Corresponding Author:

Emanuela Sirtori, CSIL Centre for Industrial Studies, Corso Monforte 15, 20122 Milan, Italy

Email: sirtori@csilmilano.com

Evaluating Business Support Measures: the Bayesian Network Approach

Francesco Giffoni

CSIL Centre for Industrial Studies, Milan, Italy

Silvia Salini

Department of Economics, Management and Quantitative Methods, University of Milan, Italy

Emanuela Sirtori

CSIL Centre for Industrial Studies, Milan, Italy

Abstract

Traditional methods to the evaluation of business support consider the firm as a 'black box' and the main interest is to see to what extent a policy has produced the intended effects. The causal mechanisms explaining how certain effects are generated are not discovered. In this paper we show the applicability of Bayesian Network Analysis in combination with theory-based evaluation as a new mixed-method approach to reveal the mechanisms, both expected and unknown, which explain the changes in firm's behaviour and economic performance due to public support. By combining graphical map analysis with statistical analysis, Bayesian Network Analysis reveals the interdependences between different drivers of change in firms so as to either confirm, deny or better qualify the theory of change of the policy.

Keywords:

Bayesian Network Analysis, Theory-based evaluation, Mixed method, Business support policy, Behavioural change

1. Introduction

Proving the effectiveness and value for money of the EU Cohesion Policy targeting enterprises development and innovation is notoriously a challenging task (Bachtler and Wren, 2007). Over the years, the European Commission and national and regional Managing Authorities have put increasing emphasis and efforts to evaluate the impact of Cohesion Policy programmes and, also, understand the mechanisms through which EU funded programmes produce their effects on supported enterprises. The set of ex-post evaluations of the European Regional Development Fund and Cohesion Fund of the 2007-2013 period included three work packages dealing with the evaluation of enterprise support, namely support to Small-Medium Enterprises (SMEs), Financial Instruments for enterprises, and support to Large Enterprises. Each of them used the Theory-Based Evaluation (TBE) approach to investigate with a fine-grained analysis the different steps and underlying assumptions that lead from inputs to intermediate and ultimate outcomes.[1]

TBE is about discovering the 'theory of change' underlying the policy intervention (Weiss, 1997; Stame, 2004; Howard, 2009; Astbury and Leeuw, 2010). Theory-based approaches are used to map the causal process which explains the contribution of a given policy intervention to the achievement of outcomes. The main objective of theory-based methods is not to estimate quantitatively how much of the effects is due to the intervention. TBE is rather focused on explaining *how* a policy contributes to produce certain effects, according to mechanisms of change and assumptions which are reflected in the intervention's theory of change. In the field of business support, for instance, Buckley (2016) has recently used Contribution Analysis to evaluate small and medium enterprise support policy and produced findings on the contribution of the multiple factors influencing firm performance.

In spite of the suitability of TBE to understand how and under what conditions public policies produce their effects and examining causality in complex interventions, it is *per se* neutral from the perspective of statistical inference. Differently from econometric models, randomised control trials or quasi-experimental studies, TBE does not aim to produce quantitative and statistically significant estimates of some parameters of interest, e.g. policy impact variables. Documentary and literature review, ethnographic observation and direct consultation are the main tools used by TBE to derive and test the theory of change of policy interventions. Empirical data can be collected and case studies implemented to assess whether the initial theory of change is confirmed or not. However, theory-based approaches do not necessarily provide quantitative measures of the impact of the policy intervention or other possible influencing factors.

[1] Specifically, the Work Package focused on SME support policy (WP2) followed the Realist Evaluation approach (Pawson and Tilley, 1997), the evaluation of programmes supporting large enterprises (WP4) applied the Contribution Analysis paradigm (Mayne, 2001; Delahais and Toulemonde, 2012) and the evaluation of financial instruments (WP3) was inspired by the Theory of Change approach (Weiss, 1995).

Mixed-methods approaches, combining qualitative and quantitative methods, could be adopted in the framework of theory-driven evaluations in order to compensate for each method's limitations (Chen, 1997). White (2009, p.15) recognises that, especially in the development field, "qualitative approaches have dominated evaluation until very recently, so a major step toward mixed methods is in fact the increased use of rigorous quantitative methods in qualitative studies". When data sets are available, evaluators should take advantage of them to quantitatively assess the impact of the intervention. In parallel, TBE could provide the understanding of how the intervention actually works and minimise the risk that the impact study is conducted by relying only on data and with no sufficient exposure to the intervention (White, 2009).

In this paper, we show how Bayesian Network Analysis (BNA) can be combined with TBE to test and confirm, quantitatively, the theory of intervention. The Bayesian method allows conclusions on the different changes produced by the policy intervention to be derived in terms of conditional probabilities (i.e. probabilities conditional on the observed data). Our ultimate goal is to strengthen the use of mixed-methods approach in the evaluation of business support policies.

Specifically, this paper presents how BNA has been used in the framework of the ex-post evaluation of Cohesion Policy programmes 2007-2013 in order to empirically test the theory of three policy instruments supporting SMEs development and innovation: i) the Polish "Technological Credit" providing a combination of grants and bank credit to support technological innovation in SMEs; ii) a measure providing grants for R&D projects conducted by enterprises in the Spanish region of Castilla y Leon, and iii) the "Title II" measure supporting investment projects implemented by micro and small enterprises in the Apulia region, in Italy. The three instruments are similar in their ultimate aim of helping SMEs overcome the market failures they face, as well as overcoming the effects of the economic crisis. However they follow different logics of intervention, as they target different enterprises (more innovative in Poland and Spain, more traditional and micro-size in Italy) and have different implementation modalities, specific objectives, eligible categories of expenditure, average value of supported projects, aid intensity and are embedded in different socio-economic contexts.[2]

Through these real case examples, the paper shows how BNA can complement a purely qualitative TBE, in order to:

- confirm the validity of the theory and disentangle the chain of effects influencing the firms' economic performance;
- capture how the policy intervention can stimulate a variety of other behavioural changes in SMEs;

[2] For more details on the logic of the three policy instruments, please refer to European Commission (2015).

- investigate the multiple variables determining the success of R&D projects, such as the degree of collaboration between firms and universities, and the intrinsic level of risk of R&D activity;
- perform scenario analysis to simulate the relative change in variables of interest upon changes in other related variables of the network.

Our objective is to explain the methodological approach adopted, show the value added of combining BNA and TBE in the context of business support evaluation, encourage its use in the same field as well as experimentation in other intervention areas.

