Social Indicators Research (2023) 170:1097–1118
https://doi.org/10.1007/s11205-023-03233-3

ORIGINAL RESEARCH

Job Satisfaction and the 'Great Resignation': An Exploratory Machine Learning Analysis

Mehmet Güney Celbiş[1,2] · Pui-Hang Wong[2,3] · Karima Kourtit[4,5] · Peter Nijkamp[4,5]

Accepted: 2 October 2023 / Published online: 30 October 2023

Abstract

Labor market dynamics is shaped by various social, psychological and economic drivers. Studies have suggested that job quit and labor market turnover are associated with job satisfaction. This study examines the determinants of job satisfaction using a large survey dataset, namely the LISS Work and Schooling module on an extensive sample of persons from the Netherlands. To handle these big data, machine learning models based on binary recursive partitioning algorithms are employed. Particularly, sequential and randomized tree-based techniques are used for prediction and clustering purposes. In order to interpret the results, the study calculates the sizes and directions of the effects of model features using computations based on the concept of Shapley value in cooperative game theory. The findings suggest that satisfaction with the social atmosphere among colleagues, wage satisfaction, and feeling of being appreciated are major determinants of job satisfaction.

Keywords Job satisfaction · Satisfaction with coworker · Pay satisfaction · Work conditions · Job attitudes

1 Introduction

Though the global health crisis has ended, its economic impacts have only started to ripple over the global labor market. In the United States, about four million workers voluntarily quit their jobs in April 2021 (Reuters, 2019). This so-called Great Resignation is observed in other advanced economies as well. In the Netherlands, it is reported that nearly one out of five people have switched their jobs in 2022 (Algemeen Dagblad, 2023). A

✉ Pui-Hang Wong
wong@merit.unu.edu

1 LAET, ENTPE, CNRS, University of Lyon, Lyon, France

2 UNU-MERIT, United Nations University, Maastricht, The Netherlands

3 Maastricht Graduate School of Governance, Maastricht University, 6211 AX Maastricht, The Netherlands

4 Faculty of Management, Open University of the Netherlands, Heerlen, The Netherlands

5 Centre for European Studies, Alexandru Ioan Cuza University, Iași, Romania

macroeconomic explanation maintains that labor shortages in some sectors lead workers to believe that they may find a better offer to compensate the drops in real wage under a tight labor market, which may explain the high turnovers (Duval et al., 2022). While this explanation has some truth in it, research on organizational behaviors may argue that the phenomenon can be driven by low job satisfaction. Unhappy workers are more likely to quit their jobs (Green, 2010), especially during and after the pandemic period. For instance, Martin et al. (2022) found that remote working has increased job stress and reduced job satisfaction in Luxembourg. Demirkaya et al. (2022) reported that the feeling of entrapment is significantly correlated with job quits in Turkey.

Departing from the job satisfaction perspective, this article evaluates an array of individual-level factors that influence job satisfaction of Dutch workers using newly available Dutch household survey data. This study does not seek to explain the occurrence of Great Resignation. It highlights that job satisfaction has escaped from the discussion. Based on the multifaceted approach of job attitudes (Judge & Kammeyer-Mueller, 2012), an economic explanation would suggest that satisfaction of pay and of work conditions are major drivers of the behaviors of workers. In this study we examine to what extent economic concerns predict overall job satisfaction during 2022 using a novel machine-learning approach. Findings from this study should help us to assess which facets of job satisfaction are associated with the 'Great Resignation' in the Dutch labor market.

2 Job Satisfaction and the Utility Approach

Job satisfaction is commonly referred to as employees' affect and attitude toward their job. Emphasized on the affective dimension, Locke (1976) defined the concept as "a pleasurable or positive emotional state resulting from the appraisal of one's job or job experiences" (p. 1300). Focusing on the cognitive aspect of job satisfaction, Weiss (2002) defined the concept as "a positive (or negative) evaluative judgment one makes about one's job or job situation" (p. 6). The affective view sees affect at work as an *indicator* of job satisfaction, while the cognitive, evaluative approach considers affective experiences on the job as a *source* of job satisfaction.

In economics, job satisfaction is frequently treated as a unidimensional variable and a function of wage (Borjas, 1979; Hamermesh, 1977).[1] An implication of this formulation is that, when members of certain social groups (e.g., women) are discriminated in the labor markets, they would experience a lower level of job satisfaction because of a lower wage or other non-wage benefits (Bartel, 1981). Interestingly, while a gender wage gap is found to exist (Blau and Kahn, 2017), studies have shown that women usually report a higher level of job satisfaction than men, a stylized fact that contradicts what the theory predicts (Blanchflower & Oswald, 1999; Clark, 1997; Sloane & Ward, 2001). However, when factors such as age and job expectations are controlled for, the gap in job satisfaction between the two genders diminishes, suggesting that job satisfaction can depend on some often-overlooked demographic factors.

The utility function approach also implies that satisfaction is, at least partly, driven by extrinsic incentives. This early conceptualization is based on the utility theory that labor

[1] A comparative study by Dolbier et al. (2005) shows the single-item measure of job satisfaction can be a psychometrically sound measurement.

supply is a rational choice after a careful deliberation to the trade-off between leisure and consumption backed by the paycheck. Although studies have found that salary is an important determinant of job satisfaction, many studies have documented that workers do value a wide range of nonpecuniary characteristics of a job, which include job security, autonomy, shorter and more flexible work hour (Berger et al., 2019; Clark, 2001; Lepinteur, 2019; Origo & Pagani, 2009). A study by Lange (2012) further showed that these job characteristics could be even more influential than individual-specific factors like personality traits and values.

If people are happy with their jobs, they should be less likely to switch jobs. Some studies use quit data to analyze impacts of pecuniary and non-pecuniary factors on job satisfaction. Clark (2001) not only showed that job satisfaction is a powerful predictor of job change, it also found that job security and pay are the most important determinants of quit. Although not been emphasized in the literature, analyses of quit data may suggest that factors contributing to satisfaction and dissatisfaction may not be identical and their impacts could also be asymmetric. These asymmetric effects, however, have not been highlighted and emphasized in the existing literature.

One major conclusion from the job satisfaction literature is that happy workers are more productive (Oswald et al., 2015). But what beneath is the belief that a well-designed reward system should improve workers productivity. The starting point of personnel economics is the principal-agent conceptualization (Laffont & Martimort, 2002) that workers may shirk (i.e., moral hazard) and productive workers are costly to recruit (i.e., adverse selection). Clever economic mechanisms are required to identify these workers ex ante and to induce their efforts ex post. From this perspective, many of the human resource management practices such as performance pay, promotion, and job autonomy can be seen as performance optimizing mechanisms. Although financial and non-pecuniary rewards are parts of the job satisfaction equation (Cassar & Meier, 2018; Cornelissen et al., 2011; Ellingsen & Johannesson, 2007; Gosnell et al., 2020; Jones et al., 2009), from this cynical view, job satisfaction is nothing more than a happy by-product of human resources policies or a means to productivity and performance.

3 Individual and Social Dimensions of Job Satisfaction

An underlying assumption behind the utility approach is that income induces satisfaction and/or happiness. Although studies have found a significant but weak relationship between the two variables, behavioral studies have challenged this fundamental assumption (Clark et al., 2005; Easterlin, 1995). While some studies suggest that the relationship may be causal (Powdthavee, 2010), three decades of economics of happiness research have contested this finding (Clark et al., 2008). One major conclusion from the literature is that subjective well-being does not always increase with income. When examined the dynamic relationship between income and subjective well-being, Easterlin (2001) argued that income increases happiness initially. But aspirations grow as one climbs the income ladder. Over a life cycle, people's level of happiness remains stable and does not increase along with salary. Increase in income has only a short-term effect on happiness. The same mechanism may explain why job satisfaction may not catch up with income.

Another explanation to the weak statistical correlation between money and satisfaction over time is related to social comparisons. Taking a geographical approach, Luttmer (2005) found that people who earn an income lower than the local average feel worse off. A large amount of economics research in job satisfaction has been testing this hypothesis. Using data on British workers, Clark and Oswald (1996) found a similar relationship between job satisfaction and *expected* income. In an experimental study, Card et al. (2012) showed that the knowledge about their earnings below the median income of their peers significantly reduces job satisfaction and increases job search intentions. People form expectations as well as inspirations based on their peers, work conditions, and wage history (Diriwaechter & Shvartsman, 2018; McBride, 2001; Poggi, 2010). When the expectation-aspiration spiral is kick-started, an increase in aspirations can negatively affect people's satisfaction levels (Mcbride, 2010). What remains unclear is whether job dissatisfaction is due to pure social comparisons or fairness concerns (Card et al., 2012; d'Ambrosio et al., 2018; Smith, 2015). Interestingly, in a study using matched employer–employee panel data in Denmark, Clark et al. (2009) found that job satisfaction is *positively* correlated with their co-workers', a finding that counters many of the existing research in this area. They interpret that this relationship can be related to people's expectations about their future earnings. A higher average salary level leads the thinking that their potential wage may increase soon. All in all, while studies have shown that expectations and aspiration matter, it remains unclear how they are formed based on social comparisons and wage profiles.

