Using Government Data and Machine Learning for Predicting Firms' Vulnerability to Economic Crisis

Euripidis Loukis[1], Niki Kyriakou[1] , Manolis Maragoudakis[2]

University of the Aegean, Greece
{eloukis, nkyr}@aegean.gr
[2] Ionian University, Greece
{mmarag}@ionio.gr

Abstract. The COVID-19 pandemic is expected to lead to a severe recessionary economic crisis with quite negative consequences for large numbers of firms and citizens; however, this is an 'old story': recessionary economic crises appear repeatedly in the last 100 years in the market-based economies, and they are recognized as one of the most severe and threatening weaknesses of them. They can result in closure of numerous firms, and decrease of activities of many more, as well as poverty and social exclusion for large parts of the population, and finally lead to political upheaval and instability; so they constitute one of the most threatening and difficult problems that governments often face. For the above reasons it is imperative that governments develop effective public policies and make drastic interventions for addressing these economic crises. Quite useful for these interventions can be the prediction of the vulnerability of individual firms to recessionary economic crisis, so that government can focus its attention as well as its scarce economic resources on the most vulnerable ones. In this direction our paper presents a methodology for using existing government data in order to predict the vulnerability of individual firms to economic crisis, based on Artificial Intelligence (AI) Machine Learning (ML) algorithms. Furthermore, a first application of the proposed methodology is presented, based on existing data from the Greek Ministry of Finance and Statistical Authority concerning 363 firms for the economic crisis period 2009-2014, which gives encouraging results.

Keywords: economic crisis, economic recession, data analytics, predictive analytics, artificial intelligence, machine learning.

1 Introduction

The COVID-19 pandemic is expected to lead to a severe recessionary economic crisis with quite negative consequences for large numbers of firms and citizens, as containment and social distancing policies adopted in most countries will on one hand 'flatten' the medical curve of virus spread (which is quite positive), but on the other hand will 'steepen' the resulting recession curve (which will have quite negative economic and social consequences) [1]. However, this is an 'old story': recessionary economic crises of various intensities and durations appear repeatedly in the last 100 years in market-based economies, and are recognized as one of the most severe and threatening weaknesses of them, having quite negative consequences for the whole economy and society [2-7]. The fluctuations that economic activity usually exhibits, and also some critical events, such as banking crises (like the one that gave rise to the 2008 Global Financial Crisis), epidemics (like the COVID-19 one), large increases in the prices of important goods (e.g. oil or gas), etc., can lead to significant economic recessions, meant as serious contractions of economic activity, resulting in economic crises. These contractions of economic activity can lead to big reductions of firms' production, procurement, investment, innovation and employment [5-7], and also closure of numerous firms, with serious social consequences, such as increase of unemployment, poverty and social exclusion, resulting in political unrest and instability.

For these reasons economic crises constitute one of the most severe and threatening problems that governments often face and have strong pressure to deal with them effectively in order at least to avoid catastrophic consequences. So it is imperative that governments develop effective public policies and make drastic interventions for addressing these economic crises. The most usual of these interventions is the provision of various kinds of support to the most vulnerable firms in the beginning of such crises, or even earlier, when an economic crisis is in sight (i.e. when there are serious forecasts that an economic crisis is going to occur, like the ones currently made by serious economists and economic research centers that the COVID-19 pandemic will give rise to a severe recessionary economic crisis [1]). This support usually includes the provision to firms of low-interest loans, subsidies, education or consulting services, co-finance of investments or employees' payroll, tax reductions, etc. However, since the available government resources (e.g. financial, human, etc.) for providing such support to firms are usually limited, much below than the ones required for meeting the high demands (from the numerous firms that submit applications for receiving this support), a selection among the applicant firms has to be made; it is based on some predefined criteria, which usually concern firm's sector and location (according to existing sectoral and regional priorities), as well as sales and profits in the last 3-5 years, market share, export activities, debts, etc. It would be very good if to these criteria could be added firm's expected vulnerability to the crisis, as an additional criterion, aiming to provide more assistance and support to the most vulnerable firms; this can increase substantially the effectiveness of such interventions, making them more focused, and improve public value generation from them towards the reduction of the negative consequences of recessionary economic crises. For the above reasons it would be quite useful if we can make a prediction of the vulnerability of individual firms to economic crisis.