The rest of the paper proceeds as follows. Section 2 presents the specificities of BNA, avoiding to use a too technical jargon; Section 3 highlights the advantages of using BNA in evaluating business support interventions. Section 4 presents concrete cases where BNA has been applied and discusses the main findings. Section 5 concludes by summarising the results, but also showing limitations and suggesting implications for future research and evaluation practice.

2. What Bayesian Network Analysis is

BNA is a statistical tool that estimates and visualises the conditional independence and dependence relationships among variables. A Bayesian Network (BN) is a probabilistic graphical model defined by a set of random variables (nodes) and directed edges connecting the variables and forming a directed acyclic graph (DAG) (Jensen and Nielsen, 2007). An edge from node *X* to node *Y* represents a statistical dependence between the two variables and indicates that the variables *X* is correlated with *Y*. Node *X* is then referred to as a 'parent' of *Y*, and conversely *Y* is referred to as 'child' or 'descendent' of *X* (Jensen, 1996; Nadkarni S and Shenoy, 2001: Kenett and Salini, 2011a, 2011b). Some software packages[3] indicate the strength of the relationship between the variables through the thickness of the arrows: the ticker the arrow, the stronger the dependence between the variables.

The graphical structure of the network is based on the estimation of conditional probability distributions of each variable entering the network, expressing the probability that a child variable takes on a certain value for each combination of values of its parents. Its graphical structure makes a BN similar to a cause-and-effect diagram and suitable to explain dependences between variables.

The computation of the conditional probability distributions requires the application of some type of data-driven learning algorithms. We estimated the BNs presented in this article by applying the Bayesian Search algorithm, which combines analysts' knowledge about the theory of intervention and quantitative statistical data (Heckerman et al., 1994). The algorithm produces an directed acyclic graph that gives the maximum score

[3] For instance the open-source software GeNIe, developed by the University of Pittsburgh. This is the software we used for our analysis.

following a hill climbing procedure (guided by a scoring heuristic) with random restarts. In other terms, the algorithm estimates the probability distributions of variables by relying on available data, and tries to find the best graphical structure that would explain dependencies between variables.

The name Bayesian Networks might be misleading and it should not be confused with the Bayesian approach. The latter is based on the Bayesian Confidence Updating method, where some hypotheses on the marginal distributions of variables are formulated by the evaluator and updated using the data through the Bayes formula.[4] Conversely, the use of BN models does not necessarily imply a commitment to Bayesian statistics and they do not typically use a full Bayesian treatment in the Bayesian statistical sense (i.e. hyper parameters and learning case by case). They do make use of Bayes Theorem during inference, and typically use priors during parameter learning from data. In fact, with BNA it is common, and it is also our case, to use frequentists methods to estimate the parameters of the conditional probability distributions (Murphy, 2001; Ben-Gal 2007). In our application, the evaluator supposes a theory, but it does not need to define initial confidence on the possibilities, i.e. *a priori* distribution on the parameter of the distribution. We use here Bayesian Networks as a confirmative approach: the network structure and the cumulative probability distribution directly derived from the data are compared with the supposed theory of intervention.

Moreover, while Bayesian Updating is usually bivariate, with fixed target variables, and the effect of the levels of the explanatory variable is tested using probabilities, the BN is multivariate and entails the estimation of *cumulative* probability distributions. The cumulative probability is a powerful tool to do evidence propagation scenario by changing the value of some variables jointly. Besides, with BNA we can represent the complex structure of dependence between variables in a simply and intuitively graphical way.

While the use of BNA is well-established in fields such as medicine, computer science, and risk analysis (Nadkarni and Shenoy, ibidem; Kenett, 2012; Horny, 2014), there is limited application in socio-economic disciplines and in the evaluation of public policies. We argue that BNA can be effectively combined with TBE methods. As Hawkins (2016) points out with reference to Realist Evaluation, but we could generalise to all TBE approaches, TBE is strong on theory and explanation but it "lack[s] adequate tests or means of validating theory". We believe that the value of TBE can be enhanced

---

[4] Examples of applications of the Bayesian approach to the evaluation of public programs are Befani and Stedman-Bryce (2017) in the field of health care policy, Busetti and Dente (2017) who used Bayesian Updating for impact evaluation of the EXPO Milano 2015 event, and Schmitt and Beach (2015), who used the Bayesian logic to show how multiple sources of evidence can be utilized to update a priori confidence in the presence/absence of parts of a causal mechanism. As far as SMEs are concerned, Majocchi et al. (2015) apply the Bayesian Analysis to test the relationship between SMEs internationalisation and their performance.

by incorporating BNA in the evaluation framework, in order to both visually examine and quantify the dependences of variables associated with a theory of change.

BNA should not be regarded as substitute of other quantitative impact evaluation methods, for instance using regression models or counterfactual techniques. In contrast, we believe that the role for BNA could be to empirically test specific mechanisms of interventions, or whole theories of change of policies which are reconstructed through a TBE approach. A similar approach was taken by Ranmuthugala et al. (2011), when using Social Network Analysis in combination with Realist Evaluation to examine the relationship between mechanisms and context factors which explain the effectiveness of a new healthcare practice. By allowing the visual examination of relationships between variables and their strengths, BNA can be a valuable tool to test the theory and answer questions such as "In light of the multiple factors influencing a result, has the intervention made a noticeable contribution to an observed outcome and in what way? Through what mechanisms?". In addition (and differently from Social Network Analysis), BNA permits to derive the "posterior" probability distribution of variables conditional on the evidence obtained from surveys, which is essential to perform scenario analyses (see below).

Since we mentioned alternative quantitative impact evaluation methods, it is worth to highlight that BNA is an approach inherently different from more traditional econometric models, and their respective results differ too. In particular, while traditional econometrics aims to estimate the parameter values of a set of individual independent variables that explain variation in the dependent variables, in BNA there is no need to distinguish between dependent variable, variables of interest and control variables, since the whole set of interdependences among variables are visualised in the DAG. Survey data are processed in order to express in probabilistic terms the (in)dependency relations among variables, allowing for multiple relations among them. For instance, if a regression may show that the SME economic performance is positively correlated with the volume of the public support received, the BNA might show instead that the economic performance is directly related to the type of investment made by the firm (e.g. acquisition of new production technologies rather than renewal of the company website), and/or other possible changes (e.g. improved the skills of employees, or increased popularity of the firm) which are in turn associated with the volume of public support. Hence, economic performance and public support are independent from each other in the network, once it is controlled for other variables. The multivariate nature of BN allows looking at and interpreting the combined behaviour of variables in the network.