One major assumption behind the neoclassical formulation of the utility function is that people gain utility mainly from consumption but not from the job per se. While the personnel economics literature covers aspects such as reward systems from a well-grounded, humanized (i.e., incentive-based) perspective, it is not unrealistic to think that under certain conditions, work can lead to a sense of achievement, which in turns shapes aspirations and hence job satisfaction (Genicot & Ray, 2020). Another stylized fact in the literature is that entrepreneurs and the self-employed are found to have a higher level of job satisfaction (Lange, 2012; Millán et al., 2013). The risk-adjusted returns and job security of entrepreneurship is known to be low. The existence of intrinsic motivations, however, offers a plausible explanation to the surprising, stylized fact (Carree & Verheul, 2012). In a laboratory setting, Ariely et al. (2008) and Chandler and Kapelner (2013) manipulated the perceived meaningfulness of a task and found that meaningfulness influences effort and labor supply behaviors. This lends support to the idea that the meaning of work could be part of the utility function in its own right (Cassar & Meier, 2018). Drakopoulos and Theodossiou (1997) considered a hierarchical utility function, in which increase in earnings, up to a certain point, ceases to induce utility, and the marginal utility of other work-related variables becomes much higher thereafter. Although the proposed modification does not directly speak to the fulfillment mechanism, the modification formally speaks to human's intrinsic motivations and the feeling of satisfaction.

If intrinsic motivations involve meaning making, a job that is connected to a person's educational background, skill sets, and competency should be more fulfilling. In fact, Nikolova and Cnossen (2020) showed that intrinsic motivations, measured by perceived job autonomy and competence, matter even more than extrinsic rewards. Feeling competent is pleasant and induces satisfaction (Loewenstein, 1999). Nevertheless, some studies have shown that the competence concern may backfire. García-Mainar and Montuenga-Gómez (2020) found that, in terms of education, overqualified workers tend to dissatisfy with their jobs. The same applies to horizontal educational mismatch—when graduates are employed in an occupation unrelated to their fields of study (Levels et al., 2014)—and skill mismatches (Vieira, 2005). However, there is only little evidence on the effects of skill obsolescence and gaps on job satisfaction (McGuinness et al., 2018).

One fundamental criticism on the study of job satisfaction is that many findings from these empirical studies may not be easily fed into neoclassical economic theories.

To some economists, if job satisfaction is nothing more than yet another term in a utility function, it can only be assumed but not explained. Accordingly, this leads to a proposal that, because job satisfaction is unobserved and is related to some volatile, external factors like economic fluctuations and labor market policies (Pilipiec et al., 2020, 2021; Ravid et al., 2017), and is unstable even within-individual (Bryson & MacKerron, 2017), instead of fixating on job satisfaction and treating it as a dependent variable, a more fruitful approach is to treat it as an explanatory variable to study worker behaviors such as quits and labor market functioning like employee turnover (Hamermesh, 2004).

Although the general working environment and coworker relationship has been an important dimension of job satisfaction in organizational psychology (Jolly et al., 2021; Kinicki et al., 2002; Smith et al., 1969), the variable has received relatively little attention in economics literature, probably due to its difficultly in incorporating into the existing theoretical framework. The same also seems to hold in organizational studies (Judge & Kammeyer-Mueller, 2012). Cassar and Meier (2018) discussed feelings of relatedness in the context of meaning of work and productivity, but they did not articulate how the concept can be related to job satisfaction. They also mentioned social comparison and fairness, which was discussed above, but they clearly did not relate it to coworker relationship. Intuitively it is easy to understand why collegial relationship may influence job satisfaction. Karlsson et al. (2004) explained that social extensions such as a family and work ameliorate feelings of inconsequentiality, and people could find meaning in one's life. But it is not obvious how it can be related to the meaning of a job (Nikolova & Cnossen, 2020). One possibility is about the pursuit of common (organizational) goals in a team setting. On a material level, cooperation makes success more likely and helps to achieve higher output or goals which cannot be completed alone. On the social level, contributing to a common goal can create a warm-glow effect when individuals consider working with or "helping" their colleagues as altruistic acts and gain utility from that (Andreoni, 1990). From this perspective, a friendly work environment can be considered as a public good. Even for pure egoists, they can gain utility from contributing to the building of a constructive work environment and creating positive spillovers simultaneously. In this regard, an affable work environment can be considered as a by-product rather than a source of job satisfaction.

In our study, the relative importance of different facets of job satisfaction will be tested using a predictive, machine learning approach. Many psychometric tests have been developed to assist in clinical diagnoses. Thus, a predictive approach is well-established in psychology. In fact, predictive validity is a core property of psychometric measures (Mulder et al., 2014). Linearity is commonly assumed in traditional statistical measures and tools such as correlation coefficients and structural equation modeling when linear algebra is used. However, higher dimension interactions and linearity are plagued in many relations. The machine learning approach employed in this study is able to capture nonlinearity which is not easily modeled using traditional regression methods (James et al., 2013). Additionally, as argued above, there are two major limitations in traditional studies using regression analysis: (1) a variable that influences job satisfaction necessarily affects dissatisfaction, and (2) the effects of a variable on job satisfaction are symmetric. An analytical advantage of the machine learning approach is its ability to reveal potential asymmetry between variables in a relationship. The notion of nonlinearity would become clearer when it is discussed in the result section.

4 Data

Data in our analysis were drawn from the Work and Schooling module of the Dutch Longitudinal Internet studies for the Social Sciences (LISS; Streefkerk & Centerdata, 2022). The LISS panel is administered by Centerdata research institute based in Tilburg University. Based on the population register of the national statistical office of the Netherlands, a random, nationally representative sample was drawn. In this study the fifteenth wave of the LISS survey was used, which is the most recent one after the pandemic. The online survey was conducted between the 4th of April 2022 and the 31st of May 2022. The cross-sectional dataset that was used consists of 420 variables with 5775 responses.

One advantage of the machine learning approach is its ability to include in an analysis a large number of variables, an approach which is seldom adopted in a typical regression analysis due to the concern of multicollinearity. Instead of pre-selecting variables, which can involve personal bias, the machine learning method takes a data-driven approach and includes as many of context relevant variables as possible. The method, however, involves a trade-off: the inclusion of additional variables usually reduces the sample size because of the missing value problem. Therefore, following Celbiş et al. (2023), an algorithmic process is implemented in order to optimize the number of observations. In each iteration, a simple regression tree analysis is conducted, and the root mean squared error (RMSE) is recorded. In the next step, the algorithm searches for the variable that, if excluded, would cause the greatest reduction in the number of observations. Then the variable is dropped from the dataset and the process is repeated. The data matrix that keeps most of the observations and has smaller impact on accuracy is used for the final analysis. Figure 6 (in the Appendix), visualizes the observation-feature trade-off of the procedure. In each iteration, represented by the x-axis, either one or more variables are dropped. Identifying the variable(s) to be dropped in each iteration is done through generating an UpSet plot developed by Lex et al. (2014).[2] In Fig. 7 (in the Appendix) we present a sample UpSet plot built in the 4th iteration as an example. According to this UpSet plot, the intersection of the features cw22o582, cw22o583, cw22o584, cw22o585, and cw22o586 account for the largest loss in observation as shown by the first vertical bar, suggesting that these features may represent connected or follow-up questions in the survey which are usually entered as missing in conjunction. Within the iteration, a regression tree is fitted into the subset of the training data which omits the above specified variables and its RMSE is noted. In the next iteration, the highest horizontal bar will correspond to the variable cw22o510, as the five features with higher bars below it will have already been dropped. In this new iteration, cw22o510 would account for the largest decrease in observations by itself unlike the earlier dropped group of variables. Therefore, the iteration will only drop cw22o510.[3] A new regression tree then is fit into this new subset of the training data and a new RMSE value will be computed. The recursive steps continue until the features of the dataset are exhausted or until the dataset has no observations left with missing values. In our case, this corresponds to iteration 54 (as shown in Fig. 6) where the y-axis represents the percentage of observations (persons) or variables that is left in this iteration. As variables responsible for high missing values are dropped, the percentage of observations retained increases. The

[2] The implementation of the UpSet plots in this study is done through the nainar package in R developed by Tierney and Cook (2023).

