Received: 23 December 2021 | Revised: 22 February 2022 | Accepted: 23 February 2022

DOI: 10.1111/rsp3.12520

ORIGINAL ARTICLE

Impacts of the COVID-19 outbreak on older-age cohorts in European Labor Markets: A machine learning exploration of vulnerable groups

Mehmet Güney Celbiş[1,2] | **Pui-hang Wong**[2,3] | **Karima Kourtit**[4,5,6] | **Peter Nijkamp**[4,5]

[1]Department of Economics, Yeditepe University, Istanbul, 34755, Turkey

[2]UNU-MERIT, Maastricht, 6211, The Netherlands

[3]Maastricht Graduate School of Governance, School of Business and Economics, Maastricht University, Maastricht, 6211, The Netherlands

[4]Faculty of Management, Open University, Heerlen, 6419, The Netherlands

[5]Centre for European Studies, Alexandru Ioan Cuza University, Iaşi, 700506, Romania

[6]School of Architecture, Planning and Design, Polytechnic University, Ben Guerir, 43150, Morocco

Correspondence
Mehmet Güney Celbiş, Department of Economics, Yeditepe University, Istanbul 34755, Turkey.
Email: celbis@merit.unu.edu

Abstract

We identify vulnerable groups through the examination of their employment status in the face of the initial coronavirus disease 2019 (COVID-19) shock through the application of tree-based ensemble machine learning algorithms on a sample of individuals over 50 years old. The present study elaborates on the findings through various interpretable machine learning techniques, namely Shapley values, individual conditional expectations, partial dependences, and variable importance scores. The structure of the data obtained from the Survey of Health, Aging and Retirement in Europe (SHARE) dataset enables us to specifically observe the *before* versus the *after* effects of the pandemic shock on individual job status in spatial labor markets. We identify small but distinct subgroups that may require particular policy interventions. We find that the persons in these groups are prone to pandemic-related job loss owing to different sets of individual-level factors such as employment type and sector, age, education, and prepandemic health status in addition to location-specific factors such as drops in mobility and stringency policies affecting particular regions or countries.

KEYWORDS
unemployment, machine learning, age, mobility, stringency

JEL CLASSIFICATION
C81, J40, J14, J23, J21

1 | INTRODUCTION

Since the outbreak of the COVID-19 crisis, a substantial worldwide effort has been made in disentangling the effects of this global economic disruption brought by the pandemic, in both academic and policy-related settings. These efforts comprise not only attempts to document past and present economic tribulations, but are also various endeavors to anticipate and counter potential impending socioeconomic challenges. A sound comprehension of how the pandemic has been affecting industries, firms, and households is paramount for providing clues on how their postpandemic behavior and choices may take shape in the immediate future. The pandemic may have significant distributional effects on the labor market, in particular, on different age cohorts. In this context, from a sectoral, individual, and locational perspective, the present study examines how working individuals in older-age cohorts have been affected by the COVID-19 pandemic with respect to their job status as a result of the initial pandemic shock in 2020. By doing so, the study aims to algorithmically detect and provide an assessment of the attributes that describe disadvantaged and vulnerable individuals.

While the economic and social upheavals caused by the COVID-19 outbreak are clearly apparent, these complex effects are still in need of being carefully disentangled and quantified (Nijkamp & Kourtit, 2022). From an economic perspective, del Rio-Chanona et al. (2020) accurately predicted that the pandemic would cost around 20% of the gross domestic product in the United States. The first COVID-19 wave hit the labor markets particularly quickly and deeply. In this regard, using a difference-in-differences approach, Bauer and Weber (2021) estimated that in Germany the effect of the pandemic accounted for up to 60% of the increase in unemployment in April 2020. A large part of the economic impact is directly related to a range of drastic government policies such as community lockdowns and social distancing. Especially the former has led to significant reductions in demand; and businesses in the catering and hospitality industry in particular were forced to shut their doors for an indefinite period (Benedetti et al., 2020; Gursoy & Chi, 2020). Even when a relaxation of stringency policies took place in summer 2020, many restaurants were able to offer only limited services with significantly reduced labor input, such as take-out and delivery services. Furthermore, international travels were restricted (Ito et al., 2020; Xue et al., 2021), causing the tourism and transportation sectors to incur significant losses, resulting in increased unemployment, further contributing to the weakening of demand and economic activity in general through a multiplier effect (Škare et al., 2021). Against this background, Barrot et al. (2021) estimated that about half of the drop in GDP can be attributed to the implementation of social distancing measures in the United States.

Even though the aforementioned effects apply to most industries and locations around the world, the economic impact of the pandemic is likely to be felt differently across social groups. For example, working mothers are found to shoulder childcare and housework burdens and suffered a greater productivity loss when schools and daycare centers were closed (Alon et al., 2020; Deryugina et al., 2021). Another similar observation on unequal impact is made by Pereira and Patel (2021), who found that self-employed racial minorities in Brazil experienced a greater decline in work hours and worse business-related outcomes than their white counterparts.

While the health and financial impact of COVID on older-age individuals is well documented (see, e.g., Wong et al., 2021), the labor market outcomes for this group remain a subject that has received relatively less attention. The present study addresses this research gap and investigates the labor market impacts of COVID-19 in the context of aging persons. We implement tree-based ensemble machine learning algorithms, namely random

forest and stochastic gradient boosting, alongside the use of Shapley values, individual conditional expectations, and partial dependences in order to increase the interpretability of these methods. As a result, by quantifying how sectoral, personal, and location-specific effects define individual-level unemployment outcomes, our findings may help with setting policy interventions and their customization for disadvantaged groups that we algorithmically detect.

This study is structured as follows. Section 2 presents the theoretical mechanisms underlying the relationship between employment outcomes and the pandemic. The data sources and the steps taken to prepare the data for the empirical analysis are outlined in Section 3, while the descriptive statistics of the variables are presented separately for each model in the subsequent section (Section 4), which describes the ensemble learning methods and interprets their results. Section 5 offers a concluding discussion and provides policy recommendations.

2 | THEORETICAL FRAMEWORK

Unemployment is related to a wide variety of demographic, socioeconomic, regional, and macroeconomic factors (Krugman, 1994; Machin & Manning, 1999; Stough et al., 2018). Despite the fact that the macroeconomic conditions affect all economic areas, gender and age alongside level of education are often identified as critical predictors of unemployment at the individual level (Breen, 2005; Martin et al., 1994; Mincer, 1991). This being said, the gender unemployment gap has been narrowing in the Organisation for Economic Co-operation and Development (OECD) countries since the early 1980s, except during recessions, when the female unemployment rate was mainly exceeded by the male unemployment rate (Albanesi & Şahin, 2018). Additionally, it has been frequently documented that young people are particularly affected by economic downturns (Choudhry et al., 2012; Verick, 2009). In this regard, during the 2008–2009 global financial crisis, the youth unemployment rate for the OECD countries peaked and reached 20.9% in 2009, compared with an average rate of 10.4% in 2005–2009 (Bell & Blanchflower, 2011). On the other hand, the effect of recessions on unemployment among the older population is more complicated; because of skills mismatch, longer unemployment spell, or labor market discrimination, older workers may choose to retire earlier than they would normally prefer and suffer social-security-related disadvantages (Bui et al., 2020; Helppie McFall, 2011; Liu et al., 2016; Neumark & Button, 2014). Such early retirement decisions may cause unemployment statistics to reflect inaccurate findings that imply milder consequences on older workers, as if they are less affected by economic adversities.

Labor market outcomes also vary by condition and type of employment. For instance, self-employed individuals were particularly affected by the pandemic in the UK (Blundell & Machin, 2020). Individuals who manage their own businesses had their work hours reduced by 50–75% per week, while over 60% of self-employed persons earned a monthly income that dropped below 1,000 pounds in April 2020 (Blundell & Machin, 2020). However, self-employment is a general category that includes persons such as entrepreneurs, partners at law firms, and workers with alternative work arrangements, including those who work in informal or gig economies (Blundell & Machin, 2020; Boeri et al., 2020).

As mentioned in the introduction, although workers in all sectors were affected by the global health crisis, the impact has been unequal across sectors. Labor-intensive service industries that require in-person interaction with their customers were hit particularly hard by the pandemic. Such industries include businesses such as restaurants, hotels, tourism, and entertainment establishments. More digitalized sectors such as the financial and telecommunications industries were less affected, as remote working was relatively more common for business operations in these areas even before the pandemic (Barrot et al., 2021). Within this framework, economic sectors can be classified into five categories, according to Fana et al. (2020), based on whether a sector provides essential goods and services to people (e.g., food production, healthcare), whether it is active (e.g., education and public administration), and whether remote working is possible. On the basis of a comparative analysis of six European economies, the authors found that the negative labor market effects of the pandemic have concentrated on the workers in low-productivity service industries.

Quite expectedly, urbanization is associated with a rise in mobility and interconnectedness (Kapitsinis, 2020), and naturally, the adverse effects of the contagious disease have been toughest among individuals living in dense urban areas, highlighting the significance of the spatial dimension of the pandemic-related socioeconomic effects (Sassen & Kourtit, 2021). In accordance with this fact, Chen et al. (2020) found that the economic impact of the pandemic is mostly captured by changes in people's mobility, through an analysis of high-frequency data from Europe and the United States. The authors have observed that the predictive power of mobility patterns is even stronger than other indicators that are indisputably related to the COVID-19 outbreak, such as death rates (Chen et al., 2020). Finally, and most relevant for the study at hand, country-level heterogeneity could be explained by the industrial composition of countries, such as those with direct implications for the extent and availability of jobs that can be performed at home, alongside confinement measures (Fana et al., 2020).

Labor market outcomes are subject to highly nonlinear and interactive effects that are subject to a significant degree of heterogeneity (see, e.g., Glaeser et al., 2020). For instance, the gender unemployment gap is not uniform, as it heavily depends on location and the institutional settings of countries, as Azmat et al. (2006) found that the unemployment rate is higher for females than it is for males in Mediterranean countries. Besides, Alonso-Villar and Del Río (2008) and (Özgüzel, 2020), using data in Spain and Turkey, respectively, showed that industry agglomeration tends to favor women in large cities. Similarly, (Celbiş, 2021c) observed that the probability of unemployment is significantly less for women in urban centers compared with their counterparts in rural areas in Turkey. Therefore, employment outcomes can be reasonably explained only when the complex interaction among individual characteristics, employment conditions, location, industrial structure, and policy environment, among other individual and place-based features, is taken into account. Such necessities can be met through the use of big data methods, thanks to their explorative and nonrestrictive structures.

Machine learning models allow the algorithmic detection and modeling of complex interactions and patterns in the data (Ge et al., 2017; Mullainathan & Spiess, 2017; Varian, 2014). This feature of machine learning models is particularly important in the presence of a large number of relevant variables and potentially complex relationships as in the research question at hand. Under such conditions, an explanatory model based on standard econometric techniques might not be practical, and a shift to artificial-intelligence-inspired methods such as machine learning can present many advantages (Celbiş, 2021a; Nijkamp et al., 2001). One of the main advantages of such techniques is that they allow the inclusion of a large set of variables in conducting an in-depth analysis without incurring bias in the selection of the explanatory variables, while allowing for the algorithmic modeling of complex interactive and nonlinear relationships (Athey, 2018; Athey & Imbens, 2019; Imbens & Athey, 2021; Harding & Hersh, 2018). Consequently, the models are algorithmically constructed in a data-driven manner (Athey, 2018). Accordingly, the present study uses a range of novel machine learning approaches to analyze the factors underlying pandemic-related unemployment among working individuals in older-age cohorts. This being said, concerns relating to causality are not internalized into most machine learning (ML) algorithms, and research needs to be designed by taking into account this potential drawback (Grimmer, 2015). It is therefore of substantial advantage that the data used in this study permit us to explicitly address this concern in relation to causality, as will be detailed in the next section.

