Artificial intelligence-based public sector data analytics for economic crisis policymaking

Euripidis N. Loukis, Manolis Maragoudakis and Niki Kyriakou
Department of Information and Communication Systems Engineering, University of the Aegean, Chios, Greece

Received 20 November 2019
Revised 22 March 2020
Accepted 26 April 2020

Abstract

Purpose – Public sector has started exploiting artificial intelligence (AI) techniques, however, mainly for operational but much less for tactical or level tasks. The purpose of this study is to exploit AI for the highest strategic-level task of government: to develop an AI-based public sector data analytics methodology for supporting policymaking for one of the most serious and large-scale challenges that governments repeatedly face, the economic crises that lead to economic recessions (though the proposed methodology is of much more general applicability).

Design/methodology/approach – A public sector data analytics methodology has been developed, which enables the exploitation of existing public and private sector data, through advanced processing of them using a big data-oriented AI technique, "all-relevant" feature selection, to identify characteristics of firms as well as their external environment that affect (positively or negatively) their resilience to economic crisis.

Findings – A first application of the proposed public sector data analytics methodology has been conducted, using Greek firms' data concerning the economic crisis period 2009–2014, which has led to interesting conclusions and insights, revealing factors affecting the extent of sales revenue decrease in Greek firms during the above crisis period and providing a first validation of the methodology used in this study.

Research limitations/implications – This paper contributes to the advancement of two emerging highly important, for the society, but minimally researched, digital government research domains: public sector data analytics (and especially policy analytics) and government exploitation of AI. It exploits an AI feature selection algorithm, the Boruta "all-relevant" variables identification algorithm, which has been minimally exploited in the past for public sector data analytics, to support the design of public policies for addressing one of the most serious and large-scale economic challenges that governments repeatedly face: the economic crises.

Practical implications – The proposed methodology allows the identification of characteristics of firms as well as their external environment that affect positively or negatively their resilience to economic crisis. This enables a better understanding of the kinds of firms that are more strongly hit by the crisis, which is quite useful for the design of public policies for supporting them; and at the same time reveals firms' practices, resources, capabilities, etc. that enhance their ability to cope with economic crisis, to design policies for promoting them through educational and support activities.

Social implications – This methodology can be very useful for the design of more effective public policies for reducing the negative impacts of economic crises on firms, and therefore mitigating their negative consequences for the society, such as unemployment, poverty and social exclusion.

Originality/value – This study develops a novel approach to the exploitation of public and private sector data, based on a minimally exploited, for such purposes, AI technique ("all-relevant" feature selection), to support the design of public policies for addressing one of the most threatening disruptions that modern economies and societies repeatedly face, the economic crises.

Keywords Data analytics, Policy analytics, Artificial intelligence, Economic crisis, Economic recession, Feature selection

Paper type Research paper

Transforming Government: People, Process and Policy
© Emerald Publishing Limited
1750-6166
DOI 10.1108/TG-11-2019-0113

1. Introduction

Data analytics represents a major shift of the focus of information and communication technologies (ICT) usage by organizations beyond the support of operations, aiming to support decision-making at the operational, tactical and even the strategic level; it can be defined as the extensive exploitation of data from various sources, using advanced quantitative analysis techniques, to support various levels of decision-making (Davenport, 2013; Davenport and Harris, 2017; Seddon *et al.*, 2017; Aydiner *et al.*, 2019). Initially, data analytics methodologies have been developed and used in the private sector, and this has given rise to the development of "business analytics" (Seddon *et al.*, 2017; Aydiner *et al.*, 2019). Its success has generated strong interest to exploit and apply the knowledge developed in this area in the public sector as well, with appropriate adaptations to its specific needs, objectives and orientations; this gives rise to the gradual development of "public sector data analytics", placing emphasis on the support of higher level strategic decisions concerning public policies for addressing the important problems of modern societies, and this has given rise to the development of "policy analytics" (Tsoukias *et al.*, 2013; Janssen and Wimmer, 2015; Daniell *et al.*, 2016; De Marchi *et al.*, 2016; Gil-Garcia *et al.*, 2018). Policy analytics can be defined as the exploitation of existing data of government agencies, possibly in combination with data from private sector firms (e.g. business information and consulting ones), using advanced quantitative analysis techniques, to support various stages of public policymaking (such as agenda setting, policy analysis, policy formulation, policy implementation, policy monitoring and evaluation) for the complex problems/needs of modern societies. The development of policy analytics constitutes one of the most important trends in the area of digital government; however, it is still in its infancy (see section 2.1); so extensive research is required to develop a wide range of methodologies for exploiting public sector data, possibly in combination with data from the private sector as well, to support public policymaking in a wide range of thematic areas of government intervention.

Another important trend in the area of digital government is the increasing exploitation of artificial intelligence (AI) techniques in government (Desouza *et al.*, 2017; Eggers *et al.*, 2017; Desouza, 2018; Sun and Medaglia, 2019; Fernandes *et al.*, 2019; Desousa *et al.*, 2019), which has also been encouraged by the first successful applications of AI in the private sector (Bean, 2018; Ransbotham *et al.*, 2019; Duan *et al.*, 2019; OECD, 2019). AI can be defined as a group of technologies that enable computers to become more intelligent, by learning from their environment, gaining knowledge from it and using it for taking or proposing/recommending action (Craglia *et al.*, 2018; OECD, 2019). There are several techniques that can provide such learning capabilities, with the most widely used among them being definitely the machine learning (ML) ones, though there are other promising AI techniques as well, whose potential in both the private and the public sectors needs further investigation (Duan *et al.*, 2019). Government agencies have started exploiting AI techniques, mainly for operational or tactical level tasks, but this trend is still in its infancy. There has been some interesting research concerning the use of AI in government, but it is mainly for the automation, support and enhancement of some operational-level tasks, and to a lower extent for the support and enhancement of tactical-level tasks (see section 2.2); however, very limited research has been conducted about the use of AI for the support and enhancement of the higher level functions of government, and especially strategic ones concerning policymaking.

Our study makes a contribution toward the advancement of these two important and promising digital government trends and research domains: public sector data analytics (and especially policy analytics) and government exploitation of AI. It develops a policyoriented data analytics (policy analytics) AI-based methodology for supporting public policymaking concerning one of the most serious and large-scale problems that governments often face, the economic crises, which repeatedly occur with varying intensities in market economies, being an inevitable trait of them, and lead to contractions of economic activity, with quite negative consequences for the society, such as unemployment, poverty and social exclusion (Diebold and Rudebusch, 1999; Keeley and Love, 2010; Knoop, 2015; Allen, 2016) (see section 2.3). In particular, the research objective of our study is as follows: to develop a public sector data analytics methodology, which exploits existing data in both the public and the private sectors, concerning, on the one hand, firms' resilience to economic crisis (such as sales evolution because of crisis), and, on the other hand, characteristics of firms and their external environment, to identify characteristics of firms and their external environment that affect positively or negatively their resilience to economic crisis.

For this purpose, we are using a "big data-oriented" AI technique, feature selection (FS), which enables identifying from a big number of potential independent variables (contained in available high-dimensionality big data sets) the ones that actually affect a specific dependent variable of interest; in particular, we are using the Boruta "all-relevant" variables identification FS algorithm (Kursa *et al.*, 2010; Kursa and Rudnicki, 2010) (see Section 2.4). FS algorithms constitute an important part of AI, which enable computers to perform one of the most critical human actions: to distinguish in their complex environments the most important elements of them with respect to their specific objectives, and then focus their attention and following actions on them; these algorithms are usually used as complements of other kinds of AI algorithms, such as the ML ones, for the improvement of their performance, but they can also be used independently (Alelyani *et al.*, 2011; Cateni *et al.*, 2013; Tang *et al.*, 2014). According to Cateni *et al.* (2013), FS algorithms are divided into traditional and AI-based ones, with the latter outperforming the former. Furthermore, our methodology provides guidance for an initial selection of characteristics of firms and their external environment to be then used as input (potential independent variables) in the above FS algorithm, based on theoretical foundations from previous IS and management science research (see Section 2.5).

The proposed policy analytics methodology can provide substantial assistance and support for the design of public policies for reducing the negative impact of an economic crisis on firms. On the one hand, it enables a better understanding of the kinds of firms (in terms of characteristics, resources, capabilities, practices, etc.) that are more strongly hit by the crisis, which is quite useful for the design of effective public policies for supporting them. On the other hand, it enables a better understanding of the kinds of firms that manage to deal successfully with the crisis, and reveal firms' characteristics, resources, capabilities and practices that enhance the firms' ability to deal successfully with economic crisis, to design policies for promoting them through educational and support activities. A first application of our methodology is presented, based on Greek firms' data for the economic crisis period 2009-2014, which leads to interesting conclusions and insights, revealing factors affecting the extent of firms' sales revenue decrease during this crisis period, and providing a first validation of our methodology.