[3] The recursive step of building sequential UpSet plots and selecting the variable or groups of variables to omit is coded by the authors.

purpose of this procedure is to identify the optimum combination of the number of observations and variables that is likely to yield the highest prediction performance in our subsequent main empirical implementations of machine learning models. However, the case at hand suggests that RMSE (not represented by any axis) is not sensitive to this trade-off as suggested by the nearly flat RMSE curve in Fig. 6. As a result, the choice of the desired observation-variable balance becomes somewhat subjective. We selected the combination which retains a balance in this trade-off such that the difference between the percentage of persons retained and the percentage of variables retained is at a minimum. This corresponds to iteration 30. Prior to implementing the above outlined steps, variables that are completely consisted of missing information, variables with no variation (i.e., same value reported for all persons), administrative variables coded into the questionnaire (e.g., start date of the interview, duration of the interview) were dropped. After extensive data cleaning and validation procedure, which involved the above algorithmic trimming of the dataset, the final dataset consists of 1878 individuals and 89 variables. 30% of this data is randomly selected and set aside as the test dataset. All models are applied on the remaining training data. The results are assessed by evaluating the root mean squared error of the models using the test data. The definitions of the top ten features selected by the model in addition to the dependent variable (job satisfaction) are presented in Table 1.

5 Empirical Models

The empirical analysis takes on two steps: prediction and interpretation. The prediction step is based on a variation of the Gradient Boosting Machine (GBM) technique by Friedman (2001, 2002). GBM is applied using the Extreme Gradient Boosting (XGBoost) algorithm by Chen and Guestrin (2016). XGBoost extends the usability of GBM by allowing regularization and adding further randomization options. The prediction also relies, to a lesser extent, on the Random Forest (Breiman, 2001) technique for clustering. Both XGBoost and Random Forest are collections of weak learners based on the binary recursive partitioning algorithm by Breiman et al. (1984). Hence, randomized (Random Forest) and sequential and randomized (XGBoost) tree-based ensemble machine learning models are used in this study. The XGBoost algorithm allows for cross validation for regularization and determining the optimum model parameters including the learning rate. We partitioned the training sample into 10 subsamples (i.e., internal validation sets) to decide parameters pertaining to tree complexity (i.e., the maximum tree depth and minimum number of observations in terminal nodes) and the learning rate through cross validation. While regression trees are normally pruned through n-fold cross validation, the random forest model produces unpruned trees. However, while cross validation is absent from the random forest proximity clustering, unbiasedness is achieved through the use of the out-of-bag (OOB) observations. The resulting Shapley Additive Explanations (SHAP) values are derived from the above mentioned cross validated gradient boosting model.

XGBoost and Random Forest present several advantages thanks to their ability to consider all possible interactions and nonlinearities as the algorithms are based on binary recursive partitioning (James et al., 2013; Varian, 2014). The aggregation of many classification trees with high variance but low bias (due to their unpruned structures), built by taking repeated samples from the training data, can significantly improve prediction accuracy while reducing the variance on the prediction function (Breiman, 1996, 2001; Friedman, 2001; Friedman et al., 2001; James et al., 2013). However, as trees built using the same

Table 1 Feature definitions

Name	Description	Scale
Work Home	Having a (partial) working-at-home day	4 categories: No, less than one day per week, about one day per week, more than one day per week
Support	The individual receives sufficient support in difficult situations	4 categories: disagree entirely, disagree, agree, agree entirely
Colleagues	Satisfaction with the general atmosphere among colleagues	0 (not at all satisfied) to 10 (fully satisfied)
Freedom	Having freedom to determine how to do work	4 categories: disagree entirely, disagree, agree, agree entirely
Work Home Hours	Hours worked from home out of average weekly working hours	Continuous
Wage Satisfaction	Satisfaction with wages or salary or profit earnings	0 (not at all satisfied) to 10 (fully satisfied)
Job Satisfaction	Satisfaction with current work	0 (not at all satisfied) to 10 (fully satisfied)
Birth Year	Year of birth	Calendar year
Income	Net monthly income in euros	13 categories: No income; EUR 500 or less; EUR 501 to EUR 1000; EUR 1001 to EUR 1500; EUR 1501 to EUR 2000; EUR 2001 to EUR 2500; EUR 2501 to EUR 3000; EUR 3001 to EUR 3500; EUR 3501 to EUR 4000; EUR 4001 to EUR 4500; EUR 4501 to EUR 5000; EUR 5001 to EUR 7500; More than EUR 7500
Work Less	Hours worked per week minus number of hours preferred to work per week (a positive value means that the person would like to work less than current hours)	Continuous
Travel Time	Minutes needed to travel between home and work	Continuous

Source: LISS Panel

training dataset are expected to be highly correlated, the benefits of using an ensemble would be limited (Aldrich & Auret, 2013; Breiman, 2001; Friedman et al., 2001). The random forest algorithm aims to cope with this correlation by introducing randomized restrictions on the feature space at each iteration (i.e., it randomly selects input features in each tree) (Breiman, 2001; Breiman & Cutler, 2020; Friedman et al., 2001; James et al., 2013). Therefore, in addition to the reduction in variance through aggregation, further reduction is made possible compared to bootstrap aggregation which is an ensemble model with correlated trees (James et al., 2013). However, gradient boosting does not decorrelate trees. Instead, each tree is a modified version of the previous one (Friedman, 2001, 2002). Nevertheless, GBM embodies randomization like Random Forest but also introduces regularization which is not present in Random Forest. As a result, a group of weak learners with low variance are chained sequentially and modified with learning steps in between leading to the bias in prediction being lowered gradually in each iteration (Friedman, 2001, 2002; Friedman et al., 2001).

A random forest with 500 unregularized regression trees is generated for predicting the individual job satisfaction level for the N persons in the training data. At each iteration, some individuals are left out of the computation, because the bootstrap aggregation algorithm on which a random forest is based draws random subsamples of 2/3 of the size of the training dataset (Breiman, 2001; James et al., 2013). Further randomization is applied by selecting a split feature from a random subset of size 1/3 of the feature set at each split (Breiman, 2001). A random forest proximity matrix ($N \times N$) is produced where the proximity score of two persons is increased by 1 each time when they are predicted to fall into the same terminal unpruned regression tree node in a random forest iteration in which they were out-of-bag (i.e., randomly left out). The matrix is divided by 500 (the number of trees) and the additive inverse is computed (Aldrich & Auret, 2013; Breiman & Cutler, 2020; Friedman, 2001).

The exploration of clusters is performed based on the random forest results. Random Forest uses the proximity scores among the observations in the training dataset in order to detect cluster structures (Aldrich & Auret, 2013; Cutler et al., 2009; Friedman, 2001). The distance measures used in conventional clustering techniques such as hierarchical and k-means clustering are prone to be dominated by uninformative features that may cloud the effects of the important model features (Cutler et al., 2012). In this regard, the main advantage of the random forest proximity matrix is due to its randomization procedure which aims the aforementioned decorrelation process. In addition, unlike classic clustering approaches, feature selection in random forest proximity plots is based on the underlying model which employs algorithmic selection (Xu & Tian, 2015). Furthermore, the random forest proximity plot used for clustering in the present study is generated using OOB observations pair-wise frequencies of sharing a terminal node, which is an internal validation procedure, leading to improvements in out-of-sample performance (Breiman & Cutler, 2020; Friedman, 2001).

The random forest proximity plots tend to detect and represent one class in one arm of a star shaped visual where pure class regions of out-of-bag observations in the training data are grouped towards the extremities of an arm due to the tree-based structure of the underlying algorithm (Hastie et al., 2009; Aldrich & Auret, 2013; Cutler et al., 2009). In this regard Friedman et al., (2001, p. 595) state that "The idea is that even though the data may be high-dimensional, involving mixed variables, etc., the proximity plot gives an indication of which observations are effectively close together in the eyes of the random forest classifier".

The interpretation of the machine learning findings is an essential part of any study in social sciences, as predictions alone—regardless of their success—cannot provide clear information and policy implications. The main interpretable machine learning tool employed in this paper is the computation and assessment of Shapley Additive Explanations (SHAP) values (Lundberg & Lee, 2017) based on the cooperative game theoretical approach by Shapley (1953). SHAP values have been introduced recently to the machine learning fields. As opposed to older approaches such as calculating variable importance scores (Lundberg & Lee, 2017; Molnar, 2019), the SHAP values approach can measure both the sizes and the directions of the relationships. The computation of SHAP values is preformed using the "SHAPforxgboost" module by Liu and Just (2020). A remarkable advantage of the SHAP value approach is that the calculations of the effect sizes are done by considering many different values and (theoretically all) combinations of model features (Celbiş, 2022; Lundberg & Lee, 2017; Molnar, 2019). Consequently, when computing the effect of a given feature for a given data instance (i.e., individual) all other variables are *not* held constant as is usually done in traditional econometric approaches. The departure from the ceteris paribus restriction leads to more realistic assessments of effect sizes, as in the real-world other factors can never be held constant in the context of a social science research. Finally, because considering all possible feature combinations and values is computationally not feasible, an approximation formulated by Štrumbelj and Kononenko (2013) is used in this study.

