demand
- How the data are obtained: Survey of businesses, Government agencies
- Classification: Occupation
- Periodicity of data availability: Annual
- Geographic detail: National
- Scope: Private sector, State and local government

[22] These estimates from three of five samples (2018/2019, 2019/2020, 2020/2021) and are considered preliminary.

publicly available data from the Wave 2 of the ORS contains job requirements for 16 physical job requirements, and 10 environmental conditions, and 12 cognitive and mental requirements. For each job requirement, ORS provides a mix of categorical and continuous measures[23], for a total of 174 variables for physical job requirements, 65 variables for environmental conditions, and 35 variables for cognitive and mental requirements. In the public-use ORS, categorical variables measure the percentage of workers in a given occupation that are subject to a given requirement, such as, for example, the percentage of workers in an occupation for which gross manipulation is required. For some job traits, ORS also provides estimates of the percent of workers subject to a given requirement for a given level of frequency: seldom, occasionally, frequently, or constantly. Continuous variables include select summary statistics by occupation of the duration of the working day for which certain job traits are required. For example, the ORS includes variables for the average number of hours spent sitting by occupation as well as the 10th, 25th, 50th, 75th, and 90th percentiles of hours spent sitting by occupation. For each requirement, the ORS reports the share of workers in an occupation whose job requires that ability. The ORS currently provides data for 390 occupations at the six-digit 2018 Standard Occupation Classification system (SOC) level.[24] Appendix 1 provides an overview of the types of variables available for each job characteristic for physical job requirements; Appendix 2 provides an overview for the environmental conditions; and Appendix 3 provides an overview for the cognitive and mental requirements.

One limitation in using the public-use ORS data was the lack of complete information for certain job characteristics and occupations. First, for 13 job characteristics, the variable containing the percentage of workers for whom a particular job characteristic is required was available (corresponding to Column 1 in Appendix 1-3) but more detailed variables containing the percentage of workers subject to a requirement "at a given frequency level" (i.e., not required/seldom/occasionally/frequently/ constantly) was unavailable (Column 2 in Appendix 1-3). Likewise, for 3 job characteristics with continuous variables (Column 3 in Appendix 1-3), the median (50th percentile) was available but not all the percentiles (10th, 25th, 50th, 75th, and 90th

[23] In the public-use ORS, categorical variables measure the percentage of workers in a given occupation that are subject to a given requirement, such as, for example, the percentage of workers in an occupation for which gross manipulation is required. ORS also provides estimates of the percent of workers subject to a given requirement for a given level of frequency: seldom, occasionally, frequently, or constantly. Continuous variables include select summary statistics by occupation of the duration of the working day for which certain job characteristics are required. For example, the ORS includes variables for the average number of hours spent sitting by occupation as well as the 10th, 25th, 50th, 75th, and 90th percentiles of hours spent sitting by occupation.

[24] Source reference: BLS Standard Occupational Classification and Coding Structure (https://www.bls.gov/soc/2018/home.htm)

percentiles). As a result, I restricted the analysis to the categorical variables in Column 1[25] and continuous variables containing mean levels in Column 3 of Appendix 1-3.

Another limitation was the missing information at the occupation level. Appendix 4-6 present the percent of occupations for which each job requirement is observed in the ORS data, for physical job requirements, environmental conditions and cognitive and mental requirements, respectively. I found a range of missing observations at the occupation level, with average rate of missing observations of 2% for environmental conditions, 14% for physical job requirements, and 21% for cognitive and mental requirements. For occupations with missing data, I imputed the mean calculated for non-missing occupations at the two-digit level (i.e., occupation groups). [26] [27] Furthermore, I excluded 4 cognitive and mental requirements for which the percent of occupations observed are below 50% (i.e., control of workload, communicating verbally, work reviewed by supervisor, problem solving). Consequently, I used the 16 physical job requirements and the 10 environmental conditions and 8 cognitive and mental requirements, and used observed and imputed data for analysis for the remainder of the analysis.

The second data source was the Occupational Information Network (O*NET) survey[28], version 26.0 released in August 2021, which supplies a comprehensive database surveying more than 200 job characteristics and worker attributes for 873 occupations coded at the six-digit level of the 2010 SOC system. Job characteristics in the O*NET include required abilities, work activities, skills, knowledge, work values, and work context (Johnson, Mermin, and Resseger 2011; Belbase, Sanzenbacher, and Gillis 2015). The O*NET database provides ratings [29], which are based on responses from workers randomly surveyed at a sample of businesses and, provides a distribution

[25] % of workers for whom the job requirement is required (1=the job requirement is required, 0=otherwise)

[26] The occupations in the SOC are classified at four levels of aggregation to suit the needs of various data users: major group, minor group, broad occupation, and detailed occupation. Each lower level of detail identifies a more specific group of occupations. The 23 major groups, listed below, are divided into 98 minor groups, 459 broad occupations, and 867 detailed occupations (For more details, see SOC 2018 Manual: https://www.bls.gov/soc/2018/soc_2018_manual.pdf).

23 Major Occupation Groups: 11-0000 Management Occupations; 13-0000 Business and Financial Operations Occupations; 15-0000 Computer and Mathematical Occupations;17-0000 Architecture and Engineering Occupations;19-0000 Life, Physical, and Social Science Occupations; 21-0000 Community and Social Service Occupations; 23-0000 Legal Occupations; 25-0000 Educational Instruction and Library Occupations; 27-0000 Arts, Design, Entertainment, Sports, and Media Occupations; 29-0000 Healthcare Practitioners and Technical Occupations; 31-0000 Healthcare Support Occupations; 33-0000 Protective Service Occupations; 35-0000 Food Preparation and Serving Related Occupations; 37-0000 Building and Grounds Cleaning and Maintenance Occupations; 39-0000 Personal Care and Service Occupations; 41-0000 Sales and Related Occupations; 43-0000 Office and Administrative Support Occupations; 45-0000 Farming, Fishing, and Forestry Occupations; 47-0000 Construction and Extraction Occupations; 49-0000 Installation, Maintenance, and Repair Occupations; 51-0000 Production Occupations; 53-0000 Transportation and Material Moving Occupations; 55-0000 Military Specific Occupations

[27] The means are similar without imputation for the limited occupation set and with imputation for the full occupation set.

[28] The O*NET data is developed under the sponsorship of the U.S. Department of Labor/Employment and Training Administration (USDOL/ETA).

[29] For all three modules each question has two parts: 1) How important is a particular ability for doing your current job? (Answer in 5-point scale; "not important" to "extremely important"); 2) What level of the ability is needed to perform the current job? (Answer in 7-point scale; anchors describing different levels of the ability at 2, 4, and 6).

of the job characteristic for an occupation including mean and standard deviation. The Abilities, Skills, and Work Activities modules of the O*NET[30], which I included in the analysis to represent job requirements as reported by workers, ask two-part questions about the Importance and Levels of a given characteristic, skill, or activity. I combined the importance and levels by multiplying them together to create a composite rating for all job characteristics. For this research, I used 52 work abilities (21 cognitive abilities, 10 psychomotor abilities, 21 physical/sensory abilities), 38 work activities (15 social activities, 9 technical work output, 14 general work activities), and 35 work skills. It is worth mentioning that, while ORS aims to understand what specific physical or cognitive capabilities are "required" to complete critical job functions of selected jobs[31], the O*NET seeks to understand what knowledge, skills, abilities, and work activities are "typical" in a particular occupation.[32] See Appendix 7 for more details.

Finally, I used the Health and Retirement Study (HRS)[33], a longitudinal household survey representing the non-institutionalized[34] U.S. population over the age of 50.[35] Respondents are surveyed every two years, allowing us to track transitions from work into retirement and/or disability status. The HRS, with its panel structure and comprehensive information on health status, functional limitations, income, assets, and benefit receipt (SSDI and SSI), has become the

---

[30] The O*NET defines an ability as "an enduring talent that can help a person do a job", a skill as "Developed capacities that facilitate learning, the more rapid acquisition of knowledge, or performance of activities that occur across jobs" and a work activity as "a set of similar actions that are performed together in many different jobs." (For more details, see https://www.onetcenter.org/dl_files/database/db_26_1_dictionary.pdf)

[31] The survey does not focus on specific capabilities or experiences that individual workers have if the employer does not require them. For example, a job may require a bachelor's degree, but workers performing the job may have more advanced degrees, such as a doctorate degree (Ph.D.). For the purposes of the ORS, the requirement is a bachelor's degree. The distinction is significant because the objective of the survey is to measure job requirements, not the characteristics of the workers.

[32] Comparison of ORS and O*NET:

- ORS: The ORS provides estimates measuring four types of occupational requirements: physical demands; environmental conditions; education, training, and experience; as well as cognitive and mental requirements. Survey estimates help to define and describe the requirements of work in the U.S. economy. The ORS is designed to explain what is required to perform critical job functions of selected jobs. The survey does not focus on specific capabilities or experiences that individual workers have if the employer does not require them. For example, a job may require a bachelor's degree, but workers performing the job may have more advanced degrees, such as a doctorate degree (Ph.D.). For the purposes of the ORS, the requirement is a bachelor's degree. The distinction is significant because the objective of the survey is to measure job requirements, not the characteristics of the workers.
- ONET: O*NET provides estimates on a different mix of knowledge, skills, and abilities required to perform a job, as well as a variety of activities and tasks involved in a job.

[33] The HRS is a nationally representative biennial panel study of Americans over age 50 years. The HRS researchers began interviewing respondents in 1992, and its sample has been replenished with new cohorts of age-eligible respondents every 6 years, adding new cohorts of 51-through 56-year-olds in 1998 and again in 2004. There are currently sixteen waves of core data available from 1992 to 2020 with about 18–23,000 participants in any given wave. It collects detailed information from age-eligible respondents and their spouses, including demographic, health, and functional and disability status measures, as well as information about income and wealth. The HRS survey is collected by the Institute of Survey Research at the University of Michigan, with funding from the National Institute on Aging and SSA (http://hrsonline.isr.umich).

[34] The civilian noninstitutional population age 16 and older is the base population group, or universe, used for Current Population Survey (CPS) statistics published by BLS. The civilian noninstitutional population excludes active duty members of the U.S. Armed Forces; people confined to, or living in, institutions or facilities such as prisons, jails, and other correctional institutions and detention centers residential care facilities such as skilled nursing homes. Included in the civilian noninstitutional population are citizens of foreign countries who reside in the United States but do not live on the premises of an embassy.