In Section 2, the background of our economic crisis policy analytics methodology is outlined. Then, in Section 3, we describe the methodology, followed by its first application in Section 4. Finally, Section 5 summarizes conclusions and proposes future research directions.

2. Background

2.1 Policy analytics

The gradual development of policy analytics constitutes one of the most important trends in the area of digital government, representing a major extension of it beyond the support of internal processes and operations, as well as transactions with citizens and firms, which were the main objectives of its first generations (Janowski, 2015; Lachana *et al.*, 2018), toward the support of the highest level function of government agencies: the public policymaking. The increasing availability of data in government agencies has a great potential to provide, after appropriate integration and processing of them, substantial support of policymaking in important domains of government intervention.

Some initial research has been conducted in the area of policy analytics, which has developed some first knowledge concerning approaches and methodologies for exploiting various sources of public sector data, using basic or more advanced quantitative analysis techniques, to support some of the stages of the policymaking cycle in some domains of government intervention, such as the economy, the social insurance, the health care, the environment, the energy provision, the justice and the management of emergency crises (Hiltz *et al.*, 2011; Baer *et al.*, 2015; Ekstrom *et al.*, 2018; Park and Johnston, 2018; Vaan den Braak and Choenni, 2018; Van Dijk *et al.*, 2018). However, this promising area of policy analytics is still in its infancy, and the potential of the large quantities of available government data for supporting policymaking has been exploited only to a limited extent; so extensive research is required to develop effective methodologies for exploiting this wealth of public sector data to the highest possible extent, and extracting valuable knowledge from them, to provide substantial support for policymaking in important domains of government intervention. Our study makes a contribution in this direction, by developing a policy data analytics methodology for exploiting public sector data, in combination with relevant private sector data, using AI (see Section 2.2) (making use of a big data-oriented AI technique, the "all-relevant" FS one – see Section 2.4), to support policymaking for one of the most serious and large-scale problems that governments repeatedly face: the economic crises (see Section 2.3).

2.2 Artificial intelligence in government

Another important trend in the area of digital government is the increasing exploitation of AI techniques by government agencies. As mentioned in Section 1, AI includes a group of techniques that enable computers to perform tasks of higher human-like intelligence, by learning from their environment, and then using the knowledge they have gained from it for taking or proposing/recommending action (Craglia *et al.*, 2018; Duan *et al.*, 2019; OECD, 2019). While the first generation of AI was based on logic and rules predefined by humans ("symbolic AI"), the second generation of it was based on logic and rules extracted automatically by computers through advanced processing of past historic data ("statistical AI"), from which models or sets of rules are constructed, that enable, on the one hand, deeper insights (e.g. concerning associations among important variables) and, on the other hand, making predictions of important variables (Duan *et al.*, 2019; OECD, 2019). In this second generation of AI, the most representative and widely used techniques are definitely the ML ones; however, there are several other promising AI techniques as well, whose potential in both the private and the public sectors needs further investigation (Duan *et al.*, 2019).

Though most of the AI technologies, and in particular the ML ones, exist for several decades, it is only recently that there has been a very high interest in their "real life" application and exploitation, initially by private sector firms, and later by government agencies as well, for a number of reasons:

- availability of large amounts of data that enable a more effective training of AI algorithms (to extract more reliable models and rules);
- advances in computing power and reduction of its cost; and
- substantial improvements of AI algorithms (Makridakis, 2017; Craglia *et al.*, 2018; Duan *et al.*, 2019; OECD, 2019; Ransbotham *et al.*, 2019).

These first "real life" applications of AI technologies in the private sector have revealed its great potential to offer important benefits: productivity improvements, sale revenue increase, higher quality decision-making and innovations in internal processes, products and services (Makridakis, 2017; Bean, 2018; Ransbotham *et al.*, 2019; Duan *et al.*, 2019; OECD, 2019).

The success stories of the "real life" AI applications in the private sector have generated high levels of interest to exploit AI techniques in the public sector as well, to automate or support more sophisticated mental tasks than the simpler routine ones automated or supported by the traditional operational information systems (IS) of government agencies (Desouza *et al.*, 2017; Eggers *et al.*, 2017; Desouza, 2018; Sun and Medaglia, 2019; Fernandes *et al.*, 2019; Desousa *et al.*, 2019). Some interesting research has been conducted concerning the exploitation of AI in a variety of public sector thematic domains for various purposes; for instance, in education (Rockoff *et al.*, 2010), social policy (Chandler *et al.*, 2011), public security (Ku and Leroy, 2014; Camacho-Collados and Liberatore, 2015), health care (mainly for supporting diseases' diagnosis and treatment planning) (Sun and Medaglia, 2019) and public transportation management (Kouziokas, 2017). However, this research has focused mainly on the exploitation of AI in government for the automation, support and enhancement of operational-level tasks, and to a much lower extent for the support and enhancement of tactical-level tasks; on the contrary, very limited research has been conducted about the exploitation of AI for the support and enhancement of the higher level functions of government, and especially the strategic ones concerning policymaking. Our study contributes to filling this gap, by developing a methodology for exploiting AI to support policymaking for addressing a very serious problem that governments repeatedly face: the economic crises (see Section 2.3).

2.3 Economic crises

One of the most serious weaknesses of market-based economies are the fluctuations of economic activity they repeatedly exhibit, which cause big problems to the economy and the society in general, so they have to be addressed by government through appropriate policies, aiming at avoiding them, if possible, and once they occur at reducing their intensities and durations and at mitigating their negative consequences for firms and citizens (Diebold and Rudebusch, 1999; Keeley and Love, 2010; Knoop, 2015; Allen, 2016). Economic crises can be defined as significant contractions of economic activity, which can be because of the "business cycles" (i.e. the fluctuations that economic activity usually exhibits in market-based economies), or can be caused by various kinds of events in the society or economy, such as big increases of the prices of important goods, especially of goods that are used as inputs in extensive production activities, such as the oil crisis in the early 1970, and also banking crises, or even epidemics (Knoop, 2015).

The economic crises have quite negative both short-term, medium- and long-term consequences for the economy and the society. The short-term consequences usually include reductions of the demand for most goods and services, resulting in serious decrease of firms' sales, production and profits, which leads to reductions of personnel employment (thus increasing unemployment and poverty) as well as materials' procurement (thus propagating the crisis from the sectors from which it started toward the sectors of their suppliers, etc., and finally to the whole of the economy, increasing the scale of the problem). Furthermore, during economic crises, firms usually reduce capital investment in production equipment, ICT, buildings, etc., and also in product, service and process innovations, which reduce the degree of renewal and improvement of their equipment, products, services and operations, as well as the exploitation of emerging new technologies, and these have serious medium- and long-term consequences on the efficiency and competitiveness of firms (Keeley and Love, 2010; Izsak *et al.*, 2013; Knoop, 2015; Allen, 2016). Therefore, it is necessary that government agencies, especially the ones having competences and responsibilities for the economy and social welfare, to design and implement public policies for reducing these negative short-term as well as medium- and long-term consequences of the economic crises. The management of the economic crises of various intensities and durations that repeatedly occur in market-based economies are among the most serious and large-scale challenges that governments face.

However, it should be noted that the above negative consequences of the economic crises are not the same for all firms: the more efficient and effective firms, offer higher value-for-money products and services, and have higher capacity to make the required adaptations to the new economic conditions the crisis gives rise to, are more resilient to the crisis and have less negative consequences on their sales revenue, and therefore on their employment, procurement as well as on their capital investment and innovation, than the less efficient and effective ones. Therefore, it is important and highly useful to develop policy analytics methodologies for identifying characteristics of firms (such as resources, capabilities and practices) and their external environment that affect positively or negatively their resilience to economic crisis; this can be quite useful for the design of more focused and effective policies for mitigating the negative impact of economic crises on firms.

This study makes a contribution in this direction: it develops an economic crisis policy analytics methodology for exploiting existing data of taxation authorities, statistical agencies and also of private sector business information and consulting firms, to identify characteristics of firms and their external environment that affect (positively or negatively) their resilience to the crisis. For this purpose, we are using an advanced AI FS algorithm, the Boruta "all-relevant" variables identification one, which is outlined in the following section.