Moreover, with BNA it is not necessary to determine ex ante the variables entering the model, but the applied learning algorithm "automatically" displays the most relevant arcs, i.e. relations between variables, discarding those for which no clear or strong influence over other variables is detected.

Nevertheless, traditional econometric analysis could be combined with BNA in two ways: by providing additional quantitative evidence on the policy impact through different models and estimation techniques; and by testing the robustness of the BNA results. By adopting a cross-check strategy, the policy analyst could use conventional methods such as regression models, to test the statistical significance of the correlation between selected variables entering the network.[5]

3. The value added of Bayesian Network Analysis

By estimating the conditional probability distribution among variables and arranging them in a DAG, BNs provide a straightforward statistical language to express relations between variables. Using BNA to complement TBE has several advantages as compared to relying on purely qualitative TBE.

First, the network structure is intuitively appealing and convenient for the representation of theories of intervention and to test such theories against empirical data. After reconstructing the theory of intervention through documentary review and interviews, the concepts entering the theory can be operationalised in variables and data can be collected on such variables. The BNA can then be employed to find interdependences among variables. Moreover, the overall network structure can provide information on the mechanisms or trajectories underlying the theory of intervention. Interestingly, the DAG nodes can represent not only random variables but also hypotheses, beliefs, and latent variables which emerge from the theory.

Second, BNA can help overcome the problem arising when a policy change is associated with multiple theories of change (Mackenzie and Blamey, 2005; Weiss, 1997). In this case, it is particularly important to test each possible theory against the evidence in order to identify which theory better reflects reality. Though BNA, different theories of change can be compared by identifying which relationship between variables predicted by theory are confirmed, as they are visible in the directed acyclical graph ('*true positive*'), which ones are not confirmed by data although being predicted by the theory ('*false negative'*), and which other links are found in spite of having not been predicted by the theory ('*false positive'*).

Third, since BNA allows for the simultaneous inclusion of several control variables defined at a higher level with respect to the unit of analysis (e.g. variables related to the regions where firms are localised), the influence of context factors can be taken into explicit account in the network. In principle, this could make BNA relevant for testing context-mechanism-outcomes configurations defined by the Realist Evaluation.

Fourth, the BN can be exploited to perform 'what-if' analysis, i.e. observing how changes in a variable entering the network affects (positively or negative and by how much), variables which are directly linked to it. Under the assumption that the

[5] We used traditional econometrics in this second way.

relationships between the variables in the network are empirically stable,[6] BNA allows to perform scenarios, which are useful, for instance, to assess how different policy options, or context conditions, directly influence the outcomes. Specifically, BNA can be used for two types of predictive support: prognostic support and diagnostic support (Kenett and Salini, ibidem). In the prognostic case, also referred to as 'top-down reasoning', a certain variable is fixed at specific values to predict the pattern of its children variables; the diagnostic scenario - or 'bottom-up reasoning' - works in the opposite direction by looking at the respective changes in the parent variables. Performing a 'what-if' analysis enables to build hypothetical worlds that the evaluator can then query and navigate. In this sense, Bayesian networks are seen as *oracles of intervention* and as a powerful technique to support strategic decisions (Glymour and Cooper, 1999; Pearl, 2000; Spirtes et al., 2001).

Finally, the hierarchical arrangement of variables (A is linked to B, which is linked to C, which is linked to D etc.) can be interpreted as a conjecture of causality between these variables (Pearl, 2000). BNA can help test in probabilistic terms the causality chains going from some input variables to outcomes. The network analysis automatically links the nodes by arrows, but the direction of the arrows needs to be validated by the policy analyst on the basis of prior knowledge on the variables, the policy instrument's theory and the relevant literature. These can provide background information to correctly interpret the causal direction between variables. For instance, if the BNA finds that variable A 'Public support' is linked to variable B 'Purchase of new technology with public aid', and variable B is linked to variable C 'Enterprise sector of activity', it is clear that the causality relation goes from A to B (the public support determines the purchase of the new technology) and from C to B (it is the sector that influences the decision to purchase the technology, and not the other way round). In those cases when no definite causal relation is known or can be assumed, we allowed these variables to be connected to each other, but with no explicit and directed relation between them.

4. BNA applied to the evaluation of business support

This section provides a set of illustrative examples of how BNs can be used in real-world evaluations to test the theory of different policy instruments in the field of business support. The examples are drawn from the ex-post evaluation of programmes cofounded by the European Regional Development Fund (ERDF) in support of SMEs development and innovation during the 2007-2013 programming period (European Commission, 2015). Specifically, three instruments were analysed in depth through theory-based impact evaluation approach:

[6] The network's stability can be determined through the "structure perturbation" method, consisting in checking the validity of the main relationships in the network by varying some part of it or marginalizing some variables (Peng and Ding, 2003; Daly et al., 2011).

- Support for technological innovation in Poland (“Technological Credit”, Measure 4.3, OP Innovative Economy 2007PL161PO001): almost 600 SMEs were provided with a grant in combination with a commercial loan to cover the purchase cost of new production technologies, both tangible (e.g. machineries, equipment, buildings) and intangible (e.g. patents and licenses). The instrument aimed to promote technological change in SMEs in order to increase their competitiveness, through a partial waiving of credit. The instrument was also intended to create awareness and experience in the delivery and use of financial instruments.
- Support for industrial R&D and innovation in Castile and León – Spain (Idea&Decide Programme, Axis I , OP Castilla y León 2007ES162PO009): 365 enterprises (out of which around 300 SMEs) benefitted from a grant co-financing the implementation of R&D projects, conducted individually or in collaboration with other firms, universities or research centres. The ultimate goal was to encourage the implementation of R&D projects particularly by SMEs.
- Aid to investment projects by micro and small enterprises in Apulia – Italy (“Title II”, Measure 6.1, OP Apulia 2007IT161PO010): a combination of interest subsidy and grant was awarded to around 3,300 micro and small enterprises The policy instrument intended to promote business modernisation and economic stabilisation of smaller firms operating in traditional sectors.

Data on 700 SMEs which benefitted from one of the above policy measures have been collected between July and September 2015 through three on-line surveys, designed so as to test the specific theory of each policy instrument. Survey responses were complemented with data on the volume of ERDF support provided, the types of projects funded and additional variables made available by the Managing Authorities on the beneficiary firms (such as their sector and size).[7] More information on the instruments, their theory of change, the sample of respondent firms, and details on all the analyses performed can be found in the European Commission report (2015) and in the related working papers by Sirtori et al. (2017) and Florio et al. (2017). Here our aim is not to discuss in detail the theory of each policy intervention and the results of our evaluation. Instead, we exploit those cases to illustrate in a selective way how BNA can be used to complement the theory-based impact evaluation and give examples of the findings it can bring. The networks obtained from the analysis of each policy instrument are presented at the end of the paper (Figure 1, 2 and 3).