Given the high importance of firm-level economic crisis vulnerability prediction it is necessary to exploit for this purpose to the highest possible extent: a) existing firms' data from economic crisis periods in various government agencies (such as Ministries of Finance and Statistical Authorities); and b) the most sophisticated algorithms for 'learning' crisis vulnerability prediction models from such data, and especially the most widely used Artificial Intelligence (AI) algorithms for such 'predictive analytics': the Machine Learning (ML) ones [8-12]. In this direction our paper presents a methodology for using existing government data in order to predict the vulnerability of individual firms to economic crisis, based on their particular characteristics (concerning strategic directions, processes, human and technological resources, structures) using ML. Furthermore, a first application of the proposed methodology is presented, based on existing data from the Greek Ministry of Finance and Statistical Authority concerning 363 firms for the economic crisis period 2009-2014, which gives encouraging results.

In the following section 2 the background of our methodology is outlined, while in section 3 the methodology is described, followed by its abovementioned application in secion 4. The final section 5 summarizes conclusions and proposes directions for future research.

2 Background

2.1 Economic Crises

As mentioned in the Introduction one of the most serious problems of market-based economies are the economic recessions that repeatedly appear, which cause significant disruptions and problems to the economy and the society in general, and have to be addressed by government through effective public policies including appropriate interventions [2-7]. The economic crises have negative both short-term as well as medium- and long-term consequences for the economy and the society. The short-term consequences include reductions of the demand for many goods and services, resulting in serious decrease of firms' sales, production and profits, leading to reductions in personnel employment (thus increasing unemployment) and materials' procurement (thus propagating the crisis towards the suppliers' sectors, etc., spreading the crisis to large parts of the economy). Furthermore, during economic crises firms usually reduce capital investment in production equipment, ICT, buildings, etc., and also in product, service and process innovations; this reduces the degree of renewal and improvement of their equipment, products, services and operation, as well as the exploitation of emerging new technologies, which causes serious medium- and long-term consequences for their efficiency and competitiveness.

Considerable research has been conducted concerning the negative impact of such recessionary economic crises on different aspects of firms' activities and performance, as well as the factors that affect the magnitude of this negative impact [13-19]. It has concluded that the above negative consequences of the economic crises differ significantly among firms, and depend critically on their individual characteristics, such as their human and technological resources, their strategic orientations, etc. Some firms have superior human and technological resources, as well as management sturtures, that allow them to be more efficient and effective during difficult crisis periods; these firms are able to offer higher value-for-money products and services during the crisis, and in general have higher capacity to make the required adaptations to the crisis conditions, which make them less vulnerable and more resilient to the crisis. Such firms have less negative consequences on their sales revenue, and therefore on their employment, procurement, as well as capital investment and innovation. For some other firms the opposite holds. Therefore it would be interesting to investigate the capabilities of AI/ML algorithms to predict firm's vulnerability to recessionary economic crisis based on its particular characteristics, e.g. concerning its human and technological resources, management structures, etc.

2.2 Artificial Intelligence

Artificial Intelligence (AI) includes a group of techniques that enable computers to perform tasks of higher intelligence, approaching the human one, by learning from their environment, and then using the knowledge they have obtained from it for taking or proposing action [20-22]. The most representative and widely used AI techniques are definitely the Machine Learning (ML) ones [8-10]. They enable exploiting historic data we possess for a number of units (e.g. individuals, firms, etc.) concerning the value of an important dependent variable (usually an outcome one), and also a set of independent variables (that might be possible causes of this outcome or factors affecting it), by processing them through various algorithms, and finally extracting knowledge from them, usually having the form of a model or a set of rules, concerning the relationships between the independent variables and the dependent one (this is usually referred to as 'model training', while the historic data used for this purpose as 'training data'). This knowledge can be used then on one hand for gaining deeper understanding and insight about these relationships, and on the other hand for predicting the value of the dependent variable for new units, based on the known values of the independent variables, which can be quite useful for supporting relevant decisions or optimizing actions.

Initially AI was used successfully in the private sector, and these 'success stories' generated high levels of interest to exploit AI techniques in the public sector as well, in order to automate or support more sophisticated mental tasks than the simpler routine ones automated or supported by the traditional operational IS of government agencies [23-27]. Some first research has been conducted concerning the exploitation of AI in a variety of public sector thematic domains, such as education, policing, justice, public health, transport, etc. [28-33]. However, this research has revealed and investigated only a small part of the large potential of AI use in government. Therefore, much more research is required in order to exploit more this potential, in order to reveal and investigate more opportunities and methodologies for AI exploitation in government, in a wide variety of public sector thematic domains. Our study makes a contribution in this direction, by developing a methodology for using AI (ML algorithms) in a highly important domain of government activity: the intervention of government for addressing the quite threatening recessionary economic crises that repeatedly appear in market-based economies and cause major disruptions and problems.