3 | THE DATA

Most of the data used in this study are from the latest version of the 'Survey of Health, Ageing and Retirement in Europe' (SHARE) project (Börsch-Supan, 2021a, 2021b). The last wave of the SHARE data (Wave 8) includes a COVID-19 questionnaire that is 'an extraordinary source of information to study health and socio-economic implications of the shock on the elderly population' (Börsch-Supan, 2021a). The regular SHARE data collection process was interrupted in March 2020 by the outbreak and resumed in June and July 2020 in the form of a special COVID-19 module. Therefore, most individuals who were previously asked the standard demographic and work-related

questions (just before the pandemic hit Europe), were subsequently asked how their job status, among other things, was affected right after the first shock, allowing us to rule out any reverse causality problem in our findings. In other words, the clear separation between the survey stages enables us to interpret our machine learning findings as causal effects of individual socioeconomic attributes on pandemic-related unemployment.

Based on the research question and the scope of this study, we use the 'Employment and Pensions' and the COVID-19 modules of SHARE Wave 8, which encompasses a group of European countries and Israel. In addition to individual factors, we also integrate information on stringency policies alongside country- and industry-specific mobility data. The stringency variable is collected by (Hale et al., 2021) and made available at 'ourworldindata.org'. The index measures the stringency level based on spatial policies applied on workplace, school, social gathering, and public transportation usage, among other categories. To further explore industry-related effects, we added the Google Covid Community Mobility Reports (Google LLC, 2021) for the work-related location categories: workplaces, retail and recreation, and groceries and pharmacies, which report the day-to-day change in the number of visitors compared with the baseline activity in those locations as observed between January 3 and February 6, 2020 (Google LLC, 2021).[1] Both the stringency and mobility values have been integrated into our data as average values over the period March to July 2020 to ensure synchronization with the measurement period of our dependent variable. Even though in this initial step the data consisted of 32,366 rows and 636 variables, most of the observations were missing values. To ensure that the research question could be properly addressed, we dropped all individuals with missing employment values or key variables of interest such as mobility, education, and industry. Excluding columns consisting only of missing or constant values yielded a dataset with 1,980 observations and 493 variables. Also, variables with high correspondence to the dependent variable (such as the change in hours worked) and other variables such as month of birth, and whether anyone in the respondent's family tested negative (as questions pertaining to testing positive and COVID-related health changes already exist), and administrative information such as whether anyone else was present during the interview were excluded.

Due to the fact that very few people per International Standard Classification of Occupations (ISCO) category exist in the SHARE data, we used the own industrial classification of the SHARE project, which consists of 14 industry classes rather than the more commonly used ISCO classification. As a result, a dataset with 1,549 rows and 417 columns was attained. However, none of the individuals in this dataset had all the 417 features recorded, and missing values rendered the data unusable for machine learning models. In the presence of many continuous, categorical, and ordinal features, imputation of such a large number of missing values would be questionable. Therefore, we were forced to take other steps to render the data usable, while retaining as much information as possible. We implemented an algorithm that detects the features or sets of features that cause the highest loss of observations in a stepwise manner, and applied a simple classification tree prediction (the classification tree algorithm is detailed in Section 4). As a result, for each step (i.e., iteration) we were able to detect the number of variables and observations that remain, together with the accuracy of a classification tree for each combination of data. The relationship is visualized in Figure 1, where the *y*-axis represents the percentage of either the number of remaining variables represented by the light-green curve, the number of persons represented by the dark-green curve, or the accuracy of the corresponding classification tree prediction on a test data (the test data concept is covered in Section 4), where the *x*-axis represents the iterations. As the accuracy of the fitted model did not significantly drop through the iterations, the largest sample, which consists of 1,549 individuals and 47 variables, was selected. On the other hand, as is often the case with unemployment rates, a small part of our sample consisted of individuals who lost their jobs because of the pandemic (260 people out of the total 1,549 persons). Unbalanced classes may cloud the workings of interpretable machine learning models, as such data will push the algorithms to make simple predictions. For instance, in the case of 260 individuals who lost jobs, a machine learning model that predicts everyone as not experiencing job loss will still be

[1]The three categories are defined as follows. Retail and recreation: concert venues, convention halls, multipurpose arenas, casinos, science centers, art centers, museums, public libraries, community centers, restaurants, beaches, marinas, and golf courses; workplaces: accommodation, nongovernmental organizations (NGOs) and associations, banks, educational institutions, hospitals and clinics, government offices, municipal services, dental, physiotherapy, and immunization offices and clinics, office towers, car dealerships, for-profit businesses, homeless shelters, food banks, and child services; grocery and pharmacy: drugstores, pharmacies, convenience stores, markets, and grocery stores (Leung et al., 2021).

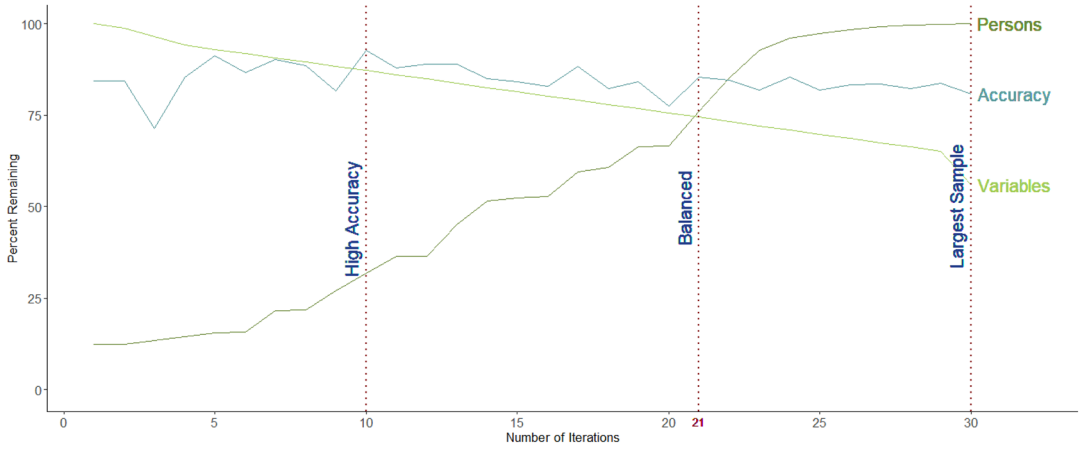


FIGURE 1 Variable-observation trade-off

approximately 80% accurate while ensuring computational efficiency and simplicity. For this reason, we selected a random sample of 260 individuals who did not lose their jobs, attaining a final dataset with 260 persons who lost their jobs because of the COVID-19 pandemic versus an equal number who retained their jobs, to allow for high interpretability of our machine learning algorithms by forcing the models to distinguish equally sized groups. Prior to this data cleaning process, the variables measuring the education level and the type of area where the individual resides were integrated from other modules of Wave 8.

4 | RESULTS OF THE TREE-BASED RANDOMIZED AND SEQUENTIAL ALGORITHMS

4.1 | Recursive binary partitioning of the training data

The target feature (i.e., dependent variable) in all the below detailed models is categorical, and the ensemble models used in this paper are based on classification trees. Prior to the prediction of employment outcomes, one-third of the individuals in our dataset are randomly sampled and set apart as the test data, and the remaining portion of the dataset (the training data) is used to conduct all machine learning procedures.[2] The single-tree framework that underlies the more complex algorithms employed in this study is based on the Classification and Regression Trees (CART) model by by Breiman et al. (1984). Also referred to as binary recursive partitioning, the classification tree algorithm is defined by Breiman et al. (1984), Friedman (2001), and James et al. (2013) as follows: each tree node indexed by r corresponds to a data region with an impurity criterion given by the Gini index I_r:

$$\begin{aligned} I_r &= \sum_{c=1}^{C} p_{rc}(1-p_{rc}) \\ \text{where } p_{rc} &= \frac{1}{N_r}\sum_{i \in A_r} \mathbf{1}(y_i = c) \end{aligned} \tag{1}$$

[2]The `R` packages used to run the machine learning algorithms and other supportive processes are as follows: `rpart` by Atkinson and Therneau (2000) for recursive binary partitioning, `randomForest` by Liaw and Wiener (2002) for the random forest model and for generating the associated proximity matrix, `xgboost` by Chen et al. (2015) for the extreme gradient boosting machine, `rpart.plot` by Milborrow (2019) for plotting the single classification, `pdp` by Greenwell (2017) for the ICE plots, and `SHAPforxgboost` by Liu and Just (2020) for computing the Shapley values and the Shapley value diagram and force plot.

In Equation 1, y_i is the outcome value for the i th individual $(i = 1,...,N)$, A_r and N_r represent respectively the set and the number of individuals who fall into node r, and c is the category index. Because the response variable consists of only two classes: 'YES' (lost job) and 'NO' (retained job) such that $c = 1,2$, then I_r can be written as $2p_{r1}(1-p_{r1})$. The purpose of each data partitioning step is to attain nodes that are as pure (i.e., homogeneous) as possible. In other words, the split attempt aims to distinguish on the basis of the smallest attainable I_r value in that step by selecting a splitting variable x_k from the input space ($X = x_1,...,x_K$ and $k = 1,...,K$) together with its specific split value v that solves:

$$\min_{k,v}\left[\frac{N_{r_1}(k,v)}{N}I_{r_1}(k,v)+\frac{N_{r_2}(k,v)}{N}I_{r_2}(k,v)\right] \tag{2}$$

where r_1 and r_2 denote the nodes resulting from the binary split; in other words, they are the 'child nodes' of r (Breiman et al., 1984; Friedman, 2001; James et al., 2013). In the case of a categorical x_k, like the variable `Industry` used in our models, the split value v simply partitions the variable into distinct classes or sets of classes when attempting the minimization of the sum of the Gini values of the two resulting nodes. If splitting the node does not contribute any reduction to node impurity, that is, the term in the square brackets in Equation 2 is not less than that of the parent node (I_r) upon minimization, then the split is not performed and r becomes a terminal node (Breiman et al., 1984; Friedman, 2001; James et al., 2013).

An unpruned classification tree will continue applying Equation 2 as long as there remain large enough data regions that can be split into smaller portions, leading to the generation of a very complex tree. However, this approach will perform poorly on out-of-sample observations such as those that were previously set apart as the test set (James et al., 2013). To prevent this overfitting issue, the complexity of the tree has to be reduced by eliminating nodes. This 'pruning' procedure is done by implementing an L-fold cross-validation algorithm, and is outlined by Friedman (2001), Sutton (2005), and James et al. (2013) as follows: a sequence of subtrees t of the highly complex unpruned tree T is generated where each $t \subseteq T$ is indexed by the complexity parameter $\alpha \geq 0$. After splitting the training data into L folds of equal size ($l = 1,...,L$), the process is repeated L times by leaving fold l out as the internal validation set in each iteration.[3] For each subtree t, the complexity level α that corresponds to the minimum weighted total misclassification error subject to a penalty term $\alpha|t|$ is identified, where $|t|$ is the number of terminal nodes $\bar{r}$:

$$\begin{aligned} E_l(\alpha) &= \sum_{\bar{r}}^{|t|} e_{\bar{r}}\frac{N_{\bar{r}}}{N}+\alpha|t| \\ \text{where } e_{\bar{r}} &= \frac{1}{N_{\bar{r}}}\sum_{i\in A_{\bar{r}}}\mathbf{1}(y_i \neq c^*_{\bar{r}}) \end{aligned} \tag{3}$$

where $E_l(\alpha)$ is the penalty-augmented error of the subtree as a function of α generated by leaving fold l out. $c^*_{\bar{r}}$ is the majority class in terminal node $\bar{r}$, $\mathbf{1}$ is the indicator function, and $\bar{r} = \{1,...,|t|\}$. Subsequently, the errors of the trees that correspond to each complexity level are averaged over the L folds of data that they leave out in prediction, and the complexity parameter α^* that gives the minimum cross-validated error is selected:

$$\alpha^* = \arg\min_{\alpha}\left[\frac{1}{L}\sum_{l=1}^{L}E_l(\alpha)\right] \tag{4}$$

[3]The terms 'test data', 'test set', and 'validation set' are usually used interchangeably in machine learning. However, to prevent any confusion between the test data (which is separated in the beginning) and the left-out folds in the cross-validation process, we refer to the latter as 'internal validation sets' of the training data.