[35] The HRS samples at the household level. In a couple household, even if only one individual is age-eligible for the study, BOTH individuals are given an interview and treated as respondents.

go-to source for research related to health, disability, and the interplay between retirement and disability benefits (Agree, & Wolf, D. A. 2017). I used the 1992–2018 waves of the RAND HRS Longitudinal File, merged with the ORS and O*NET measures of job characteristics and occupational characteristics using the restricted version of the HRS Industry and Occupation Data, which includes occupation information at the 4-digit Census code level[36] (equivalent to 6-digit Standard Occupational Classification). Previous research has shown that over 90% of workers from the HRS are able to be merged occupational information from the O*NET (Johnson, Mermin, and Resseger 2011; Belbase, Sanzenbacher, and Gillis 2015). For the respondents in my analytic sample, 94% of respondents in the HRS sample who reported occupations in the analytic sample (see Table 2) were merged occupation information from the ORS and 96% of the respondents were merged occupation information from the O*NET. Table 1 provides a list of key variables used by the current study from the HRS, ORS and O*NET:

[36] United States Census Bureau Industry and Occupation Classification (Source: https://www.census.gov/programs-surveys/cps/technical-documentation/methodology/industry-and-occupation-classification.html)

Table 1 Key Variables Selected from Three Survey Programs into the Current Study

Data[1)]	Category	Key Variables
Occupational Requirement Survey (ORS)	Job requirements	• Physical job requirements (16 job characteristics), environmental working conditions (10 job characteristics), cognitive/mental requirements (8 job characteristics)
Occupational Information Network Survey (O*NET)	Job characteristics	• Cognitive abilities (21 variables), psychomotor abilities (10 variables), physical/sensory abilities (21 variables), social activities (15 variables), technical work activities (9 variables), general work activities (14 variables), work skills (35 variables)
Health and Retirement Study (HRS)	Work arrangements	• Work hours (full/part-time,1 variable), types of employment (employed/self-employed, 1 variable), workplace benefits (health insurance covered by employer, pension from current job, 2 variables), union contract/employee contract (1 variable)
	Employer characteristics and accommodations[37]	• Allowed flexibility of work hours; changed job content; modified job tasks to suit workers' abilities; helped workers learn new job skills and provided training; provided assistance at tasks when needed; emotional support; medical care provided (or arranged for); paid medical leave; time off when needed (unpaid); parking made easier/closer; adapted/changed working environment; monetary compensation (including payment for medical expenses); offered early retirement; helped get workers' compensation/disability (1 variable) • Firm size (1 variable)
	Self-reported job demands	• Current job requires lots of physical effort (1 variable); lifting heavy loads (1 variable); stooping, kneeling, or crouching (1 variable); good eyesight (1 variable); involves lots of stress (1 variable)
	Health/Disability	• Self-reported health, chronic diseases[2)], health issues (before age 16) • Self-reported health, chronic diseases, health issues (in the 50s-specifically in wave prior to disability onset) • Disability status (Work-limiting health condition) • Health-related behaviors/lifestyle (physical activity, drinking, smoking, preventive behaviors) • BMI, height, weight
	Healthcare utilization	• Hospital stays, medical care utilization (doctor visits, out of pocket expenditures)
	Benefits	• Workers' Compensation, Social Security Disability Insurance (SSDI), Supplemental Security Income (SSI), Veterans Administration Disability Compensation • SSI/SSDI episodes (application, denials, approvals, reapplication), worker's compensation and veteran's administration disability benefits
	Demographics	• Age, gender, race, marital status, birthplace, census region/division, veteran status
	Socioeconomic Status (SES)	• Education, income (individual/household earnings, labor and non-labor income, benefits), wealth (household assets, financial assets, housing assets, debts, mortgages), occupations/industry

Notes. For details of the variables, see codebook

1. Data Source: Health and Retirement Study 1992-2018, ORS reference year 2021 complete dataset, ONET version 26.0

2. Chronic diseases and health issues include: cancer, diabetes, lung diseases, asthma, respiratory disorder, speech impairment, allergic condition, heart trouble, ear problems, epilepsy or seizures, high blood pressure, depression, drugs/alcohol problems, past injuries, childhood disability.

[37] If the respondent reported that their employer did something special to help them out, they were then asked more detailed questions about what types of things the employer did.

The total number of the analytic sample from the final dataset (i.e., HRS matched with ORS and O*NET) was N=18,730 (corresponding to N=261,366 person-year observations). Table 2 lists the sample restrictions I applied to construct the analytic sample and the sample size after each restriction:

Table 2 Sample Size and Restrictions Selected from Three Survey Programs into the Current Study

Sample restrictions[1), 2)]	**N**
(a) Entered the panel without reporting a work disability	18,730
• (b) Did not not experience work-limiting health condition before normal retirement age[38]	14,101
• (c) Experienced work-limiting health condition before normal retirement age	4,629
- (d) Not working at the time of disability onset	505
- (e) Self-employed at the time of disability onset	1,635
- (f) Employed at the time of disability onset	2,489
· (g) Accommodated	720
· (h) Not accommodated	1,662

Notes.
1. Data Source: Health and Retirement Study 1992-2018, ORS reference year 2021 complete dataset, ONET version 26.0
2. Sample (a) – (h):
(a) HRS respondents entering panel without work disability and whose age was below 55 at the entering wave;
(b) Among respondents (a), respondents who do NOT report work-limiting health conditions before normal retirement age;
(c) Among respondents (a), respondents who reported experiencing work-limiting health conditions before normal retirement age;
(d) Among respondents (c), respondents who were employed at disability onset;
(e) Among respondents (c), respondents who were NOT employed at disability onset;
(f) Among respondents (c), respondents who were self-employed at disability onset;
(g) Among respondents (d), response who reported being accommodated at disability onset;
(h) Among respondents (d), response who reported not being accommodated at disability onset.

Of the 18,730 HRS respondents who entered the panel without reporting a work disability (i.e., work-limiting health condition), 4,629 or 25% reported that they experienced work-limiting health condition while still in working-age years (i.e., before they become eligible to claim full Social Security benefits in row (c) in Table 2). I further restricted the sample to individuals who were employed at the time their health first began to limit their ability work, [39] i.e., 4,124 respondents (including the self-employed in row (e) and the employed in row (f)). Next, the respondents were asked if their employer did anything special to help them out so that they could stay at work at the time their health started to limit the ability to work, to which possible responses are "yes," "no," "left immediately," "self-employed" and (starting in 1998 survey) "no help needed." I excluded

[38] Age 66 for respondents who were born in 1943-1954. For more details, see "Social Security Administration: Full Retirement and Age 62 Benefit By Year Of Birth" (https://www.ssa.gov/benefits/retirement/planner/agereduction.html)
[39] Respondents who answered "yes" to the question "Were you employed at the time your health began to limit your ability to work"

all respondents who gave an answer other than yes or no. Out of 2,480 respondents who were employed at the time of disability onset, 720 or 29% reported being accommodated by their employers.

4. Part 1. Descriptive Analysis

Table 3 and Table 4 present summary statistics for the main sample of 18,730 respondents, overall, by disability status, employment type when they first entered the HRS survey, and by accommodation status at disability onset if they ever experience work limiting health condition before reaching the normal retirement age.

First, in columns (b) and (c) of Table 3, I compared demographics, health and socioeconomic conditions, as well as job characteristics of respondents who never experience disability ("no disability group") and those who experience disability while still in working-age years ("disability group") and evaluated whether they are significantly different from each other (p-value that compares these two groups for statistical difference). The average number of years of education was significantly higher for the "No disability group" group than the "disability group" (13.0 years vs. 12.2 years). The no disability group also had fewer female respondents (49% vs. 52%), a higher marriage rate (70% vs. 65%), and significantly higher household assets (324,000 USD vs. 182,000 USD) and earnings (37,000 vs. 24,000 USD). The "no disability group" also reported a significantly higher number of chronic conditions already in their early 50s (1.2 vs. 0.8) compared to the "disability group", including high blood pressure, diabetes, and arthritis, which are identified as major risk factors for more severe health conditions such as heart disease, stroke, and kidney diseases (Nanayakkara et al., 2021, Barrett-Connor et al., 2018, Pikula et al., 2018, Moonesinghe et al, 2019, Wong et al., 2016). Their occupations were reported to be more physically demanding (ORS physical activity index, 0.18 vs. -0.17; O*NET physical abilities index -0.89 vs. -1.18) with greater exposure to physically demanding work environment (ORS physical environment index, 0.49 vs. 0.35), but less cognitively or socially burdensome (ORS cognitive requirements, -0.24 vs. 0.14; O*NET cognitive abilities index, -1.52 vs. -0.50) as shown by the average standardized indices[40] for job requirements and job abilities from ORS and O*NET in Table 4. The self-reported job stress was not significantly different between the two groups (96% vs. 96%).

[40] Job demand indices are standardized with a mean of 0 and a standard deviation (SD) of 1.

Second, columns (g) and (h) in Table 3 and Table 4 compared respondents whose employers accommodated their disabilities and those whose employers did not accommodate their disabilities, with a p-value to compare these two groups for statistical differences. In my sample, 25% of the respondents who entered the survey without any prior work disability experienced work-limiting health conditions before they reached the normal retirement age, and only 29% of newly disabled workers received some form of employer accommodation upon becoming disabled.[41] These respondents who reported that their employer did something to help them out also provided information on what types of accommodations their employer provided and I categorized their responses into four types of accommodation following a previous study on employer accommodation using the HRS data (Maestas et al. 2017) : 1) time accommodation (allowing more breaks, allowing different arrival or departure times or shortening the work day), reported by 29% of accommodated respondents; 2) provision of special equipment or transportation (getting special equipment, arranging special transportation), reported by 13% of respondents; (3) work changes (changing the job, helping to learn new job skills), reported by 40% of respondents; and 4) other types of accommodation (getting someone to help, emotional support; medical care provided (or arranged for); paid medical leave; adapted/changed working environment; monetary compensation, including payment for medical expenses; offered early retirement; helped get workers' compensation/disability), reported by 43%. These accommodations are not mutually exclusive (See Appendix 8 for a more detailed breakdown of the types of accommodation). The group difference test showed that, except for some minor differences in education, earnings, and the number of chronic conditions in their early 50s, individuals whose employers accommodated their disabilities were not significantly different from those whose employers did not accommodate their disabilities in terms of demographics (gender, race/ethnicity, married with spouse), health conditions (ever diagnosed chronic conditions) and occupational characteristics (average job demands). While newly disabled workers who were accommodated had higher education (12.7 vs. 12.2 years) and earnings (32,000 vs. 29,000 USD), I found no systematic evidence that workers with certain job characteristics or healthier workers were more likely to be accommodated at disability onset.

[41] While this statistic is similar with other work using national surveys, it is notable that a probability-based survey of private- and federal-sector employers found much higher accommodation rates, in the range of 60–70% (e.g., Bruyere, 2000). One possible explanation for the discrepancy is that the HRS (and other surveys) asks about accommodation only when respondents say they have a work-limiting health problem (Hills et al., 2016). This conditioning sequence will skip people who have been accommodated if they no longer consider themselves work-limited (perhaps because the accommodation was successful).

Nonetheless, it is worth nothing that accommodated workers continued to work significantly more following disability onset as indicated in Table 5. While 48% of newly disabled workers who were accommodated continued to work full-time in 2 years after disability onset, only 34% of those who were not accommodated still worked in 2 years after disability onset. Finally, the analysis showed that 34.1% of newly disabled workers applied for Social Security Disability Insurance (SSDI) or Supplemental Security Income (SSI) or both within 2 years of disability onset and workers whose employers accommodated their disabilities were observed to have a lower rate of disability claiming in both SSDI and SSI.

Table 3 Summary Statistics: Demographics and Health Conditions (as reported by survey respondents)

Variable	(a) All	Disability Status			Employment Status at Disability Onset			Employer Accommodation Status		
		(b) No disability	(c) Disability	P-value	(d) Employed	(e) Not employed	(f) Self-employed	(g) Accommodated	(h) Not accommodated	P-value
Demographics										
Age	55.0	55.1	54.6	<0.001	54.3	55.0	54.6	54.0	54.5	<0.001
Education (in years)	12.8	13.0	12.2	<0.001	12.4	11.7	13.0	12.7	12.2	<0.001
Female (%)	49%	49%	52%	<0.001	50%	62%	33%	50%	50%	0.99
Black (%)	20%	20%	21%	0.13	20%	24%	13%	20%	21%	0.66
Hispanic (%)	14%	14%	13%	0.15	12%	16%	8%	11%	12%	0.35
Married with partners (%)	69%	70%	65%	<0.001	64%	64%	71%	65%	64%	0.71
Household Assets (1,000 current USD)	289	324	182	<0.001	135	171	445	146	128	0.099
Self-reported Health conditions										
Body Mass Index (mean)	27.9	27.7	28.7	<0.001	29.0	28.2	28.4	29.5	28.8	0.017
Smoker (%)	23%	20%	31%	<0.001	31%	33%	29%	30%	31%	0.58
Ever diagnosed (%)										
– Diabetes	10%	9%	14%	<0.001	14%	16%	9%	15%	14%	0.41
– High Blood Pressure	34%	32%	40%	<0.001	40%	42%	29%	43%	40%	0.093
– Cancer	4%	4%	5%	<0.001	6%	6%	5%	7%	5%	0.19
– Lung Disease	3%	2%	5%	<0.001	5%	5%	3%	7%	5%	0.021
– Arthritis	25%	21%	37%	<0.001	38%	37%	32%	41%	38%	0.14
– Heart Problem	6%	5%	9%	<0.001	9%	9%	11%	10%	10%	0.49
– Stroke	1%	1%	2%	<0.001	2%	3%	1%	3%	2%	0.27
– Psych Problem	6%	5%	10%	<0.001	9%	11%	8%	11%	8%	0.004
Number of chronic conditions (mean)	0.9	0.8	1.2	<0.001	1.2	1.3	1.0	1.4	1.3	0.002
Self-reported health rating in wave prior to onset (1=excellent ~ 5=poor)	2.4	2.3	2.8	<0.001	2.8	2.9	2.6	2.8	2.8	0.96
No. Obs (individuals)	**18,730**	**14,101**	**4,629**	**-**	**2,489**	**1,635**	**505**	**720**	**1,662**	**-**