4.1 Disentangling the chain of effects on firms’ performance

In general terms and in different ways, each of the three policy instruments expected to improve the economic performance and competitiveness of supported firms.

[7] When this information was not already available at the Managing Authority, it was collected through the survey.

Their underlying theory was tested by directly asking beneficiary SMEs their opinion about the extent to which the project, implemented thanks to EU support, allowed them to improve their economic performance. While the theory of change of the policy instruments, emerging from a review of policy and programming documents and interview to policy makers and implementing bodies, turned out to be quite simplistic, opinions on firms performance were asked for a wider and more detailed number of variables, including: increase in turnover, increase in export, increase in the number of clients, increase in the variety of clients, increase in entrepreneurs' personal income, reduction of production cost, and improved resilience to the crisis. Answers were given on an ordinary 5-point Likert-type scale (from 'nil or very limited effect', to 'very high effect').

We then processed survey responses with BNA in order to examine the following aspects:

- *Which specific types of investment projects were more likely related to specific effects on firms' performance*. For instance, the BNA of the Polish policy instrument suggests that, conditional to our data, the main mechanisms through which firms' performance improves is via the increase in export, achieved by accessing to a new foreign market, which is in turn associated with the purchase of more modern technology with the Technological Credit.
- *The way how different economic performance effects are linked to each other*. The Castilla y Leon instrument is associated with an increase in sales for most of beneficiary firms, which in turn is strongly linked with an increase in the types of clients. Firms, which declared having enlarged the variety of their clients, also declared that they managed to resist the effect of the economic crisis. This variable in turn is associated with positive expectations on future economic performance.
- *To what extent the initial theory managed to foresee and make explicit all the possible drivers of firm's economic performance*. Still in the case of Castilla y Leon, the thickness of the arrows reveals that a significant change directly activated by the R&D grant within the SME consists in the improvement of the enterprise's reputation. This change did not directly cause any observable economic effects for most of the beneficiaries, but it is however associated with positive expectations about the firms' future performance. The link between R&D project, the resulting improvement of company reputation and future improvement of economic results is an unexpected mechanisms of change emerging from the survey responses, not explicitly foreseen by the theory.

BNA turned out to be useful to test the causal mechanism going from inputs to outcomes, validate or reject causal links predicted by the theory, but also to better qualify the theory of change, disentangling the different factors determining which outcomes are achieved and how, and even findings unexpected mechanisms of change which were not explicitly predicted by the theory.

Even if, in our evaluation study, outcomes were defined according to firms' perspective, it is important to underline that the analysis in principle could accommodate also the inclusion of quantitative financial indicators, such as turnover change derived from the firms' balance sheets or other official monitoring indicators. Firm-level real financial data could be entered in the network and analysed together with other firm-related survey data.

4.2 Capturing behavioural change

The empirical literature suggests that some of the effects produced by policy interventions do not necessarily or immediately translate into improvement in economic performance, but could be in the form of changes in attitudes and approaches in doing business for beneficiary enterprises (Lam, 2005; Jensen et al., 2007; Amara et al., 2008; Damanpour and Aravind, 2012; Parrilli et al., 2016). We define these changes as behavioural changes, as opposed to economic results such as changes in turnover, profits, etc. As it emerges from the overall evaluation study (European Commission, 2016), behavioural changes could potentially shift SMEs from their initial trajectories and produce deep structural effects.

Each of the three instruments analysed were aimed to trigger some sorts of behavioural change in beneficiary SMEs, some of them more easily observable and measurable (such as employing a young researcher, or purchasing technologically more advanced equipment), others pertaining to the entrepreneur's mind set, for instance his/her willingness to take risks and innovate. The surveys to beneficiary enterprises were thus used to assess whether the expected behavioural changes actually took place. Through the BNA, we investigated the following aspects:

- *Whether the expected behavioural changes actually occurred and how* (theory testing). For instance, the Apulian instrument aimed to induce smaller size enterprises to invest in their business. 73% of the surveyed firms admitted that, after benefitting from Title II, started to consider the idea of implementing new investment projects never considered before. The BNA reveals that this variable is linked to other changes in the mind-set of entrepreneurs (e.g. higher importance attached to hiring more skilled employees) as well as to future intention to apply for public support. Instead, it is not directly related to any specific type of investment already made by firms and supported by Title II aid. As another example, the R&D grants in Castilla y Leon have encouraged beneficiary SMEs to increase their level of R&D expenditure and stimulate the implementation of increasingly collaborative and complex projects. The BNA confirms the theory and indicates that neither the size or sector of the firms influence to a significant extent the occurrence of these behavioural changes. Instead, other factors played a role, as described in Section 4.3.
- *The relationship between behavioural changes and economic performance variables*. Conditional to our dataset, behavioural change variables are usually not

directly related to economic performance outcomes already achieved, but they may be mediated by the types of investment project implemented. For instance, firms in Castilla y Leon which thanks to the R&D grants managed to increase the range of products offered or to enter new foreign markets are those which declared that will likely further increase their R&D expenditure in the future.

- *The relationship between behavioural changes and other possible outcomes*. We tested which variables would have influenced the firms' willingness to apply for other public funding in the future and we found that this is strongly connected with the firm's satisfaction about how the instrument was implemented (e.g. time required to receive the public contribution, or assistance received during the application phase), and not so much with changes in economic performance or other behavioural changes. This is particularly clear in the analysis of the Apulian and Polish instruments.

Our experience shows the value of using survey data to find mechanisms of change in firm's behaviour, and explore to what extent and how they are related to immediate economic performance outcomes or to firm's future intentions. BNA provided the opportunity to investigate these aspects and enrich the theory of change underpinning the policy instrument.

4.3 Revealing the role of collaboration and risk on R&D outcomes

In this section we dig a bit more in the analysis of one policy instrument, in order to show how BNA can help observe and interpret specific mechanisms of change. We focus on the Castilla y Leon policy instrument and analyse how collaboration with other firms or research centres influences the effectiveness of R&D projects implemented by SMEs, and how multiple risk factors associated with the implementation of R&D projects can determine the firms' behaviour and project results.