3 The Proposed Methodology

As mentioned in section 2.1 previous research concerning the negative impact of recessionary economic crises on firms has found that its magnitude differs significantly among firms, depending on their particular characteristics, such as their human and technological resources, management structures, etc. Furthermore, previous management science research has identified the main elements of a firm that determine its performance; so we can expect that these main elements might determine to a considerable extent firm's performance during the crisis in coping with the difficult recession conditions as well, and therefore the degree of its resilience/vulnerability to the crisis. Several conceptualizations of the main elements of a firm have been developed, the most widely validated and used among them being definitely the 'Leavitt's Diamond' framework [34]. According to it the most important elements of a firm that determine its performance are: a) task (= the strategies as well as the administrative and production processes of the firm); b) people (= the skills of firm's human resources of the firm); c) technology (= the technologies used for implementing the above processes); and d) structure (= the organization of the firm in departments, and the communication and coordination patterns them). An extension of it has been developed subsequently, which analyses the above 'task' element into 'strategy' and 'processes' [35]. Therefore we expect that firm's characteristics concerning the above five main elements (strategy, processes, people, technology and structure) might be good predictors of its crisis resilience/vulnerability. Government traditionally collects large amounts of data concerning firms' economic performance (the Taxation Authorities) and also firms' chacteristics concerning human resources, use of various technologies (e.g. ICT or various production technologies), organization, innovation and other strategic orientations, etc (the Statistical Authorities). It is quite useful to exploit these existing government data from economic crisis periods in order to construct firm-level crisis resilience/vulnerability prediction models based on firm's characteristics.

In particular, using government firm-level data from an economic crisis period concerning:

- on one hand firms' resilience/vulnerability to the crisis (e.g. data from Taxation Authorities concerning decrease of sales revenue, profitability, employment, etc. during the crisis)
- and on the other hand characteristics of the same firms concerning the above main elements (e.g. relevant data collected from Statistical Authorities),

We can train/construct prediction models of the resilience/vulnerability of a firm to recessionary economic crisis (to be used as dependent variable Y), based on the above characteristics of it (to be used as independent variables Xi); for this purpose we can use several different alternative ML algorithms [8-10], such as Decision Trees, Random Forests, Gradient Boosted Trees, Support Vector Machines, Generalized Linear Modelling, etc., and finally select among them the one providing the highest prediction accuracy. The constructed model can be used for predicting the value of the resilience/vulnerability of an individual firm to future economic crisis Y (taking into account that all economic crises, despite their differences, have in common their main characteristics: economic recession, contraction of economic activity and decrease of demand for products and services of most sectors) for which we have the values of the abovementioned characteristics Xi. The structure of the proposed methodology is shown in Fig.1.