Notes.
1. Data Source: Health and Retirement Study 1992-2018
2. Sample (a) – (h):
(a) HRS respondents entering panel without work disability and whose age was below 55 at the entering wave;
(b) Among respondents (a), respondents who do NOT report work-limiting health conditions before normal retirement age;
(c) Among respondents (a), respondents who reported experiencing work-limiting health conditions before normal retirement age;
(d) Among respondents (c), respondents who were employed at disability onset;
(e) Among respondents (c), respondents who were NOT employed at disability onset;
(f) Among respondents (c), respondents who were self-employed at disability onset;
(g) Among respondents (d), response who reported being accommodated at disability onset;
(h) Among respondents (d), response who reported not being accommodated at disability onset.
3. Statistical significance indicated by *p < 0.05. **p < 0.01. ***p < 0.001

Table 4 Summary Statistics: Work and Job Characteristics (as reported by survey respondents)

Variable	(a) All	Disability Status			Employment Status at Disability Onset			Employer Accommodation Status		
		(b) No disability	(c) Disability	P-value	(d) Employed	(e) Not employed	(f) Self-employed	(g) Accommodated	(h) Not accommodated	P-value
Work										
Work hours (hours per week)	41.5	41.7	40.8	<0.001	40.6	39.8	43.0	40.8	40.6	0.67
Years worked	25.9	25.8	26.2	0.096	28.5	21.1	30.8	28.3	29.0	0.16
Job involved a lot of stress	96%	96%	96%	0.37	96%	96%	94%	96%	97%	0.67
Job involves a lot of physical effort	69%	67%	76%	<0.001	76%	75%	76%	77%	76%	0.66
Job involves a lot of lifting heavy	47%	45%	54%	<0.001	54%	54%	58%	53%	54%	0.79
Job involves a lot of stooping	65%	63%	71%	<0.001	70%	71%	73%	73%	70%	0.09
Job involves a lot of good eyesight	96%	96%	97%	0.14	97%	97%	95%	96%	97%	0.21
Firm size										
– Less than 15	27%	27%	25%	0.037	22%	12%	79%	23%	22%	0.122
– 15-24	5%	5%	5%	0.48	7%	2%	3%	8%	7%	0.098
– 25-499	29%	31%	26%	0.06	38%	13%	8%	39%	37%	0.231
– 500+	10%	11%	9%	0.12	13%	6%	1%	13%	13%	0.53
– Missing Information	28%	26%	36%	0.01	20%	68%	9%	17%	21%	0.12
Earnings (1,000 current USD)	34	37	24	<0.001	30	15	22	32	29	0.032
DB/DC Pension	43%	44%	38%	<0.001	55%	20%	10%	60%	54%	0.005
Employer Health Insurance. (Own)	55%	57%	49%	<0.001	65%	33%	24%	69%	64%	0.024
Employer Health Insurance. (Spouse)	20%	20%	21%	0.46	17%	25%	24%	17%	18%	0.44
Average Job Demands (Standardized)										
ORS										
– Physical Activity	-0.09	-0.17	0.18	<0.001	0.15	0.09	0.65	0.18	0.15	0.89
– Physical Environment	0.39	0.35	0.49	0.064	0.47	0.41	0.81	0.39	0.52	0.54
– Cognitive Requirements	0.05	0.14	-0.24	<0.001	-0.24	-0.54	0.76	-0.20	-0.26	0.80
O*NET										
– Physical Abilities	-1.11	-1.18	-0.89	<0.001	-0.92	-0.87	-0.85	-0.91	-0.91	1.00
– Cognitive Abilities	-0.75	-0.50	-1.52	<0.001	-1.53	-1.66	-1.02	-1.12	-1.79	0.02
– Social Activities	0.34	0.50	-0.16	<0.001	-0.12	-0.49	0.70	-0.08	-0.19	0.73
Employer Accommodation										
% Received accommodation	-	-	-	-	-	-	-	100%	0%	-
– Time accommodation	-	-	-	-	-	-	-	29%	0%	-
– Equipment/transportation	-	-	-	-	-	-	-	13%	0%	-
– Change job/new skills	-	-	-	-	-	-	-	40%	0%	-
– Other help	-	-	-	-	-	-	-	43%	0%	-
No. Obs (individuals)	**18,730**	**14,101**	**4,629**	**-**	**2,489**	**1,635**	**505**	**720**	**1,662**	**-**

Notes.
1. Data Source: Health and Retirement Study 1992-2018, ORS reference year 2021 complete dataset, ONET version 26.0
2. Sample (a) – (h):
(a) HRS respondents entering panel without work disability and whose age was below 55 at the entering wave;
(b) Among respondents (a), respondents who do NOT report work-limiting health conditions before normal retirement age;
(c) Among respondents (a), respondents who reported experiencing work-limiting health conditions before normal retirement age;
(d) Among respondents (c), respondents who were employed at disability onset;
(e) Among respondents (c), respondents who were NOT employed at disability onset;
(f) Among respondents (c), respondents who were self-employed at disability onset;
(g) Among respondents (d), response who reported being accommodated at disability onset;
(h) Among respondents (d), response who reported not being accommodated at disability onset.
3. Statistical significance indicated by *p < 0.05. **p < 0.01. ***p < 0.001
4. For the list of variables used to create job indices, see appendix 1-3 (ORS), appendix 7 (O*NET)

Table 5 Summary Statistics: Work Status and Benefit Claiming after Disability Onset (as reported by survey respondents)

Variable	Employer Accommodation Status			
	(a) Employed at Disability Onset	(b) Accommodated	(c) Not accommodated	P-value
Work Status within 2 years of Disability Onset (Working for pay)				
− Working	37%	48%	34%	<0.001
− Full-time working	25%	33%	22%	<0.001
− Part-time working	12%	15%	12%	<0.001
Benefit Claiming Decision within 2 Years of Disability Onset				
Benefit Claiming	34%	31%	37%	<0.001
− Social Security Disability Insurance (SSDI)	20%	17%	22.%	<0.001
− Supplemental Security Income (SSI)	2%	1%	2%	<0.001
− Both	3%	2%	3%	<0.001
− SSDI or SSI (Either, but don't know)	8.4%	5.6%	10.1%	<0.001
No. Obs (individuals)	**2,489**	**720**	**1,662**	

Notes.
1. Data Source: Health and Retirement Study 1992-2018, ORS reference year 2021 complete dataset, ONET version 26.0
2. Sample (a) – (c):
(a) Among HRS respondents entering panel without work disability and whose age was below 55 at the entering wave, respondents who experience work-limiting health conditions before normal retirement age and were employed at disability onset
(b) Among respondents (a), response who were accommodated at disability onset;
(c) Among respondents (a), response who were NOT accommodated at disability onset.
3. Statistical significance indicated by *p < 0.05. **p < 0.01. ***p < 0.001

5. Part 2: Predicting Disability Onset during Late Working Years

In this section, I built machine learnings models to predict disability onset (dependent variable) before normal retirement ages based on demographics, health, and socioeconomic conditions, as well as job characteristics of respondents who did not report any work-limiting health conditions when they first entered the HRS survey in their early 50s (see Table 1 for a list of predictor variables used in the model). Compared to traditional statistics, machine learning models allow us to optimally use the predictive information embedded in a high-dimensional matrix of predictors (also called input variables or features), maximizing predictive power by utilizing different functional forms (polynomials, interactions, etc)[42] while also conducting separate tests with different estimation samples in order to avoid overfitting (Allen 1974, Stone 1974, Geisser 1975, Mitchell 1997).[43] The main objective was to build a model that performs well on out-of-sample data, i.e., data that was unseen before and used only to produce the prediction/forecast on it. To do so, I trained and compared three different models for binary classification, including random forests (Breiman et al., 1984; Breiman, 1996, 2001; Hastie, Tibshirani and Friedman, 2001), gradient boosted trees (Friedman, 2001, 2002; Hastie, Tibshirani and Friedman, 2001), and ℓ_1-penalized logistic regression (also called Logistic Least Absolute Shrinkage and Selection

[42] I include additional non-linear transformations of the base input variables as model inputs, up-to a third-order polynomial of all continuous variables and all second-order interactions of these and the original variables to improve performance.
[43] Overfitting is a condition that occurs when a machine learning or deep neural network model performs significantly better for training data than it does for new data.

Operator (LASSO) regression) (Tibshirani, 1996; Hastie, Tibshirani and Friedman, 2001) among many others, as they are flexible binary classification algorithms that can deal with high dimensional data and prevent overfitting (See Table 6 for a more details of each machine learning model).

Table 6 Machine Learning Models

Model	Methods
Random forests	A random forest consists of a large number of individual decision trees that operate as an ensemble i.e., it relies on the ensemble algorithm bagging (bootstrap aggregating, Breiman, 1996, 2001), where a base learner (decision tree[1)]) is fit on a with-replacement bootstrap sample of the original sample. This process is repeated multiple times, and the predictions of the base learner (each individual tree estimator) across the different bootstrap samples are then aggregated. Since individual tree estimators tend to overfit, averaging their predictions through random forests substantially reduces variance at a negligible cost of bias.
Gradient boosted trees	Gradient boosting works by sequentially adding shallow tree classifiers to the ensemble. Each new tree is fit to the residuals of the previous one, partially correcting the predecessor's errors and improving overall predictive performance. By sequentially combining models, boosting can substantially improve upon the prediction of the simple base model and explain large parts of the residual error.
ℓ_1- penalized logistic regression	Logistic Lasso (ℓ_1 −penalized logistic regression) is conceptually similar to a simple linear model and classical regression. Lasso stands for Least Absolute Shrinkage and Selection Operator. It shrinks the regression coefficients toward zero by penalizing the regression model with a penalty term (called L1-norm), which is the sum of the absolute coefficients. Since the penalty function is based on the ℓ_1 norm, some coefficients are shrunk exactly to zero, leading to a more parsimonious model. This means that, lasso can be also seen as an alternative to the subset selection methods for performing variable selection in order to reduce the complexity of the model.

Notes.
1. Source: Mitchell 1997, edited by author.
2. Tree-based models are a class of nonparametric algorithms that work by partitioning the feature space into a number of smaller (non-overlapping) regions with similar response values using a set of splitting rules. Predictions are obtained by fitting a simpler model (e.g., a constant like the average response value) in each region. Such divide-and-conquer methods can produce simple rules that are easy to interpret and visualize with tree diagrams.

To compare the performance of these models, I reserved 20% of the data as a test set[44], which were randomly selected and never used to train the models. Splitting the dataset into training and testing datasets rather than using all available data for training the model guarantees an unbiased comparison of the performance as well as generalization capability of different machine learning models since the testing data represent unseen data that were not used for training. To

[44] Holdout data is important in supervised machine learning to verify that the model that was trained and validated on historical data will produce similar performance when using new data while in operation. Holdout data should be kept separate from the training and validation data sets, and only used in the final assessment of the model's performance. This independence is important to prevent bias and to properly represent the behavior of the model with new data input going forward. https://c3.ai/glossary/data-science/holdout-data/

further avoid overfitting and ensure good out-of-sample performance, I applied 10-fold cross-validation[45] to train all three models. Once I built and trained the models, I used classification accuracy, which measures the percentage of correct predictions out of the total predictions made, as the primary performance metric, along with Area Under the Curves (AUC), which provides a natural tool to select optimal models across all thresholds of sensitivity (true positive rates) and specificity (true negative rates) to measure the performance of three different machine learning models (Fawcett 2006). The predictive performance was compared for the 20% hold-out sample.