6 Findings

Among the included variables, both satisfaction with coworker relations and the pay are important features in predicting job satisfaction. The random forest proximity plot shown in Fig. 1 visualizes the clusters based on proximities in prediction between individuals.[4] The plot suggests the existence of about three clusters based on the roles of model features in explaining job satisfaction.

- highly satisfied by colleagues atmosphere (yellow).
- moderately satisfied by colleagues atmosphere (green).
- poorly satisfied by colleagues atmosphere (dark green).

In the plot, larger-sized circles indicate individuals with lower wage satisfaction. These individuals are grouped towards the intersection of the "arms", suggesting that the model has a harder time in distinguishing them (i.e., in iterations in which they were out of sample, people with low wage satisfaction often fell into same terminal nodes when run down the tree). The clear formation of the arms as separate clusters suggests a successful differentiation of the individuals through their inherent similarities. We also observe that people who are poorly satisfied with their jobs are not all part of the same cluster; suggesting that job dissatisfaction may arise from a diverse set of factors.

Regarding the SHAP analysis, a grid search procedure yielded the following optimal parameters for the XGBoost model except for subsample and number of trees, where the former is decided based on the finding by Friedman (2002) and the latter on computational restrictions:

[4] The matrix dimensions are reduced using metric multidimensional scaling (Friedman, 2001).

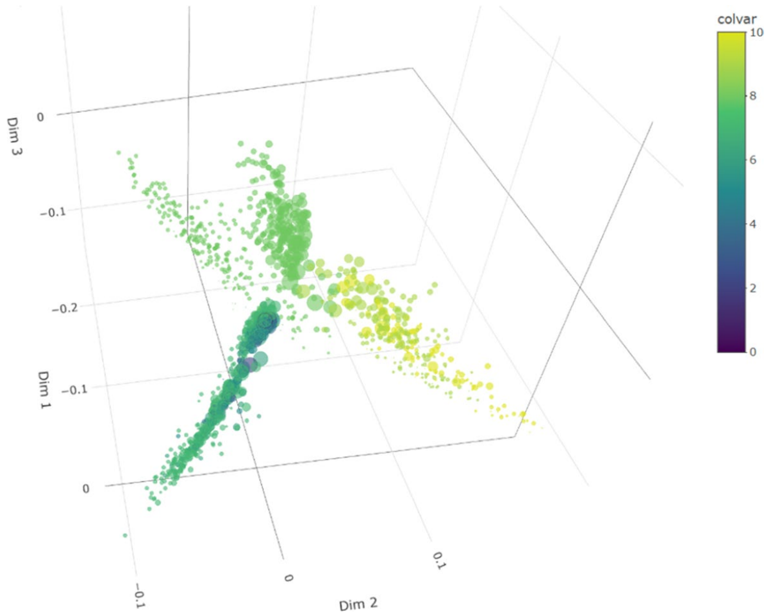

Fig. 1 Clusters detected by random forest predictions

- Learning rate: 0.01
- Maximum tree depth: 10
- Minimum number of observations in terminal nodes: 1
- Subsample ratio in each iteration: 0.5
- Feature subsample ratio in each iteration: 0.5
- Number of Trees: 10,000

The model is run on the training data, which consists of 70% of the observations randomly sampled from the full dataset. The job satisfaction levels of the individuals in the remaining test data are predicted with a RMSE of 1.14. The SHAP values that are computed based on the XGBoost predictions are visualized in Fig. 2, where each dot represents an individual. The values show the contribution of each feature value—where features are on the y-axis and higher values are represented with darker colors—on the deviation of a specific individual's predicted value from the mean prediction (the point 0 on the x-axis). The top ten features with the highest SHAP importance values, listed next to the variable name, are presented in the figure. The features that affect job satisfaction the strongest are the first three variables, as the SHAP importance values of the remaining features are all less than 0.1.

The SHAP analysis summarized in Fig. 2 suggests that a high satisfaction of the atmosphere among colleagues has a positive effect on job satisfaction. Furthermore, this variable has the highest importance in the prediction of job satisfaction. The relationship is shown in more detail in the SHAP dependence plot in Fig. 3 (a slight amount of jitter is used for better representation). A value of 10 for this variable alone can account to up to more than

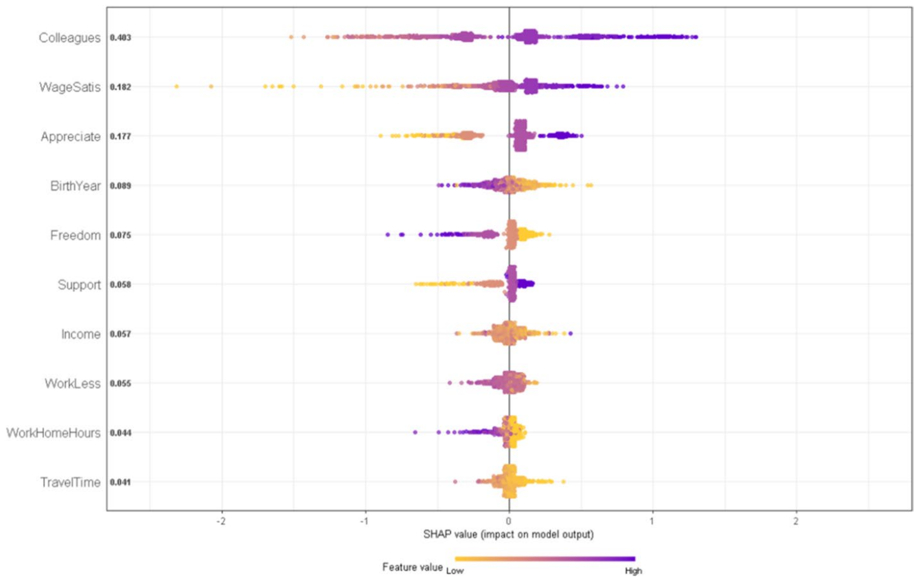

Fig. 2 SHAP values

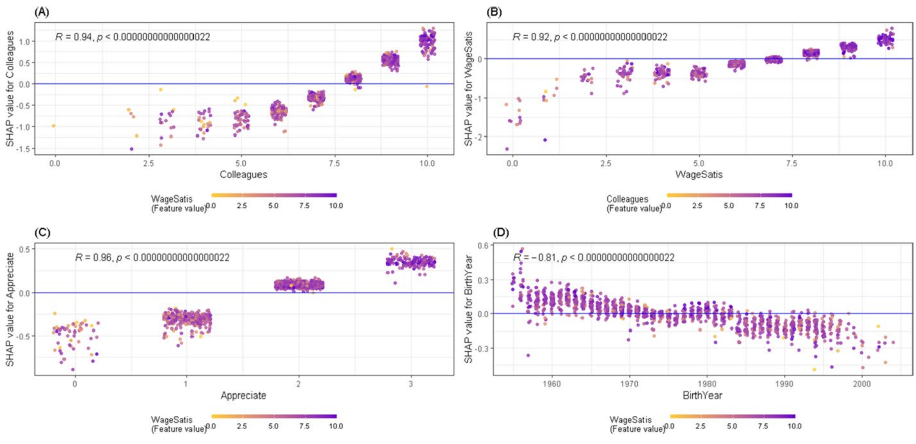

Fig. 3 SHAP feature dependence plots

1 point (in Likert scale) positive deviation from the mean job satisfaction value in the training data.