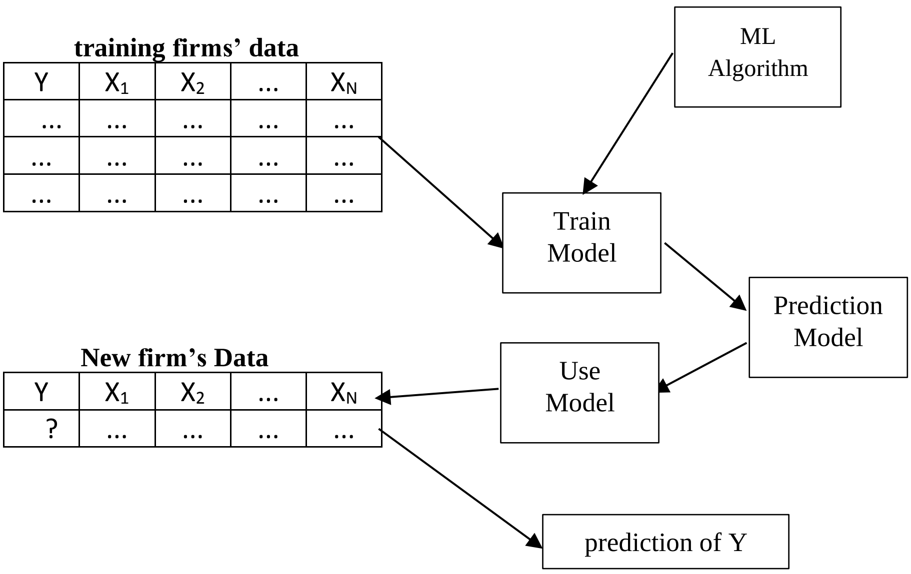

Figure 1: Structure of the proposed methodology

Our methodology will use existing government data from two main sources:
I) Data from Taxation Authorities for a large number of firms concerning the decrease of sales revenue, profits or other measures of economic performance during the economic crisis
II) Data from Statistical Authorities about characteristics of the same firms concerning the above five main elements of them:
- <u>Strategic orientations</u>: characteristics concerning the degree of adoption of the main strategies described in relevant strategic management literature [36], such as cost leadership, differentiation, focus, innovation, export, etc.
- <u>Processes:</u> characteristics of firm's processes, such as complexity, efficiency, formality, flexibility, etc.
- <u>Human Resources:</u> characteristics concerning the general education/skills level of firm's human resources (e.g. shares of firm's personnel having tertiary education, vocational/technical education, etc.), as well as the possession of specific skills concerning important technologies (e.g. concerning ICTs or various production technologies), the provision of various kinds of training, etc.
- <u>Technology:</u> characteristics concerning the use of the main enterpsise ICTs, such as Enterprise Resource Planning (ERP), Customer Relationships Management (CRM), Supply Chain Management (SCM), Business Intelligence/Business Analytics (BI/BA) systems, Collaboration Support (CS), e-sales, etc. systems, and also social media, cloud computing, or other technologies, or the use of various production technologies.

- Structure: characteristics concerning various aspects of the management structure of the firm, such as its main structural design (functional, product/service based, geographic, matrix), degree of differentiation, specialization, centralization-decentralization, or the use of organic structural forms (such as teamwork or job rotation) [37-38]. Furthermore, we can also include general data about each firm, such as sector, degree of competiveness in this sector (in comparison with the other competitor firms), etc.

4 Application

A first application of the proposed methodology has been made, using firm-level data for the period 2009-2014 from the Ministry of Finance – Taxation Authorities as well as the Statistical Authority of Greece, in order to construct a prediction model of the most important measure of firm's vulnerability to the crisis, the reduction of sales revenue due to economic crisis, based on firm's characteristics. In particular, we used data provided by the Ministry of Finance – Taxation Authorities concerning the percentage of sales revenue reduction due to the economic crisis in the period 2009 – 2014 (dependent variable) of 400 Greek firms; these data were discretized by the Ministry of Finance (in order to avoid providing detailed firm's sales revenue data) into a variable with 13 possible discrete values (SALREV_RED): increase by more than 100%; increase by 80-100%; increase by 60-80%; increase by 40-60%; increase by 20-40%; increase by 1-20%; unchanged sales; decrease by 1-20%; decrease by 20-40%; decrease by 40-60%; decrease by 60-80%; decrease by 80-100%; decrease by more than 100%. For 363 of these firms from the Statistical Authority we found data concerning 43 characteristics of them (independent variables), which concern their strategic orientations (12 variables), human resources (9 variables), technology (16 variables), structure (1 variable) and also general characteristics (5 variables). The definitions – corresponding questions for our dependent and independent variables are shown in the Appendix. These 363 firms cover a wide range of sectors and sizes: 40.2% of them were from manufacturing sectors, 9.4% from constructions, and 50.4% from services sectors; also, 52.6% of them were small, 36.1% medium and 11.3% large ones.

Using the above data we constructed prediction models of SALREV_RED based on the abovementioned 43 firm characteristics. Since the dependent variable is ordinal, with a large number of values (13) and equal distances between them, it approaches a continuous variable, we have constructed and compared prediction models of SALREV_RED using the main alternative ML algorithms proposed by relevant literature [8-10] for the case that the dependent variable is continuous: the classical regression (termed as 'Generalized Linear Model') and also four alternative more advanced statistical methods:

- Decision Trees, which splits the initial training set into more homogeneous sub-sets using the most appropriate independent variables, and finally construct a decision tree with internal 'decision nodes' concerning some of the independent variables (the most useful as predictors) – actually tests concerning these independent variables – as well as leaf 'terminal nodes' that include predicted values of the dependent variable [9, 40];
- Random Forests, which creates a pool of decision trees based on su-sets of the training

data set, and then gets the prediction from each of them and finally selects the best prediction by means of ‘voting’ of these decision trees [39-40];
- Gradient Boosted Trees, which is based on boosting prediction accuracy through the combination of a learning algorithm in a series, in order to achieve a ‘strong learner’ from many sequentially connected ‘weak learners’; usually the the weak learners are decision trees, and consequtive decision trees are constructed, with each new tree attempting to minimize the errors of the previous tree [41];
- Support Vector Machines, which mqp each training data item as a point in n-dimensional space, where n is number of independent variables; then, the algorithm tries to find hyper-plane that best fits the available training data [40, 42].