2.4 Artificial intelligence feature selection – the Boruta algorithm

FS algorithms constitute an important class of "big data-oriented" AI algorithms, which aim at determining from a big number of features – potential independent variables – the ones that affect a dependent variable of interest (Tang *et al.*, 2014). The exponentially increasing availability of big data today leads to massive data sets, which are characterized, on the one hand, by large sample size, i.e. contain data about a large number of units (e.g. individuals, firms), and, on the other hand, by high dimensionality, i.e. contain data about a large number of features of these units. This high dimensionality makes it difficult to make sense of such massive data and understand which of these multiple features are relevant affecting the dependent variable we are studying (usually an important outcome variable), and which features are not relevant. Furthermore, the use of large numbers of features, both relevant and nonrelevant ones, as independent variables in ML algorithms leads to reduction of their performance with respect to accuracy in the prediction of the dependent variable for new units, which is highly important in "predictive analytics" (Alelyani *et al.*, 2011; Tang *et al.*, 2014). For these reasons, the FS algorithms are quite important:

- for extracting valuable knowledge and insight from big data sets as to which their features affect our dependent variable; and
- for an initial selection of relevant features to be used as inputs in ML algorithms, to have higher levels of accuracy in the prediction of the dependent variable.

These FS AI algorithms can be divided into two main categories:

(1) the "minimal – optimal" ones, which determine a small, minimal set of features affecting the dependent variable, which can provide optimal prediction accuracy for it (most traditional FS algorithms belong to this category); and

(2) the "all-relevant" ones, which determine all the features that affect the dependent variable, and not only the nonredundant ones, as it happens in "minimal – optimal" FS algorithms (there is a smaller number of novel algorithms belonging to this category) (Kursa *et al.*, 2010; Kursa and Rudnicki, 2010; Alelyani *et al.*, 2011; Tang *et al.*, 2014).

Therefore, if there are a number of features that affect the dependent variable, which are to some extent redundant (i.e. there is some degree of association among them), the "minimal – optimal" FS algorithms will select only some of them (a minimal subset), which have low levels of redundancy (association); however, it will not select some other features, which affect the dependent variable, but have high levels of association with the selected ones (as these other features do not increase further the accuracy of the prediction of the dependent variable, beyond the accuracy achieved based on the initially selected features). On the contrary, the "all-relevant" FS algorithms will select all the features that affect the dependent variable, regardless of possible associations among them. Therefore, the "all-relevant" FS algorithms are appropriate if our objective is to extract knowledge and insight from big data sets as to which the features actually affect the dependent variable, whereas the "minimal – optimal" FS algorithms are appropriate if our objective is to predict the value of a dependent variable for a new unit from the corresponding values of the independent variables for this unit (which are known).

Because the objective of this study is the former (to extract knowledge and insight concerning characteristics of a firm and its external environment that affect the degree of its sales revenue reduction because of economic crisis), we are using an "all-relevant" FS algorithm. In particular, we are using the Boruta "all-relevant" variables identification algorithm (Kursa *et al.*, 2010; Kursa and Rudnicki, 2010; Alelyani *et al.*, 2011; Tang *et al.*, 2014), which is an FS approach, particularly useful when one is interested in understanding the "mechanisms" related to a dependent variable of interest, rather than just building a "black box" predictive model of it with good prediction accuracy. The basic idea of the Boruta algorithm is that based on the original feature set, another artificial set of features is created, which consists of shuffled copies of all features, which are called "shadow features". This "shadow features" set is then merged with the original one; a Random Forest classifier is constructed based on the merged data set; and for each feature, an importance measure is calculated (usually the "mean decrease impurity" of the feature), to evaluate the importance of the feature. At each iteration, Boruta FS algorithm evaluates one real feature, by assessing whether it has a higher importance than the best of the shadow features, and if this does not happen, the feature is removed (as it is un-important). Finally, the algorithm stops when all features get either confirmed or removed, or it reaches a specified limit of runs. This FS algorithm offers three crucial advantages over other techniques that might be used for the same purpose (e.g. various kind of regression):

- It can handle large numbers of features without performance and reliability deterioration, so it is appropriate for exploiting really "big data" sets; this does not happen in other techniques that might be used for the same purpose, such as regression analysis, in which when the number of independent variables increases, the confidence intervals of the bi-coefficient estimations increase as well, so some statistically significant ones may be incorrectly estimated as insignificant.

- If there are associations – correlations between some of the features (which is quite usual), this algorithm will not omit features that affect the dependent variable if they are associated – correlated with other selected features, again this does not happen in other techniques that might be used for the same purpose; for instance, in regression analysis, if some independent variables that actually affect the dependent variable have high levels of correlation, then for some of them, their bi-coefficients might be correctly estimated as statistically significant, but for some others, their bi-coefficients might be incorrectly estimated as statistically insignificant (multi-collinearity problem).
- Also, the Boruta FS algorithm can identify not only the features that have linear effects on the dependent variable, but also the ones having nonlinear effects on it.

It should be noted that limited research has been conducted on the exploitation of advanced AI algorithms for supporting policymaking, though the latter quite often necessitates distinguishing from complex economic and social environments their most important elements with respect to government agencies' specific objectives, to focus their attention and scarce personnel and economic resources on them; this usually requires analyzing big data sets and identifying their most relevant features/variables for a specific problem/need that has to be addressed, which have substantial impact on the targeted outcome variables. Furthermore, most of these limited studies use FS AI algorithms as complements of MS algorithms, to improve the performance of the latter by selecting the most relevant independent variables (Sethi and Jain, 2010; Chang and Tsai, 2017). Our study contributes to filling this research gap, by investigating the independent use of an FS AI algorithm for supporting policymaking for one of the biggest problems of market-based economies: the economic crises.

2.5 Conceptualizations of main elements of a firm and external environment

Though the Boruta FS algorithm, as mentioned above, can process large numbers of variables without performance and reliability deterioration, to select all the relevant ones that affect the dependent variable (which in our study is firm's resilience to economic crisis), it is useful to make initially a preselection of reasonable and meaningful potential independent variables, among the numerous variables that might be available in the quite big data sets of government (e.g. of statistical agencies, taxation authorities, etc.). This can be done by preselecting from these big data sets appropriate characteristics of firms and their external environment that, according to previous management and IS literature, are expected to affect their performance (as they might affect firms' performance in coping with the crisis as well, and therefore their resilience to the difficult crisis conditions). Highly useful for this purpose as theoretical foundations can be conceptualizations of the main elements of a firm, and also conceptualizations of firm's external environment, that determine firm's performance, developed in previous management science and IS research; they can provide direction for the initial filtering of the numerous variables that the existing big data sets might include and preselection of a set of reasonable and meaningful potential independent variables to be processed by the Boruta FS algorithm. This is highly important as previous literature has emphasized the risks posed by purely "big data-driven" research, lacking theoretical direction and foundation (so it does not exploit the valuable knowledge that relevant theories incorporate), which can lead to the production of meaningless or spurious conclusions that cannot be explained and understood, and constitute unsafe bases for action (Gonzalez-Bailon, 2013; Mazzocchi, 2015; Rai, 2016). According to Rai (2016), big data and theory are mutually reinforcing and synergistic for producing highly reliable knowledge, as theory provides basis for collecting/using the appropriate data, as well as for making sense of them.

The most widely recognized and used conceptualization of the main elements of a firm is definitely the classical "Leavitt"*s* Diamond" framework (Leavitt, 1964); according to it, the four main elements of a firm that constitute the main determinants of its performance are as follows:

- task (= the strategies, as well as the administrative and production processes of the firm);
- people (= the skills of firm's human resources of the firm);
- technology (= the technologies used for implementing the above processes); and
- structure (= the organization of the firm in departments, and the communication and coordination patterns of them).

However, this fundamental framework focuses mainly on firm's resources (e.g. human, technological, organizational), but subsequent strategic management research has revealed the high importance for the performance of a firm not only of its resources but also of their exploitation for developing capabilities to performing efficiently and effectively the most important functions of the firm (Johnson *et al.*, 2017); gradually, firm's capabilities have been widely recognized as important elements of a firm, and the most critical determinants of its performance. Because of the increasing penetration and use of ICT by firms, among the most important capabilities of a firm are definitely its ICT-related capabilities (Lu and Ramamurthy, 2011; Chen *et al.*, 2014), which concern firm's ability for efficient and effective mobilization and exploitation of ICT resources to support and enhance firm's activities.