Finally, the particular $t \subseteq T$ with $\alpha = \alpha^*$ is chosen as the optimal tree Friedman (2001); James et al. (2013).[4] In addition to the identification of α^*, parameter tuning for determining the maximum depth and minimum size of a tree node is made through a grid search in the range of 1 to 10 for the number of observations and 5 to 15 for tree depth, by increments of one – similar to the application by Boehmke (2020).

The resulting classification tree is shown in Figure 2, where the values in the terminal nodes represent the share of people who lost their jobs (YES), those who did not (NO), and the portion of the data that falls into the leaf. The topmost partitioning is made using the variable `EmpType`, which sets aside the individuals who are employed in the public sector. Unless the public sector employee did not rank her/his prepandemic health level (`HealthBefore`) as worse than 'excellent', "good', or 'very good', the classification tree strongly predicts no job loss. Therefore, the classification tree identifies a small but crucially vulnerable group of public sector employees (4% of our sample) who are predicted to experience job loss due to their less favorable initial health levels. There is a health-related inequality among public sector employees who are generally seen as more protected, financially, from the negative effects of the pandemic.

Turning to the persons who either work in the private sector or manage their own business, we observe a strong prediction of job loss if mobility in the retail and recreation category (`Retail_Recreat`) dropped by more than a score of 31. The tree predicts retained employment for a group of persons that are in locations where the drop in mobility is less than 31 in the following industry categories (`Industry`) that appear to be more resilient to the pandemic effect: 'agriculture, hunting, forestry, fishing', 'mining, quarrying', 'financial intermediation', 'real estate, renting and business activities'. Individuals working in the industries 'manufacturing', 'electricity, gas, and water supply', 'construction', 'wholesale and retail trade; repair of motor vehicles, motorcycles, and personal and household goods', 'transport, storage, and communication', 'other community, social, and personal service activities' (all industry codes and variables are defined in Table 1, and their descriptive statistics are presented in Table 2) are predicted to become unemployed if they were born before 1962 and live in places where the workplace mobility (`Workplaces`) fell less than the value 27, as opposed to locations with higher drops in mobility. As a result, industry, mobility, and age-related inequality is more prominent in relation to the labor market outcomes for the persons who own businesses or work in the private sector, whereas for those who work in the public sector, health-related effects are more powerful. From this point of view, the tree identifies yet another small and vulnerable group of individuals through the variable `BirthYear`: relatively older individuals (i.e., those born before

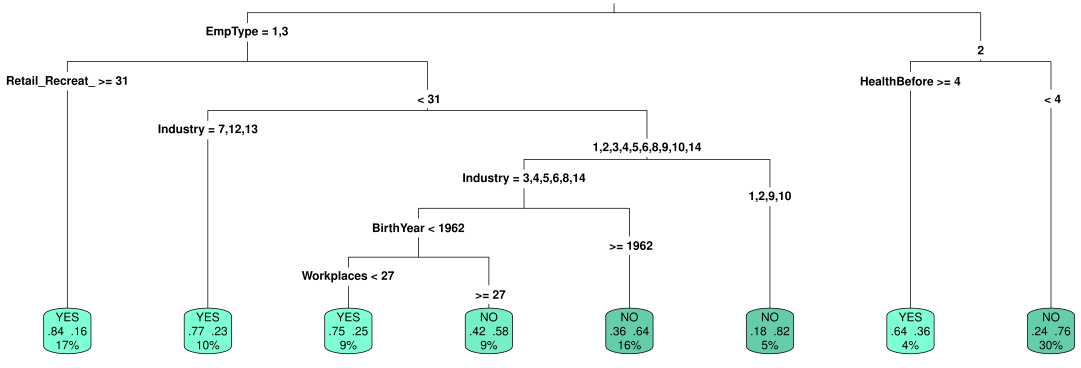


FIGURE 2 Sample classification tree

[4]The reason why '$\subseteq$' is used rather than '$\subset$' is because α^* may be equal to zero, and therefore, the penalty term $\alpha|t|$ in Equation 3 would also be zero, resulting with the generation of the unrestricted classification tree T (James et al., 2013).

TABLE 1 Features Selected by the Single Classification Tree

Name	Description	Values
`BirthYear`	The year of birth of the participant.	Quantitative variable.
`EmpType`	Employment category.	**1** = Private sector, **2** = Public sector, **3** = Self-employed.
`HealthBefore`	Perception of health status before the outbreak.	A 1-to-5 scale where 1 = Excellent, and 5 = Poor.
`Industry`	Industry category.	**1** = Agriculture, hunting, forestry, fishing, **2** = Mining and quarrying, **3** = Manufacturing, **4** = Electricity, gas, and water supply, **5** = Construction, **6** = Wholesale and retail trade; repair of motor vehicles, motorcycles, and personal and household goods, **7** = Hotels and restaurants, **8** = Transport, storage, and communication, **9** = Financial intermediation, **10** = Real estate, renting and business activities, **11** = Public administration and defense; compulsory social security, **12** = Education, **13** = Health and social work, **14** = Other community, social, and personal service activities.
`Retail_Recreat`	Average change in the number of visitors and time spent in places in the retail and recreation category between July and March 2020, compared with a prepandemic baseline value. Quantitative variable.	Additive inverse is used so that higher values indicate higher decreases in mobility.
`Workplaces`	Average change in the number of visitors and time spent in places in the workplaces category between July and March 2020, compared with a prepandemic baseline value. Quantitative variable.	Additive inverse is used so that higher values indicate higher decreases in mobility.

1962) working in the manufacturing, energy supply, construction, transportation/communication, and community/personal service areas are more likely to lose their jobs as a result of the drop in mobility in workplaces. Interestingly, if the drop is more than the score of 27, the tree predicts that there will not be job losses, but if the mobility drop is less, then job loss is expected. A possible explanation of this finding can be hypothesized to be based on remote working opportunities: in workplaces where a transition to remote work is more possible, older individuals may have been more likely to retain their jobs, whereas this group of individuals may have been the first to be laid off in workplaces that were not able to considerably reduce their activity. Clearly, the aforementioned algorithmic finding calls attention to the need for in-depth case studies, and points at new research avenues. The predictive accuracy of this unitary classification tree is 69.8%. The code used to generate this single tree was programmed to repeat by making new random splits of train versus test data until a tree with an accuracy above 50% that is not just a stump (i.e., consisting of a single split) is generated.[5]

[5]The `R` code will be made available to reviewers if requested.

TABLE 2 Descriptive Statistics of the Features Selected by the Classification Tree

Variable	Median or mean	SD	Mode	Min	Max
BirthYear	1960	5.15	1960	1942	1970
EmpType	–	–	1	1	3
HealthBefore	3	–	3	1	5
Industry	–	–	13	1	14
Retail_Recreat	23.02	10.46	–	3.67	49.12
Workplaces	28.13	4.78	–	20.30	37.31
Number of observations = 520 for all variables					

Note: The central tendency measure is reported using the median for ordinal variables, and the mean for quantitative variables.

4.2 | Random forest results

Despite the implementation of *L*-fold cross-validation to avoid overfitting, the unitary classification tree outlined above lacks robustness to data replacements, and is susceptible to omitting relevant features that have high correlations with other predictors (Athey & Imbens, 2019; James et al., 2013). The random forest technique Breiman (2001) improves predictive accuracy and decreases the variance of the prediction by retaining the low bias structure of an overfit single tree, while reducing variance thanks to averaging. Based on the bootstrap aggregation algorithm developed by Breiman (1996), a random forest model generates an ensemble of unpruned classification trees (i.e., different versions of *T*) by adding randomization into feature selection (Breiman, 2001; Friedman, 2001). The 'forest' comprises $|T|$ number of trees *T*, each generated by using a random sample of size *N* from the training set ($T = 1,...,|T|$). The issue pertaining to the omission of relevant variables discussed earlier for a classification tree may persist in the ensemble approach if all trees in the ensemble use the same predictor space, as is the case in bootstrap aggregation models (Friedman, 2001; James et al., 2013). On the other hand, the random forest algorithm considers features $Z \subset X$ where *Z* is sampled randomly at each node split, with size $\sqrt{|X|}$ (Breiman, 2001), thereby de-correlating the trees in the forest (James et al., 2013).

Making predictions alone, regardless of how accurate they are, does not lead to useful results when it comes to answering socioeconomic and policy-relevant questions. Since the predictions by ensemble models are made through the contribution of a large number of trees, plotting a single tree is clearly not practical. In this regard, interpretable machine learning techniques come forth as highly useful approaches for deriving valuable socioeconomic information from predictions. These techniques involve the retrospective assessment of the prediction process by explaining how the algorithm reached the final predictions, identifying the important variables, and showing how these variables contributed to predictions by assessing the magnitudes and the directions of effects. Various elegant and informative techniques exist to discover important socioeconomic and policy-relevant patterns and mechanisms. For this purpose, we employ novel techniques such as the computation of conditional expectations, partial dependences, and deriving Shapley values that are rooted in cooperative game theory.

One tool of interpreting the results of a ML model is the variable importance measure, which is used to assign a score and rank to the inputs used in ensemble model. Variable importances are based on calculating how useful an input was for prediction: when selected as a split variable in any given tree in the ensemble, a feature leads to a reduction ΔI in node impurity, calculated for each split in a tree as shown in Equation 5 and averaged over all $|T|$ trees in the ensemble (James et al., 2013).

$$\Delta I = I_r - \left[\frac{N_{r_1}}{N} I_{r_1} + \frac{N_{r_2}}{N} I_{r_2}\right] \tag{5}$$

and subsequently scaled into a value between one and zero. In Equation 5, N_{r_1} and I_{r_1} denote the size and the Gini index of the first child node resulting from the binary split respectively, while their counterparts indexed by 2 refer to the second child node, and I_r is the Gini index of the parent node as defined in Equation 1.

The variable importances computed by the random forest model are presented in the left panel of Figure 3, and the features selected by the model that were not selected by the single tree are defined in Table 3 alongside their descriptive statistics in Table 4. In the variable importance graphs, we present only the 20 most important features. Results suggest that mobility (particularly in the retail and recreation category), employment type, industry, the stringency level, and previous health condition are features that lead to high decreases in mean node impurity. These findings are mostly consistent with the results of the single tree, which, however, did not identify the stringency effects that the random forest model discerned. Education is the fourth most important individual-level feature, prompting us to further investigate its role in the subsequent steps in the form of conditional expectation and Shapley value analysis. As suggested by the unitary tree in Figure 2, while the prepandemic health situation of a person is among the relatively more important features, other health-related individual- or household-level predictors `Died`, `Hospitalized`, and `HadTreatment` have lower importance levels, but jointly suggest the relevance of personal and family health status to employment outcomes in addition to `HealthBefore`. Gender is also selected by the random forest model as an important feature. Finally, we observe several country-specific effects, which we further examine by focusing on country-specific partial dependences.