For each of these five ML algorithms we calculated its mean absolute error, which is according to relevant ML literature the recommended measure of prediction accuracy for the case that the dependent variable is continuous, using the k-fold cross-validation procedure (we selected k=10), which includes the following steps [8-10, 40]:
a) the data set is divided randomly in k equally sized partitions,
b) then we estimate the prediction model k times, in each of them using one of these segments as test data, and the remaining ones as training data,
c) and finally the absolute prediction error is estimated as the mean of the prediction errors over the test sets of the above k runs.

In Fig. 2 we can see the estimated mean absolute prediction errors of these five ML algorithms. We can see that the examined algorithms have similar levels of prediction performance/error (ranging from 1.422 to 1.572), with the Decision Trees algorithm exhibiting the lowest mean absolute error (1.422). This is a satisfactory prediction performance, given the small size of the dataset we have used (data from 363 firms), so using a larger dataset (which is feasible, as Ministries of Finance – Taxation Authorities and Statistical Authorities have such data for quite large numbers of firms) can result in a smaller mean absolute error, and therefore an even more accurate prediction of the sales revenue reduction of a firm during an economic crisis.

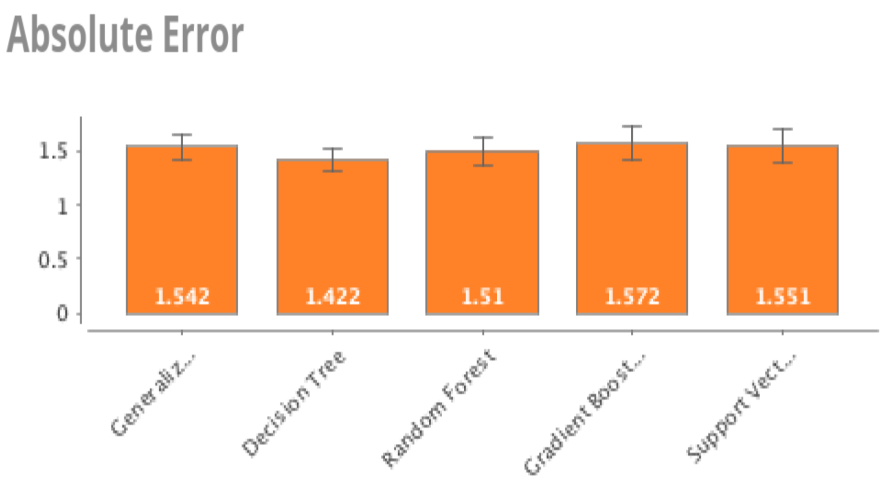

Figure 2: Mean absolute prediction errors of the five ML algorithms

5 Conclusions

In the previous sections has been described the development of a methodology for exploiting existing government data for economic crisis periods, in order to construct prediction models of individual firms' vulnerability to economic crises, based on Artificial Intelligence (AI) Machine Learning (ML) algorithms. In particular, it uses existing data for economic crisis periods, on one hand from Taxation Authorities (concerning firms' sales revenue, profits, employment, etc. decrease), and on the other hand from Statistical Authorities (concerning human and technological resources, structures, processes, strategic orientations, etc. of these firms); from these data prediction models can be trained/constructed, which enable the prediction of the vulnerability of a firm to future economic crises (e.g. in terms of sales revenue, profitability, employment, etc. decrease), based on firm's chatacteristics (concerning human and technological resources, structures, processes, strategic orientations, etc.). The prediction of the degree of vulnerability of a firm to economic crisis can be useful as a criterion in firm support programmes of government agencies, for the selection of the applicant firms that will receive support, as part of multi-criteria evaluation of applications, in combination with other selection criteria usually used in such programmes, such as sector, geographic location, sales, profits, market share, export activities, debts, etc.

The research described in this paper has interesting implications for both research and practice. With respect to research it creates new knowledge in the emerging areas of government data analytics and government AI exploitation, concerning a highly important domain of economic government activity/intervention, aiming to address the quite threatening recessionary economic crises that repeatedly appear in market-based economies and cause major disruptions and problems. With respect to practice it proposes a methodology of using existing historic government data for constructing prediction models of the vulnerability that individual firms will exhibit to future recessionary economic crises. These prediction models can be very useful to government agencies that implement firm support programs in the beginning of economic crises, or even earlier when crises are in sight, in order to use more effectively their available financial resources.