Furthermore, strategic management research has revealed that beyond the above "ordinary" capabilities, in dynamic fast-changing business environments their "dynamic" capabilities are of critical importance for firms' performance as well (Teece, 2007; Drnevich and Kriauciunas, 2011); dynamic capabilities are defined as firm's abilities for sensing changes in its external environment that create opportunities and threats, seizing the opportunities, and making the required reconfigurations of firm's resources base (e.g. modifications, extensions, upgrades and new combinations of firm's resources) to adapt to and remain competitive in the new external conditions. As the economic crises give rise to big changes in firm's external environment [e.g. decrease of demand for their products and services, changes in customers' needs and preferences (toward lower cost and higher value for money products and services), new products and service offerings by competitors, etc.], we expect their dynamic capabilities to be highly important for adapting to and coping with the new conditions, and finally exhibiting higher resilience to crises. According to relevant literature, the most important dynamic capabilities of a firm for its performance are as follows:

- firm's absorptive capacity (ACAP) (Cohen and Levinthal, 1990; Camisón and Forés, 2010), defined as its ability to recognize and acquire useful new knowledge from the external environment, assimilate it, integrate/combine it with its existing internal knowledge and then exploit it to make valuable innovations in its processes, products and services; and
- organizational agility (Sherehiy *et al.*, 2007; Lu and Ramamurthy, 2011), defined as the ability to detect changes in market environment and respond to them.

With respect to firm's external environment, the most widely recognized and used conceptualization of it is Porter's Five Forces Framework (Johnson *et al.*, 2017), which defines the main elements of firm's external sectoral micro-environment that shape the "generalized competition" it faces, and finally affect significantly its performance; these main elements are competition (price and non-price), bargaining power of buyers, bargaining power of suppliers, threat of new entrants and threat of substitutes. Furthermore, as mentioned above, strategic management research has revealed the importance also of the degree of dynamism of firm's external environment with respect to products and services, technologies and customers' preferences (Teece, 2007; Chen *et al.*, 2014).

Therefore, from the numerous variables that might be available in big government and private firms' data sets, we can make an initial preselection of variables concerning the abovementioned main elements of a firm, which, according to previous management and IS research literature, determine its performance: task (strategic orientations and processes), human resources, technology (and especially use of ICT), structure and also ordinary capabilities (and especially ICT capabilities) as well as dynamic capabilities (and especially ACAP and agility); and also variables concerning the abovementioned main elements of a firm's external environment that also affect significantly its performance: price and non-price competition, bargaining power of buyers, bargaining power of suppliers, threat of new entrants and threat of substitutes, as well as various aspects of environment dynamism. These preselected variables can be used as potential independent variables and processed by the Boruta FS algorithm, to identify all the relevant ones that affect the dependent variable of this study: firm's resilience to the crisis.

3. Proposed methodology

The proposed economic crisis policy analytics methodology aims to identify characteristics of a firm and its external environment that affect its resilience to economic crisis, focusing initially on the most important measure of it: the reduction of firms' sales revenue because of the crisis (however, our methodology can be used for any other measure of firm's resilience to economic crisis). Therefore, the dependent variable of our methodology is the degree of firm's sales revenue reduction because of the crisis. The capabilities and advantages offered by the abovementioned advanced Boruta AI FS algorithm (outlined in Section 2.4) allow us to examine a large number as well as a wide thematic range of potential independent variables, to identify all the relevant and influential ones. This enables us to examine not only simpler characteristics of the firm, but also a wide range of more sophisticated ones, concerning important ordinary and dynamic capabilities of the firm, as previous management and IS research has revealed their importance as determinants of firm's performance (so they might affect firms' performance in coping with the crisis as well, and therefore their resilience to the difficult crisis conditions); however, it is not possible to find data about these capabilities in government data sets, so we have to complement them with relevant data sets of private sector firms' (e.g. business information or consulting firms) that include such data.

Our methodology can provide a substantial support for the design of public policies for reducing the negative impact of economic crisis on firms, which can impair significantly their both short-term and medium- and long-term performance. In particular, it enables a better understanding of the kinds of firms (in terms of characteristic, resources, capabilities, practices, etc.) that experience more negative consequences from the crisis, to design appropriate targeted public policies and programs for supporting them; and also the kinds of firms that have been more resilient and resistant to the crisis, to learn from them which can be quite useful in order to; and it enables the identification of characteristics, resources, capabilities, practices, etc. that enhance firms' ability to cope with economic crisis, so that we can promote them through educational and support activities.

In particular, our methodology exploits existing data from the following sources:

- From taxation authorities, we can exploit data they possess concerning firms' sales revenue before and during the economic crisis, from which sales revenue reduction because of economic crisis can be calculated.
- From statistical agencies, and also from private sector business information and consulting firms, we can exploit data they possess concerning various characteristics of firms (e.g. resources, capabilities, practices, etc.) and their external environment.

The proposed methodology includes two stages of processing these data, which are shown with their theoretical foundations in Figure 1:

(1) An initial filtering of the numerous variables that might be available in the big government and private sector data sets we are using, and preselection of meaningful potential independent variables among them, based on the theoretical foundations outlined in Section 2.5.

(2) Processing of the preselected variables using an AI "all-relevant" FS algorithm, such as the abovementioned Boruta one, to determine/identify all the relevant variables that actually affect the degree of sales revenue reduction because of the crisis.

In particular, based on the theoretical foundations outlined in Section 2.5, we preselect potential independent variables that belong to the following eight categories:

(1) *Strategic orientations*: this category can include variables concerning the degree of adopting the main strategies described in relevant strategic management literature (Johnson *et al.*, 2017), such as cost leadership, differentiation, focus, innovation and export.

(2) *Processes*: it can include various characteristics of firm's processes, such as complexity, efficiency, formality and flexibility.

(3) *Human resources*: it can include variables 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

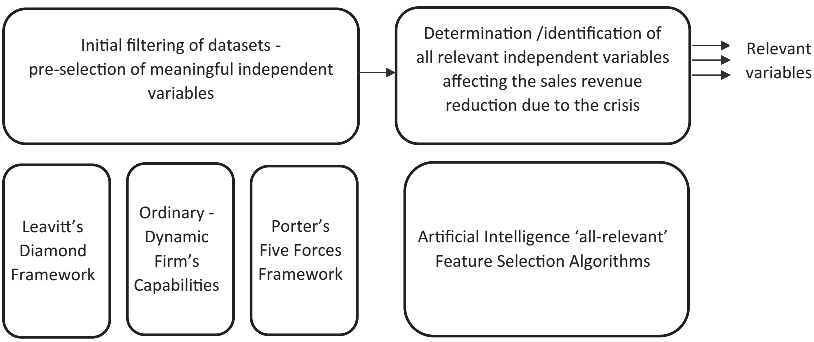

Figure 1.
Stages of economic crisis policy analytics methodology and their foundations

ICTs or various production technologies), the provision of various kinds of training, etc.

(4) *Technology*: variables concerning the use of various important ICTs [such as Enterprise Resource Planning (ERP) systems, Customer Relationships Management (CRM) systems, Supply Chain Management (SCM) systems, Business Intelligence/ Business Analytics (BI/BA) systems, Collaboration Support (CS) systems, e-sales, social media, cloud computing, etc.], or the use of various production technologies.

(5) *Structure*: variables concerning various aspects of the structure of the firm, such as main structural design (functional, product/service based, geographic, matrix), degree of differentiation, specialization, centralization/decentralization, use of organic structural forms, etc. (Sherehiy *et al.*, 2007; Jones, 2013).

(6) *Ordinary capabilities*: variables concerning the levels of firms' capabilities to perform efficiently and effectively the main firm's functions, such as the ones proposed by Porter's value chain model (inbound logistics, operations, outbound logistics, marketing and sales, service (primary ones) and human resource management, technology development, procurement, infrastructure) (Johnson *et al.*, 2017); and also the levels of various ICT capabilities of the firm (Lu and Ramamurthy, 2011; Chen *et al.*, 2014).

(7) *Dynamic capabilities*: variables concerning the main aspects of firm's ACAP (such as recognition and acquisition or relevant external knowledge; assimilation of it; integration/combination of it'; and exploitation for innovations in its processes, products and services) (Cohen and Levinthal, 1990; Camisón and Forés, 2010), as well as agility (e.g. agility with respect to emergence of new technologies, new suppliers, new products and services as well change of prices by competitors, etc.) (Sherehiy *et al.*, 2007; Lu and Ramamurthy, 2011).

(8) *External environment*: variables concerning the intensity of the five aspects of the "generalized competition" proposed by Porter's Five Forces Framework (Johnson *et al.*, 2017): price and non-price competition, bargaining power of buyers, bargaining power of suppliers, threat of new entrants and threat of substitutes; and also variables concerning various aspects of dynamism of firm's external environment.