The theory expected the policy instrument to accompany SMEs along a process of behavioural change, by encouraging innovative SMEs to increase their level of R&D expenditure and stimulating the implementation of increasingly collaborative and challenging projects. Survey data confirm that, over the period from 2000 to 2013, more than 75% of beneficiary SMEs have increased the complexity and level of ambition of the R&D projects undertaken. The overall budget spent for R&D and the propensity to collaborate with other enterprises or with universities/research centres have increased for around half of the surveyed enterprises.

The BNA allows us to better understand the role played by R&D collaboration. Two main findings can be derived. First, the volume of R&D grant, and hence the size of the investment project,[8] increases if the firm had already previous collaboration experience with research centres or universities. This is also related to the level of education of the entrepreneur: the higher the level of education the more probable they

[8]Aid intensity is proportional to the investment volume.

are to have already collaborated with universities on other R&D projects. The same correlation cannot be found with reference to collaboration with other enterprises. Second, we found that about half of respondents believe that their level of collaboration with both universities ad enterprises might increase in the future and this goes along with expectations of an overall increase in R&D expenditure. This is especially true for firms which thanks to the R&D grant received during the 2003-2017 period have succeeded to enlarge their range of products and/or enter new foreign markets.

Against a vast strand of the literature arguing that collaboration in R&D projects is an important determinant of firm (especially SME) performance (Mowery, 1983; Cohen and Levinthal, 1990; Lipnack and Stamps, 2000; Agrawal, 2001; Bozeman and Gaughan, 2007; de Jong and Freel, 2009; Ebrahim et al, 2010; Cunningham and Gök, 2012), our analysis confirms this fact, but indicates that collaboration is not a direct determinant of economic performance outcomes. Instead, it affects the firm's probability to implement larger (i.e. more costly) investment projects. In turn, this is strongly related with the level of education of the firm's manager, suggesting that his/her skills have a role to play.

As far as risk factors are concerned, the network shows that the probability to increase sales and exports after the project implementation is strongly connected with the risk of not fully achieving the research objectives, the uncertainty about the potential for commercialization of the R&D outputs, uncertainty about future market conditions due to the economic crisis and the fear of having insufficient managerial experience and skills in the enterprise to achieve/maximise the project objectives. The combination of these risks is strongly related to the generation of economic outcomes, specifically the increase of sales and exports. Interestingly, the strongest risk factor affecting the economic results of beneficiary SMEs is not the market risk associated with the ongoing economic crisis (although it also played a role), but the possibility of not having sufficient skills and experience to complete the project.

The risks of not finding complementary external financial resources to start the project, or that the project would turn out to be more costly than forecasted do not directly affect the project results, but are linked to the volume of the public grant received. The BNA reveals that the volume of the R&D grant is larger for SMEs which perceived a higher finance risk.[9]

This information has practical relevance. It shows that, even if the policy intervention contributes at stimulating R&D expenditure and encouraging SMEs to embark in R&D projects, an additional component of risk related to market and management issues would remain to affect the probability of accomplishment of research results. If the goal is to stimulate the realisation of research and innovation activities by SMEs, this risk should be taken into account by the policy maker and explicitly incorporated into the theory of intervention.

[9] Another finding is that the volume of the grant, and then of the project implemented, is higher for firms born as spinoff from universities or other companies.

4.4 ‘ What-if’ scenario analysis

A peculiar feature of BNA is that it allows to perform scenario analysis. The values of certain variables can be manually changed in order to see the resulting change in the probability distribution function of its respective ascendant or descendent variables. In our context, we performed scenario analysis in order to identify the variables which were most strongly influencing other variables of the network and understand to what extent changes in one variable could change others.

In the Polish example, as already mentioned, we found that the variable that mediated the impact of the policy measure on the firms’ economic performance is the firm’s level of export at the moment when the public support was received. More specifically, the higher the export share in the year of application of the instrument, the more likely is for the firm to use the public support to implement an investment that would increase sales in the new foreign markets and, thereby, further increase its exports as a share of turnover. We conducted a scenario analysis on the Bayesian Network in order to understand:

- How the outcome variable (increase in the export share, defined as a discrete ordered distribution that can take six states: ‘I do not know’, ‘Not at all’, ‘Little’, ‘Enough’ ‘Appreciably’ or ‘Very much) is influenced by its ‘parent’ variables, i.e. the initial export share (ordered variable distributed over five classes, from null to more than 50% of export/sales ratio) and whether the investment helped the enterprise enter new foreign markets (binary variable);
- How variations in exports influence the effect on sales increase (ordered variables).

We hypothetically increased the share of SMEs which declare that they have increased exports ‘enough’, ‘appreciably’ and ‘very much’. Changing the distribution of this variable automatically provokes a change in the distribution of the other variables. The prior and posterior distributions of the variables according to the new scenario is illustrated in Figure 4.

The scenario analysis shows that:

- A higher increase in exports is associated with enterprises having a higher initial level of export share (more specifically, above 10%);
- In a scenario where all SMEs enjoy at least an ‘enough’ level of export increase thanks to the investment realised, the share of SMEs which declare that they have entered new foreign markets would increase (from 26% in the real sample to 36% in the hypothetical scenario);
- If more SMEs experience an increase in exports, the share of SMEs which undergo an increase in sales would increase too. It is interesting to highlight that the relationship between exports and sales is not linear: the higher the increase in

exports the more probable the distribution of the sales variable would concentrate on the states 'enough' 'appreciably', while less SMEs would select 'very much'.

Scenario analysis can be a valuable tool for policy making. It helps test alternative hypothetical scenarios related to the policy theory of intervention and better design and target the policy instrument. For instance, assuming that the firm's size class is a relevant variable to explain the effectiveness of the policy instrument, the scenario analysis could show to what extent the outcome variable would change assuming that all beneficiary firms were, say, micro-size enterprises. Also, it could be used to show how a change in aid intensity (ratio of support over volume of the investment) would affect other variables in the network.

5. Conclusions

The starting point for our work is the acknowledgement that theory-based approaches are good at explaining mechanisms of effectiveness of policy interventions, usually by means of interviews, case studies and, in general, qualitative analysis, but they have limited statistical power. At the same time, traditional econometric analysis, as well as experimental or quasi-experimental methods can provide quantitative indication of the policy impact, without however exploring the complex mechanisms of policy performance. We argue that BNA can be a valuable tool to complement theory-based evaluation of business support measures and empirically test the theory of policy instruments and the multiple and complex relationships among variables. BNA allows finding conditional probability distributions characterising sample data. We tested the use of BNA in combination with TBE in the framework of the ex-post evaluation of Cohesion Policy programmes targeting SMEs during the 2007-2013 period.