Future research is required towards: i) further application of the proposed methodology using larger datasets, in various different national contexts experiencing economic crises of different intensities, and also at sectoral level (using also sector specific independent variables); ii) investigation of the prediction performance of other ML algorithms, and especially Deep Learning ones; iii) analysis of the legal aspects of the practical application of the proposed methodology by government agencies, and especially with respect to the new EU GDPR (as the data used by our methodology have been collected by government for quite different purposes).

References

1. Baldwin, R. and Di Mauro, B. W.: Mitigating the COVID economic crisis: Act fast and do whatever it takes. Center of Economic Policy Research Press, London (2020)

2. Keeley, B, Love, P.: From Crisis to Recovery - The Causes, Course and Consequences of the Great Recession. OECD Publishing, Paris (2010).
3. Knoop, T. A.: Recessions and Depressions: Understanding Business Cycles -2nd edition. Praeger Santa Barbara, California (2015).
4. Allen, R. E.: Financial Crises and Recession in the Global Economy - Fourth Edition. Cheltenham, Edward Elgar Publications (2016).
5. Santana, M., Valle, R, Galanb, J. L.: Turnaround strategies for companies in crisis: Watchout the causes of decline before firing people. Business Research Quarterly 20(3), 206---211 (2017).
6. Izsak, K., Markianidou, P., Lukach, R., and Wastyn, A.: The impact of the crisis on research and innovation policies. European Commission, DG Research, Brussels (2013).
7. Arvanitis, S., Woerter, M.: Firm Characteristics and the Cyclicality of R&D Investments. Industrial and Corporate Change 23(5), 1141-1169 (2014).
8. Russell, S., Norvig, P.: Artificial Intelligence: A Modern Approach - 3rd Edition. Pearson, Essex, UK (2020).
9. Murphy, K. P.: Machine learning : a probabilistic perspective. MIT Press, Cambridge, Mass. (2013)
10. Witten, I. H., Frank, E., Hall, M. A.,, Pal, C. J.: Data mining - practical machine learning tools and techniques. Morgan Kaufmann, Amsterdam, London (2017).
11. Blum, A., Hopcroft, J.,, Kannan, R.. Foundations of Data Science. Cambridge University Press, Cambridge (2020).
12. Siegel, E.: Predictive Analytics. Wiley, New Jersey (2013).
13. Latham, S.: Contrasting Strategic Response to Economic Recession in Start-Up versus Established Software Firms. Journal of Small Business Management 47(2), 180–201 (2009).
14. Latham, S., Braun, M.: Economic recessions, strategy, and performance: a synthesis. Journal of Strategy and Management 4(2), 96-115 (2011).
15. Wickramasinghe, V., Perera, G.: HRM practices during the global recession (2008-2010) - Evidence from globally distributed software development firms in Sri Lanka. Strategic Outsourcing: An International Journal 5(3), 188-212 (2012).
16. Ahmed, M. U., Kristal, M. M., Pagell, M.: Impact of operational and marketing capabilities on firm performance: Evidence from economic growth and downturns. International Journal of Production Economics 154, 59-71 (2014).
17. Lai, Y., Saridakis, G., Blackburn, R., Johnstone, S.: Are the HR responses of small firms different from large firms in times of recession? Journal of Business Venturing 31(1), 113-131 (2016).
18. Burger, A., Damijan, J. P., Kostevc, C., Rojec, M.: Determinants of firm performance and growth during economic recession: The case of Central and Eastern European count. Economic Systems, 41(4), 569-590 (2017).
19. Arvanitis, S., Loukis, E.: Factors explaining ICT expenditure behaviour of Greek firms during the economic crisis 2009-2014. 7th International Conference on eDemocracy, Athens, Greece (2017).
20. Craglia M. (Ed.), Annoni A., Benczur P., Bertoldi P., Delipetrev P., De Prato G., Feijoo C., Fernandez Macias E., Gomez E., Iglesias M., Junklewitz H, López Cobo M., Martens B., Nascimento S., Nativi S., Polvora A., Sanchez I., Tolan S., Tuomi I., Vesnic Alujevic L.: Artificial Intelligence - A European Perspective, EUR 29425 EN. EU Publications Office, Luxembourg (2018).
21. Duan, Y., Edwards, J. S., Dwivedi, Y. K.. Artificial intelligence for decision making in the era of Big Data – evolution, challenges and research agenda. International Journal of Information Management 48, 63–71 (2019).