4. Application

An application of the economic crisis policy analytics methodology described in the previous section has been made for the identification of characteristics of Greek firms as well as their external environment that affect the degree of their sales revenue reduction because of the long and intensive economic crisis that Greece experienced (Gourinchas *et al.*, 2016). For this purpose, we have used existing Greek firm's data for the period 2009–2014 from three sources:

(1) the Ministry of Finance – Taxation Authorities;

(2) the Hellenic Statistical Authority; and

(3) the ICAP S.A., a well-known business information and consulting firm.

In particular, we have used data from these three sources for 363 Greek firms: 40.2% of them were from manufacturing sectors, 9.4% from constructions and 50.4% from services sectors; 52.6% of them were small, 36.1% medium and 11.3% large ones.

Our dependent variable was the percentage of sales revenue reduction because of the economic crisis in the period 2009-2014, which was discretized by the Ministry of Finance (to avoid providing too detailed data about this critical topic and violating data protection regulations) 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%; and decrease by more than 100%.

We selected the following 61 independent variables, from 7 out of the 8 categories described in the previous section (because we did not find any variable in the available data sets concerning the "Processes" category):

(1) *Strategic orientations*: degree of adopting a cost leadership strategy (STRAT_CL), a differentiation strategy (STRAT_DIF), a product/service innovation strategy (STRAT_INNOV) (five levels ordinal variables); introduction of product/service innovations in the past three years (INNOV_PRS), introduction of process innovations in the past three years (INNOV_PROC) (binary variables); percentage of sales revenue coming from new products/services introduced in the past three years (NEW_PS_P), percentage of sales revenue coming from products/services significantly improved in the past three years (IMPR_PS_P) (continuous variables); introduction of innovations in the production processes or in the services delivery processes (INN_PRSD), introduction of innovations in the sales, shipment or warehouse management processes (INN_SSWM), introduction of innovations in support processes (such as equipment maintenance) (INN_SUPP); conduct of R&D (R&D) (binary variables); and exports as percentage of firm's sales revenue (EXP_P) (continuous variable).

(2) *Human resources*: number of employees (EMPL); percentage of firm's employees having tertiary education (EMPL_TERT), vocational/technical education (EMPL_VOCT), high school education (EMPL_HIGH), elementary school education (EMPL_ELEM); percentage of firm's employees using computer for their work (EMPL_COM), firm's intranet (EMPL_INTRA), internet (EMPL_INTER); and ICT personnel as a percentage of firm's total workforce (EMPL_ICT) (continuous variables).

(3) *Technology*: degree of ERP system use (D_ERP), CRM system use (D_CRM), SCM system use (D_SCM), BI/BA use (D_BIBA), CS systems use (D_CS) (five levels ordinal variables); conduct of e-sales (E-SAL) (binary variable); use of social media for sales' promotion (SM_SPRO) for collecting customers' opinions and complaints about firm's products and services (SM_OPCO), for collecting ideas for improving products and services (SM_IMPS), for finding personnel (SM_PERS), for internal co-operation within the firm (SM_INTC), for information exchange with other partner firms (SM_IPAR) (three levels of ordinal variables); use of cloud computing (CLOUD) (binary variable); and degree of using cloud computing IAAS (CL_IAAS), cloud computing PAAS (CL_PAAS), cloud computing SAAS (CL_SAAS) (five levels ordinal variables)

(4) *Structure*: use of organic structural forms (such as teamwork and job rotation) (ORG) (binary variable).

(5) *Ordinary capabilities*: six variables concerning the main ICT-related capabilities identified in relevant literature (Lu and Ramamurthy, 2011; Chen *et al.*, 2014) concerning ICT strategic planning/alignment (ICT_STRPL), cooperation between ICT and business units (ICT_BUSC), cooperation with ICT vendors (ICT_VENDC), development of ICT applications (ICT_ADEV), modification of ICT applications

(ICT_AMOD) and integration of ICT applications (ICT_AINT) (five levels ordinal variables).

(6) *Dynamic capabilities*: four variables concerning the main aspects of ACAP (Cohen and Levinthal, 1990; Camisón and Forés, 2010): firm's ability for external relevant knowledge recognition and acquisition (ACAP_ACQ), dissemination and analysis (ACAP_DIS), assimilation and integration in firm's knowledge base (ACAP_INT) and exploitation for process, products and services innovations (ACAP_EXP); and six variables concerning the main aspects of organizational agility (Sherehiy *et al.*, 2007; Lu and Ramamurthy, 2011) with respect to firm's reaction ability to respond to the introduction of new products and services by competitors (AG_NPS), new pricing policies of them (AG_NPR), changes in the demand for its products and services (AG_CDE), to customize its products and services to customers' special needs (AG_CUST), to expand to new markets (AG_EXPM) and to change suppliers to reduce cost and increase quality (AG_CSUP) (five levels of ordinal variables).

(7) *External environment*: number of competitors (N_COMP) (continuous variable); intensity of price competition (INT_PCOM), non-price competition (INT_NPCOM); and also four environmental dynamism variables concerning the extent of changes in products and services (DYN_PRS), technologies (DYN_TECH), demand for products/services (DYN_PRS) and competitors' movements (DYN_COMP) (Sherehiy *et al.*, 2007; Lu and Ramamurthy, 2011) (five levels of ordinal variables).

(8) *General*: sector (SECT) (binary variable: manufacturing/services).

The results from processing the above variables using the Boruta FS AI algorithm are shown in Table 1, in which we can see "all-relevant" variables identified in order of importance for the dependent variable. In particular, ten variables have been identified that affect the degree of sales revenue reduction because of the crisis (SALREV_RED). For each of them, we examined then whether it has a positive and negative effect: for each binary and ordinal variable, this was done by calculating and comparing the averages of SALREV_RED for all its discrete values; for the continuous variables, we did the same after discretizing them (initially we recoded them into corresponding binary variables based on the median value: values lower than the median were recoded into 0, whereas values higher than the median were recoded into 1; and then we recoded them similarly into corresponding four levels of variables based on the quartile values).

Table 1. Relevant variables affecting the degree of sales revenue reduction because of the crisis

Variable	Impact
Use of organic structural forms (teamwork, job rotation) (ORG)	Negative
Percentage of sales revenue coming from new products/services introduced in the past three years (NEW_PS_P)	Negative
Introduction of innovations in support processes (such as equipment maintenance) (INN_SUPP);	Negative
Introduction of innovations in the sales, shipment or warehouse management processes (INN_SSWM)	Negative
Number of employees (EMPL)	Negative
Degree of ERP systems use (D_ERP)	Negative
Percentage of personnel having vocational/technical education (EMPL_VOCT)	Positive
Exports as percentage of firm's sales revenue (EXP_P)	Negative
Percentage of personnel having tertiary education (EMPL_TERT)	Negative
Capability for integration of ICT applications (ICT_AINT)	Negative

We remark that the most important of the examined variables (firm characteristics) for the dependent variable (degree of sales revenue reduction because of the crisis) is the use of organic structural forms (such as teamwork and job rotation), which has negative impact on SALREV_RED, so it reduces the negative consequences of the crisis on firm's sales revenue. The economic crises give rise to big changes in firms' external environment (e.g. decrease of demand for their products and services, changes in customers' needs and preferences, new lower cost products and service offerings by competitors, etc.); they increase its complexity; the adoption of organic structures (such horizontal teams) allows a more intensive exchange and synthesis of information and knowledge among employees from different functions and departments, which enables:

- a better and more holistic understanding of these environmental changes/complexities; and
- a more effective design and implementation of actions for responding to them, such as new products/services with lower cost or higher value-for-money, new pricing policies, expansions to new markets (both domestic and foreign ones), etc.

Furthermore, we remark that four out of the ten identified relevant variables belong to the category "strategic orientations," all of them having negative impact on SALREV_RED, so they represent strategies that increase firm's resilience to economic crisis. Three of them concern innovation strategies: percentage of sales revenue coming from new products and services, introduction of innovations in support processes (such as equipment maintenance), as well as introduction of innovation in sales, shipment, warehouse management processes. Therefore, the introduction of new products and services creates new markets and sales opportunities, which generate new sales revenue that compensates to some extent for the reduction of sales revenue from the "traditional" products and services of the firm because of the crisis; additionally, it is possible to introduce new products and services, which are lower cost and higher value-for-money variants of previous ones, so they are highly attractive for most customers during the crisis. Furthermore, the above process innovations increase firm's efficiency, which is quite useful for coping with the difficult crisis conditions and becoming more competitive to gain a larger share of the sharply declining market during the crisis. The fourth "strategic orientations" related variable concerns the adoption of an export strategy; this indicates that exports generate sales revenue from foreign markets (not suffering from economic crisis), so they reduce firm's reliance on its domestic market that suffers from economic crisis and declining demand, and decrease the negative consequences of the latter on firm's overall sales revenue.