Table 7 presents overall statistics describing the proportion of respondents who experienced disability in their working years before reaching their normal retirement age, and it shows that there is a sufficiently large number of employees (25%) who reported experiencing disability for us to estimate the predictive models.

Table 7 Outcome Variable for Predictive Model

	Respondents who reported never experiencing disability by normal retirement age	Respondents who reported experience disability by normal retirement age	Total
N	14,101	4,629	18,730 [1)]
%	75%	25%	

Notes. Out of 18,730 individuals who entered the survey without reporting a prior work disability, 14,984 individuals (80%) were used to train the model and 3,746 individuals (20%) were used to test the model performance.

Next, I present the diagnostics for the predictive models, which show how well the models allowed us to classify individuals into groups that are different in their probability of experiencing disability before their normal retirement age. Table 8 shows the results for the predictive models utilizing a full set of predictors available in Table 1 (including demographic characteristics, socioeconomic conditions, health conditions and behaviors, healthcare utilization and job characteristics), and Table 9 shows the results using a limited set of predictors, including basic demographic characteristics and socioeconomic conditions (i.e., age, education, marital status, earnings, and household assets) and job characteristics. The models in Table 9 did not utilize any prior information on respondents' health or healthcare utilization in predicting disability onset and

[45] Cross-validation is a resampling procedure used to evaluate machine learning models on a limited data sample. The procedure has a single parameter called k that refers to the number of groups that a given data sample is to be split into. As such, the procedure is often called k-fold cross-validation. When a specific value for k is chosen, it may be used in place of k in the reference to the model, such as k=10 becoming 10-fold cross-validation. Cross-validation is primarily used in applied machine learning to estimate the skill of a machine learning model on unseen data. That is, to use a limited sample in order to estimate how the model is expected to perform in general when used to make predictions on data not used during the training of the model. Summarize the skill of the model using the sample of model evaluation scores (Allen (1974), Stone (1974), Geisser (1975)).

were built to test whether these machine learning models had enough predictive power even without any information on respondents' prior health.

Table 8 Model Performance (Full Set of Predictors)

Full Set of Predictors

		Logistic LASSO		***Random Forest***		***Gradient Boosted Trees***	
		Actual		**Actual**		**Actual**	
		No Disability	Disability	No Disability	Disability	No Disability	Disability
Prediction	No Disability	2124	518	2111	531	2140	502
	Disability	112	992	90	1014	94	1010
Overall Accuracy		83%		83%		84%	
Sensitivity		95%		96%		96%	
Specificity		66%		66%		67%	
AUC		0.85		0.86		0.86	

Notes. 1. Model performance was tested for the 20% test set (N=3,746). 2. Outcome variables: Disability Onset before Normal Retirement Age. 3. Predictor variables: See Table 1. 4. AUC: Area Under the Curve

Table 9 Model Performance (Limited Set of Predictors)

Limited Set of Predictors

		Logistic LASSO		***Random Forest***		***Gradient Boosted Trees***	
		Actual		**Actual**		**Actual**	
		No Disability	Disability	No Disability	Disability	No Disability	Disability
Prediction	No Disability	2124	524	1992	677	2016	653
	Disability	182	895	92	985	11	966
Overall Accuracy		81%		79%		80%	
Sensitivity		92%		95%		95%	
Specificity		63%		59%		60%	
AUC		0.83		0.78		0.82	

Notes. 1. Model performance was tested for the 20% test set (N=3,746). 2. Outcome variables: Disability Onset before Normal Retirement Age. 3. Predictor variables: See Table 1. 4. AUC: Area Under the Curve

Overall, the best option from each class of models achieved about 80% accuracy whether they were based on a full set of predictors or a limited set of predictors (without prior information on health or healthcare utilization). This means that these models correctly classified disability onset in late working years for four out of every five people who did not report any work disability when they entered the HRS data. I found that the different models showed very similar performance. The accuracy of the gradient-boosted trees in Table 8 was 84%, relative to which the logistic LASSO or the random forest achieved about one percentage point better accuracy. In Table 9 based on a limited set of predictors, the regularized logistic performed slightly better than the ensemble methods such as random forest or gradient-boosted trees but the difference was not substantial.

However, I found that the models did not perform equally well in identifying the true positive (an outcome where the model correctly predicts the positive class, i.e., respondents who experience disability onset) or the true negative (an outcome where the model correctly predicts the negative class, i.e., respondents who do not experience disability onset). For all three models, the true positive rate (sensitivity) was very high, with about 95% of the respondents experiencing disability were correctly identified as such across the models, whereas the true negative rate (specificity) was lower, with between 66-67% of respondents not experiencing disability onset were correctly identified in Table 8 and between 59 and 63% in Table 9. Specificity was also what distinguished the different models in Table 9. While ensemble methods including random forest and gradient boosted trees had higher sensitivity rates as shown in Table 8 and Table 9, ℓ_1-penalized logistic regression (i.e., logistic LASSO) showed higher specificity and was about 3 percentage points more likely to correctly identify those who did not experience disability onset. In addition, AUC[46] is a summary measure that represents how well prediction are ranked across the true positives and true negatives, and as a rule of thumb, an AUC above 0.85 means high classification accuracy, one between 0.75 and 0.85 moderate accuracy, and one less than 0.75 low accuracy (D'Agostino et al., 2018). All three models achieved high classification accuracy when using a full set of predictors as shown in Table 8, and moderately high accuracy when using a limited set of predictors (excluding prior information on health and healthcare utilization) as shown in Table 9 based on the AUC measure.

[46] An area of 1.0 represents a model that made all predicts perfectly. An area of 0.5 represents a model as good as random.

Finally, Table 10 shows the list of variables that enhanced predictive power the most from the gradient-boosted trees, which has shown the best predictive performance when using a full set of predictors among the three models. Similarly, Table 11 shows the list of most predictive variables from the ℓ_1- penalized logistic regression, which has shown the best performance based on a limited of set of predictors without health information among the three models. I ranked the variables by the order of descending influence using model-specific Variable Importance (VI)[47] computed for the training sample. In Table 10 and Table 11, I show the relative importance[48], which is defined as the percent improvement with respect to the most important predictor, for interpretability.[49] First, looking at the 20 most influential variables that are non-job characteristics in Table 10, I found that healthcare utilization such as the number of hospital/doctor visits is the most important predictor for disability onset, which is not surprising that these are common indicators for potential health issues.[50] Following are the employees' age (61.4) and economic conditions such as earnings (53.4) and household income (35.9), and health conditions such as BMI (29.8), cognition (15.7), experience of arthritis (14.3) as well as other health-related behaviors such as amount of drinking (10.5). Among the 20 most influential job characteristics that were predictive of disability onset, a combination of physical and psychomotor activities such as low postures (kneeling (8,7), crouching (8.7), stopping (8.7)), fine manipulation (8.4) and leg movements (8.2) were identified as the most important predictors, followed by technical operations (8.2) and social activities such as "Performing for or Working Directly with the Public" (7.6) and "Communicating with Supervisors, Peers, or Subordinates" (7.6). Overall, the relative importance of job characteristics was lower than that of healthcare information and demographic characteristics.

Second, in the prediction model without the respondents' prior health or healthcare information, I found that earnings were the most predictive variable, followed by life satisfaction (24.2), key demographic variables such as race (White 20.3, Hispanic 15.8, Other race 6.1) and

[47] Variable importance refers to how much a given model "uses" that variable to make accurate predictions. The more a model relies on a variable to make predictions, the more important it is for the model (Inglis et al., 2021)

[48] Relative importance is calculated by dividing each variable importance score by the largest importance score of the variables, then multiply by 100% (Inglis et al., 2021)

[49] For instance, in Table 10, the relative importance of the hospital visit is set to 100 because it is the most important predictor for disability onset. Next, CESD scores is the second important predictor in this model and its contribution to the predictive ability is 85.7% of that of hospital visit. Therefore, the relative variable importance of CESD scores is 85.7.

[50] The fact that prior information on health or healthcare utilization has important predictive power in the machine learning methods used in this paper are important for fully harnessing that power. However, the implications of these results are unclear. It could be the case that healthcare utilization helps identify groups with more. vs. less potentials for future disabilities. It could also be the case that use of healthcare services has a causal impact on whether an individual's health condition improves and does not experience sever disability in the future. Disentangling these selection and treatment effects represents a key area of future research.

gender (13.3), and health behaviors including drinking (10.9) and smoking (8.9), as shown in Table 11. Among the job characteristics that were predictive of disability onset, handling/moving objects (14.1) and performing general physical activities (12.3) had the most predictive power, followed by other physical job activities such as crouching (9.8) and controlling legs or foots (7.9). Furthermore, similar to the model using a full set of predictors, technical operations and monitoring activities (7.1) were also predictive of disability onset. In sum, the machine learning models indicated that non-healthcare information or predictors that are not directly related to prior health conditions were still highly predictive of disability onset as indicated by the overall accuracy rates in Table 9, and physical and technical job characteristics were the most predictive characteristics that enhanced the predictive power.

Table 10 Important Predictors (Gradient Boosted Trees Using a Full Set of Predictors)

Non Job Characteristics				Job Characteristics			
	Category	**Variable**	**Relative Importance**		**Category**	**Variable**	**Relative Importance**
1	Healthcare	Hospital Visit	100.0	**1**	Physical	Kneeling (ORS)	8.7
2	Health	CESD Score	85.7	**2**	Physical	Crouching (ORS)	8.7
3	Healthcare	Doctor Visit	76.7	**3**	Physical	Stooping (ORS)	8.7
4	Demographics	Age	61.4	**4**	Cognitive	Time Sharing (O*NET)	8.4
5	SES	Earnings	53.4	**5**	Physical	Fine manipulation, one or both hands (ORS)	8.4
6	SES	Household Income	35.9	**6**	Physical	Push/Pull Feet or Legs, with Hands (ORS)	8.2
7	Health	BMI	29.8	**7**	Technical	Operations Monitoring (O*NET)	8.2
8	Health	Weight	24.9	**8**	Cognitive	Perceptual Speed (O*NET)	7.8
9	Work	Years of Working	24.4	**9**	Social	Performing for or Working Directly with the Public (O*NET)	7.6
10	Healthcare	Out-of-Pocket Expenditures	24.4	**10**	Technical	Inspecting Equipment, Structures, or Materials (O*NET)	7.6
11	Demographics	Length of Marriage	20.0	**11**	Social	Communicating with Supervisors, Peers, or Subordinates (O*NET)	7.6
12	SES	Assets (Primary Residence)	19.3	**12**	Social	Resolving Conflicts and Negotiating with Others (O*NET)	7.3
13	Others	Spouse Age	18.6	**13**	Physical	Climb ramps (structure related) (ORS)	7.3
14	SES	Debt	18.2	**14**	Physical	Hearing Remotely (ORS)	7.3
15	Health	Cognition	15.7	**15**	Cognitive	Performing Administrative Activities (O*NET)	7.3
16	Health	Ever Had Arthritis	14.3	**16**	Physical	Performing General Physical Activities (O*NET)	7.3
17	Health	Self-reported Health	13.0	**17**	Physical	Far Vision (O*NET)	7.1
18	Healthcare	Number of Private Health Insurance	10.8	**18**	Physical	Monitor Processes, Materials, or Surroundings (O*NET)	7.1
19	Health Behaviors	Amount of Drinking	10.5	**19**	Cognitive	Scheduling Work and Activities (O*NET)	6.7
20	Others	Life Satisfaction	10.1	**20**	Social	Communicating with Persons Outside Organization (O*NET)	6.7