22. OECD. Artificial Intelligence in Society. OECD Publishing, Paris (2019).
23. Eggers, W. D., Schatsky, D., Viechnicki, P.: AI-augmented government. Using cognitive technologies to redesign public sector work. Deloitte University Press (2017).
24. Desouza, K. C., Krishnamurthy, R., Dawson, G. S.: Learning from public sector experimentation with artificial intelligence. Brookings, Washington D.C. (2017).
25. Desouza, K. C.: Delivering Artificial Intelligence in Government: Challenges and Opportunities. IBM Center for The Business of Government, Washington D.C. (2018).
26. Sun, T. Q., Medaglia, R.: Mapping the challenges of Artificial Intelligence in the public sector: Evidence from public healthcare. Government Information Quarterly 36(2), 368-383 (2019).
27. DeSousa, W. G., DeMelo, E. R. P., De Souza Bermejo, P. H., Sous Farias, R. A., Gomes, A. O: How and where is artificial intelligence in the public sector going? A literature review and research agenda. Government Information Quarterly (in-press) (2019).
28. Rockoff, J. E., Jacob, B. A., Kane, T. J., Staiger, D. O.: Can you recognize an effective teacher when you recruit one? Education Finance and Policy 6(1), 43–74 (2010).
29. Chandler, D., Levitt, S. D., List, J. A.: Predicting and preventing shootings among at-risk youth. The American Economic Review 101(3), 288–292 (2011).
30. Kang, J. S., Kuznetsova, P., Luca, M., Choi, Y.: Where not to eat? Improving public policy by predicting hygiene inspections using online reviews. In: Proceedings of Empirical Methods in Natural Language Processing Conference 2013, 1443–1448 (2013).
31. Camacho-Collados, M. Liberatore, F.: A Decision Support System for predictive police patrolling. Decision Support systems 75, 25-37 (2015).
32. Ku, C. H., Leroy, G.: A decision support system: Automated crime report analysis and classification for e-government. Government Information Quarterly 31(4), 534-544 (2014).
33. Kouziokas, G. N.: The application of artificial intelligence in public administration for forecasting high crime risk transportation areas in urban environment. Transportation Research Procedia 24, 467-473 (2017).
34. Leavitt, H. J.: Applied organization change in industry: structural, technical, and human approaches. In Cooper, S., Leavitt, H. J., Shelly, K. (Eds.) New perspectives in organizational research, 55–71. Wiley, Chichester (1964).
35. Scott-Morton, M.S. (1991). The Corporation of the 1990s.: Oxford University Press, New York.
36. Johnson, G., Whittington, R., Angwin, D., Regner, P., Scholes, K.: Exploring Corporate Strategy. 10th edn. Financial Times/ Prentice Hall, UK (2014).
37. Jones, G. R.: Organizational Theory, Design, and Change. 7th edn. Pearson Education Limited, UK (2013).
38. Donaldson, L.: The Contingency Theory of Organization. Sage Publications, London, UK (2001).
39. Breiman, L.: Random Forests.Machine Learning 45 (1), 5–32 (2001).
40. Tan, P. N., Steinbach, M., Karpatne, A., Kumar, V. : Introduction to Data Mining - 2nd Edition. Pearson, NY, USA (2019).
41. Friedman, J. H.: Stochastic Gradient Boosting. Computational Statistics & Data Analysis 38 (4), 367–78 (2002).
42. Hastie, T. Tibshirani, R., Friedman, J.: The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer, New York (2011).

Appendix. Definitions/questions of the dependent and independent variables

Dependent Variable	
SALREV_RED	Total percentage of change of your sales (increase or decrease) during the economic crisis of 2009-2014
Independent Variables – Strategic Orientations	
STRAT_CL	To what extent does your business strategy include low prices in comparison with the competition? (five levels ordinal variable)
STRAT_DIF	To what extent does your business strategy include high quality of products/services in comparison with the competition? (five levels ordinal variable)
STRAT_INNOV	To what extent does your business strategy include introduction of new products /services (with significant innovations) ? (five levels ordinal variable)
INNOV_PRS	Over the last three years did your firm introduce product innovations (=new or significantly improved products)? (binary)
INNOV_PROC	Over the last three years did your firm introduce process innovations (=new or significantly improved processes)? (binary)
NEW_PS_P	What percentage of your total sales revenue (turnover) in 2014 came from new products/services that were introduced in the market during the three previous years ? (continuous)
IMPR_PS_P	What percentage of the total sales revenue (turnover) in 2014 came from products /services that you had introduced before 2012, but were improved significantly over the last three years? (continuous)
INN_PRSD	Did you introduce methods/process innovation in the goods production or services' delivery processes in the last three years? (binary)
INN_SSWM	Did you introduce methods/process innovations in the sales, shipment or warehouse management processes? (binary)
INN_SUPP	Did you introduce methods/process innovations in the support processes (e.g. in the equipment maintenance ones) (binary)
R&D	Did your firm conduct R&D (Research and Development) in the last three years? (binary)
EXP_P	Percentage of exports in firm's sales revenue in 2014 (contin.)
Independent Variables – Human Resources	
EMPL	Number of employees at the end of 2014 (including any temporary employees, part-time, etc., who should be counted as full time equivalents) (continuous)
EMPL_TERT	Percentage of tertiary education graduates in the personnel of your firm (continuous)