Furthermore, there are two of the identified relevant variables that concern the exploitation of ICTs in the firm, both of them having negative impact on SALREV_RED: the use of ERP systems and the capability for integration of existing ICT applications; this reveals two important technological characteristics that increase firm's resilience to economic crisis. The use of ERP systems provides comprehensive and integrated electronic support of all firm's functions, so it enhances their efficiency, which is quite useful for coping with the crisis and gaining a large share of the declining demand during the crisis (for which there is strong competition). Also, a high capability for integrating existing ICT applications enables the interconnection of isolated "islands of automation" (belonging to the same or different departments), and their evolution toward an integrated ICT infrastructure, enabling data and functionality of one ICT application to be exploited by others as well; this improves cooperation between firm's departments, and enhances firm's efficiency, increasing its resilience to crisis.

Finally, there are three of the identified relevant variables that concern firm's "human resources." The employment of personnel with tertiary education has a negative impact on SALREV_RED, so it increases firm's ability to cope with the crisis; on the contrary, the employment of personnel with lower vocational/technical level education (though less costly) has the opposite effects. This indicates that the employment personnel with high level of education increases firm's ability to understand better the big changes as well as complexities in firm's environment that the crisis gives rise to, and then design and implement more effective response actions, such as new products/services with lower cost or higher value-for-money, new pricing policies and expansions to new markets (both domestic and foreign ones). The number of firm's employees has also a negative impact on SALREV_RED, indicating that larger firms have lower reductions of sales revenue because of the crisis.

It is interesting that most of the examined ordinary capabilities (with the only exception of the abovementioned capability for integration of existing ICT applications), as well as all the examined dynamic capabilities (concerning firm's ACAP and organizational agility), are not included among the ones found as "relevant" (i.e. affecting firm's resilience in the crisis). This indicates that firms do not exploit these capabilities effectively (possibly with appropriate adaptations of them) to cope with the economic crisis.

It should be noted that the above findings are highly reliable, because of the three crucial advantages that the FS AI algorithm we have used offers over other techniques that might be used for the same purpose (e.g. various kind of regression), which have been mentioned in Section 2.4. These findings are not sensitive to variations of the potential independent variables set used, as the subset of the most important independent variables, which affect substantially the dependent one, is extracted based on the influence of each of them on the values of the dependent variable. So, we believe that our findings are useful for the design of effective public policies for reducing the negative consequences of economic crisis on firms. In particular, our findings indicate that Greek government agencies, to reduce the negative consequences of the economic crisis on firms, should design and implement effective public policies (such as legislation, financial support and provision of training and consulting) for promoting firms' innovation and export activities. Furthermore, it is necessary to design and implement effective public policies (including financial support, as well as provision of training and consulting) for promoting the adoption of ERP systems, organic structural forms (complementing their hierarchical structures with horizontal teamwork) and for employing personnel of higher educational level. These public policies should be focused on small and medium firms, as they seem to be less resilient to the crisis.

5. Conclusions

In the previous sections, a public sector data analytics methodology has been presented, which aims to support the highest strategic-level task of government, the development of public policy, based on advanced AI techniques. In particular, it exploits existing big data sets of the public as well as the private sector, and processes them with an advanced big data-oriented AI FS algorithm, to identify characteristics of a firm (e.g. resources, capabilities, practices, etc.) as well as its external environment that affect (positively or negatively) its resilience to economic crisis. This provides valuable support for policymaking concerning one of the most serious and large-scale problems that governments repeatedly face: the economic crises. It enables the development of appropriate and focused public policies for reducing the negative impact of economic crisis on firms, and therefore on the economy and the society. Furthermore, because the big data sets of the public and the private sector that we can use for the above purposes usually include quite large numbers of variables, our methodology provides also guidance for an initial filtering of them and preselection of meaningful and reasonable independent variables to be processed by the above AI algorithm, based on sound theoretical foundations from previous management science and IS research. This enables avoiding the risks posed by the purely "big data-driven" research, lacking theoretical direction and foundation, which have been identified in previous relevant literature (Gonzalez-Bailon, 2013; Mazzocchi, 2015; Rai, 2016). The AI FS technique our methodology is based on provides serious advantages in comparison with other alternative techniques that can be used for the same purpose (such as various kinds of regression analysis): it can handle large numbers of features, even highly correlated ones, without the performance and reliability deterioration exhibited by the other alternative techniques. A first application of this methodology has been presented, which leads to interesting findings and insights, and provides a first validation of this methodology.

Our public sector data analytics methodology is of wider applicability, as it can be used after some minor adaptations in many thematic domains of government intervention for identifying characteristics of firms that affect positively or negatively the adoption of various positive behaviors/activities (such as export, expansion to other countries and adoption of new technologies) or negative ones (such as reduction of personnel employment and disinvestment), by combining and exploiting multiple sources of public and private sector data. This can be quite useful for the design of effective design of public policies for promoting various positive firms' behaviors/activities and reducing negative ones.

Our research has interesting implications for research and practice. With respect to research, it creates new knowledge in two emerging, highly important for the society and the economy, digital government research domains:

- public sector data analytics, focusing on policy analytics, by developing an approach for exploiting big public and private sector data to support policymaking concerning one of the most serious problems of the market-based economies (though our approach is of much more general applicability in a wide range of thematic domains of government intervention); and
- AI exploitation in government, by developing an approach for the exploitation of an AI "all-relevant" FS algorithm, which has not been exploited in the past for public sector data analytics, to support public policymaking on such a critical economic challenge.

With respect to practice, it provides support to government agencies for designing policies for reducing the negative impact on firms of one of the most important problems of market-based economies: the economic crises. It enables a better understanding of the kinds of firms that are more strongly hit by the crisis, which is quite useful for the design of effective and focused public policies for supporting them; and at the same time, it reveals firms' characteristics, resources, capabilities, practices, etc. that enhance their ability to cope with economic crisis, so that appropriate policies for promoting them through educational and support activities can be designed and implemented. Furthermore, as mentioned above, the proposed methodology is of wider applicability for supporting policymaking in many other thematic domains of government intervention, aiming at the promotion of various kinds of firms' positive behaviors/activities, and the reduction of negative ones.

However, further research is required in the above directions:

- further application of our methodology in various national contexts, at both national and sectoral level, for various kinds of economic crises, using a wider range of potential independent variables;

- improvement of the methodology based on the results of such applications;
- extension of the methodology with additional AI techniques, such as clustering ones;
- application of the methodology in other thematic domains of government intervention for supporting the design of various kinds of policies; and
- use of the above ten "relevant" variables we have identified for the construction of a prediction model of firms' resilience to crisis (which can be useful in future economic crises); also it would be useful to examine to what extent the use of all the initial 61 independent variables – or a smaller subset of them identified using a "minimal optimal" FS algorithm – would improve prediction accuracy.

References

Alelyani, S., Wang, L. and Liu, H. (2011), "The effect of the characteristics of the dataset on the selection stability", in *Proceedings of the 23rd IEEE International Conference on Tools with Artificial Intelligence*.

Allen, R.E. (2016), *Financial Crises and Recession in the Global Economy – Fourth Edition*, Edward Elgar Publications, Cheltenham.

Aydiner, A.S., Tatoglu, E., Bayraktar, E., Zaim, S. and Delen, D. (2019), "Business analytics and firm performance: the mediating role of business process performance", *Journal of Business Research*, Vol. 96, pp. 228-237.

Baer, D., Wimmer, M., Glova, J., Papazafeiropoulou, A. and Brooks, L. (2015), "Analysis of policy cases in the field of energy policy", in Janssen, M., Wimmer, N., Deljoo, A. (Eds), *Policy Practice and Digital Science – Integrating Complex Systems, Social Simulation and Public Administration in Policy Research'*, pp. 355-378. Springer International Publishing AG – Public Administration and Information Technology, Switzerland.

Bean, R. (2018), "How big data and AI are driving business innovation in 2018. MIT Sloan management review, February 2018".

Camacho-Collados, M. and Liberatore, F. (2015), "A decision support system for predictive police patrolling", *Decision Support Systems*, Vol. 75, pp. 25-37.

Camisón, C. and Forés, B. (2010), "Absorptive capacity new insights for its conceptualization and measurement", *Journal of Business Research*, Vol. 63 No. 7, pp. 707-715.

Cateni, S., Vannucci, M., Vannocci, M. and Colla, V. (2013), "Variable selection and feature extraction through artificial intelligence techniques", in Leandro Valim De Freitas and Ana Paula Barbosa Rodrigues De Freitas (Eds), *Multivariate Analysis in Management, Engineering and the Sciences*, IntechOpen, pp. 103-118.