Notes. 1. Data Source: Health and Retirement Study 1992-2018, ORS reference year 2021 complete dataset, ONET version 26.0.
2. Variable importance is computed as the number of times the variable is split on weighted by the depth of the split.
3. Outcome variables: Disability Onset before Normal Retirement Age.
4. Predictor variables: See Table 1. 4
5. Acronyms: SES (Socioeconomic Status), CESD (Center for Epidemiologic Studies Depression Scale)
6. Variable categories were prescribed by the survey programs

Table 11 Important Predictors (Logistic LASSO Model Using a Limited Set of Predictors)

Non Job Characteristics				Job Characteristics			
	Category	**Variable**	**Relative Importance**		**Category**	**Variable**	**Relative Importance**
1	SES	Earnings	100	**1**	Physical	Handling and Moving Objects (O*NET)	14.1
2	Others	Life Satisfaction	24.2	**2**	Physical	Performing General Physical Activities (O*NET)	12.3
3	Demographics	White	20.3	**3**	Physical	Crouching (ORS)	9.8
4	Work	Years of Working	18.1	**4**	Physical	Require Physical Effort (self-report) (HRS)	8.6
5	Health Behaviors	Smoking Now	18.1	**5**	Physical	Foot and Leg Controls (ORS)	7.9
6	Demographics	Hispanic	15.8	**6**	Technical	Operations Analysis (O*NET)	7.1
7	Demographics	Female	13.3	**7**	Physical	Sound Localization (O*NET)	6.5
8	Health	Health as a child	13.1	**8**	Technical	Judging the Qualities of Things, Services, or People (O*NET)	6.1
9	Work	Self-employed	11.4	**9**	Technical	Installation (O*NET)	6.0
10	Health Behaviors	Frequency of Drinking	10.9	**10**	Physical	Hearing in Person (ORS)	5.6
11	Health Behaviors	Ever Drinking	10.1	**11**	Job Control	Ability to Pause Work (ORS)	5.3
12	Health Behaviors	Ever Smoking	8.9	**12**	Technical	Controlling Machines and Processes (O*NET)	5.3
13	Health Behaviors	Amount of Drinking	8.2	**13**	Technical	Monitor Processes, Materials, or Surroundings (O*NET)	5.0
14	SES	Household Assets	7.2	**14**	Cognitive	Problem Sensitivity (O*NET)	4.9
15	Others	Length of Marriage	6.6	**15**	Social	Assisting and Caring for Others (O*NET)	4.9
16	Work	Full Time Working	6.2	**16**	Physical	Strength (ORS)	4.8
17	Demographics	Other Race	6.1	**17**	Physical	Rate Control (O*NET)	4.8
18	Demographics	Married	4.6	**18**	Physical	Extent Flexibility (O*NET)	4.7
19	Demographics	Age	4.4	**19**	Physical	Fine Manipulation (ORS)	4.4
20	SES	Education (Some College)	4.3	**20**	Physical	Push/Pull Feet or Legs, with Hands (ORS)	4.2

Notes. 1. Data Source: Health and Retirement Study 1992-2018, ORS reference year 2021 complete dataset, ONET version 26.0.
2. Variable importance is computed as the number of times the variable is split on weighted by the depth of the split.
3. Outcome variables: Disability Onset before Normal Retirement Age.
4. Predictor variables: See Table 1. 4
5. Acronyms: SES (Socioeconomic Status), CESD (Center for Epidemiologic Studies Depression Scale)
6. Variable categories were prescribed by the survey programs

6. Part 3: Effects of Employer Accommodation on Labor Supply and Benefit Claiming

Lastly, I estimated the effect of employer accommodation on labor supply and benefit claiming decisions to evaluate whether employer accommodation might be effective in preventing or slowing labor force exit and/or SSDI claiming. I used two econometric methods: Ordinary Least Squares (OLS) as a baseline model and Propensity Score Matching (PSM) to estimate casual effects. The outcome variables of interests included: 1) whether an individual is working or not in the following waves (2 years later vs. 4 years later) after disability onset (1=work, 0=otherwise); and 2) whether an individual applies for disability benefits such as Social Security Disability Benefits (and/or Supplemental Security Income) in the following waves (2 years later vs. 4 years later) after disability onset (1=claim benefits, 0=otherwise).

To be able to estimate causal relationships between employer accommodation and these outcome variables, employer accommodation should be randomly assigned to respondents, i.e., individuals do not self-sort into employers who provide accommodation. If the assignment of employer accommodation is not random and therefore individuals who receive accommodation and those who do not are systematically different[51], affecting their labor supply or benefit claiming decisions, these estimates cannot be interpretable as causal. This key assumption that treatment assignment is independent of potential outcomes (called Conditional Independence) is commonly used but fundamentally untestable (David, 1979). However, the descriptive analysis in Table 3 - Table 5 showed that except for only a few socioeconomic conditions, individuals whose employers accommodated their disabilities were not systematically different from those whose employers did not accommodate their disabilities.[52] In order to adjust any further differences between the treatment group (i.e., individuals whose employers accommodate their disabilities) and the control group (i.e., those whose employers do not accommodate their disabilities), I used propensity score matching (Rosenbaum, 1983), which estimates each individual's propensity to receive a binary treatment (in this case, employer accommodation), via a probit or logit regression, as a function of observable characteristics and matches individuals with similar propensities.[53] Essentially, this

[51] Selection bias is when participants in a program (treatment group) are systematically different from non-participants (control group) (Altonji et al, 2005)
[52] Furthermore, I conduct a test of selection bias proposed by Altonji et al. (2005) and the results indicate that the degree of selection on unobservables would have to be 9.5 times the degree of selection on observables in order for the unobservables to have the same effects as the observables. Considering the number of control variables included in my model, selection on unobservables of this magnitude is well outside the range of plausibility.
[53] Estimated propensity scores are used to reweight the distribution of covariates X in the control group to match the distribution of X observed in the treated group.

method places more weight on individuals who are not accommodated but are similar to individuals who are accommodated based on observable characteristics and less weight on unaccommodated individuals that are less similar, so that these two groups of individuals are more comparable. Therefore, one can use the reweighted control group to estimate the counterfactual distribution of the outcome for the treated group had they never been treated.

I implemented propensity score matching estimation as follows. First, I estimated the propensity score function $p(X_i)$ using a probit regression of employer accommodation (treatment) on individual and job characteristics X_i[54] observed in the wave prior to onset. I then constructed the following estimator for the average treatment effect on the treated (ATET)[55]:

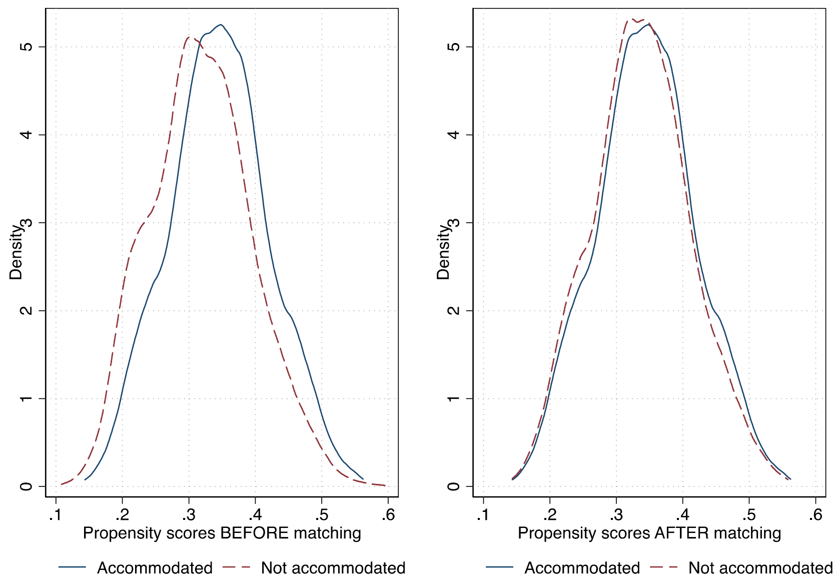
$$\Delta^{ATET} = \frac{1}{N^T} \sum_{i=1}^{N^T+N^C} \left(D_i Y_i - (1 - D_i) \frac{p(X_i)}{1 - p(X_i)} Y_i \right)$$

where $D_i = 1$ if individual i was accommodated by their employer at disability onset and $D_i = 0$ otherwise, N^T is the number of individuals who were accommodated and N^C the number of individuals who were not accommodated, and Y is the outcome of interest (i.e., labor supply and benefit claiming decisions in the following waves).

Propensity score methods typically assume a common support (also called overlap) (Imbens, 2004)[56], i.e., the range of propensities to be treated must be the same or similar for treated and control units even if the density functions have quite different shapes. To test this assumption, Figure 1 shows the distribution of propensity scores by accommodation status before and after propensity score matching. As expected, a substantial portion of the accommodated and unaccommodated group overlapped after matching, and the PSM reduced the difference between these two groups and the distributions overlap substantially indicating good covariate balance.

[54] Control variables include: gender, age groups, education, race, age difference with the spouse, indicator for whether in a couple, indicator for whether the spouse is working, indicator for poor health, cognitive test scores, annual wage (log), type of employer-sponsored pension plan (DB, DC, or DB/DC), existence of employer-provided health insurance (respondent and spouse), and time fixed effects.
[55] ATET is an average effect of some treatment on the group of individuals that received treatment (as opposed to, for example, the effect of the treatment averaged across all individuals in a study regardless of whether or not they received the treatment). Previous research has shown that this estimator in the equation is a consistent estimator of ATET (Dehejia and Wahba, 1999; DiNardo, Fortin and Lemieux, 1996).
[56] Overlap assumptions states that all individuals have a positive probability of receiving treatment and anyone who would always or never get the treatment should not be included in the study. Another way of stating this is that the treated and untreated distributions need to overlap. Intuitively, if the treated and untreated individuals do not overlap, it means they are very different, and we will not be able to extrapolate the effect of one group to the other (e.g., testing a new drug in an experiment where only men receive the treatment and then assume women will respond to it equally well)

Figure 1 Distribution of Propensity Score: Before vs. After Matching



Notes. Figure 1 shows each individual's propensity to receive a binary treatment (employer accommodation), as a function of observable characteristics and matches individuals with similar propensities.

Finally, Table 12-14 present estimates of the effects of employer accommodation on labor supply outcomes and benefit claiming decisions in the following waves after disability onset (in 2 years vs. 4 years), using both OLS[57] and PSM. In Table 12, employer accommodation at disability onset increased the probability that an individual was working in 2 years (i.e., the first wave the respondent reports that his health limits his ability to work in some way) following disability onset overall. However, the effect size was bigger after implementing PSM, with a 13.6 percentage points increase over the work participation rate at disability onset compared to a 10.9 percentage points increase when using OLS. The effect size reduced to 7.7 percentage points two years later (i.e., up to four years after disability onset) in PSM and 5.3 percentage points in OLS, which were still statistically significant, suggesting that employer accommodation had lasting effects on labor supply decisions of newly disabled workers although the effect size diminished over time. Furthermore, a more detailed breakdown by the type of accommodation indicates that workers who receive accommodations related to changing job tasks or learning new skills were 19.8 percentage points more likely to work within 2 years following disability onset than workers who receive no accommodation at all based on the PSM. However, time accommodation such as

[57] Same outcome and control variables used in the OLS and PSM models.

shortening work days, allowing arrival or departure change, and allowing more breaks or rest periods as well as getting special equipment or someone to help the workers maintained statistically significant results four year later after disability onset.