The combination of TBE and BNA proved to be a valuable and informative methodology of analysis which deserves to be further developed in the evaluation of public programmes and individual policy instruments. From an ex-post perspective, it contributed to guiding the evaluator towards an in depth understanding of the object of analysis and the identification of the causal links, thus leading to clear answers to the evaluation questions.

BNA was found to be rather intuitive to use, very flexible and providing added value to the evaluation as compared to a purely qualitative TBE. It was crucial to properly test the theory and find hidden or unexpected mechanisms of change. In combination with other analytical methodologies, including regression models, the BNA could ensure that robust results are obtained and lead to a clear idea of whether the policy instrument is effective and how.

Some limitations and challenges in the application of BNA can also be identified. First, since the network structure and conditional probability distribution are based on available data, the quality of data clearly affects the results and, therefore, appropriate efforts should be put in place to minimise the risk of bias. If data originate from surveys to beneficiaries, as in our case, possible sources of bias could derive from the self-selection of respondents and, consequently, differences between the sample and population's distributions. The inability of interviewed people to properly understand the question, or their strategic behaviour, leading them not to reveal their true opinion or position, are other possible sources of bias.[10]

The sample size is also a relevant. In order to ensure a certain robustness of the network, the practice shows that the sample size should be proportional to the number of variables entering the network. In general, the larger the sample size, the more likely it is to build large and complex network which are stable against possible perturbations.

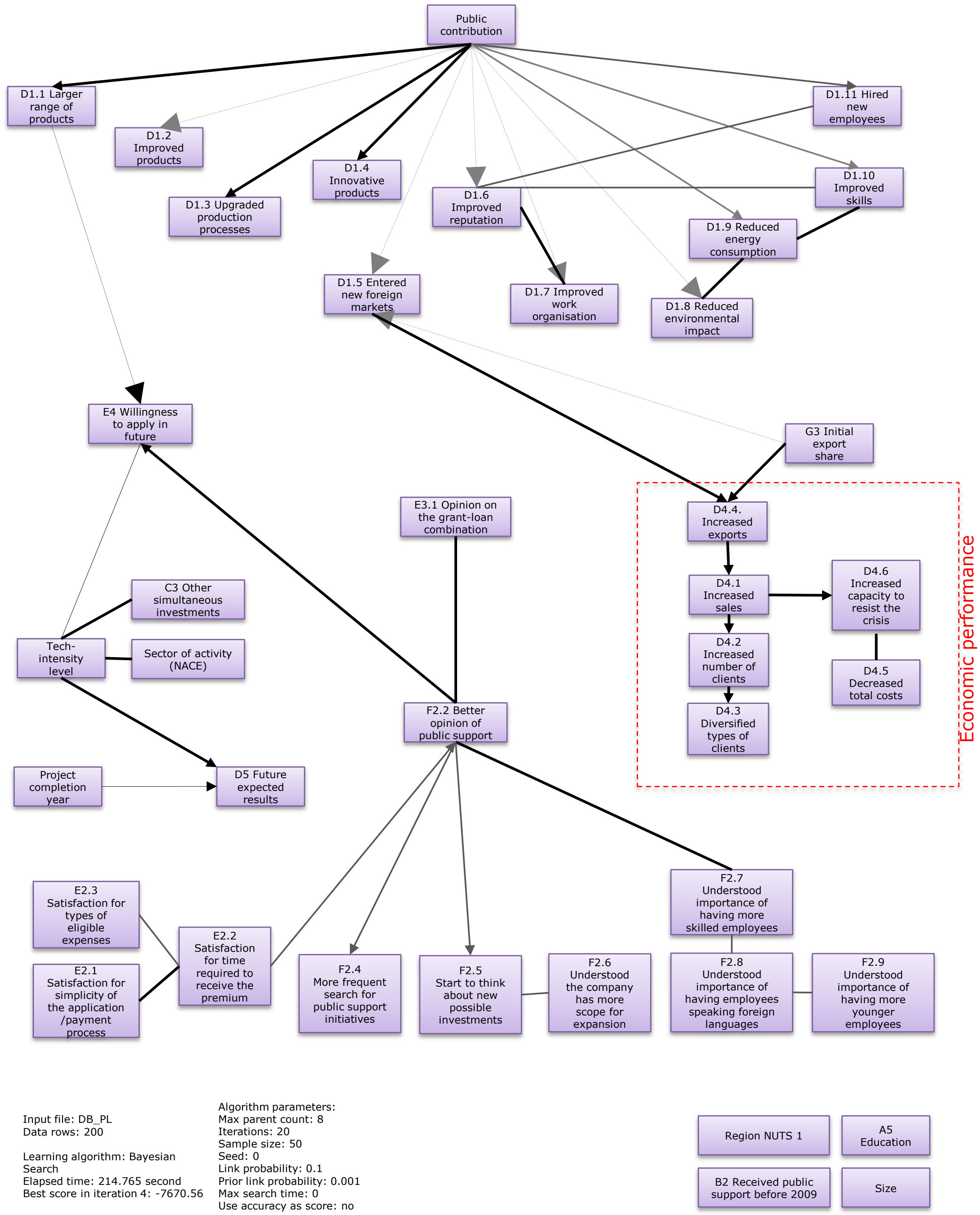
The second limitation concerns the possibility to infer causal relationships from the DAG. The majority of statistical studies can only infer 'correlation' or 'association' among variables, unless randomized experimental trials are performed. In order to be able to interpret the links between interacting variables in the Bayesian network in causal terms, we need to introduce two assumptions: that there are no latent variables acting as confounding factors, and that each variable of the network is conditional independent to its non-effects, both direct and indirect, given its direct causes (the causal Markov assumption). In our application of BNA, it is the theory of change resulting from TBE and the evaluator's background knowledge and interpretation that allow direct relationships among variables to be interpreted as conjectures of direct causal relations (Pearl, 2000; Williamson 2005).

Future research directions include the application of BNA to evaluate other types of policy instruments targeted to SME, but also in other fields of intervention characterised by a rather clear theory of intervention and a sufficient number of beneficiaries.

Furthermore, it would be interesting to experiment the combination of BNA with other quantitative data analysis methods, including those using counterfactual techniques, where data and opinions not only of beneficiary firms but also of potential beneficiaries are considered. The quasi-experimental approach could help verify the causal links between input and output variables. In principle the counterfactual approach could also be built in the BN itself. Changes in business practice and economic performance could be investigated by surveying beneficiary enterprises and a suitable control group of non-beneficiary firms. The variable related to the volume of public support received (which is nil for non-beneficiary firms) would enter the network. It would then be possible to assess whether beneficiary firms have followed a different trajectory than non-beneficiary ones and observe the mechanisms of change triggered by the policy intervention.