It is possible to look further into the random forest findings by visualizing the random forest proximity matrix in the form of a two- or three-dimensional proximity plot (Breiman & Cutler, 2020; Friedman, 2001). Plotting similarity scores among the data instances (observations) is an effective approach for detecting cluster structures (Aldrich & Auret, 2013; Friedman, 2001). Because each tree in the random forest draws its own random sample, some observations are left out. The outcome for these out-of-bag (OOB) observations are predicted using the corresponding unrestricted tree and recorded into an initially empty $N \times N$ proximity matrix $\mathbf{S}$ each time any terminal node $A_{\bar{r},T}$ of iteration T includes both OOB individuals i and j ($j = 1,\ldots N, i \neq j$) such that s_{ij} is increased by 1 (Aldrich & Auret, 2013; Breiman & Cutler, 2020; Friedman, 2001). Next, $\mathbf{S}$ is scaled by dividing with $|T|$ and converted to an $N \times N$ dissimilarity matrix $\mathbf{D} = 1 - \mathbf{S}$ (Aldrich & Auret, 2013). Finally, through multidimensional scaling, dissimilarities can be expressed in two or three dimensions that can be represented in the random forest proximity plot (Friedman, 2001), as shown in Figure 3. In the classification context, the proximity plots are generally star shaped, with each arm representing different classes. Observations in relatively pure class regions (for instance, $A_{\bar{r}}$'s that contain mostly only employed individuals) fall into the extremities of an arm, whereas the center of the star consists of observations that the model cannot distinguish sufficiently well (Aldrich & Auret, 2013; Friedman, 2001).

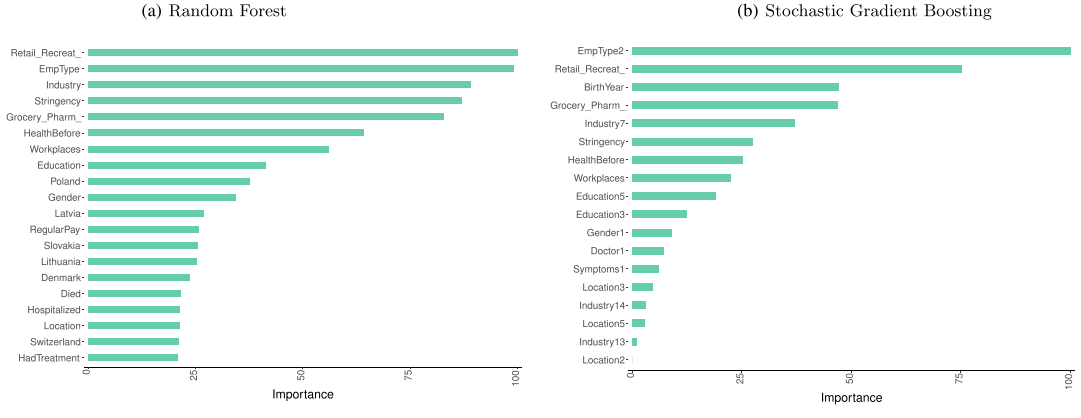


FIGURE 3 Variable importance and ICE plots – random forest

TABLE 3 Features Selected by the Ensemble Models

Name	Description	Values
`Country (Name)`	Binary variables.	Equals one if the country is the one specified in the variable name, and zero otherwise.
`Doctor`	Binary variable.	Equals zero if the responded has visited a doctor/medical facility other than hospital since outbreak, and 1 otherwise.
`Died`	Binary variable.	Equals zero if anyone in the respondent's household died from COVID-19, and 1 otherwise.
`Gender`	Binary variable.	Equals one if female, zero if male.
`Grocery_Pharm`	Average change in the number of visitors and time spent in places in the groceries and pharmacies category between July and March 2020, compared with a prepandemic baseline value. Quantitative variable.	Additive inverse is used so that higher values indicate higher decreases in mobility.
`Hospitalized`	Binary variable.	Equals zero if the respondent has been treated in hospital since outbreak, and 1 otherwise.
`Education`	Respondent's highest ISCED-97 level of education	A 0-to-6 scale where 0 indicates no education, and 6 indicates doctoral or equivalent level.
`Location`	Size classification of the residence are of the respondent	A 1-to-6 scale where 1 indicates 'a big city' and 5 indicates 'a rural area or village'.
`RegularPayNo`	Binary variable	Equals one if the respondent does not receive any regular payment, and zero otherwise.
`Stringency`	The COVID-19 stringency index.	Composite score derived from the strictness of the stringency policies with regard to school and workplace closures, cancellation and restriction of public events, gatherings, public transport, information campaigns, internal and international movement, travel controls, vaccination, and face covering policies. A score of 100 indicates the strictest level, whereas zero is the minimum score.
`Symptoms`	Binary variable.	Equals zero if anyone in the respondent's household had COVID-19 symptoms, and 1 otherwise.

[6]See Table 19 in the UNESCO ISCED classifications report available at https://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf.
[7]See https://ourworldindata.org/grapher/covid-stringency-index?tab=table for more information on the stringency index developed by Hale et al. (2021).

The random forest proximities are visualized in two and three dimensions in Figure 4, where larger circles represent individuals in places with higher mobility drops in the retail and recreation category. Light-blue circles represent individuals working in the private sector, dark-blue circles correspond to persons who are self-employed, and gray circles represent individuals in the public sector. Clear employment type clusters are identified by the random forest based on the prediction of unemployment status. A cluster of unemployed individuals is located on the far end of the arm consisting of the self-employed and those who are working in the private industry (a web-based version of these figures where one can hover above the dots to see the employment status and mobility score can be provided upon request). Aside from the aforementioned cluster, we observe the self- or private employed versus the public

TABLE 4 Descriptive Statistics of the Features Selected by the Classification Tree

Variable	Median or mean	SD	Mode	Min	Max
Doctor	0.69	–	–	0	1
Died	0.98	–	–	0	1
Gender	0.62	–	–	0	1
Grocery_Pharm	4.78	8.31	–	-9.5	24.6
Hospitalized	0.93	–	–	0	1
Education	3	–	3	0	6
Location	4	–	5	1	5
RegularPayNo	0.96	–	–	0	1
Stringency	56.8	5.8		48.7	72.6
Symptoms	0.85	–	–	0	1
Number of observations = 520 for all variables					

Note: The central tendency measure is reported using the median for ordinal variables, and the mean for quantitative and binary variables.

sector employees clustered into separate arms as expected, further highlighting the role of job type in pandemic-related employment outcomes. In conjunction with our earlier findings, we can state that there exist two distinct subgroups (i.e., private/self-employed and public) in which there are more specific subgroups – defined by personal, industrial, and locational attributes – that consist of vulnerable individuals who may require different sets of policy interventions (e.g., health-related versus industry-related policy approaches).

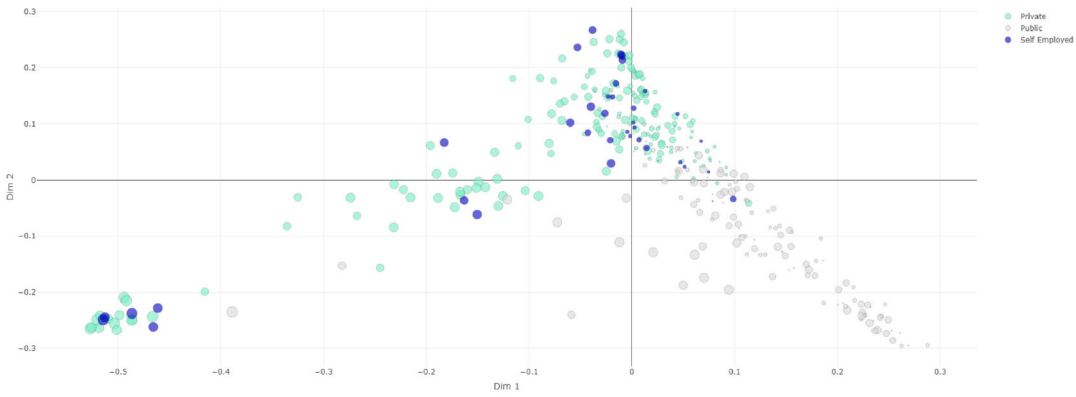
4.3 | Gradient boosted trees

Random forests build trees independently, with the aim to de-correlate them. On the other hand, the gradient boosting machine algorithm (GBM, Friedman et al., 2001; Friedman, 2002) generates sequential classification trees, each learning from the prediction errors made by its preceding counterpart. The GBM can be extended to include stochasticity, similar to the random forest approach, in the form of a stochastic gradient boosting machine (SGBM, Friedman et al., 2001; Friedman, 2002), and the adoption of stochasticity can be further extended computationally through the application of an extreme gradient boosting machine algorithm (XGBoost, Chen & Guestrin, 2016). The steps of the GBM algorithm in a classification context established by Friedman (2001); Friedman et al. (2001), and Friedman (2002) are adapted into our data as follows: the algorithm is initiated by a constant initial prediction $\hat{y}_{i,1}$, which is the same for every person i, where the second subscript indexes the first of the sequential trees (i.e., a single leaf) where $t = (1,...,T)$.[8] This initial prediction is approximated by minimizing the expected value of the loss function, the negative likelihood $L(y_i, f(X))$, with respect to $f(X)$, the predicted log odds of becoming unemployed because of the pandemic:

$$\hat{y}_{i,1} = \arg\min_{f(X)} \sum_{i=1}^{N} L(y_i, f(X)) \tag{6}$$

[8]While the same index t was earlier used to denote the subtrees in cross-validation, henceforth t denotes the classification trees built by the gradient boosting algorithm.