Table 12 Effects of Employer Accommodation on Labor Supply in 2 years vs. 4 years

	Working in 2 Years after Disability Onset		**Working in 4 Years after Disability Onset**	
	(1) OLS [1]	(2) PSM[2]	(3) OLS	(4) PSM
Employer Accommodation at Disability Onset (All)	0.109*** (0.011)	0.136*** (0.031)	0.053*** (0.011)	0.077** (0.032)
- Shorten work days	0.060*** (0.017)	0.140*** (0.051)	0.023 (0.018)	0.100* (0.053)
- Allow arrival or departure change	0.089*** (0.015)	0.178*** (0.043)	0.007 (0.015)	0.081* (0.045)
- Allow more breaks or rest periods	0.094*** (0.015)	0.160*** (0.041)	0.035** (0.015)	0.095** (0.043)
- Arrange special transportation	0.144*** (0.036)	0.150 (0.107)	0.030 (0.036)	0.109 (0.111)
- Get special equipment for job	0.067*** (0.019)	0.168*** (0.054)	0.037* (0.019)	0.122** (0.056)
- Change the job to something they could do	0.141*** (0.016)	0.198*** (0.044)	0.041** (0.016)	0.034 (0.046)
- Help learn new skills	0.103*** (0.020)	0.163*** (0.056)	-0.005 (0.020)	0.025 (0.058)
- Get someone to help you	0.075*** (0.015)	0.152*** (0.042)	0.090*** (0.015)	0.153*** (0.043)
- Assist you in receiving rehabilitative services	0.012 (0.020)	0.085 (0.056)	-0.012 (0.020)	0.062 (0.057)
Observations[3]	2,309		2,169	

Notes. 1) Ordinary Least Squares; 2) Propensity Score Matching; 3) The Sample is HRS respondents who reported a new work-limiting health condition before their normal retirement age, who were employed at disability onset (excluding the self-employed). Columns (1), (3) reports OLS estimates and columns (2), (4) reports estimates after propensity score matching. Control variables are from the HRS and include: gender, age groups, education, race, age difference with the spouse, indicator for whether in a couple, indicator for whether the spouse is working, indicator for poor health, cognitive test scores, annual wage (log), type of employer-sponsored pension plan (Defined Benefits (DB), Defined Contribution (DC), or DB/DC), existence of employer-provided health insurance (respondent and spouse), and time fixed effects. Statistical significance indicated by *p < 0.05. **p < 0.01. ***p < 0.001. Standard errors are reported in parenthesis

Next, Table 13 present the effects of employer accommodation on claiming SSDI benefits and Table 14 shows the results for both SSDI and SSI combined. The results showed a significant but relatively small effects (2 percentage points) of employer accommodation on benefit claiming within two years of disability onset. Overall, employer accommodation at disability onset decreased the probability that an individual claims SSDI in 2 years after disability onset by about 2 percentage points in both OLS and PSM. However, the effects disappeared in the subsequent wave, i.e., within four years of disability onset, and I found no evidence that employer accommodation reduced subsequent SSDI or SSI applications.

Table 13 Effects of Employer Accommodation on Benefit Claiming in 2 years vs. 4 years (SSDI Only)

	Claim Benefit in 2 Years after Disability Onset		**Claim Benefit in 4 Years after Disability Onset**	
	(1) OLS [1]	(2) PSM[2]	(3) OLS	(4) PSM
Employer Accommodation at Disability Onset (All)	-0.019*** (0.004)	-0.020*** (0.004)	-0.008* (0.005)	-0.015 (0.014)
- Shorten work days	-0.025*** (0.007)	-0.022*** (0.007)	0.002 (0.008)	-0.013 (0.023)
- Allow arrival or departure change	-0.025*** (0.006)	-0.024*** (0.006)	-0.011* (0.007)	-0.017 (0.020)
- Allow more breaks or rest periods	-0.016*** (0.006)	-0.013** (0.006)	0.003 (0.006)	-0.004 (0.019)
- Arrange special transportation	-0.038** (0.015)	-0.040*** (0.015)	-0.035** (0.016)	-0.050 (0.053)
- Get special equipment for job	-0.044*** (0.008)	-0.043*** (0.008)	0.013 (0.008)	0.005 (0.024)
- Change the job to something they could do	-0.026*** (0.007)	-0.021*** (0.006)	0.008 (0.007)	0.014 (0.020)
- Help learn new skills	-0.027*** (0.008)	-0.020** (0.008)	0.015* (0.009)	0.025 (0.025)
- Get someone to help you	-0.031*** (0.006)	-0.024*** (0.006)	0.001 (0.007)	0.005 (0.019)
- Assist you in receiving rehabilitative services	-0.046*** (0.008)	-0.040*** (0.008)	0.023*** (0.008)	0.027 (0.025)
Observations	2,309		2,169	

Notes. 1) Ordinary Least Squares; 2) Propensity Score Matching; 3) The Sample is HRS respondents who reported a new work-limiting health condition before their normal retirement age, who were employed at disability onset (excluding the self-employed). Columns (1), (3) reports OLS estimates and columns (2), (4) reports estimates after propensity score matching. Control variables are from the HRS and include: gender, age groups, education, race, age difference with the spouse, indicator for whether in a couple, indicator for whether the spouse is working, indicator for poor health, cognitive test scores, annual wage (log), type of employer-sponsored pension plan (Defined Benefits (DB), Defined Contribution (DC), or DB/DC), existence of employer-provided health insurance (respondent and spouse), and time fixed effects. Statistical significance indicated by *p < 0.05. **p < 0.01. ***p < 0.001. Standard errors are reported in parenthesis

Table 14 Effects of Employer Accommodation on Benefit Claiming in 2 years vs. 4 years (SSDI or SSI)

	Claim Benefit in 2 Years after Disability Onset		**Claim Benefit in 4 Years after Disability Onset**	
	(1) OLS [1]	(2) PSM[2]	(3) OLS	(4) PSM
Employer Accommodation at Disability Onset (All)	-0.019*** (0.005)	-0.021*** (0.005)	-0.003 (0.005)	-0.008 (0.014)
- Shorten work days	-0.015** (0.007)	-0.013* (0.007)	0.007 (0.008)	-0.004 (0.024)
- Allow arrival or departure change	-0.026*** (0.006)	-0.026*** (0.006)	-0.017** (0.007)	-0.020 (0.020)
- Allow more breaks or rest periods	-0.013** (0.006)	-0.010 (0.006)	0.003 (0.007)	0.001 (0.019)
- Arrange special transportation	-0.040*** (0.015)	-0.043*** (0.015)	-0.040** (0.016)	-0.054 (0.055)
- Get special equipment for job	-0.047*** (0.008)	-0.046*** (0.008)	0.007 (0.008)	0.002 (0.025)
- Change the job to something they could do	-0.019*** (0.007)	-0.015** (0.007)	0.011 (0.007)	0.020 (0.021)
- Help learn new skills	-0.027*** (0.008)	-0.019** (0.008)	0.009 (0.009)	0.022 (0.026)
- Get someone to help you	-0.035*** (0.006)	-0.027*** (0.006)	-0.005 (0.007)	0.001 (0.019)
- Assist you in receiving rehabilitative services	-0.049*** (0.008)	-0.043*** (0.008)	0.017* (0.009)	0.024 (0.026)
Observations	2,309		2,169	

Notes. 1) Ordinary Least Squares; 2) Propensity Score Matching; 3) The Sample is HRS respondents who reported a new work-limiting health condition before their normal retirement age, who were employed at disability onset (excluding the self-employed). Columns (1), (3) reports OLS estimates and columns (2), (4) reports estimates after propensity score matching. Control variables are from the HRS and include: gender, age groups, education, race, age difference with the spouse, indicator for whether in a couple, indicator for whether the spouse is working, indicator for poor health, cognitive test scores, annual wage (log), type of employer-sponsored pension plan (Defined Benefits (DB), Defined Contribution (DC), or DB/DC), existence of employer-provided health insurance (respondent and spouse), and time fixed effects. Statistical significance indicated by *p < 0.05. **p < 0.01. ***p < 0.001. Standard errors are reported in parenthesis

7. Conclusion

In this paper, I used data on newly disabled workers[58] from the publicly available survey data[59], including the Health and Retirement Study, Occupational Requirement Survey, and Occupational Information Network Survey (O*NET) to provide a machine learning model to identify workers and job characteristics that are predictive of disability onset in later working years (before the normal retirement age) and present new evidence on the short and long-term effects of employer accommodation on labor supply and benefit claiming behaviors of newly disabled workers.

First, the descriptive analysis showed that the average number of years of education was significantly higher for individuals who never experienced a work disability ("No disability group") than individuals who experienced a work disability before their normal retirement age ("Disability group"), and the "No disability group" also had fewer female respondents, a higher marriage rate, and significantly higher household assets and earnings. They also reported a significantly higher number of chronic conditions already in their early 50s compared to the "Disability group", including high blood pressure, diabetes, and arthritis, and their occupations were reported to be more physically demanding, with greater exposure to physically demanding work environment, but less cognitively or socially burdensome. While the "Disability group" were at a more socioeconomic disadvantage compared with the "No disability group", however, I found no evidence that respondents whose employers accommodated their disabilities and those whose employers did not were significantly different. The analysis showed that, except for some minor differences in education, earnings, and the number of chronic conditions in their early 50s, accommodated individuals were not systematically different from unaccommodated individuals in most observable characteristics, such as demographics, health conditions and occupational characteristics. What is unknown, however, is the extent to which firm and job characteristics, worker characteristics and their disability types explain the variation in provision of disability accommodation at the local level (e.g., variation within state, city, company, etc), as well as heterogeneity in accommodation rates by industry, which could be further investigated using a rich administrative database in the future research.

---

[58] Individuals who were not disabled when they enter the panel but who subsequently report a work disability that began when they were employed

[59] 1992–2018 waves of the RAND HRS Longitudinal File, merged with the ORS and O*NET measures of job characteristics and occupational characteristics using the restricted version of the HRS Industry and Occupation Data, which includes occupation information at the 4-digit Census code level (equivalent to 6-digit Standard Occupational Classification)

Second, the results from the machine learning models suggest that relying on publicly available survey data such as HRS, ORS and O*NET and machine learning techniques can be useful in predicting disability onset of older workers in their later life. I found that, while prior information on employees' health or healthcare utilization enhanced the predictive performance of the models overall, a combination of key demographic and socioeconomic variables, along with additional information on job characteristics and work environment, was powerful enough for identifying individuals who were likely to have disability with about 80 percent accuracy. The ability to do this could be valuable for identifying workers at risk of experiencing work limiting health conditions and targeting resources for accommodation to workers and occupations where the probability of developing disability is higher for workers with certain socioeconomic conditions or in a job that requires more physical activities. Furthermore, performing a more detailed analysis on what specific job characteristics matter more for which occupations would allow us to better target higher-risk occupations and industries.

Finally, I found that workers who experienced disability onset and were accommodated by their employers were 13.6 percentage points more likely to stay in the labor force in 2 years after the disability onset, and 7.7 percentage points more by the next survey wave, i.e., up to four years after onset. Although most types of accommodation evaluated were effective (except for arranging special transportation and assisting workers in receiving rehabilitative services), I found that accommodations involving a work change (i.e., changing the job to something they could do) were most effective in delaying retirement in the wave immediately following, suggesting this was a relatively more effective form of accommodation in the short term. However, accommodations involving a time change and allowing more flexibility (i.e., shortening workdays, allowing arrival or departure change) showed more lasting effects four year later after disability onset in keeping workers in the labor force. Employer accommodation had also short-term effects of reducing subsequent SSDI or SSI applications, although the effect size was small at about 2 percentage points, and I found no evidence that accommodation reduced SSDI or SSI applications in the long-term. These results imply that, if accommodation rates can be increased, more workers would remain in the labor force, at least temporarily. However, the study shows that whether the individuals apply for disability benefits or not is less likely to be affected by employer accommodation compared with the effect on deciding to continue working or not and therefore encouraging employer accommodation of disabilities is less likely to affect the growing number of SSDI beneficiaries.

References

Agree, Emily M., Douglas A.W., 2017. "Disability Measurement in the Health and Retirement Study." *Forum for health economics & policy* 21.1 (2017): n. pag. Web.

Alan et al., 2022. Visualizing Variable Importance and Variable Interaction Effects in Machine Learning Models, *Journal of Computational and Graphical Statistics*, 31:3, 766-778, DOI: 10.1080/10618600.2021.2007935

Altonji, J.G., Elder, T.E., Taber, C.R., 2005. An evaluation of instrumental variable strategies for estimating the effects of catholic schooling. *J. Hum. Resour*. 40 (4), 791–821.