Focusing on wage satisfaction, we observe that high wage satisfaction has a positive effect on job satisfaction. While this is not surprising, it should be highlighted that the importance of wage satisfaction is less than half of that of satisfaction in colleagues. The interaction and dependence between these two top variables are further visualized in the form of a two-way partial dependence plot (PDP) in panel A of Fig. 4. Unlike the SHAP dependence plots, the PDPs represent joint predictions by holding constant all other features, except one or two features of interest (Friedman, 2001). The predictions are

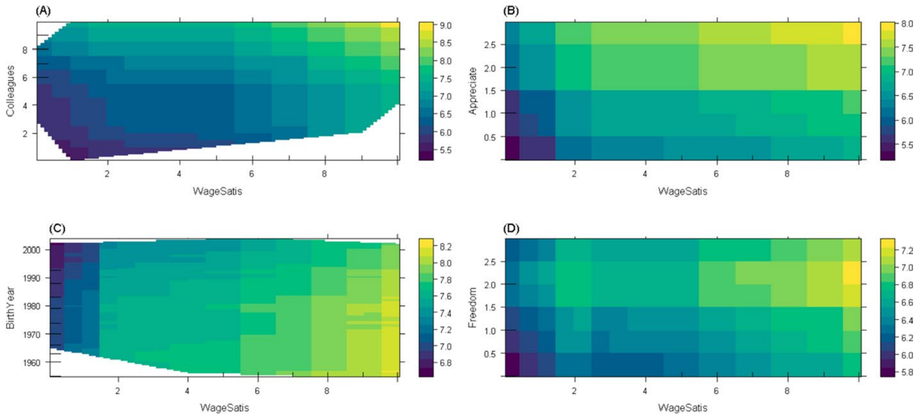

Fig. 4 Two-way patrial dependence plots

recalculated for each value of the features(s) of interest and averaged over the individuals in the training data (Friedman, 2001). However, some predictions suggested by the PDPs may be implausible for features that have high correlation with variables that are held constant (Friedman, 2001; Molnar, 2019). The PDP for *Colleagues* and *Wage Satisfaction* suggests that high wage satisfaction, while making a difference, does not predict truly high job satisfaction levels in the absence of high satisfaction in the atmosphere with colleagues. A similar outcome can be deducted from panel B of Fig. 4 which shows that high wage satisfaction without feeling of appreciation does not predict high job satisfaction. We also observe in panel C of Fig. 4 that higher wage satisfaction is associated with high job satisfaction, but the effect is stronger for older individuals. Finally, panel D suggests that high wage satisfaction without work freedom does not predict high job satisfaction.

The feature *Appreciate* is ranked third in Fig. 2. This result suggests that the perception that people get the appreciation they deserve for their work has a positive effect on job satisfaction. Feeling unappreciated has a slightly stronger negative effect than the positive effect of feeling appreciated.

The SHAP values of the variable *Birth Year* are quite symmetrically spread about zero and indicate that younger individuals in the dataset tend to be less satisfied with their jobs. The effect of this variable and the remaining ones are relatively small compared to those of *Colleagues*, *Wage Satisfaction*, and *Appreciate*.

The remaining features in the set of top ten variables with the highest SHAP importance values do not determine job satisfaction to a high extent individually, but collectively they affect the prediction. We briefly summarize the suggestions of their SHAP values. Lack of freedom in organizing one's work has a negative effect on job satisfaction. However, the effect is not symmetric: the positive effect of high freedom is less than the negative effect of lower freedom. Furthermore, the perception of getting enough support in difficult situations has a positive effect. Similar to the freedom variable, the effect of perceived support is also not symmetric.

The income variable is among the top used variables by the algorithm in predicting satisfaction. However, its effect is small and not clear. In elaborating personal economic gain, we instead focus on the wage satisfaction feature.

Preferring to work less hours than current is negatively associated with job satisfaction, but the effect is not very pronounced. Another feature in relation to work hours, *Work*

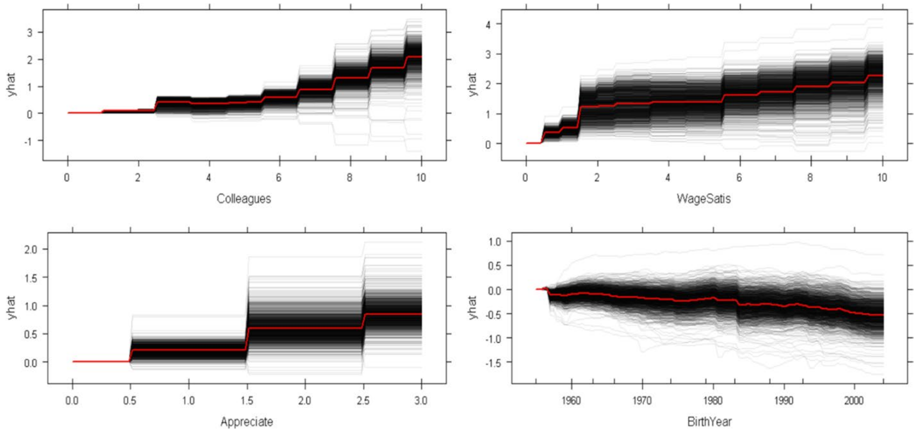

Fig. 5 Individual conditional expectation plots

Home Hours suggest that more hours worked from home has a negative effect on job satisfaction. Again, the effect is not symmetric. Finally, for travel time the effect is weak, but it is evident that long travel duration has a negative impact on job satisfaction.

Further information is given by the SHAP dependence plots. In panels A and B of Fig. 3 we observe that the positive effect of *Colleagues* and *Wage Satisfaction* begin after about a score of 7, and that there exist individuals for whom their high satisfaction with their colleagues make a big impact on their job satisfaction despite low wage satisfaction. On the other hand, panel C shows that feeling appreciated has a positive effect that becomes stronger if the person agrees with this opinion stronger. Finally, in panel D we observe that younger persons (born after around 1985) tend to have lower job satisfaction even if they may have high wage satisfaction. The effect is more negative the younger the person is.

Whereas the above discussed partial dependence computations visualize the mean predictions of individual job satisfaction level, individual conditional expectation (ICE) plots (Goldstein et al., 2015) visualize the predicted change in everyone's job satisfaction level by plotting individual curves. Centered ICE plots provide a more explicit presentation by anchoring each individual curve at a given y-intercept value (Goldstein et al., 2015; Molnar, 2019). The ICE plots shown in Fig. 5 visualize the predicted paths for each individual in the training data where each line represents an individual. The plots for *Colleagues*, *Wage Satisfaction*, *Appreciate*, and *Birth Year* respectively suggest that the direction of the effects is mostly similar for individuals, although a small amount of heterogeneity in expectations exists.

7 Discussion and Conclusion

Welfare maximization is generally, in standard economic textbooks, regarded as a respectable economic objective or driver in any society. But the empirical measurement of welfare (including life satisfaction) is still fraught with many hurdles and uncertainties. In practice, GDP per capita—or, in a labor market context, wage rates—are often regarded as signposts for economic performance. However, this measuring rod has many serious shortcomings, such as the neglect of distributional and equity aspects, the exclusive focus on income to the detriment of essential consumption categories (such as human health, food, safety, education, green environment, quality of life), the bias caused by the presence of negative externalities or social costs (e.g., climate change, social stress, environmental decay), or the omission of the welfare implications of the worker's balance between leisure time and working time. The welfare of a society depends among other aspects on how satisfied laborers in that society are with their work and work environments. The present study has aimed to shed further light on the individual dimension of social welfare.

The present study used the most recent wave available of the Work and Schooling module of the LISS survey on individuals in the Netherlands. The empirical analysis was mostly founded on tree-based sequential ensemble prediction algorithms. The predictions were elaborated in detail using interpretable machine learning techniques to quantify the strengths and directions of the relationships between the survey features and the level of job satisfaction of an individual.

The main result is that wage satisfaction alone is not sufficient to ensure job satisfaction for the analyzed sample of individuals from the Netherlands. Being satisfied with the atmosphere among one's colleagues and feeling appreciated are also essential for job satisfaction. While low wage satisfaction can have a strong negative effect on job satisfaction, high satisfaction with colleagues has a stronger potential positive effect on job satisfaction compared to the effect of wage satisfaction. Among other results, we also observe that younger people are less satisfied with their jobs.

It is believed that facet-based measures such as the Job Descriptive Index (JDI) predict overall job satisfaction well (Judge & Klinger, 2008). These measures cover similar dimensions of job satisfaction (Dunham et al., 1977; Kinicki et al., 2002). For instance, the JDI examines job satisfaction in five dimensions: work, supervision, coworkers, pay, and promotion. The Minnesota Satisfaction Questionnaire (MSQ; Weiss et al., 1967) measures job satisfaction in terms of compensation, advancement, coworkers, and supervisor human relations. The Index of Organizational Reactions (IOR; Smith, 1976) looks at supervision, the kind and amount of work, finance, coworkers, physical conditions, career prospects, and company identification. Most if not all related variables are featured in our machine learning analysis. Results from our machine learning approach generally support the construct of these popular job satisfaction measurements. However, based on the SHAP importance values of features, using the value of 0.1 as a cut-off point, it is found that three features predict job satisfaction particularly well: coworker atmosphere, pay satisfaction, and recognition. And among the top three factors, the coworker dimension performs the best. Based on results from our study, simple measures like the JDI and MSQ perform reasonably well. Nevertheless, the lower predictive power of some features suggests that looking only at few features may not be sufficient capturing the overall picture. More is not necessarily better given the lower predictive performance of some other indicators.