(a) Two Dimensions



(b) Three Dimensions

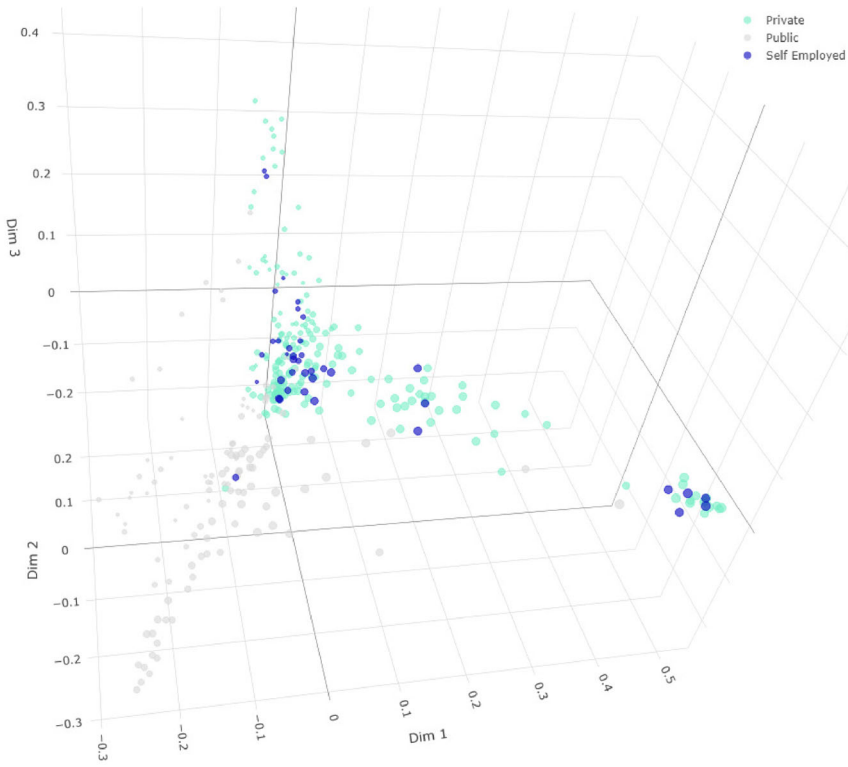


FIGURE 4 Job type clusters based on mobility – random forest

The GBM improves on the errors made by the preceding trees. To carry out this sequential improvement in the predictions, GBM computes the pseudo-residuals ε_{it} for each individual i resulting from the predictions made by the preceding tree $t-1$ as the negative of the derivative of the loss function with respect to the predicted log odds:

$$\varepsilon_{it} = -\left[\frac{\partial L(y_i, f(X_i))}{\partial f(X_i)}\right]_{f=f_{t-1}} \text{for } i = 1,\ldots,N \tag{7}$$

Next, a regression tree is used to predict the targets ε_{it}.[9] The regression tree yields terminal nodes $\bar{r}$ that consist of the set of residuals corresponding to all i observations $X_i \in A_{\bar{r}}$ (similar to the notation used in the earlier outline of a single-tree framework, a tree t in a GBM iteration has $|t|$ number of terminal nodes $\bar{r}$). The output value $\varepsilon_{\bar{r}t}$ for node $\bar{r}$ is computed by minimizing the loss function including the correction term ε (Friedman, 2001; Friedman et al., 2001):

$$\varepsilon_{\bar{r}t} = \arg\min_{\varepsilon} \sum_{x_i \in A_{\bar{r}t}} L(y_i, f_{t-1}(X_i) + \varepsilon) \tag{8}$$

Finally, the prediction for person i's employment status is updated such that the new tree fractionally learns from the preceding tree's predictions f_{t-1}:

$$\hat{y}_{it} = f_{t-1}(X_i) + \delta\varepsilon_{\bar{r}t}\mathbf{1}(X_i \in A_{\bar{r}t}) \tag{9}$$

where δ is the 'learning rate' ($0 < \delta < 1$), $\mathbf{1}$ is the indicator function, and $\delta\varepsilon_{\overline{mt}}$ is the step size taken to improve predictions (Friedman, 2001; Friedman et al., 2001). In other words, each tree 'corrects' the errors made by the preceding tree by revising the prediction for i based on the magnitude of the error, but with some doubt, as 'corrections' themselves may be wrong, leading to new errors that need to be handled by new trees. The 'doubt' is reflected by a learning rate less than 1, and it has been shown that small δ values and a large T (often decided by an early-stopping rule) lead to significant improvements in prediction (Friedman, 2001; Friedman et al., 2001). In addition to the above framework, we apply the stochastic GBM by (Friedman, 2002) through the use of the extreme gradient boosting (XGBoost) machine. This framework allows the introduction of stochasticity in two ways; firstly, by selecting a new random subsample of the training data at each iteration; and secondly, by sampling a subset of features $Z \subset X$ at each node of each tree as in the random forest model. XGBoost algorithm by Chen and Guestrin (2016) is a popular computational extension used in the implementation of GBM, as it allows one to control further parameters for regularization (i.e., tree pruning as discussed in the context of a single tree), and implements the earlier-discussed tenfold cross-validation process, leading to greater generalization capability in addition to improving computational efficiency.[10]

The XGBoost algorithm encodes all categorical features into binary values prior to making predictions, allowing specific feature classes to be scored separately on the basis of their importance, as can be seen in the right panel of Figure 3. The stochastic gradient boosting model outlined above is applied using the following parameter values that result from a grid search that yields the parameter values that correspond to the minimum test error estimate: $\delta = 0.3$, a maximum depth $= 15$, a variable subsample ratio of 0.5 at each split, a minimum node size of 5, and a number of trees of 9. While the grid search did not suggest subsetting observations at each iteration, the second-lowest test error was only 0.003 higher than the first (0.645 and 0.648, respectively), we use the parameter values corresponding to the former for higher computational speed.

In line with the findings of the random forest model and the single classification tree, the SGBM results show that being a public sector employee (`EmpType2`) – as opposed to being self-employed or working in the private sector – being in a country experiencing large decreases in mobility, age, working in industry 7 (hotels and restaurants),

[9]A regression tree, which also is a part of the CART framework of Breiman et al. (1984), partitions the data on the basis of a squared error loss function as opposed to a classification tree, which uses node impurity. As a regression tree is essentially the continuous dependent variable version of a classification tree, the details of the former are not covered in this study.

[10]Chen and He (2015) and Adam-Bourdarios et al. (2015) demonstrated how XGBoost can be used for discovering the Higgs boson on the basis of data generated by the the Large Hadron Collider.

stringency policy level, prepandemic health status, education, and gender rank among the variables with the highest importance. Furthermore, additional health- and industry-related factors are also selected. The main difference of the SGBM results from the random forest findings is that the former finds age as an important determinant of pandemic-related unemployment status, prompting us to investigate the role of this feature – which was also highlighted by the single tree model earlier – in the following stages of our empirical analysis. The SGBM results are also in accordance with the random forest findings in relation to the importances of health-related outcomes despite selecting different health variables. Finally, the SGBM attributes some importance, albeit negligibly low, to the location type where the individual resides.

4.4 | Applications of interpretable ML tools

4.4.1 | Partial dependences and individual conditional expectations

The variable importance plots presented in Figure 3 are useful for detecting the factors that lead to different employment outcomes. They also serve as a starting point to decide on the factors for which it is worthy to delve in further. We elaborate on the association between unemployment status and the remaining variables in our dataset using three interpretation tools: individual conditional expectations, partial dependences, and Shapley values, outlined below. The partial dependence approach introduced by Friedman (2001) allows a plotting of the predicted outcome for the values of a specific predictor variable x_k averaged over all individuals i in the training data. If we set apart a group of variables of interest X_s (usually one or two due to limitations relating to visualization and computation) from the predictor space $X=(x_1,\ldots,x_K)$ versus the remaining input features X_C, then the partial dependence of the predicted outcome on X_s can be represented according to Friedman (2001) and Aldrich and Auret (2013):

$$f_S(X_S)=\frac{1}{N}\sum_{i=1}^{N} f(X_S,X_{i,C}) \tag{10}$$

where the $X_{i,C}$ are actual values of features C for individuals in the training data that are 'held constant', while predictions for each value of X_S are remade for each individual and subsequently averaged. This approach is conceptually, but not technically, analogous to 'ceteris paribus' effects in economics. Partial dependences may lead to unrealistic conclusions for a given feature in X_S that has a high degree of correlation with a feature or a number of features in X_C (Friedman, 2001; Molnar, 2019). Furthermore, heterogeneous effects are not represented in partial dependence values (Molnar, 2019). A useful approach to cope with these disadvantages is constructing individual conditional expectation (ICE) plots, as formulated by Goldstein et al. (2015). An ICE representation disaggregates the global partial dependence of the response feature on X_C, and plots the conditional expectation curves for each individual (Goldstein et al., 2015; Molnar, 2019). However, when individual curves are plotted in this fashion, the outcome will be stacked curves with different intercepts because predictions for the first values of the variables of interest will differ (Goldstein et al., 2015; Molnar, 2019). By anchoring curves $f_{i,S}$ at a specific value of the variable of interest, it is possible to conveniently represent the heterogeneity by plotting the ICE curves $f_{i,S}^{centered}$ that share the same intercept. In other words, the the differences in levels due to different $X_{i,C}$'s are removed by using an anchor point x^* (Goldstein et al., 2015):

$$f_{i,S}^{centered}=f_{i,S}-f(x^*,X_{i,C}) \tag{11}$$

The ICE plots for the three mobility categories alongside the policy stringency variable are shown in Figure 5, 6 where the random forest and SGBM results are shown in the left and right panels, respectively. The SGBM plots display a larger number of overlapping curves compared with the random forest ICE diagrams, due to the fact that the

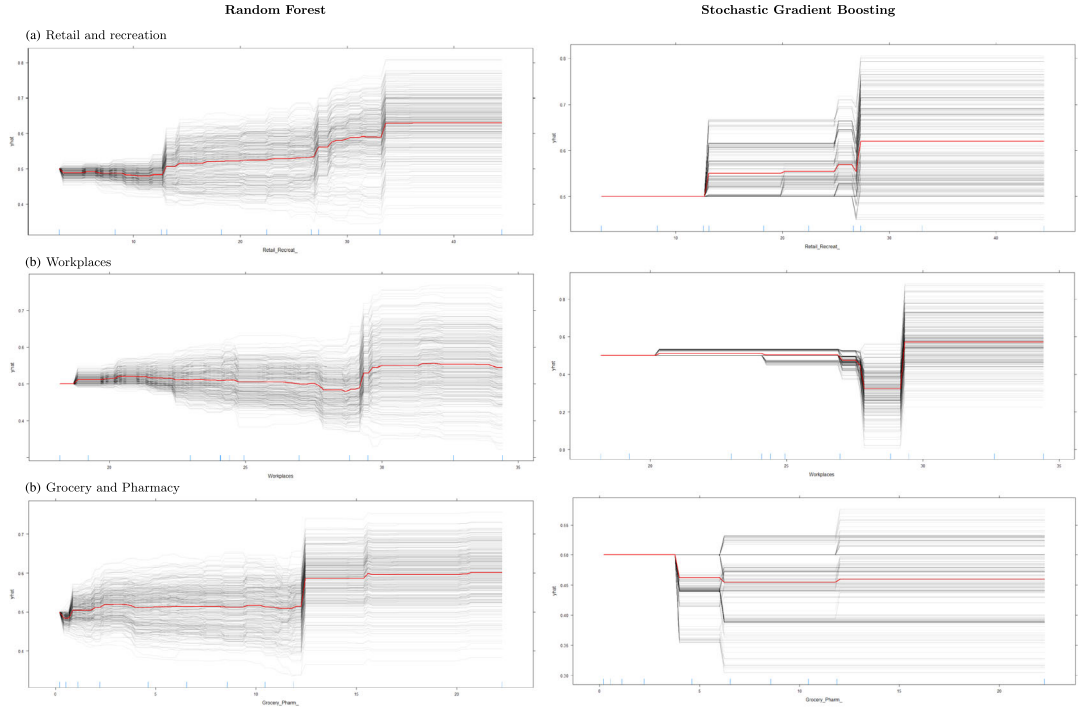


FIGURE 5 Individual conditional expectation plots – mobility

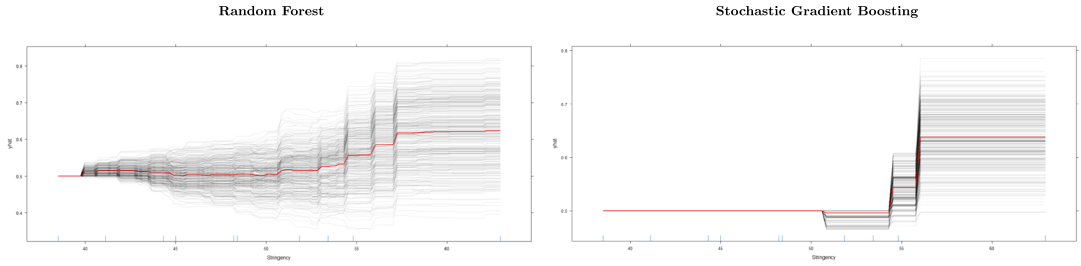


FIGURE 6 Individual conditional expectation plots – stringency

latter method, as outlined earlier, generates unpruned trees. The ICE plots in the first row of Figure 5 suggest that, while larger drops in mobility in the retail and recreation category lead to higher probabilities of job loss, the effect is heterogeneous. The heterogeneity is also observable in the SGBM counterpart of the plot at the right-hand side. In general, the ICE plots hint that individuals who are subject to scores higher than about 30 are particularly more likely to experience pandemic-related job loss. While a similar outcome is observed for the drop in mobility in workplaces, the SGBM results for mobility in the grocery and pharmacy category do not suggest a clear pattern, consistent with the fact that the SGBM ranked the aforesaid category at a lower importance level. Finally, the ICE plots for policy stringency for both models suggest that higher stringency levels are associated with higher probabilities of job loss, while highlighting that some individuals are less affected by stringency policies.