[10] We explained the way how we addressed all these limitations in European Commission (2015).

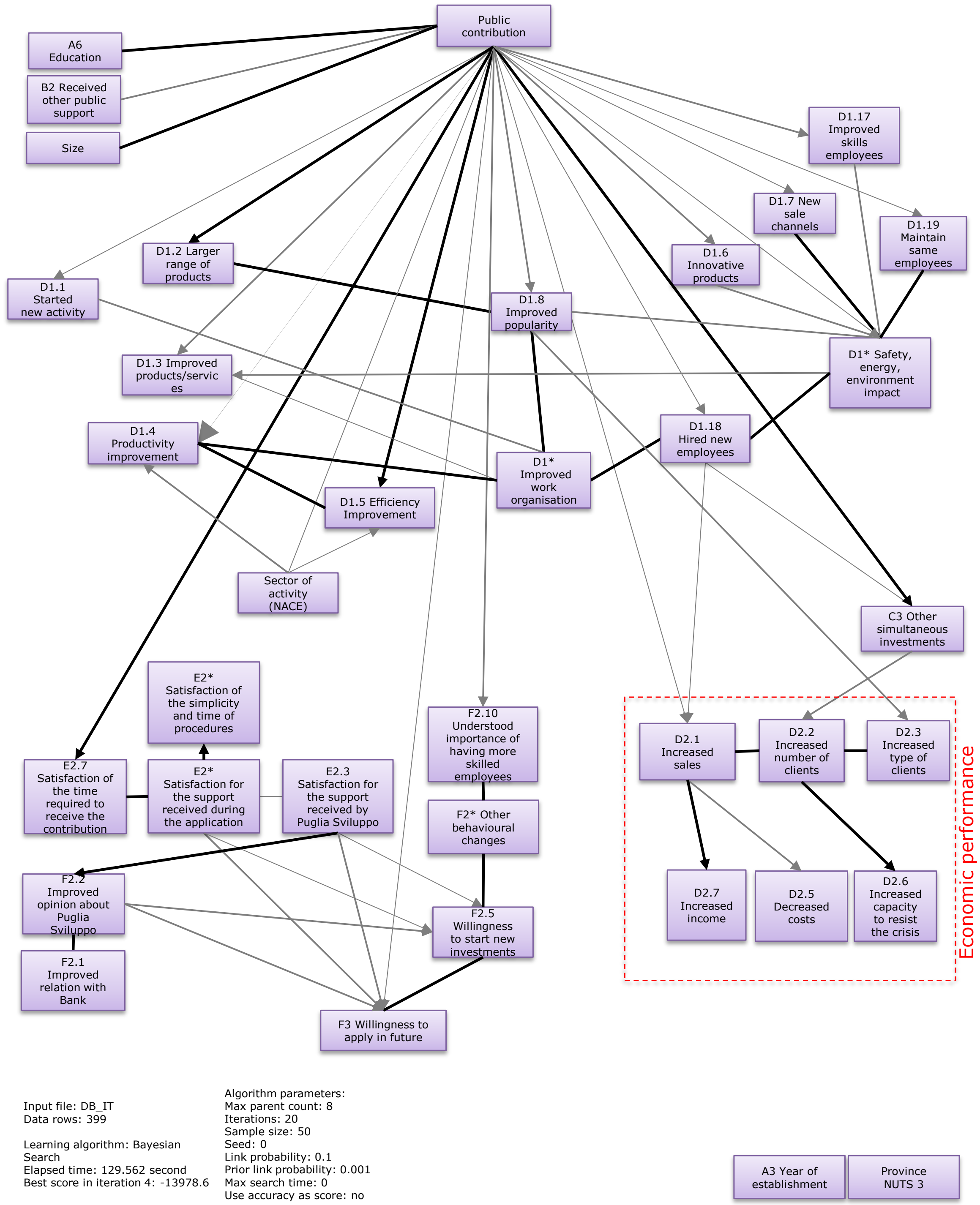
Figure 1. Bayesian Network Analysis of the Technological Credit (Poland)



Note: Software used for the BNA: GeNIe. The graph includes some variables (bottom right) that, in spite of having been controlled for during the construction of the model, do not result to be strongly linked to any other particular variable.

Source: European Commission (2015) based on CSIL analysis.

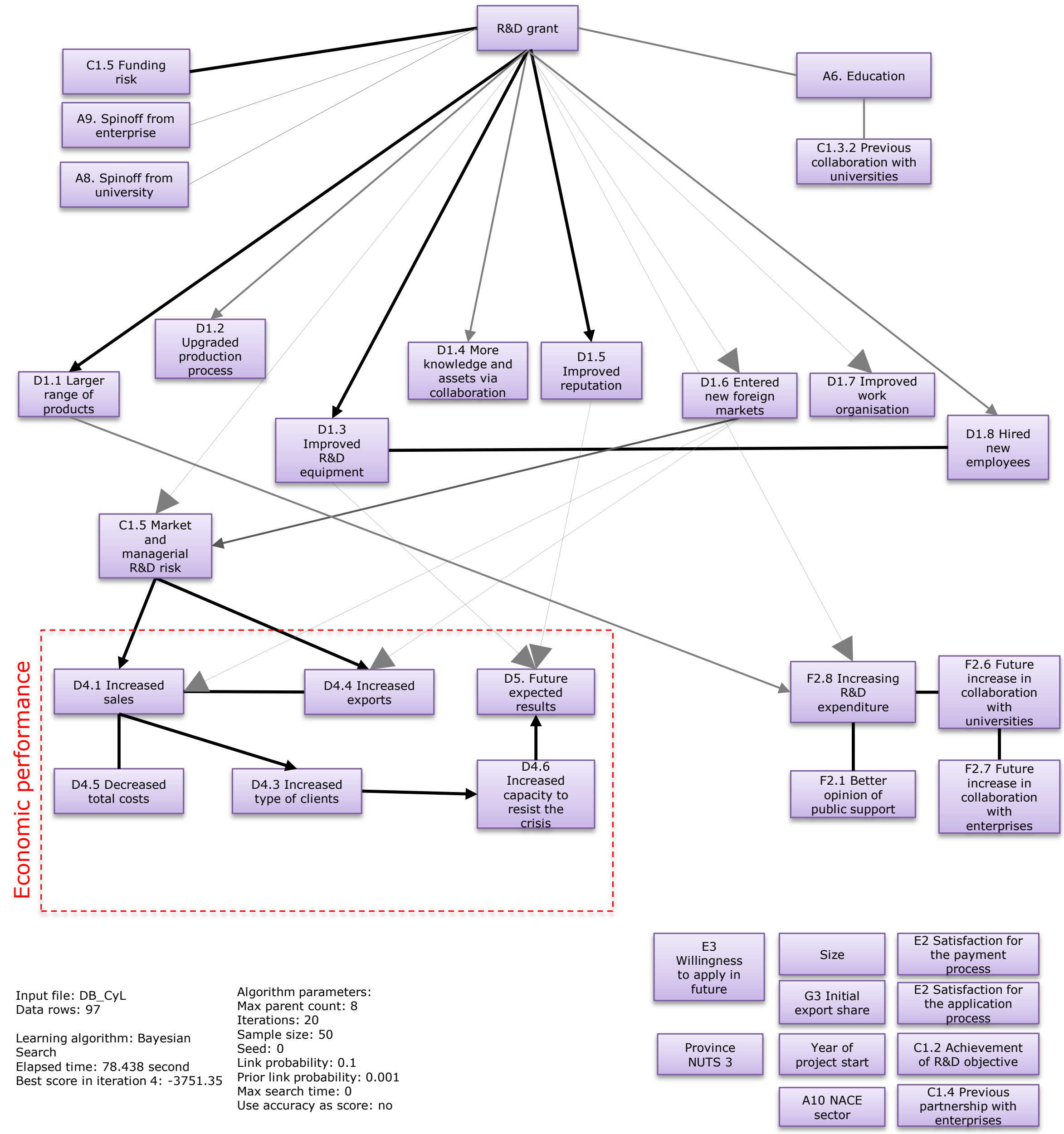
Figure 2. Bayesian Network Analysis of the Title II policy instrument (Apulia, Italy)