Barrett-Connor E, Wingard D, Wong N, Goldberg R. 2018. Chapter 18: Heart disease and diabetes (PDF, 1.07 MB) . In: Cowie CC, Casagrande SS, Menke A, et al, eds. Diabetes in America, 3rd ed. NIH Pub No. 17-1468. National Institutes of Health; 2018:18.1–18.30

Belbase, Anek, Geoffrey T Sanzenbacher, and Christopher M Gillis. 2015. Beyond Blue and White Collar: Age-Related Decline, Occupation, and Retirement Timing.

Ben-Shalom Y, Schimmel Hyde J. Opportunities for Early Intervention to Avoid Prolonged Work Disability: Introduction to the Special Section. *J Occup Rehabil*. 2018 Dec;28(4):569-573.PMID: 30298239.

Blekesaune, M., & Solem, P. E., 2005. Working Conditions and Early Retirement: A Prospective Study of Retirement Behavior. Research on Aging, 27(1), 3–30. https://doi.org/10.1177/0164027504271438

Bruce, D. M, Mok, Wallace K.C., 2019. Disability, earnings, income and consumption, Journal of Public Economics, Volume 171, 2019, Pages 51-69, ISSN 0047-2727, https://doi.org/10.1016/j.jpubeco.2018.06.011

Burkhauser, R.V., Stapleton, D.C., 2003. The Decline in Employment of People with Disabilities: A Policy Puzzle. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

D'agostino, J.V., Rodgers, E., & Mauck, S.A., 2018. Addressing Inadequacies of the Observation Survey of Early Literacy Achievement. *Reading Research Quarterly, 53*, 51-69.

David H.A, Mark G. D., 2003. The Rise in the Disability Rolls and the Decline in Unemployment, *The Quarterly Journal of Economics*, Volume 118, Issue 1, February 2003, Pages 157–206.

Dehejia, R.H., Wahba, S., 1999. Causal effects in nonexperimental studies: reevaluating the evaluation of training programs. *J. Am. Stat. Assoc*. 94, 1053–1062.

Dawid, A. P., 1979. "Conditional Independence in Statistical Theory". *Journal of the Royal Statistical Society*, Series B. 41 (1): 1–31.

DiNardo, J., Fortin, N.M., Lemieux, T., 1996. Labor market institutions and the distribution of wages, 1973-1992: a semiparametric approach. *Econometrica* 64, 1001–1044.

Dushi, I., Rupp, K., 2013. Disability Shocks Near Retirement Age and Financial Well-Being. Social Security Bulletin, Vol. 73, No. 3, 2013

Fawcett, T, 2006. "An Introduction to ROC Analysis" (PDF). Pattern Recognition Letters. 27 (8): 861–874. doi:10.1016/j.patrec.2005.10.010.

Imbens, G.W., 2015. Matching methods in practice: three examples. *J. Hum. Resour*. 50 (2), 373–419.

Geisser, S, 1975. The predictive sample reuse method with applications. J Am Stat Assoc. 1975, 70 (350): 320-328. 10.1080/01621459.1975.10479865.

Johnson, R. W., Mermin, G. B. T., Murphy, D., 2007. The impact of late-career health and employment shocks on Social Security and other wealth. *CRRB Working Paper* WP 2007-26. Center for Retirement Research at Boston College.

Johnson, R. W, Mermin, G. B. T., Matthew R., 2011. Job Demands and Work Ability at Older Ages. *Journal of Aging & Social Policy* 23 (2): 101–18.

Jonathan G., Jeffrey D. K., Disability insurance rejection rates and the labor supply of older workers, *Journal of Public Economics*, Volume 64, Issue 1, 1997, Pages 1-23, ISSN 0047-2727, https://doi.org/10.1016/S0047-2727(96)01590-3.

Lahelma E, Laaksonen M, Lallukka T, Martikainen P, Pietiläinen O, Saastamoinen P, Gould R, Rahkonen O. Working conditions as risk factors for disability retirement: a longitudinal register linkage study. BMC Public Health. 2012 Apr 26;12:309. doi: 10.1186/1471-2458-12-309. PMID: 22537302; PMCID: PMC3438015.

Nanayakkara N, Curtis A.J, Heritier S, et al. 2021. Impact of age at type 2 diabetes mellitus diagnosis on mortality and vascular complications: systematic review and meta-analyses. Diabetologia. 2021; 64(2):275–287. doi:10.1007/s00125-020-05319-w

Nichols et al., 2020. Synthesis of Evidence about Stay-at-Work/Return to-Work (SAW/RTW) and Related Programs. Abt Associates. Prepared for: Chief Evaluation Office and Disability Employment Policy.

Lopez Garcia, I., Maestas, N., and Mullen, K.J., 2019. "Latent Work Capacity and Retirement Expectations." Ann Arbor, MI. University of Michigan Retirement and Disability Research Center (MRDRC) Working Paper; MRDRC WP 2019-400. https://mrdrc.isr.umich.edu/publications/papers/pdf/wp400.pdf

Maestas, N., Mullen, K.J., Powell,D., Wachter,T.V., Wenger, J.B., 2017. Working Conditions in the United States Results of the 2015 American Working Conditions Survey. RAND Report.

Maestas, N., Mullen, K.J., Strand, A., 2013. Does Disability Insurance Receipt Discourage Work? Using Examiner Assignment to Estimate Causal Effects of SSDI Receipt. *American Economic Review*. 103 (5), 1797–1829.

Mitchell, T, 1997. Machine Learning. New York: McGraw Hill. ISBN 0-07-042807-7. OCLC 36417892.

Moonesinghe R, Yang Q, Zhang Z, Khoury MJ. 2019. Prevalence and cardiovascular health impact of family history of premature heart disease in the United States: analysis of the National Health and Nutrition Examination Survey, 2007–2014. Journal of the American Heart Association. 2019;8(14):e012364. doi:10.1161/JAHA.119.012364

Hill, M.J., Maestas, N., Mullen, K.J., 2016. Employer accommodation and labor supply of disabled workers, *Labour Economics*, Volume 41, Pages 291-303, ISSN 0927-5371.

Imbens, G.W., 2014. Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review. *Review of Economics and Statistics*. 2004; 86(1): 4–29.

Leuven, E, Sianesi, B., 2003. "PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing"

Rosenbaum, P. R., Rubin, D.B., 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects". *Biometrika*. 70 (1): 41–55. doi:10.1093/biomet/70.1.41.

Schimmel J., Stapleton D.C., 2017. Using the Health and Retirement Study for Disability Policy Research: A Review. *Forum for Health Economics and Policy*; 20(2). PMID: 31419905.

Schimmel J., Stapleton D.C., 2012. The Financial Repercussions of New Work-Limiting Health Conditions for Older Workers. *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*. 2012; 49(2):141-163. doi:10.5034/inquiryjrnl_49.02.04

Schimmel, J., Wu, A. Y., & Livermore, G., 2022. Responding to Disability Onset in the Late Working Years: What do Older Workers do? *Research on Aging*.

Smith, J. P., 2003. Consequences and Predictors of New Health Events. NBER working paper. No. 10063. National Bureau of Economic Research.

Stapleton et al., 2015. Targeting Early Intervention to Workers Who Need Help to Stay in the Labor Force. https://www.dol.gov/sites/dolgov/files/ODEP/research/saw-rtw/TargetingEarlyIntervention_FINAL_2015.pdf

Stone, M, 1974. Cross-validatory choice and assessment of statistical predictions. J Roy Stat Soc B Met. 1974, 36, No. 2: 111-147.

Suijkerbuijk. Y.B, Schaafsma, F.G, van Mechelen, J.C, Ojajärvi, A, Corbière, M, Anema, J.R. 2017. Interventions for obtaining and maintaining employment in adults with severe mental illness, a network meta-analysis. Cochrane Database Syst Rev. 2017 Sep

12;9(9):CD011867. doi: 10.1002/14651858.CD011867.pub2. PMID: 28898402; PMCID: PMC6483771.

Susan C., Wilbert V.D.K., 2008. The work disincentive effects of the disability insurance program in the 1990s, *Journal of Econometrics*, Volume 142, Issue 2, 2008, Pages 757-784, ISSN 0304-4076, https://doi.org/10.1016/j.jeconom.2007.05.016.

Pikula A, Howard BV, Seshadri S. 2018. Chapter 19: Stroke and diabetes (PDF, 493.93 KB). In: Cowie CC, Casagrande SS, Menke A, et al, eds. Diabetes in America, 3rd ed. NIH Pub No. 17-1468. National Institutes of Health; 2018:19.1–19.23

Wong ND, Zhao Y, Patel R, et al. 2016. Cardiovascular risk factor targets and cardiovascular disease event risk in diabetes: a pooling project of the Atherosclerosis Risk in Communities Study, Multi-Ethnic Study of Atherosclerosis, and Jackson Heart Study. Diabetes Care. 2016;39(5):668–676. doi:10.2337/dc15-2439

Wu, A. Y., & Schimmel H, J., 2019. The Postretirement Well-Being of Workers With Disabilities. *Journal of Disability Policy Studies*, *30*(1), 46-55. https://doi.org/10.1177/1044207318793161

Appendices

Appendix 1: Types of variables for Each Physical Job Requirement as Reported by Survey Respondents 1)

		Type of Variable		
Name of Job Requirement		(1) Percent of Workers Job Characteristics Required 2)	(2) Frequency: Category levels 3)	(3) Duration: Mean/Percentile 4)
1	Gross manipulations	X4)	X	-
2	Fine manipulation	X	X	-
3	Foot or leg controls	X	X	-
4	Standing	-	-	X
5	Sitting	-	-	X
6	Keyboarding	X	X	-
7	Speaking	X	X	-
8	Lifting and carrying	-	-	X
9	Driving	X	-	-
10	Climbing			
	Structural ramps or stairs	X	X	-
	Work-related ramps or stairs	X	X	-
	Ladders, ropes, or scaffolds	X	X	-
11	Low postures	X	X	
	Crawling	X	X	-
	Crouching	X	X	-
	Stooping	X	X	-
	Kneeling	X	X	-
12	Reaching			
	Reaching at or below the shoulder	X	X	-
	Reaching overhead	X	X	-
13	Pushing and pulling			
	With feet/legs	X	X	-
	With hands/arms	X	X	-
14	Strength level			
	Sedentary	X	-	-
	Light work	X	-	-
	Medium work	X	-	-
	Heavy work	X	-	-
	Very heavy work	X	-	-
15	Vision			
	Far	X	-	-
	Near	X	-	-
	Peripheral	X	-	-
16	Hearing			
	In person speech	X	-	-
	Remote speech	X	-	-
	Telephone	X	-	-
	Other sounds	X	-	-

Notes. 1) Data Source: ORS reference year 2021 complete dataset (390 occupations), 2) % of workers for whom the job requirement is required (1=the job requirement is required, 0=otherwise), 3) Frequency of the job requirement. Frequency levels include: not required/seldom/occasionally/frequently/constantly; 4) Duration of the job requirement. Expressed in mean and percentiles: 10th, 25th, 50th, 75th, and 90th percentiles (categorical variable); 5) "X" indicates that the type of variable is available for the job requirement. "-" indicates that the type of variable is not available for the job requirement.

Appendix 2: Types of Variables for Each Environmental Working Condition Reported by Survey Respondents [1)]

Name of Job Requirement		Type of Variable		
		(1) Percent of Workers Job Characteristics Required [2)]	(2) Frequency: Category levels [3)]	(3) Duration: Mean/Percentile [4)]
1	Humidity	X[4)]	X	-
2	Extreme cold	X	X	-
3	Extreme heat	X	X	-
4	Heavy vibrations	X	X	-
5	High, exposed places	X	X	-
6	Hazardous contaminants	X	X	-
7	Proximity to moving mechanical parts	X	X	-
8	Wetness	X	X	-
9	Outdoors	X	X	-
10	Noise [6)]	X	-	-

Notes. 1) Data Source: ORS reference year 2021 complete dataset (390 occupations), 2) % of workers for whom the job requirement is required (1=the job requirement is required, 0=otherwise), 3) Frequency of the job requirement. Frequency levels include: not required/seldom/occasionally/frequently/constantly; 4) Duration of the job requirement. Expressed in mean and percentiles: 10th, 25th, 50th, 75th, and 90th percentiles (categorical variable); 5) "X" indicates that the type of variable is available for the job requirement. "-" indicates that the type of variable is not available for the job requirement. 6) The percentage of workers exposed to noise is categorized by three levels of intensity levels: "quiet", "moderate", and "loud".