Industry-based heterogeneity may be addressed through the use of two-way partial dependence plots (PDP). While the usefulness of ICE plots is due to their ability to display individual curves, the advantage of two-way PDPs is the possibility to view how two features affect the predictions jointly. Each individual is represented by a pixel in

the two-way PDPs shown in Figure 7, where lighter-colored pixels correspond to individuals with higher probability of experiencing unemployment due to the pandemic as predicted by the random forest model. Each row in the plot in panel A of Figure 7 represents the industries as defined in Table 1 in succession, the *x*-axis represents the stringency level, and the legend on the left presents the color scale corresponding to probability predictions. The seventh row from the bottom that corresponds to the 'hotels and restaurants' industry (industry 7) mostly consisted of lighter pixels compared with those that represent the other industry categories. In other words, regardless of the stringency level, individuals in this industry have higher probability of losing their jobs because of the COVID-19 pandemic. Above a stringency score of about 55, individuals working in most other industries are also more likely to lose their jobs except for those who work in mines, quarries, and in the public administration sectors, which are quite resilient to stringency-related measures. The aforementioned finding regarding the vulnerability of the individuals working in the 'hotels and restaurants' category, and the resilience of those who work in mines, quarries, and public administration, persists in all the other PDP diagrams except for panel B of Figure 8. More specifically, in panels B, C, and D of Figure 7, which plot the industry categories against the decrease in mobility in the 'retail and recreation', 'workplaces', and 'grocery and pharmacy' categories, respectively, all highlight the same aforementioned industries as either highly vulnerable or highly resilient. In panel A of Figure 8, we observe that older individuals, particularly those born before around 1955, are predicted to have higher probability of job loss. In line with earlier findings, the persons in the public administration sector, alongside those in the mining and quarry industries, are almost unaffected, while the effect is more pronounced for those who work in hotels and restaurants.

The heterogeneity that may be due to education levels is visualized in panel B of Figure 8. We observe that people born before around 1955 with relatively lower education levels constitute the group with the highest unemployment probability. On the other hand, people in this age group with at least an upper secondary education level degree are found to be more likely to retain their jobs. This finding becomes even more relevant when the unequal impacts of the pandemic on access to education are considered: a lack of education due to the pandemic may lead to further disadvantages in similar future crises (Türk, 2021).

The industry-specific predictions are shown in Figure 9, where the height and the value on top of each bar is the probability prediction. Consistent with our earlier findings, the 'hotels and restaurants' category has the highest value, followed by 'other community, social and personal service activities', while the most resilient industries are 'mining and quarrying', 'public administration and defense, compulsory social security', 'financial intermediation', and 'agriculture, hunting, forestry, fishing'.

FIGURE 7 Industry, mobility, and stringency – random forest

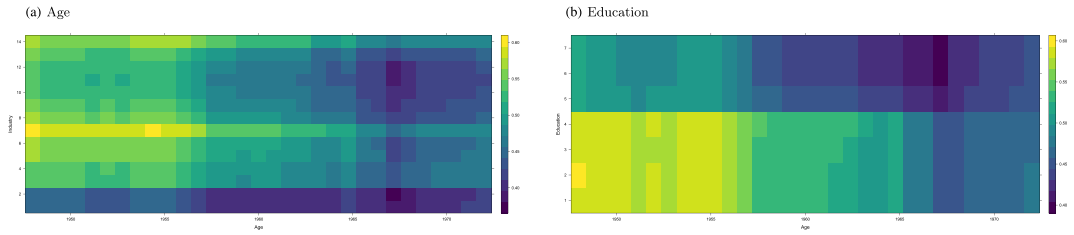


FIGURE 8 Industry, age, and education – random forest

FIGURE 9 Industry partial dependences – random forest

A similar categorical visualization is shown for the countries in the sample in Figure 10 where darker-colored bars indicate countries with higher probabilities of pandemic-related unemployment (also numerically shown at the top of the bars), and higher bars represent higher stringency levels. While countries with both high stringency and high job loss probabilities seem to cluster together, there also exist countries with higher job loss probabilities despite lower stringency levels. It should be noted that the probabilities are based on partial dependences, and while numerically they are very similar in terms of country averages, the effects can be more pronounced for certain individuals, as hinted by the earlier-discussed ICE plots, further underlying the benefits of employing tree-based ensemble methods for prediction in the presence of heterogeneous and nonlinear relationships.

4.4.2 | Shapley value explanations

In addition to variable importances, partial dependences, and individual conditional expectations, extracting further explanatory information from the ensemble models can be carried out by calculating the Shapley values, which is an explanatory method introduced in cooperative game theory by Shapley (1953). Following the study by Lundberg and Lee (2017), the use of Shapley values recently became a popular approach in machine learning for the purpose of better understanding how each feature contributes to individual predictions (Sundararajan & Najmi, 2020).

For the purpose of avoiding the overuse of the term 'value', we henceforth use the terms 'Shapley value' and 'Shapley score' interchangeably in this study. The Shapley score ϕ_h for a specific feature value h of variable x_k for a given data instance (i.e., observation/person) i is presented in Equation 12 (Lundberg & Lee, 2017; Molnar, 2019):

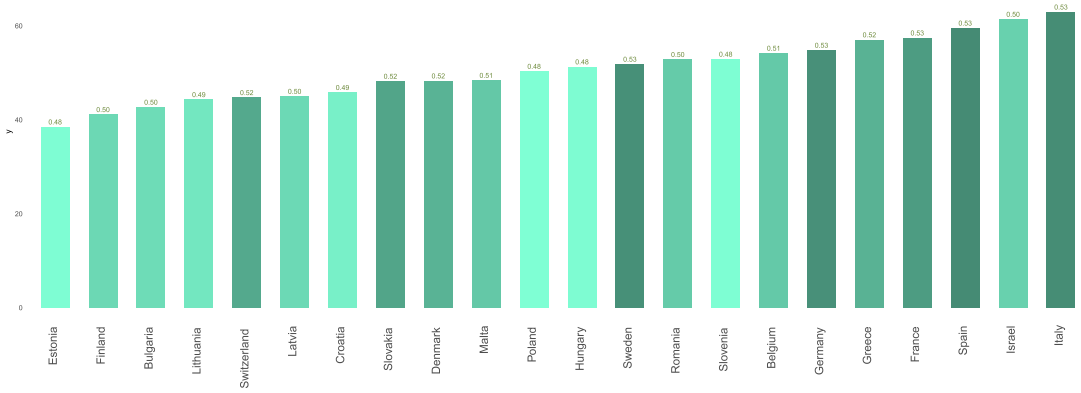


FIGURE 10 Country partial dependences and stringency levels – random forest

$$\phi_h = \sum_{S \subseteq X \setminus \{h\}} \frac{|S|!(|X| - |S| - 1)!}{|X|!} [g(S \cup \{h\}) - g(S)] \tag{12}$$

where S is a subset of feature values, and g is the expected value of multiple predictions by the GBM using randomly drawn values for features not in a given subset S minus the average prediction for all observations (as shown in Equation 13 below).[11] In sum, the Shapley method replaces the features of a given observation with those of a randomly sampled observation from the data, except for a randomly defined subset of the feature values of the observation, and subsequently computes the difference in the deviations in predictions from the average prediction when the feature value equals h versus when it does not – as shown in the last term of Equation 12 (Lundberg & Lee, 2017; Molnar, 2019). When h is excluded, the corresponding feature value from the randomly selected observation is included instead. Therefore, the contribution of the specific feature value h to prediction is averaged over all possible combinations of the subsets $S \subseteq X \setminus \{h\}$ (Lundberg & Lee, 2017; Molnar, 2019).[12] As mentioned above, for each possible subset S, the feature values of random observations from the data are used as replacements for the remaining features and the contribution of h is calculated multiple times and integrated, and the resulting deviation from the average prediction is (Molnar, 2019):

$$g(S) = \int f(X) dP_{x \notin S} - E(f(X)) \tag{13}$$

However, evaluating Shapley values for all possible subsets with versus without h has an exponential computational time complexity (Molnar, 2019; Štrumbelj & Kononenko, 2013). Štrumbelj and Kononenko (2013) present the following approximation:

$$\overline{\phi}_h = \frac{1}{M} \sum_{m=1}^{M} \left[f^* (\hat{y}(i)^m_{+h}) - f^* (\hat{y}(i)^m_{-h}) \right] \tag{14}$$

where $f^* (\hat{y}(i)^m_{+h})$ and $f^* (\hat{y}(i)^m_{-h})$ are the GBM predictions for individual i made by replacing a random number of feature values with those of a randomly selected individual j from the training data. In $f^* (\hat{y}(i)^m_{-h})$, the feature value of interest h is also taken from j. The calculation is done in M iterations ($m = 1, \ldots, M$) and $\overline{\phi}_h$ is the mean value of all

[11] In fact, the cardinality of X as written in Equation 13 is equal to K based on our earlier definition of X such that $|X| = K$.

[12] An alternative expression of the Shapley value is $\phi_h = \frac{1}{|X|!} \sum_{S \subseteq X \setminus \{h\}} \sum_{\lambda} [f(S_\lambda \cup \{h\}) - f(S_\lambda)]$, where λ indexes the order that the feature value of interest h can be included in the calculation Vedder (2020).

iteration results (Molnar, 2019; Štrumbelj & Kononenko, 2013). We execute our Shapley value analysis using the `SHAPforxgboost` algorithm by Liu and Just (2020), which allows reporting of Shapley values globally – for each variable for every individual – as presented in Figure 11. In the Shapley value diagram in Figure 11, each variable is ordered on basis of its Shapley importance level, which is the sum of the absolute values of its Shapley scores over each individual *i* (Molnar, 2019):

$$\Phi_h = \sum_{i=1}^{n} |\phi_{i,h}| \tag{15}$$

for the *i* th person in the training data ($i = 1, \ldots, N$). On the other hand, the Shapley force plot shown in Figure 12 visualizes the ϕ_h values for each *i* in a stacked manner.

The feature effect sizes and directions are represented in the Shapley value diagram in Figure 11, where each dot represents an individual, and darker colors indicate higher feature values. The features on the *x*-axis are ordered in a descending manner based on their Shapley importance scores as given by Equation 15, and the Shapley scores for specific feature values can be traced to the *x*-axis. It follows that binary features are represented in two colors rather than a spectrum. The Shapley value diagram shows that not being in the public sector is associated with higher contributions to predictions of higher unemployment probabilities. Higher decreases in mobility in workplaces, retail and recreation, and grocery and pharmacy categories are also associated with higher probabilities of pandemic-related job loss. For workplaces, discussed earlier in particular in the context of a single tree, there are individuals subject to high but not severe drops in mobility who are predicted to keep their jobs, an observation that may be related to the fact that remote working is common in offices, banks, and similar locations that are in the workplaces category. High-stringency policy values, earlier birth years, being female, and working in industry 7 (hotels and restaurants) contribute to individual-specific predictions of unemployment. While having a postsecondary, nontertiary education is found to increase the chances of job retention, individuals with lower secondary level of education are somewhat more vulnerable to pandemic-related job loss. Regarding health variables, having worse health status prior to the pandemic, having visited a doctor, and having experienced COVID-19 symptoms are associated with higher chances of unemployment. Finally, the Shapley value results also find a very small association between working in the 'health and social work' category and unemployment.