EMPL_VOCT	Percentage of vocational/technical education graduates in the personnel of your firm (continuous)
EMPL_HIGH	Percentage of high school graduates in the personnel of your firm (continuous)
EMPL_ELEM	Percentage of elementary school graduates in the personnel of your firm (continuous)
EMPL_COM	What percentage of the employees of your firm use computer in their work (e.g. PC, terminal, or laptop)? (continuous)
EMPL_INTRA	What percentage of the employees of your firm uses the intra-net (internal network) of the firm in their work the? (continu-ous)
EMPL_INTER	What percentage of the employees of your firm uses Internet in their work? (continuous)
EMPL_ICT	Percentage of qualified ICT personnel in the workforce of your firm (continuous)
Independent Variables – Technology	
D_ERP	To what extent are Enterprise Resource Planning (ERP) sys-tems used in your firm? (five levels ordinal variable)
D_CRM	To what extent are Customer Relationship Management (CRM) sys-tems used in your firm? (five levels ordinal variable)
D_SCM	To what extent are Supply Chain Management (SCM) systems (= systems that support the electronic exchange of information with customers, suppliers and business partners, such as inven-tory levels, orders, production, shipments, invoices, etc.) used in your firm? (five levels ordinal variable)
D_BIBA	To what extent are Business Intelligence/Business Analytics systems (= systems that support advanced forms of processing business data, which lead to the creation of useful reports, as well as various types of models that aim at the support of deci-sion-making – this can be either a separate software, or a mod-ule of an ERP or CRM system) used in your firm? (five levels ordinal variable)
D_CS	To what extent are Collaboration support systems (= systems that support the internal collaboration between employees of the firm, and/or external collaboration with customers, suppliers and partners, offering capabilities of sharing various forms of content (e.g. text files, images), forum, instant messaging (and other forms of communication), project management, etc.) used in your firm? (five levels ordinal variable)
E-SAL	Do you conduct online sales of products/services through the Internet? (binary)
SM_SPRO	To what extent do you use social media for sales promotion? (five levels ordinal variable)
SM_OPCO	To what extent do you use social media in order to collect cus-

	tomers' opinions, comments and complaints about your prod-ucts or services? (five levels ordinal variable)
SM_IMPS	To what extent do you use social media in order to collect ideas for improvements or innovations in your product or services? (five levels ordinal variable)
SM_PERS	To what extent do you use social media in order to search for and find personnel? (five levels ordinal variable)
SM_INTC	To what extent do you use social media in order to support the internal exchange of information and co-operation among the employees of your firm? (five levels ordinal variable)
SM_IPAR	To what extent do you use social media in order to support the external exchange of information and co-operation with other firms (e.g. part-ners, suppliers, customers, etc.)? (five levels or-dinal variable)
CLOUD	Do you use cloud computing? (binary)
CL_IAAS	To what extent you use IaaS (Infrastructure as a Service = use of remote computing power and storage through the Internet)?
CL_PAAS	To what extent you use PaaS (Platform as a Service = remote use of the above plus database management systems and appli-cation develop-ment tools/ environments/languages through the Internet)? (five levels ordinal variable)
CL_SAAS	To what extent you use SaaS (Software as a Service = use through the Internet of remote application software that run on provider's systems)? (five levels ordinal variable)
Independent Variables – Structure	
ORG	Over the last three years did your firm use organic structural forms of work organization (such as teamwork and job rota-tion) ? (binary)
Independent Variables – General	
SECT	Firm's sector
COMP_PROF	How good was the financial performance of your firm over the last three years in comparison with your competitors in terms of profitability? (five levels ordinal variable)
COMP_SALR	How good was the financial performance of your firm over the last three years in comparison with your competitors in terms of sales revenue? (five levels ordinal variable)
COMP_MS	How good was the financial performance of your firm over the last three years in comparison with your competitors in terms of market share? (five levels ordinal variable)
COMP_ROI	How good was the financial performance of your firm over the last three years in comparison with your competitors in terms of ROI (return on investment)? (five levels ordinal variable)