Relatedly, features that capture job characteristics do not perform particularly well. This challenges the job characteristics model (Hackman & Oldham, 1976), the dominant approach in job satisfaction research (Judge et al., 2017). Although the social dimension of job satisfaction is well recognized and is featured in all major measures, coworker relations is understudied when compared with other dimensions such as work conditions and pay satisfaction (Judge & Kammeyer-Mueller, 2012). Our findings also suggest that if an objective of the many traditional human resources management policies is to improve job satisfaction, some of the focus on, for example, skills mismatch and training, could be less important than the cultivation of a supportive and collegial working environment. Future research should focus more on the social aspect. There are several prominent theories related to the social environment in workplace (Jolly et al., 2021). The conservation of resources theory (Halbesleben et al., 2014; Hobfoll, 1989) maintains that, as a resource, (perceived) social supports from colleague and job supervisors helps workers to regulate resources in times of difficulty to prevent (mental) resource loss such as burnout. Focusing on job performance and engagement, the job demands–resources theory (Bakker & Demerouti, 2007; Gerich & Weber, 2020) suggests that social resources in workplace could help workers to improve their performance. Given the general interpretation of the survey question, it remains unclear in which way coworker atmosphere in our analysis entails. Future research is still required to disentangle the mechanism behind the finding. Furthermore, since support has been picked up by a separate top-ranked variable, and work pressure, while included, was not a highly relevant variable in predicting job satisfaction. One interpretation is that work pressure has been captured by other variables, which have lower importance, one may favor the job demands–resources theory which emphasizes the role of support on performance. Another interpretation is that work pressure has a lower importance because part of it has been captured by social support. If this interpretation is correct, the conservation of resources theory might be more important. Which explanation of the finding is correct remains a topic of future research.

What do our findings suggest about the role of job satisfaction in the Great Resignation in the Netherlands? The importance of pay satisfaction is partly consistent with the explanation about the tightness of the labor market in the pandemic era. Under high inflation, workers are predictably unsatisfied with their wages. A tight labor market would favor workers to shift jobs and to ask for a high bid. However, the predictive power of collegial atmosphere out-weights that of wage satisfaction by a lot. It is likely that job satisfaction, mainly through its impact on coworker relations, plausibly related to remote working and home office, has a greater impact on job shifts when compared with the wage factor. As the Netherlands is a developed economy, the hierarchy of needs for individuals are likely to be different from middle and low income countries. The importance of coworker relationship may have relatively less importance compared to wages and other features pertaining to living standards in countries when working individuals are considerably more concerned about their socioeconomic well-being. While our evidence could shed light on the research question at hand with respect to high-income countries, empirical research on data from lower income countries may point towards different results.

Appendix

See Figs. 6 and 7.

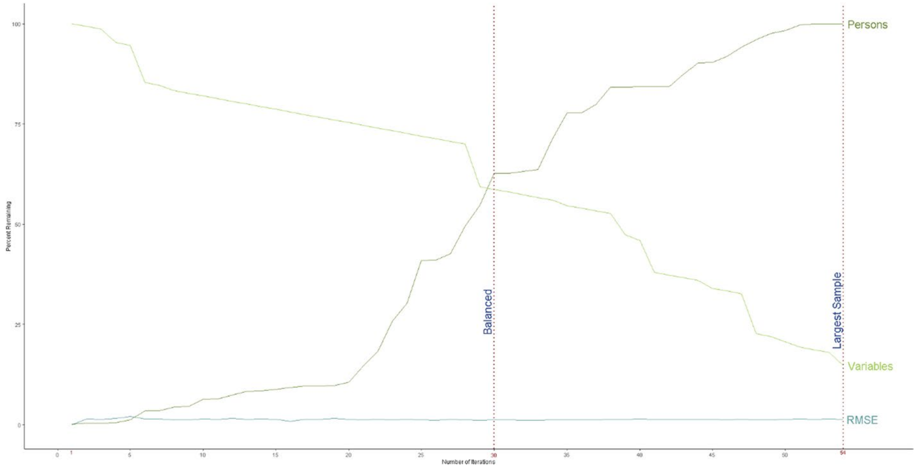

Fig. 6 Iterations and the feature-observation trade-off

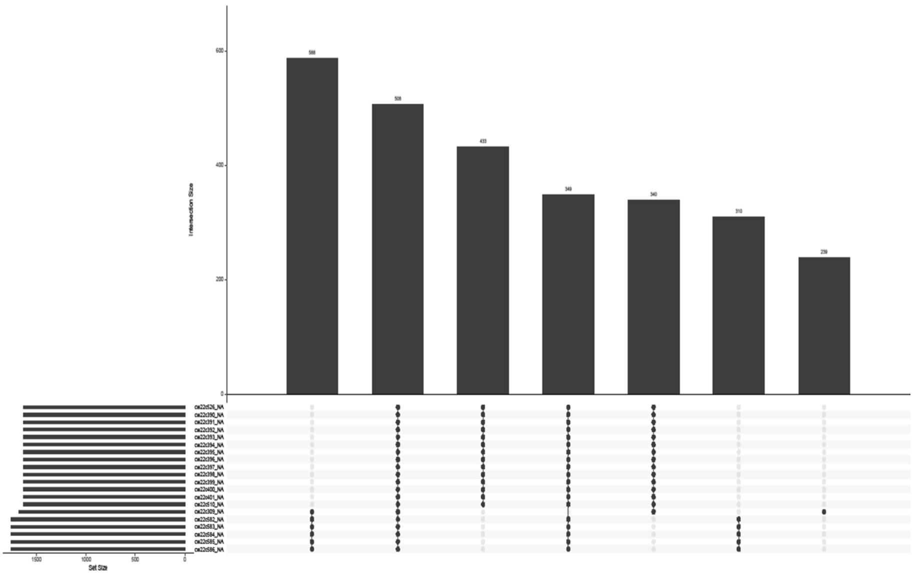

Fig. 7 A sample UpSet plot from the fourth iteration

Funding The authors acknowledge the grant from the Romanian Ministry of Research, Innovation and Digitisation, CNCS—UEFISCDI, project number PN-III-P4-PCCE-2021-1878, within PNCDI III, project—Institutions, Digitalisation and Regional Development in the EU.

Declarations

Conflict of interest The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

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

Algemeen Dagblad. (2023). Grote ontslaggolf waait over uit amerika, 1 op de 5 werknemers wisselt van baan. Retrieved 23 January 2023 from https://www.ad.nl/economie/grote-ontslaggolf-waait-over-uit-amerika-1-op-de-5-werknemers-wisselt-van-baan ad658da1/.

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *Economic Journal, 100*(401), 464–477. https://doi.org/10.2307/2234133

Ariely, D., Kamenica, E., & Prelec, D. (2008). Man's search for meaning: The case of Legos. *Journal of Economic Behavior & Organization, 67*(3–4), 671–677. https://doi.org/10.1016/j.jebo.2008.01.004

Bakker, A. B., & Demerouti, E. (2007). The job demands-resources model: State of the art. *Journal of Managerial Psychology, 22*(3), 309–328.

Bartel, A. P. (1981). Race differences in job satisfaction: A reappraisal. *Journal of Human Resources, 16*(2), 294–303. https://doi.org/10.2307/145514

Berger, T., Frey, C. B., Levin, G., & Danda, S. R. (2019). Uber happy? Work and well-being in the 'gig economy.' *Economic Policy, 34*(99), 429–477. https://doi.org/10.1093/epolic/eiz007

Blanchflower, D. G., & Oswald, A. J. (1999). Well-being, insecurity and the decline of American job satisfaction. NBER working paper, 7487.

Blau, F. D., & Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature, 55*(3), 789–865. https://doi.org/10.1257/jel.20160995

Borjas, G. J. (1979). Job satisfaction, wages, and unions. *Journal of Human Resources, 14*(1), 21–40. https://doi.org/10.2307/145536

Breiman, L. (1996). Bagging predictors. *Machine Learning, 24*(2), 123–140. https://doi.org/10.1007/BF00058655

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Breiman, L., & Cutler, A. (accessed February 1, 2020). Random Forests. https://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm.