A final visualization of the Shapley results is presented by the Shapley force plot in Figure 12, where a certain cluster is magnified in the lower panel as an example. The variables with the ten highest Shapley importance values are represented in the diagram, where darker colors represent higher Shapley scores. The rest of the variables are grouped into the 'rest variables' category marked in yellow. Each bar in the Shapley force plot represents a person in the training data; higher bars indicate higher unemployment probabilities predicted for those individuals as opposed to lower (or negative) values. As a result, the force plot displays how each feature contributed to person-specific

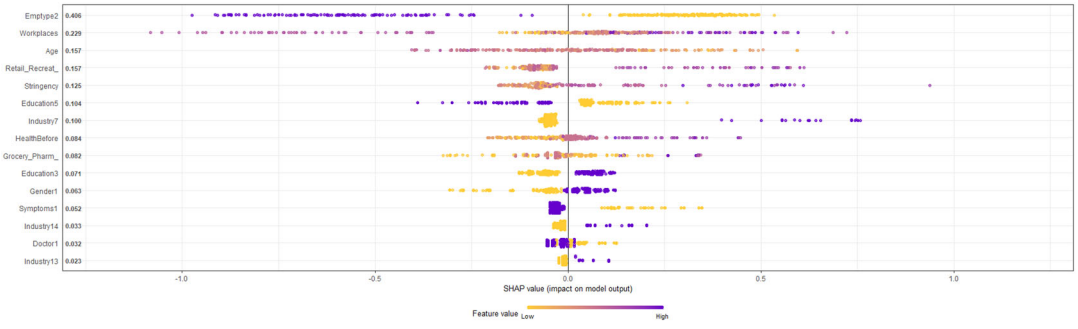


FIGURE 11 Shapley values

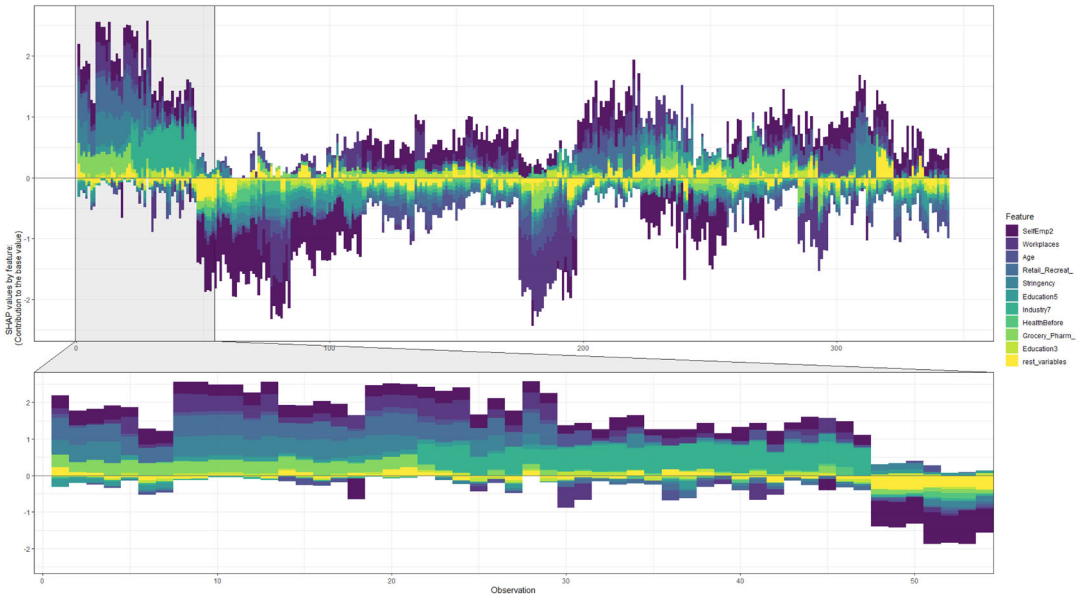


FIGURE 12 Shapley force plot

predictions. The low area covered in yellow for most individuals suggests that the SGBM-based Shapley procedure has been able to explain most of the reasons behind prediction outcomes using a set of variables similar to those found important by the random forest and the single-classification tree models.

The above detailed series of tree-based models, and the interpretable machine learning techniques used to elaborate on their findings, have allowed us to gain new insights from the SHARE dataset, which would not be possible to derive using conventional statistical methods. The variables were selected in an algorithmic manner, and subsequently ranked, scored, and assessed using repeated predictions with iterative procedures. The results for the industry, employment type, mobility, and stringency variables are consequently robust to the inclusion of a large set of features, interactions, and nonlinear relationships. Furthermore, through the utility of Shapley values, we have been able to evaluate the magnitudes and directions of the effects of specific feature values on the prediction of employment status.

5 | CONCLUDING REMARKS AND POLICY IMPLICATIONS

In this study, we observed how the COVID-19 pandemic, in its initial stage, has affected the individuals in older-age cohorts unequally in terms of their job status in labor markets in Europe. We have identified vulnerable subgroups, that is, those who are less likely to retain their jobs in the face of pandemic-related effects, and uncovered the personal-, industrial-, and country-level conditions that distinguish these groups. Our findings add to the pool of evidence showing that the pandemic impacts have not been uniform across different groups. This finding is not new in human history. For instance, the first bubonic plague pandemic (AD 541–549), also referred to as 'the plague of Justinian' after the Eastern Roman emperor who contracted and survived the disease, hit the poorer persons the hardest (Shrewsbury, 2005). In fact, Shrewsbury (2005) argues that the bubonic plague was 'chiefly a disease of the poor' in the British Isles, spread by the rats and fleas that nest in wattle-and-daub hovels in which the poor live as opposed to manors and castles made of stone and wood. The Black Death (AD 1346–1353) caused significant population decrease in Europe, and in its aftermath population growth shifted power to the landowners in labor and land

markets, who bought peasant property, which in turn led to increased land rents and decreased wages despite the efforts by the rulers to alleviate these harsh conditions (Findlay & Lundahl, 2020). The more recent pandemics, the Spanish Flu (1918–1920) and the Swine Flu (2009–2010) breakouts have particularly affected the socioeconomically disadvantaged groups such as young widows with many children, eventually leading to new pension schemes in South Africa (Mamelund, 2017). The effects of the Spanish Flu have even been observed in the long term, as in the United States individuals born from pregnancies during the height of the pandemic experienced significantly higher unemployment and school dropout rates and disabilities (Mamelund, 2017). In Oslo, mortality rates were negatively correlated with income and quality of residence (Mamelund, 2017).

There is clearly a wealth of anecdotal findings on pandemic-related effects. The present research also presents contributions to the historical evidence on unequal pandemic effects. Our tree-based machine learning models, thanks to being rooted in recursive binary partitioning techniques, have been instrumental for extracting novel information from the data by permitting all possible interactions and nonlinearities, leading to a segmentation of small but highly vulnerable subgroups within a sample of working individuals over 50 years of age. For instance, we discovered that, despite being known as working in professions with relatively high job security, public sector employees with lower prepandemic health levels have been prone to job loss. On the other hand, while we observed that initial health status is also important in the private sector, there have been distinct attributes related to age, industry, and mobility/stringency effects – particularly in the retail and recreation category – that had even stronger effects on employment outcomes. Such stringency policies may lead to long-term impacts for the individuals in the self-employed group as trust in governments and institutions are crucial factors that support entrepreneurial behavior (Celbiş, 2021b). Education alongside health are found to be important determinants for persons born before 1955. It seems also plausible that a lower labor mobility of older-age cohorts makes these groups more vulnerable in cases of a shock like the COVID-19 pandemic. The existence of such distinct groups calls for in-depth case studies for the purpose of designing special sets of policy interventions based on distinct target groups. Against this background, our findings can open new avenues for policy-relevant explorations. The discoveries made by the machine learning methods can be complemented by qualitative and quantitative follow-up studies to better understand the socioeconomic and spatial labor market mechanism of the identified subgroups and their special needs together with the policies that can meet those needs. Finally, the results of the present study may also hint at changes in future job preferences based on industry resilience in the postpandemic world.

ORCID

Mehmet Güney Celbiş https://orcid.org/0000-0002-2790-6035

REFERENCES

Adam-Bourdarios, C., Cowan, G., Germain, C., Guyon, I., Kégl, B., & Rousseau, D. (2015). The Higgs boson machine learning challenge. In *Nips 2014 workshop on high-energy physics and machine learning*, pp. 19–55.

Albanesi, S., & Şahin, A. (2018). The gender unemployment gap. *Review of Economic Dynamics*, *30*, 47–67.

Aldrich, C., & Auret, L. (2013). *Unsupervised process monitoring and fault diagnosis with machine learning methods*. London: Springer.

Alon, T., Doepke, M., Olmstead-Rumsey, J., & Tertilt, M. (2020). The impact of COVID-19 on gender equality. *NBER Working Paper*, *37*, 26947.

Alonso-Villar, O., & Del Río, C. (2008). Geographical concentration of unemployment: A male–female comparison in Spain. *Regional Studies*, *42*(3), 401–412.

Athey, S. (2018). The impact of machine learning on economics, *The economics of artificial intelligence: An agenda*. Chicago: University of Chicago Press, pp. 507–547.

Athey, S., & Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, *11*(1), 685–725.

Atkinson, E. J., & Therneau, T. M. (2000). *An introduction to recursive partitioning using the rpart routines*: Mayo Foundation, Rochester.

Azmat, G., Güell, M., & Manning, A. (2006). Gender gaps in unemployment rates in OECD countries. *Journal of Labor Economics*, *24*(1), 1–37.

Barrot, J.-N., Grassi, B., & Sauvagnat, J. (2021). Sectoral effects of social distancing. *American Economic Review Papers and Proceedings*, *111*, 277–81.

Bauer, A., & Weber, E. (2021). COVID-19: How much unemployment was caused by the shutdown in Germany? *Applied Economics Letters*, *28*(12), 1053–1058.

Bell, D. N. F., & Blanchflower, D. G. (2011). Young people and the Great Recession. *Oxford Review of Economic Policy*, *27*(2), 241–267.

Benedetti, R., Piersimoni, F., Pignataro, G., & Vidoli, F. (2020). Identification of spatially constrained homogeneous clusters of COVID-19 transmission in italy. *Regional Science Policy & Practice*, *12*(6), 1169–1187.

Blundell, J., & Machin, S. (2020). Self-employment in the COVID-19 crisis. In *CEP Covid-19 Analysis Paper No.003. London, Centre for Economic Performance, London School of Economics and Political Science*.

Boehmke, B. (2020). University of Cincinnati Business Analytics R Programming Guide. https://uc-r.github.io/

Boeri, T., Giupponi, G., Krueger, A. B., & Machin, S. (2020). Solo self-employment and alternative work arrangements: A cross-country perspective on the changing composition of jobs. *Journal of Economic Perspectives*, *34*(1), 170–95.