Appendix 3: Types of Variables for Each Cognitive and Mental Requirements Reported by Survey Respondents [1)]

Name of Job Requirement		Type of Variable		
		(1) Percent of Workers Job Characteristics Required [2)]	(2) Frequency: Category levels [3)]	(3) Duration: Mean/Percentile [4)]
1	Interaction with general public	X[4)]	-	-
2	Working around crowd	X	-	-
3	Supervisory duties	X	-	-
4	Supervisor is present	X	-	-
5	Basic people skills	X	-	-
6	Telework available	X	-	-
7	Ability to pause work	X	-	-
8	Control of workload	X	-	-
9	Communicating verbally	X	X	-
10	Work reviewed by supervisor	X	X	-
11	Problem solving	X	X	-
12	Work pace	X	X	-

Notes. 1) Data Source: ORS reference year 2021 complete dataset (390 occupations), 2) % of workers for whom the job requirement is required (1=the job requirement is required, 0=otherwise), 3) Frequency of the job requirement. Frequency levels include: not required/seldom/occasionally/frequently/constantly; 4) Duration of the job requirement. Expressed in mean and percentiles: 10th, 25th, 50th, 75th, and 90th percentiles (categorical variable); 5) "X" indicates that the type of variable is available for the job requirement. "-" indicates that the type of variable is not available for the job requirement. 6) The percentage of workers exposed to noise is categorized by three levels of intensity levels: "quiet", "moderate", and "loud".

Appendix 4: Percent of Occupations Observed for Physical Job Requirements Reported by Survey Respondents

	Name of Job Requirements	(1) Percentage of Occupations Observed (%) [2]
1	Gross manipulation	100%
2	Fine manipulation	99%
3	Foot or leg controls	92%
4	Standing	82%
5	Sitting	79%
6	Keyboarding	97%
7	Verbal communication	97%
8	Lifting and carrying	81%
9	Driving	79%
10	Climbing	
	Structural ramps or stairs	80%
	Work-related ramps or stairs	93%
	Ladders, ropes, or scaffolds	95%
11	Low postures	
	Crawling	93%
	Crouching	87%
	Stooping	87%
	Kneeling	89%
12	Reaching	
	Reaching at or below the shoulder	93%
	Reaching overhead	84%
13	Pushing and pulling	
	With feet/legs	92%
	With hands/arms	91%
14	Strength level	50%
15	Vision	
	Far	83%
	Near	92%
	Peripheral	80%
16	Hearing	
	In person speech	77%
	Remote speech	80%
	Telephone	75%
	Other sounds	76%

Notes. 1) Data Source: ORS reference year 2021 complete dataset (390 occupations), 2) Percentage of occupations for which the data is available in the survey (out of 390 occupations)

Appendix 5: Percent of Occupations Observed for Environmental Working Conditions Reported by Survey Respondents

	Name of Job Characteristics	Percentage of Occupations Observed (%)[2)]
1	Humidity	98%
2	Extreme cold	99%
3	Extreme heat	99%
4	Heavy vibrations	98%
5	High, exposed places	97%
6	Hazardous contaminants	96%
7	Proximity to moving mechanical parts	96%
8	Wetness	96%
9	Outdoors	97%
10	Noise [3)]	99%

Notes. 1) Data Source: ORS reference year 2021 complete dataset (390 occupations), 2) Percentage of occupations for which the data is available in the survey (out of 390 occupations), 3) The percentage of workers exposed to noise is categorized by three levels of intensity levels: "quiet", "moderate", and "loud".

Appendix 6: Percent of Occupations Observed for Cognitive and Mental Requirements Reported by Survey Respondents

	Name of Job Characteristics	Percentage of Occupations Observed (%) [2)]
1	Interaction with general public	79%
2	Working around crowd	77%
3	Supervisory duties	98%
4	Supervisor is present	81%
5	Basic people skills	84%
6	Telework available	81%
7	Ability to pause work	80%
8	Control of workload	48%
9	Communicating verbally	46%
10	Work reviewed by supervisor	47%
11	Problem solving	39%
12	Work pace	64%

Notes. 1) Data Source: ORS reference year 2021 complete dataset (390 occupations), 2) Percentage of occupations for which the data is available in the survey (out of 390 occupations), 3) Job characteristics with sample size below 50% are not included in the analysis (control of workload, communicating verbally, work reviewed by supervisor, problem solving).

Appendix 7: O*NET Measures of Abilities, Skills, and Work Activities

	O*NET Module	Variable label
1	Cognitive abilities	Oral Comprehension
2	Cognitive abilities	Written Comprehension
3	Cognitive abilities	Oral Expression
4	Cognitive abilities	Written Expression
5	Cognitive abilities	Fluency of Ideas
6	Cognitive abilities	Originality
7	Cognitive abilities	Problem Sensitivity
8	Cognitive abilities	Deductive Reasoning
9	Cognitive abilities	Inductive Reasoning
10	Cognitive abilities	Information Ordering
11	Cognitive abilities	Category Flexibility
12	Cognitive abilities	Mathematical Reasoning
13	Cognitive abilities	Number Facility
14	Cognitive abilities	Memorization
15	Cognitive abilities	Speed of Closure
16	Cognitive abilities	Flexibility of Closure
17	Cognitive abilities	Perceptual Speed
18	Cognitive abilities	Spatial Orientation
19	Cognitive abilities	Visualization
20	Cognitive abilities	Selective Attention
21	Cognitive abilities	Time Sharing
22	Psychomotor abilities	Arm-Hand Steadiness
23	Psychomotor abilities	Manual Dexterity
24	Psychomotor abilities	Finger Dexterity
25	Psychomotor abilities	Control Precision
26	Psychomotor abilities	Multi-limb Coordination
27	Psychomotor abilities	Response Orientation
28	Psychomotor abilities	Rate Control
29	Psychomotor abilities	Reaction Time
30	Psychomotor abilities	Wrist-Finger Speed
31	Psychomotor abilities	Speed of Limb Movement
32	Physical abilities	Static Strength
33	Physical abilities	Explosive Strength
34	Physical abilities	Dynamic Strength
35	Physical abilities	Trunk Strength
36	Physical abilities	Stamina
37	Physical abilities	Extent Flexibility
38	Physical abilities	Dynamic Flexibility
39	Physical abilities	Gross Body Coordination
40	Physical abilities	Gross Body Equilibrium
41	Sensory abilities	Near Vision
42	Sensory abilities	Far Vision

43	Sensory abilities	Visual Color Discrimination
44	Sensory abilities	Night Vision
45	Sensory abilities	Peripheral Vision
46	Sensory abilities	Depth Perception
47	Sensory abilities	Glare Sensitivity
48	Sensory abilities	Hearing Sensitivity
49	Sensory abilities	Auditory Attention
50	Sensory abilities	Sound Localization
51	Sensory abilities	Speech Recognition
52	Sensory abilities	Speech Clarity
53	Work activities - Interacting with Others	Interpreting the Meaning of Information for Others
54	Work activities - Interacting with Others	Communicating with Supervisors, Peers, or Subordinates
55	Work activities - Interacting with Others	Communicating with Persons Outside Organization
56	Work activities - Interacting with Others	Establishing and Maintaining Interpersonal Relationships
57	Work activities - Interacting with Others	Assisting and Caring for Others
58	Work activities - Interacting with Others	Selling or Influencing Others
59	Work activities - Interacting with Others	Resolving Conflicts and Negotiating with Others
60	Work activities - Interacting with Others	Performing for or Working Directly with the Public
61	Work activities - Interacting with Others	Coordinating the Work and Activities of Others
62	Work activities - Interacting with Others	Developing and Building Teams
63	Work activities - Interacting with Others	Training and Teaching Others
64	Work activities - Interacting with Others	Guiding, Directing, and Motivating Subordinates
65	Work activities - Interacting with Others	Coaching and Developing Others
66	Work activities - Interacting with Others	Provide Consultation and Advice to Others
67	Work activities - Interacting with Others	Staffing Organizational Units
68	Work activities - General	Monitoring Processes, Materials, or Surroundings
69	Work activities - General	Identifying Objects, Actions, and Events
70	Work activities - General	Inspecting Equipment, Structures, or Materials
71	Work activities - General	Estimating the Quantifiable Characteristics of Products, Events, or Information
72	Work activities - General	Judging the Qualities of Objects, Services, or People
73	Work activities - General	Processing Information
74	Work activities - General	Evaluating Information to Determine Compliance with Standards
75	Work activities - General	Analyzing Data or Information
76	Work activities - General	Making Decisions and Solving Problems
77	Work activities - General	Thinking Creatively
78	Work activities - General	Updating and Using Relevant Knowledge
79	Work activities - General	Developing Objectives and Strategies
80	Work activities - General	Scheduling Work and Activities
81	Work activities- General	Organizing, Planning, and Prioritizing Work
82	Work activities - Work Output	Performing General Physical Activities
83	Work activities - Work Output	Handling and Moving Objects
84	Work activities - Work Output	Controlling Machines and Processes
85	Work activities - Work Output	Operating Vehicles, Mechanized Devices, or Equipment

86	Work activities - Work Output	Working with Computers
87	Work activities - Work Output	Drafting, Laying Out, and Specifying Technical Devices, Parts, and Equipment
88	Work activities - Work Output	Repairing and Maintaining Mechanical Equipment
89	Work activities - Work Output	Repairing and Maintaining Electronic Equipment
90	Work activities - Work Output	Documenting/Recording Information
91	Skills	Reading Comprehension
92	Skills	Active Listening
93	Skills	Writing
94	Skills	Speaking
95	Skills	Using mathematics to solve problems
96	Skills	Using scientific rules and methods to solve problems
97	Skills	Critical Thinking
98	Skills	Active Learning
99	Skills	Learning Strategies
100	Skills	Monitoring
101	Skills	Social Perceptiveness
102	Skills	Coordination
103	Skills	Persuasion
104	Skills	Negotiation
105	Skills	Instructing
106	Skills	Service Orientation
107	Skills	Complex Problem Solving
108	Skills	Operations Analysis
109	Skills	Technology Design
110	Skills	Equipment Selection
111	Skills	Installation
112	Skills	Programming
113	Skills	Operations Monitoring
114	Skills	Operation and Control
115	Skills	Equipment Maintenance
116	Skills	Troubleshooting
117	Skills	Repairing
118	Skills	Quality Control Analysis
119	Skills	Judgment and Decision Making
120	Skills	Systems Analysis
121	Skills	Systems Evaluation
122	Skills	Time Management
123	Skills	Management of Financial Resources
124	Skills	Management of Material Resources
125	Skills	Management of Personnel Resources

Notes. Data Source: O*NET version 26.0 (2021), 873 occupations

Appendix 8 Types of Accommodation

Category	Types of accommodations provided
Any time accommodation	• Allowed flexibility of work hours (arrival or departure change) • Allow more breaks or rest periods • Shorten workdays
Change job/new skills	• Changed job tasks to suit workers' abilities • Helped workers learn new job skills and provided training;
Any equipment /transportation	• Get special equipment for job • Arrange special transportation
Other help	• Emotional support; medical care provided (or arranged for); paid medical leave; time off when needed (unpaid); parking made easier/closer; adapted/changed working environment; monetary compensation (including payment for medical expenses); offered early retirement; helped get workers' compensation/disability; assist in receiving rehabilitative services; provided assistance at tasks when needed; get someone to help

Notes. 1) Data Source: Health and Retirement Study 1992-2018, 2) Categories of employer accommodation from Maestas et al. 2017