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

Bryson, A., & MacKerron, G. (2017). Are you happy while you work? *Economic Journal, 127*(599), 106–125. https://doi.org/10.1111/ecoj.12269

Card, D., Mas, A., Moretti, E., & Saez, E. (2012). Inequality at work: The effect of peer salaries on job satisfaction. *American Economic Review, 102*(6), 2981–3003. https://doi.org/10.1257/aer.102.6.2981

Carree, M. A., & Verheul, I. (2012). What makes entrepreneurs happy? Determinants of satisfaction among founders. *Journal of Happiness Studies, 13*, 371–387. https://doi.org/10.1007/s10902-011-9269-3

Cassar, L., & Meier, S. (2018). Nonmonetary incentives and the implications of work as a source of meaning. *Journal of Economic Perspectives, 32*(3), 215–238. https://doi.org/10.1257/jep.32.3.215

Celbiş, M. G. (2022). Unemployment in rural Europe: A machine learning perspective. *Applied Spatial Analysis and Policy*. https://doi.org/10.1007/s12061-022-09464-0

Celbiş, M. G., Wong, P., Kourtit, K., & Nijkamp, P. (2023). 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, 15*(3), 559–584. https://doi.org/10.1111/rsp3.12520

Chandler, D., & Kapelner, A. (2013). Breaking monotony with meaning: Motivation in crowdsourcing markets. *Journal of Economic Behavior & Organization, 90*, 123–133. https://doi.org/10.1016/j.jebo.2013.03.003

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). ACM.
Clark, A. E. (1997). Job satisfaction and gender: Why are women so happy at work? *Labour Economics, 4*(4), 341–372. https://doi.org/10.1016/S0927-5371(97)00010-9
Clark, A. E. (2001). What really matters in a job? Hedonic measurement using quit data. *Labour Economics, 8*(2), 223–242. https://doi.org/10.1016/S0927-5371(01)00031-8
Clark, A., Etilé, F., Postel-Vinay, F., Senik, C., & Van der Straeten, K. (2005). Heterogeneity in reported well-being: Evidence from twelve European countries. *Economic Journal, 115*(502), C118–C132. https://doi.org/10.1111/j.0013-0133.2005.00983.x
Clark, A. E., Frijters, P., & Shields, M. A. (2008). Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles. *Journal of Economic Literature, 46*(1), 95–144. https://doi.org/10.1257/jel.46.1.95
Clark, A. E., Kristensen, N., & Westergård-Nielsen, N. (2009). Job satisfaction and co-worker wages: Status or signal? *Economic Journal, 119*(536), 430–447. https://doi.org/10.1111/j.1468-0297.2008.02236.x
Clark, A. E., & Oswald, A. J. (1996). Satisfaction and comparison income. *Journal of Public Economics, 61*(3), 359–381. https://doi.org/10.1016/0047-2727(95)01564-7
Cornelissen, T., Heywood, J. S., & Jirjahn, U. (2011). Performance pay, risk attitudes and job satisfaction. *Labour Economics, 18*(2), 229–239. https://doi.org/10.1016/j.labeco.2010.09.005
Cutler, A., Cutler, D. R., & Stevens, J. R. (2009). Tree-based methods. In X. Li & R. Xu (Eds.), *High-dimensional data analysis in cancer research* (pp. 83–102). Springer. https://doi.org/10.1007/978-0-387-69765-9_5
Cutler, A., Cutler, D. R., & Stevens, J. R. (2012). Random forests. In C. Zhang & Y. Ma (Eds.), *Ensemble machine learning: Methods and applications* (pp. 157–175). Springer.
D'Ambrosio, C., Clark, A. E., & Barazzetta, M. (2018). Unfairness at work: Well-being and quits. *Labour Economics, 51*, 307–316. https://doi.org/10.1016/j.labeco.2018.02.007
Demirkaya, H., Aslan, M., Güngör, H., Durmaz, V., & Şahin, D. R. (2022). Covid-19 and quitting jobs. *Frontiers in Psychology, 13*, 916222. https://doi.org/10.3389/fpsyg.2022.916222
Diriwaechter, P., & Shvartsman, E. (2018). The anticipation and adaptation effects of intra-and interpersonal wage changes on job satisfaction. *Journal of Economic Behavior & Organization, 146*, 116–140. https://doi.org/10.1016/j.jebo.2017.12.010
Dolbier, C. L., Webster, J. A., McCalister, K. T., Mallon, M. W., & Steinhardt, M. A. (2005). Reliability and validity of a single-item measure of job satisfaction. *American Journal of Health Promotion, 19*(3), 194–198. https://doi.org/10.4278/0890-1171-19.3.194
Drakopoulos, S. A., & Theodossiou, I. (1997). Job satisfaction and target earnings. *Journal of Economic Psychology, 18*(6), 693–704. https://doi.org/10.1016/S0167-4870(97)00030-5
Dunham, R. B., Smith, F. J., & Blackburn, R. S. (1977). Validation of the index of organizational reactions with the JDI, the MSQ, and faces scales. *Academy of Management Journal, 20*(3), 420–432. https://doi.org/10.5465/255415
Duval, R., Oikonomou, M., & Tavares, M. M. (2022). Tight jobs market is a boon for workers but could add to inflation risks. Retrieved 23 January 2023 from https://www.imf.org/en/Blogs/Articles/2022/03/31/tight-jobs-market-is-a-boon-for-workers-but-could-add-to-inflation-risks.
Easterlin, R. A. (1995). Will raising the incomes of all increase the happiness of all? *Journal of Economic Behavior & Organization, 27*(1), 35–47. https://doi.org/10.1016/0167-2681(95)00003-B
Easterlin, R. A. (2001). Income and happiness: Towards a unified theory. *Economic Journal, 111*(473), 465–484. https://doi.org/10.1111/1468-0297.00646
Ellingsen, T., & Johannesson, M. (2007). Paying respect. *Journal of Economic Perspectives, 21*(4), 135–149. https://doi.org/10.1257/jep.21.4.135
Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics, 5*, 1189–1232.
Friedman, J., Hastie, T., & Tibshirani, R. (2001). *The elements of statistical learning*. Springer.
Friedman, J. H. (2002). Stochastic gradient boosting. *Computational Statistics & Data Analysis, 38*(4), 367–378. https://doi.org/10.1016/S0167-9473(01)00065-2
García-Mainar, I., & Montuenga-Gómez, V. M. (2020). Over-qualification and the dimensions of job satisfaction. *Social Indicators Research, 147*, 591–620. https://doi.org/10.1007/s11205-019-02167-z
Genicot, G., & Ray, D. (2020). Aspirations and economic behavior. *Annual Review of Economics, 12*, 715–746. https://doi.org/10.1146/annurev-economics-080217-053245
Gerich, J., & Weber, C. (2020). The ambivalent appraisal of job demands and the moderating role of job control and social support for burnout and job satisfaction. *Social Indicators Research, 148*, 251–280. https://doi.org/10.1007/s11205-019-02195-9