Chandler, D., Levitt, S.D. and List, J.A. (2011), "Predicting and preventing shootings among at-risk youth", *American Economic Review*, Vol. 101 No. 3, pp. 288-292.

Chang, Y.W. and Tsai, C.Y. (2017), "Apply deep learning neural network to forecast number of tourists", *Proceedings of IEEE 31st International Conference on Advanced Information Networking and Applications (AINA), Taipei, Taiwan*.

Chen, Y., Wang, Y., Nevo, S., Jin, J., Wang, L. and Chow, W.S. (2014), "IT capability and organizational performance: the roles of business process agility and environmental factors", *European Journal of Information Systems*, Vol. 23 No. 3, pp. 326-342.

Cohen, W.M. and Levinthal, D.A. (1990), "Absorptive capacity: a new perspective on learning and innovation", *Administrative Science Quarterly*, Vol. 35 No. 1, pp. 128-152.

Craglia, M., Annoni, A., Benczur, P., Bertoldi, P., Delipetrev, P., De Prato, G., Feijoo, C., Fernandez Macias, E., Gomez, E., Iglesias, M. and Junklewitz, H. (2018), *Artificial Intelligence – A European Perspective, EUR 29425 EN*, EU Publications Office, Luxembourg.

Daniell, K.A., Morton, A. and Rios Insua, D. (2016), "Policy analysis and policy analytics", *Annals of Operations Research*, Vol. 236 No. 1, pp. 1-13.

Davenport, T. (Ed.) (2013), "Enterprise analytics – optimize performance, process and decisions through big data", *International Institute for Analytics*, – Pearson Education, Inc., Upper Saddle River, NJ.

Davenport, T. and Harris, J. (2017), "Competing on analytics – the new science of winning", *Revised Edn*, Harvard Business Review Press, Boston, MA.

De Marchi, G., Lucertini, G. and Tsoukias, A. (2016), "From evidence-based policy making to policy analytics", *Annals of Operations Research*, Vol. 236 No. 1, pp. 15-38.

DeSousa, W.G., DeMelo, E.R.P., De Souza Bermejo, P.H., Sous Farias, R.A. and Gomes, A.O. (2019), "How and where is artificial intelligence in the public sector going? A literature review and research agenda", *Government Information Quarterly (in-Press)*., Vol. 36 No. 4,

DeSouza, K.C. (2018), *Delivering Artificial Intelligence in Government: Challenges and Opportunities*, IBM Center for The Business of Government, Washington, DC.

DeSouza, K.C., Krishnamurthy, R. and Dawson, G.S. (2017), *Learning from Public Sector Experimentation with Artificial Intelligence*, Brookings, Washington, DC.

Diebold, F.X. and Rudebusch, G.D. (1999), *Business Cycles: Durations, Dynamics, and Forecasting*, Princeton University Press, NJ.

Drnevich, P.L. and Kriauciunas, A.P. (2011), "Clarifying the conditions and limits of the contributions of ordinary and dynamic capabilities to relative firm performance", *Strategic Management Journal*, Vol. 32 No. 3, pp. 254-279.

Duan, Y., Edwards, J.S. and Dwivedi, Y.K. (2019), "Artificial intelligence for decision making in the era of big data – evolution, challenges and research agenda", *International Journal of Information Management*, Vol. 48, pp. 63-71.

Eggers, W.D., Schatsky, D. and Viechnicki, P. (2017), *AI-Augmented Government. Using Cognitive Technologies to Redesign Public Sector Work*, Deloitte University Press.

Ekstrom, J.A., Lau, G.T. and Law, K.H. (2018), "Policy analytics tools to identify gaps in environmental governance", In: Gil-Garcia, J.R., Pardo, T., Luna-Reyes, L.F. (Eds), *Policy Analytics, Modeling, and Informatics – Innovative Tools for Solving Complex Social Problems*, pp. 289-314. Springer International Publishing AG – Public Administration and Information Technology, Switzerland.

Fernandes, E., Holanda, M., Victorino, M., Borges, V., Carvalho, R. and Van Erven, G. (2019), "Educational data mining: predictive analysis of academic performance of public school students in the capital of Brazil", *Journal of Business Research*, Vol. 94, pp. 335-343.

Gil-Garcia, J.R., Pardo, T.A. and Luna-Reyes, L.F. (2018), "Policy analytics: definitions, components, methods, and illustrative examples", In: Gil-Garcia, J.R., Pardo, T.A., Luna-Reyes, L.F. (Eds), *Policy Analytics, Modelling, and Informatics*, pp. 1-16. Springer International Publishing AG – Public Administration and Information Technology, Switzerland.

Gonzalez-Bailon, S. (2013), "Social science in the era of big data", *Policy and Internet*, Vol. 5 No. 2, pp. 147-160.

Gourinchas, P.O., Philippon, T. and Vayianos, D. (2016), "The analytics of the Greek crisis", *Nber Macroeconomics Annual*, Vol. 31 No. 1, pp. 1-81.

Hiltz, S.R., Diaz, P. and Mark, G. (2011), "Introduction: social media and collaborative systems for crisis management", *ACM Transactions on Computer-Human Interaction*, Vol. 18 No. 4, pp. 1-6.

Izsak, K., Markianidou, P., Lukach, R. and Wastyn, A. (2013), *The Impact of the Crisis on Research and Innovation Policies*, European Commission, DG Research, Brussels.

Janowski, T. (2015), "Digital government evolution: from transformation to context-ualization", *Government Information Quarterly*, Vol. 32 No. 3, pp. 221-236.

Janssen, M. and Wimmer, M. (2015), "Introduction to policy-making in the digital age", in Janssen, M., Wimmer, M., Deljoo, A. (Eds), *Policy Practice and Digital Science – Integrating Complex Systems, Social Simulation and Public Administration in Policy Research*, 1-14. Springer International Publishing AG – Public Administration and Information Technology, Switzerland.

Johnson, G., Whittington, R., Scholes, K., Angwin, D. and Regner, P. (2017), *Exploring Corporate Strategy – Text and Cases*, 11th ed., Pearson Education Limited.

Jones, G.R. (2013), *Organizational Theory, Design, and Change*, 7th ed., Pearson Education Limited.

Keeley, B. and Love, P. (2010), "From crisis to recovery – the causes", *Course and Consequences of the Great Recession*, OECD Publishing, Paris.

Knoop, T.A. (2015), *Recessions and Depressions: Understanding Business Cycles*, 2nd ed., Praeger Santa Barbara, CA.

Kouziokas, G.N. (2017), "The application of artificial intelligence in public administration for forecasting high crime risk transportation areas in urban environment", *Transportation Research Procedia*, Vol. 24, pp. 467-473.

Ku, C.H. and Leroy, G. (2014), "A decision support system: automated crime report analysis and classification for e-government", *Government Information Quarterly*, Vol. 31 No. 4, pp. 534-544.

Kursa, M.B. and Rudnicki, W.R. (2010), "Feature selection with the Boruta package", *Journal of Statistical Software*, Vol. 36 No. 11, pp. 1-13.

Kursa, M.B., Jankowski, A. and Rudnicki, W.R. (2010), "Boruta – a system for feature selection", *Fundamenta Informaticae*, Vol. 101 No. 4, pp. 271-285.

Lachana, Z., Alexopoulos, C., Loukis, E. and Charalabidis, Y. (2018), "Identifying the different generations of e-government – an analysis framework", *Proceedings of 12th Mediterranean Conference on Information Systems (MCIS) 2018, Corfu, Greece*.

Leavitt, H.J. (1964), "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.

Lu, Y. and Ramamurthy, K.R. (2011), "Understanding the link between information technology capability and organizational agility: an empirical examination", *MIS Quarterly*, Vol. 35 No. 4, pp. 931-954.

Makridakis, S. (2017), "The forthcoming artificial intelligence (AI) revolution: its impact on society and firms", *Futures*, Vol. 90, pp. 46-60.

Mazzocchi, F. (2015), "Could big data be the end of theory in science? – a few remarks on the epistemology of data-driven science", *EMBO Reports*, Vol. 16 No. 10, pp. 1250-1255.

OECD (2019), *Artificial Intelligence in Society*, OECD Publishing, Paris.

Park, C.H. and Johnston, E.W. (2018), "An event-driven lens for bridging formal organizations and informal online participation: how policy informatics enables just-in-Time responses to crises", in Gil-Garcia, J.R., Pardo, T., Luna-Reyes, L.F. (Eds), *Policy Analytics, Modeling, and Informatics – Innovative Tools for Solving Complex Social Problems*, pp. 343-336. Springer International Publishing AG – Public Administration and Information Technology, Switzerland.