Börsch-Supan, A. (2021a). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 8. COVID-19 Survey 1. Release version: 1.0.0. SHARE-ERIC. Data set. https://doi.org/10.6103/SHARE.w8ca.100

Börsch-Supan, A. (2021b). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 8. Release version: 1.0.0. SHARE-ERIC. Data set. https://doi.org/10.6103/SHARE.w8.100

Breen, R. (2005). Explaining cross-national variation in youth unemployment. *European Sociological Review*, *21*(2), 125–134.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Monterey, CA: Wadsworth and Brooks.

Breiman, L. (1996). Bagging predictors. *Machine Learning*, *24*(2), 123–140.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32.

Breiman, L., & Cutler, A. (2020). Random forests. https://www.stat.berkeley.edu/∼breiman/RandomForests/cc_home.htm

Bui, T. T. M., Button, P., & Picciotti, E. G. (2020). Early evidence on the impact of coronavirus disease 2019 (COVID-19) and the recession on older workers. *Public Policy & Aging Report*, *30*(4), 154–159.

Celbiş, M. G. (2021a). Applications of machine learning models in regional and demographic economic analysis: A literature survey. In Cochrane, W., Cameron, M. P., & Alimi, O. (Eds.), *New frontiers in regional science: Asian perspectives*: Springer Singapore, pp. 219–229.

Celbiş, M. G. (2021b). A machine learning approach to rural entrepreneurship. *Papers in Regional Science*, *100*(4), 1079–1104.

Celbiş, M. G. (2021c). Social networks, female unemployment, and the urban–rural divide in Turkey: Evidence from tree-based machine learning algorithms. *Sosyoekonomi*, *29*(50), 73–94.

Chen, S., Igan, D. O., Pierri, N., Presbitero, A. F., Soledad, M., & Peria, M. (2020). Tracking the economic impact of COVID-19 and mitigation policies in Europe and the United States. (*125*). Washington.

Chen, T., & Guestrin, C. (2016). xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794.

Chen, T., & He, T. (2015). Higgs boson discovery with boosted trees. In *NIPS 2014 workshop on high-energy physics and machine learning*, *69-80*.

Chen, T., He, T., Benesty, M., Khotilovich, V., & Tang, Y. (2015). xgboost: Extreme gradient boosting. R package version 0.4-2, 1-4.

Choudhry, M. T., Marelli, E., & Signorelli, M. (2012). Youth unemployment rate and impact of financial crises. *International Journal of Manpower*, *33*(1), 76–95.

del Rio-Chanona, R. M., Mealy, P., Pichler, A., Lafond, F., & Farmer, J. D. (2020). Supply and demand shocks in the COVID-19 pandemic: An industry and occupation perspective. *Oxford Review of Economic Policy*, *36*(Supplement_1), S94–S137.

Deryugina, T., Shurchkov, O., & Stearns, J. (2021). COVID-19 disruptions disproportionately affect female academics. *American Economic Review Papers and Proceedings*, *111*, 164–68.

Fana, M., Pérez, S. T., & Fernández-Macías, E. (2020). Employment impact of COVID-19 crisis: From short term effects to long terms prospects. *Journal of Industrial and Business Economics*, *47*(3), 391–410.

Findlay, R., & Lundahl, M. (2020). Demographic shocks and the factor proportions model: from the plague of Justinian to the Black Death, *The Economics of the Frontier* (pp. 125–172). Berlin: Springer.

Friedman, J., Hastie, T., & Tibshirani, R. (2001). *The elements of statistical learning*, Vol. 1. New York: Springer.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, *5*, 1189–1232.

Friedman, J. H. (2002). Stochastic gradient boosting. *Computational Statistics & Data Analysis*, *38*(4), 367–378.

Ge, Z., Song, Z., Ding, S. X., & Huang, B. (2017). Data mining and analytics in the process industry: The role of machine learning. *IEEE Access*, *5*, 20590–20616.

Glaeser, E., Kourtit, K., & Nijkamp, P. (Eds.) (2020). *Urban Empires: Cities as global rulers in the New Urban World Edited by* Glaeser, E., Kourtit, K., & Nijkamp, P. New York: Routledge.

Goldstein, A., Kapelner, A., Bleich, J., & Pitkin, E. (2015). Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation. *Journal of Computational and Graphical Statistics*, *24*(1), 44–65.

Google LLC (2021). Google COVID-19 Community Mobility Reports. http://www.google.com/covid19/mobility/ Accessed: 12 November 2021.

Greenwell, B. M. (2017). pdp: An R package for constructing partial dependence plots. *The R Journal*, *9*(1), 421–436.

Grimmer, J. (2015). We are all social scientists now: How big data, machine learning, and causal inference work together. *PS: Political Science & Politics*, *48*(1), 80–83.

Gursoy, D., & Chi, C. G. (2020). Effects of COVID-19 pandemic on hospitality industry: Review of the current situations and a research agenda. *Journal of Hospitality Marketing & Management*, *29*, 527–529.

Hale, T., Angrist, N., Goldszmidt, R., Kira, B., Petherick, A., Phillips, T., Webster, S., Cameron-Blake, E., Hallas, L., Majumdar, S., & Tatlow, H. (2021). A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nature Human Behaviour*, *5*(4), 529–538.

Harding, M., & Hersh, J. (2018). Big data in economics. In *IZA World of Labor, (451)*. https://doi.org/10.15185/izawol.451

Helppie McFall, B. (2011). Crash and wait? The impact of the Great Recession on the retirement plans of older Americans. *American Economic Review*, *101*(3), 40–44.

Imbens, G., & Athey, S. (2021). Breiman's two cultures: A perspective from econometrics. *Observational Studies*, *7*(1), 127–133.

Ito, H., Hanaoka, S., & Kawasaki, T. (2020). The cruise industry and the COVID-19 outbreak. *Transportation Research Interdisciplinary Perspectives*, *5*, 100136.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning*, Vol. 112. New York: Springer.

Kapitsinis, N. (2020). The underlying factors of the COVID-19 spatially uneven spread. initial evidence from regions in nine EU countries. *Regional Science Policy & Practice*, *12*(6), 1027–1045.

Krugman, P. (1994). Past and prospective causes of high unemployment. In *Reducing unemployment: Current issues and policy options*, pp. 49–80.

Leung, K. H. B., Alam, R., Brooks, S. C., & Chan, T. C. Y. (2021). Public defibrillator accessibility and mobility trends during the COVID-19 pandemic in Canada. *Resuscitation*, *162*, 329–333.

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News*, *2*(3), 18–22.

Liu, K., Salvanes, K. G., & Sørensen, E. O. (2016). Good skills in bad times: Cyclical skill mismatch and the long-term effects of graduating in a recession. *European Economic Review*, *84*, 3–17.

Liu, Y., & Just, A. (2020). Shapforxgboost: Shap plots for 'xgboost'. R package version 0.1.0.

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Proceedings of the 31st international conference on neural information processing systems*, *4768-4777*.

Machin, S., & Manning, A. (1999). The causes and consequences of long-term unemployment in Europe, *Handbook of Labor Economics*: Elsevier, pp. 3085–3139.

Mamelund, S.-E. (2017). Profiling a pandemic, Natural History Magazine, (September), 6-10.

Martin, J. et al. (1994). The extent of high unemployment in OECD countries. *Reducing Unemployment: Current Issues and Policy Options*, 1–40.

Milborrow, S. (2019). Plot 'rpart' models: An enhanced version of 'plot.rpart'. Retrieved 21 from https://cran.r-project.org/package=rpart.plot

Mincer, J. (1991). Education and unemployment: National Bureau of Economic Research working paper 3838.

Molnar, C. (2019). Interpretable machine learning. https://christophm.github.io/interpretable-ml-book/

Mullainathan, S., & Spiess, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives*, *31*(2), 87–106.

Neumark, D., & Button, P. (2014). Did age discrimination protections help older workers weather the great recession? *Journal of Policy Analysis and Management*, *33*(3), 566–601.

Nijkamp, P., & Kourtit, K. (2022). Place-specific corona dashboards for health policy: Design and application of a "Dutchboard". *Sustainability*, Forthcoming.

Nijkamp, P., Poot, J., & Vindigni, G. (2001). Spatial dynamics and government policy: An artificial intelligence approach to comparing complex systems. In Fischer, M. M., & Frolich, J. (Eds.), *Knowledge, complexity and innovation systems*. Boston, MA: Springer, pp. 369–401.

Özgüzel, C. (2020). Agglomeration Effects in a Developing Country: Evidence from Turkey. Mimeo, Paris School of Economics halshs-02878368.

Pereira, I., & Patel, P. C. (2021). Impact of the COVID-19 pandemic on the hours lost by self-employed racial minorities: evidence from Brazil. *Small Business Economics*, 1–37.

Sassen, S., & Kourtit, K. (2021). A post-corona perspective for smart cities: 'Should I stay or should I go?' *Sustainability*, *13*(17), 9988.

Shapley, L. S. (1953). A value for n-person games. In H.W., K., & A.W., T. (Eds.), *Contributions to the Theory of Games II, Annals of Mathematics Studies no. 28*. Princeton, New Jersey: Princeton University Press, pp. 307–317.

Shrewsbury, J. F. D. (2005). *A history of bubonic plague in the British Isles*: Cambridge University Press.

Škare, M., Soriano, D. R., & Porada-Rochoń, M. (2021). Impact of COVID-19 on the travel and tourism industry. *Technological Forecasting and Social Change*, *163*, 120469.

Stough, R. R., Kourtit, K., Nijkamp, P., & Blien, U. (Eds.) (2018). *Modelling aging and migration effects on spatial labor markets Edited by* Stough, R. R., Kourtit, K., Nijkamp, P., & Blien, U. Berlin: Springer.

Štrumbelj, E., & Kononenko, I. (2013). Explaining prediction models and individual predictions with feature contributions. *Knowledge and Information Systems*, *41*(3), 647–665.

Sundararajan, M., & Najmi, A. (2020). The many Shapley values for model explanation. In *Proceedings of the 37th International Conference on Machine Learning, PMLR*, PMLR, pp. 119:9269–9278.

Sutton, C. D. (2005). Classification and regression trees, bagging, and boosting. In Rao, C. R., Wegman, E. J., & Solka, J. L. (Eds.), *Data mining and data visualization*, Handbook of Statistics, Vol. 24: Elsevier, pp. 303–329.

Türk, U. (2021). A multilevel analysis of the contextual effects in distance education outcomes during COVID-19. *Eastern Journal of European Studies*, *12*, 149–169.

Varian, H. R. (2014). Big data: New tricks for econometrics. *Journal of Economic Perspectives*, *28*(2), 3–28.

Vedder, K. (2020). From Shapley values to explainable AI. In *University of Pennsylvania, GRASP Game Theory Seminar*. Downloadable from http://vedder.io/misc/FromShapleyValuesToExplainableAISlides.pdf

Verick, S. (2009). Who is hit hardest during a financial crisis? The vulnerability of young men and women to unemployment in an economic downturn.

Wong, P.-H., Kourtit, K., & Nijkamp, P. (2021). The ideal neighbourhoods of successful ageing: A machine learning approach. *Health & Place*, *72*, 102704.

Xue, D., Liu, Z., Wang, B., & Yang, J. (2021). Impacts of COVID-19 on aircraft usage and fuel consumption: A case study on four Chinese international airports. *Journal of Air Transport Management*, *95*, 102106.

How to cite this article: Celbiş, M. G., Wong, P., Kourtit, K., & Nijkamp, P. (2022). Impacts of the COVID-19 outbreak on older-age cohorts in European Labor Markets: A machine learning exploration of vulnerable groups. *Regional Science Policy & Practice*, 1–26. https://doi.org/10.1111/rsp3.12520