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. https://doi.org/10.1080/10618600.2014.907095
Gosnell, G. K., List, J. A., & Metcalfe, R. D. (2020). The impact of management practices on employee productivity: A field experiment with airline captains. *Journal of Political Economy, 128*(4), 1195–1233. https://doi.org/10.1086/705375
Green, F. (2010). Well-being, job satisfaction and labour mobility. *Labour Economics, 17*(6), 897–903. https://doi.org/10.1016/j.labeco.2010.04.002
Hackman, J. R., & Oldham, G. R. (1976). Motivation through the design of work: Test of a theory. *Organizational Behavior and Human Performance, 16*, 250–279. https://doi.org/10.1016/0030-5073(76)90016-7
Halbesleben, J. R., Neveu, J. P., Paustian-Underdahl, S. C., & Westman, M. (2014). Getting to the "COR" understanding the role of resources in conservation of resources theory. *Journal of Management, 40*(5), 1334–1364. https://doi.org/10.1177/0149206314527130
Hamermesh, D. S. (1977). *Economic aspects of job satisfaction*. Wiley.
Hamermesh, D. S. (2004). Subjective outcomes in economics. *Southern Economic Journal, 71*(1), 1–11. https://doi.org/10.1002/j.2325-8012.2004.tb00619.x
Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning*. Springer.
Hobfoll, S. E. (1989). Conservation of resources: A new attempt at conceptualizing stress. *American Psychologist, 44*(3), 513–524. https://doi.org/10.1037/0003-066X.44.3.513
James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning*. Springer.
Jolly, P. M., Kong, D. T., & Kim, K. Y. (2021). Social support at work: An integrative review. *Journal of Organizational Behavior, 42*(2), 229–251. https://doi.org/10.1002/job.2485
Jones, M. K., Jones, R. J., Latreille, P. L., & Sloane, P. J. (2009). Training, job satisfaction, and workplace performance in Britain: Evidence from WERS 2004. *Labour, 23*, 139–175. https://doi.org/10.1111/j.1467-9914.2008.00434.x
Judge, T. A., & Kammeyer-Mueller, J. D. (2012). Job attitudes. *Annual Review of Psychology, 63*, 341–367. https://doi.org/10.1146/annurev-psych-120710-100511
Judge, T. A., & Klinger, R. (2008). Job satisfaction: Subjective well-being at work. In M. Eid & R. J. Larsen (Eds.), *The science of subjective well-being* (pp. 393–413). Press.
Judge, T. A., Weiss, H. M., Kammeyer-Mueller, J. D., & Hulin, C. L. (2017). Job attitudes, job satisfaction, and job affect: A century of continuity and of change. *Journal of Applied Psychology, 102*(3), 356–374. https://doi.org/10.1037/apl0000181
Karlsson, N., Loewenstein, G., & McCafferty, J. (2004). The economics of meaning. *Nordic Journal of Political Economy, 30*(1), 61–75.
Kinicki, A. J., McKee-Ryan, F. M., Schriesheim, C. A., & Carson, K. P. (2002). Assessing the construct validity of the job descriptive index: A review and meta-analysis. *Journal of Applied Psychology, 87*(1), 14.
Laffont, J., & Martimort, D. (2002). *The theory of incentives*. Princeton University Press.
Lange, T. (2012). Job satisfaction and self-employment: Autonomy or personality? *Small Business Economics, 38*, 165–177. https://doi.org/10.1007/s11187-009-9249-8
Lepinteur, A. (2019). The shorter workweek and worker wellbeing: Evidence from Portugal and France. *Labour Economics, 58*, 204–220. https://doi.org/10.1016/j.labeco.2018.05.010
Levels, M., Van der Velden, R., & Allen, J. (2014). Educational mismatches and skills: New empirical tests of old hypotheses. *Oxford Economic Papers, 66*(4), 959–982. https://doi.org/10.1093/oep/gpu024
Lex, A., Gehlenborg, N., Strobelt, H., Vuillemot, R., & Pfister, H. (2014). UpSet: Visualization of intersecting sets. *IEEE Transactions on Visualization and Computer Graphics, 20*(12), 1983–1992. https://doi.org/10.1109/TVCG.2014.2346248
Liu, Y., & Just, A. (2020). SHAPforxgboost: SHAP Plots for 'XGBoost'. R package version 0.1.0.
Locke, E. A. (1976). *The nature and causes of job satisfaction*. Rand McNally College Publishing Company.
Loewenstein, G. (1999). Because it is there: The challenge of mountaineering… for utility theory. *Kyklos, 52*(3), 315–343. https://doi.org/10.1111/j.1467-6435.1999.tb00221.x
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* (pp. 4768–4777).
Luttmer, E. F. (2005). Neighbors as negatives: Relative earnings and well-being. *Quarterly Journal of Economics, 120*(3), 963–1002. https://doi.org/10.1093/qje/120.3.963
Martin, L., Hauret, L., & Fuhrer, C. (2022). Digitally transformed home office impacts on job satisfaction, job stress and job productivity COVID-19 findings. *PLoS ONE, 17*(3), e0265131. https://doi.org/10.1371/journal.pone.0265131
McBride, M. (2001). Relative-income effects on subjective well-being in the cross-section. *Journal of Economic Behavior & Organization, 45*(3), 251–278. https://doi.org/10.1016/S0167-2681(01)00145-7

McBride, M. (2010). Money, happiness, and aspirations: An experimental study. *Journal of Economic Behavior & Organization, 74*(3), 262–276. https://doi.org/10.1016/j.jebo.2010.03.002

McGuinness, S., Pouliakas, K., & Redmond, P. (2018). Skills mismatch: Concepts, measurement and policy approaches. *Journal of Economic Surveys, 32*(4), 985–1015. https://doi.org/10.1111/joes.12254

Millán, J. M., Hessels, J., Thurik, R., & Aguado, R. (2013). Determinants of job satisfaction: A European comparison of self-employed and paid employees. *Small Business Economics, 40*, 651–670. https://doi.org/10.1007/s11187-011-9380-1

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

Mulder, H., Hoofs, H., Verhagen, J., van der Veen, I., & Leseman, P. P. (2014). Psychometric properties and convergent and predictive validity of an executive function test battery for two-year-olds. *Frontiers in Psychology, 5*, 733. https://doi.org/10.3389/fpsyg.2014.00733

Nikolova, M., & Cnossen, F. (2020). What makes work meaningful and why economists should care about it. *Labour Economics, 65*, 101847. https://doi.org/10.1016/j.labeco.2020.101847

Origo, F., & Pagani, L. (2009). Flexicurity and job satisfaction in Europe: The importance of perceived and actual job stability for well-being at work. *Labour Economics, 16*(5), 547–555. https://doi.org/10.1016/j.labeco.2009.02.003

Oswald, A. J., Proto, E., & Sgroi, D. (2015). Happiness and productivity. *Journal of Labor Economics, 33*(4), 789–822. https://doi.org/10.1086/681096

Pilipiec, P., Groot, W., & Pavlova, M. (2020). A longitudinal analysis of job satisfaction during a recession in The Netherlands. *Social Indicators Research, 149*, 239–269. https://doi.org/10.1007/s11205-019-02233-6

Pilipiec, P., Groot, W., & Pavlova, M. (2021). The causal influence of increasing the statutory retirement age on job satisfaction among older workers in the Netherlands. *Applied Economics, 53*(13), 1498–1527. https://doi.org/10.1080/00036846.2020.1827136

Poggi, A. (2010). Job satisfaction, working conditions and aspirations. *Journal of Economic Psychology, 31*(6), 936–949. https://doi.org/10.1016/j.joep.2010.08.003

Powdthavee, N. (2010). How much does money really matter? estimating the causal effects of income on happiness. *Empirical Economics, 39*, 77–92. https://doi.org/10.1007/s00181-009-0295-5

Ravid, O., Malul, M., & Zultan, R. (2017). The effect of economic cycles on job satisfaction in a two-sector economy. *Journal of Economic Behavior & Organization, 138*, 1–9. https://doi.org/10.1016/j.jebo.2017.03.028

Reuters. (2019). U.S. job openings, quits hit record highs in April. Retrieved 23 January 2023 from https://www.reuters.com/business/us-trade-deficit-narrows-april-2021-06-08/.

Shapley, L. S. (1953). A value for n-person games. In H. Kuhn & A. Tucker (Eds.), *Contributions to the theory of games II, annals of mathematics studies* (Vol. 28, pp. 307–317). Princeton University Press.

Sloane, P. J., & Ward, M. E. (2001). Cohort effects and job satisfaction of academics. *Applied Economics Letters, 8*(12), 787–791. https://doi.org/10.1080/13504850110045733

Smith, F. J. (1976). The Index of Organizational Reactions (IOR). *JSAS Catalog of Selected Document in Psychology*, Vol. 6 (i976), Ms. No. 1265.

Smith, J. C. (2015). Pay growth, fairness, and job satisfaction: Implications for nominal and real wage rigidity. *Scandinavian Journal of Economics, 117*(3), 852–877. https://doi.org/10.1111/sjoe.12091

Smith, P. C., Kendall, L. M., & Hulin, C. L. (1969). *The measurement of satisfaction in work and retirement: A strategy for the study of attitudes*. Rand Mcnally.

Streefkerk, M., & Centerdata (2022). LISS panel—work and schooling—wave 15.

Štrumbelj, E., & Kononenko, I. (2013). Explaining prediction models and individual predictions with feature contributions. *Knowledge and Information Systems, 41*(3), 647–665. https://doi.org/10.1007/s10115-013-0679-x

Tierney, N., & Cook, D. (2023). Expanding tidy data principles to facilitate missing data exploration, visualization and assessment of imputations. *Journal of Statistical Software, 105*(7), 1–31. https://doi.org/10.18637/jss.v105.i07

Varian, H. R. (2014). Big data: New tricks for econometrics. *Journal of Economic Perspectives, 28*(2), 3–28. https://doi.org/10.1257/jep.28.2.3

Vieira, J. A. C. (2005). Skill mismatches and job satisfaction. *Economics Letters, 89*(1), 39–47. https://doi.org/10.1016/j.econlet.2005.05.009

Weiss, D. J., Dawis, R. V., England, G. W., & Lofquist, L. (1967). *Minnesota studies in vocational rehabilitation. Manual for the Minnesota Satisfaction Questionnaire*. University of Minnesota.

Weiss, H. M. (2002). Deconstructing job satisfaction: Separating evaluations, beliefs and affective experiences. *Human Resource Management Review, 12*(2), 173–194. https://doi.org/10.1016/S1053-4822(02)00045-1

Xu, D., & Tian, Y. A. (2015). comprehensive survey of clustering algorithms. *Annals of Data Science, 2*, 165–193. https://doi.org/10.1007/s40745-015-0040-1