Rai, A. (2016), "Synergies between big data and theory", *MIS Quarterly*, Vol. 40 No. 2, pp. iii-iix.

Ransbotham, S., Khodabandeh, S., Fehling, R., LaFountain, B. and Kiron, D. (2019), *Winning with AI*, MIT Sloan Management Review and Boston Consulting Group, Boston, MA.

Rockoff, J.E., Jacob, B.A., Kane, T.J. and Staiger, D.O. (2010), "Can you recognize an effective teacher when you recruit one?", *Education Finance and Policy*, Vol. 6 No. 1, pp. 43-74.

Seddon, P., Constantinidis, D., Tamm, T. and Dod, H. (2017), "How does business analytics contribute to business value", *Information Systems Journal*, Vol. 27 No. 3, pp. 237-269.

Sethi, P. and Jain, M. (2010), "A comparative feature selection approach for the prediction of healthcare coverage", *Proceedings of 4th International Conference on Information Systems, Technology and Management (ICISTM 2010), Bangkok, Thailand*.

Sherehiy, B., Karwowski, W. and Layer, J.K. (2007), “A review of enterprise agility: concepts, frameworks, and attributes”, *International Journal of Industrial Ergonomics*, Vol. 37 No. 5, pp. 445-460.

Sun, T.Q. and Medaglia, R. (2019), “Mapping the challenges of artificial intelligence in the public sector: evidence from public healthcare”, *Government Information Quarterly*, Vol. 36 No. 2, pp. 368-383.

Tang, J., Alelyani, S. and Liu, H. (2014), “Feature selection for classification: a review”, *Data Classification: Algorithms and Applications*, pp. 37-64.

Teece, D. (2007), “Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance”, *Strategic Management Journal*, Vol. 28 No. 13, pp. 1319-1350.

Tsoukias, A., Montibeller, G., Lucertini, G. and Belton, V. (2013), “Policy analytics: an agenda for re-search and practice”, *EURO Journal on Decision Processes*, Vol. 1 No. 1-2, pp. 115-134.

Vaan den Braak, S., (2018), and Choenni, S. “Development and use of data-centric information systems to support policymakers: applied to criminal justice systems”, in Gil-Garcia, J.R., Pardo, T., Luna-Reyes, L.F. (Eds), *Policy Analytics, Modeling, and Informatics – Innovative Tools for Solving Complex Social Problems*, pp. 343-361, Springer International Publishing AG – Public Administration and Information Technology, Switzerland.

Van Dijk, V., Kalidien, S. and Choenni, S. (2018), “Smart monitoring of the criminal justice system”, *Government Information Quarterly*, Vol. 35 No. 4, pp. 24-32.

Appendix

	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?
STRAT_DIF	To what extent does your business strategy include high quality of products/services in comparison with the competition?
STRAT_INNOV	To what extent does your business strategy include introduction of new products/ services (with significant innovations)?
INNOV_PRS	Over the past three years did your firm introduce product innovations (=new or significantly improved products)?
INNOV_PROC	Over the past three years did your firm introduce process innovations (=new or significantly improved processes)?
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?
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 past three years?
INN_PRSD	Did you introduce methods/process innovation in the goods production or services' delivery processes in the past three years?
INN_SSWM	Did you introduce methods/process innovations in the sales, shipment or warehouse management processes?
INN_SUPP	Did you introduce methods/process innovations in the support processes (e.g. in the equipment maintenance ones)?
R&D	Did your firm conduct R&D (Research and Development) in the past three years?
EXP_P	Percentage of exports in firm's sales revenue in 2014
	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)
EMPL_TERT	Percentage of tertiary education graduates in the personnel of your firm
EMPL_VOCT	Percentage of vocational/technical education graduates in the personnel of your firm
EMPL_HIGH	Percentage of high school graduates in the personnel of your firm
EMPL_ELEM	Percentage of elementary school graduates in the personnel of your firm
EMPL_COM	What percentage of the employees of your firm use computer in their work (e.g. PC, terminal or laptop)?
EMPL_INTRA	What percentage of the employees of your firm uses the intranet (internal network) of the firm in their work?
EMPL_INTER	What percentage of the employees of your firm uses internet in their work?
EMPL_ICT	Percentage of qualified ICT personnel in the workforce of your firm
	Independent variables – technology
D_ERP	To what extent are Enterprise Resource Planning (ERP) systems used in your firm?
D_CRM	To what extent are Customer Relationship Management (CRM) systems used in your firm?
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 inventory levels, orders, production, shipments and invoices) used in your firm?
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 decision-making – this can be either a separate software, or a module of an ERP or CRM system) used in your firm?

(*continued*)

Table A1. Definitions/questions of the dependent and independent variables

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?
E-SAL	Do you conduct online sales of products/services through the internet?
SM_SPRO	To what extent do you use social media for sales promotion?
SM_OPCO	To what extent do you use social media to collect customers' opinions, comments and complaints about your products or services?
SM_IMPS	To what extent do you use social media to collect ideas for improvements or innovations in your product or services?
SM_PERS	To what extent do you use social media to search for and find personnel?
SM_INTC	To what extent do you use social media to support the internal exchange of information and co-operation among the employees of your firm?
SM_IPAR	To what extent do you use social media to support the external exchange of information and co-operation with other firms (e.g. partners, suppliers, customers, etc.)?
CLOUD	Do you use cloud computing?
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 application development tools/environments/ languages through the internet)?
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)?
	Independent variables – structure
ORG	Over the past three years did your firm use organic structural forms of work organization (such as teamwork and job rotation)?
	Independent variables – ordinary (ICT) capabilities
ICT_STRPLAL	To what extent does your firm have ICT strategies and plans, which are connected with the overall strategies and plans of the firm (ICT business strategic alignment)?
ICT_BUSC	To what extent in your firm there is good cooperation, mutual understanding and trust between the ICT unit/personnel and the business units/personnel who use ICT for their work?
ICT_VENDC	To what extent does your firm have good cooperation, trust and exchange of information with its ICT vendors (of hardware/software/networks), as well as support from them for solving all ICT-related problems?
ICT_ADEV	To what extent does your firm have capability of rapid internal development of new software applications (by the ICT personnel of your firm) to meet new needs?
ICT_AMOD	To what extent does your firm have capability of rapid internal implementation of various modifications in your application software (by the ICT personnel of your firm) to meet requirements' changes?
ICT_AINT	To what extent does your organization have capability of rapid internal implementation of various interconnections/integrations of existing applications (by the ICT personnel of your firm)?
	Independent variables – dynamic capabilities
ACAP_ACQ	To what extent does your firm have capability of recognizing, identifying and acquiring relevant data useful for the firm knowledge from its external environment (suppliers, partners, research centers, universities)?
ACAP_DIS	To what extent does your firm have internal practices and procedures for the internal dissemination and analysis of this external knowledge?

(*continued*)

Table A1.

ACAP_INT	To what extent does your firm can absorb/assimilate this external knowledge and integrate it into the existing knowledge base of the firm?
ACAP_EXP	To what extent does your firm have capability to use/exploit this knowledge for process, products and services innovations?
AG_NPS	To what extent can your firm easily and quickly react to the introduction of new products and services by competitors?
AG_NPR	To what extent can your firm easily and quickly introduce new pricing policies in response to competitors' prices changes?
AG_CDE	To what extent can your firm easily and quickly respond to changes in the demand for your products or services?
AG_CUST	To what extent can your firm easily and quickly adapt/customize its products and services to meet specific customers' needs?
AG_EXPM	To what extent can your firm easily and quickly expand into new markets within the country or abroad?
AG_CSUP	To what extent can your firm easily and quickly change suppliers to achieve supply costs' reductions and quality improvements?
	Independent variables – external environment
N_COMP	Number of your main competitors
INT_PCOM	How intensive is the competition you face from other firms with respect to price (price competition)?
INT_NPCOM	How intensive is the competition you face from other firms with respect to other competition dimensions (non-price competition), such as quality of products/services, customization of them and services' provision?
DYN_PRS	To what extent in the external environment/markets of your firm products and services quickly become obsolete?
DYN_TECH	To what extent in the external environment of your firm there are often changes in the technologies of your products and services?
DYN_PRS	To what extent in the external environment of your firm there are unpredictable changes in the demand for your products and services?
DYN_COMP	To what extent in the external environment of your firm there are unpredictable movements of the competitors?
	Independent variables – general
SECT	Firm's sector

Table A1.

Corresponding author
Euripidis N. Loukis can be contacted at: eloukis@aegean.gr