Note: Software used for the BNA: GeNIe. The graph includes some variables (bottom right) that, in spite of having been controlled for during the construction of the model, do not result to be strongly linked to any other particular variable. Asterisks indicate principal component variables. For more details, see European Commission (2015).

Source: European Commission (2015) based on CSIL analysis.

Figure 3. Bayesian Network Analysis of the Castilla y Leon policy instrument (Spain)

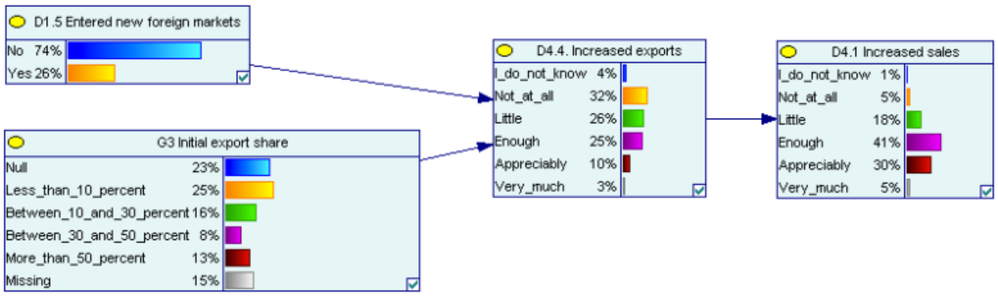


Note: Software used for the BNA: GeNIe. The graph includes some variables (bottom right) that, in spite of having been controlled for during the construction of the model, do not result to be strongly linked to any other particular variable. Asterisks indicate principal component variables. For more details, see European Commission (2015).

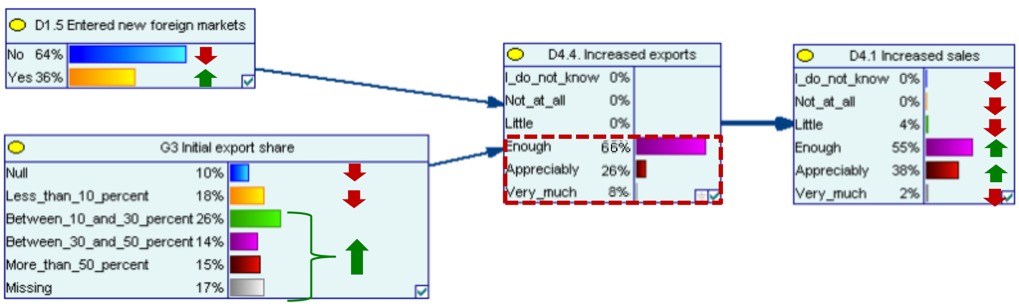
Source: European Commission (2015) based on CSIL analysis.

Figure 4. Scenario analysis

Baseline network



Posterior network (evidence propagation scenario)



Source: European Commission (2015) based on CSIL analysis.

Acknowledgements

This paper draws from the study ‘Ex-post evaluation of Cohesion Policy programmes 2007-2013 financed by the European Regional Development Fund (ERDF), and Cohesion Fund (CF) - Work Package 2: Support to SMEs - Increasing research and innovation in SMEs and SME development’. The study was carried out by CSIL, selected by the Evaluation Unit, DG Regional and Urban Policy, European Commission, through a call for tenders by open procedure (Contract Number: 2014CE16BAT002). This is one among the sixteen studies of whole ex-post evaluation of the ERDF and Cohesion Fund investigating the key outcomes of Cohesion Policy in 2007-2013, and one among the three studies dealing with enterprise support (the remaining two being on Financial Instruments -WP3, and Large Enterprises -WP4). Intermediate and final reports of the study are available at the following link:
http://ec.europa.eu/regional_policy/en/policy/evaluations/ec/2007-2013/#3.
The authors are grateful for the very helpful insights from the EC staff and other members of the Steering Group. The authors are responsible for any remaining errors or omissions.

Bios of authors

Francesco Giffoni is researcher at CSIL (Milan) where he is responsible for statistical and economic analysis in the field of industrial and regional development policies.

Silvia Salini is associate professor in Statistics and Research methods at the Department of Economics, Management and Quantitative methods, University of Milan. She is an expert of survey data analysis.

Emanuela Sirtori is partner and senior researcher at CSIL (Milan), specialised in a range of qualitative and quantitative evaluation methods, including BNA and theory-based evaluation.

References

Agrawal A (2001) University-to-industry knowledge transfer: literature review and unanswered questions. *International Journal of Management Reviews*, 3: 285-302.

Amara N, Landry R, Becheikh N and Ouimet M (2008) Learning and novelty of innovation in established manufacturing SMEs. *Technovation* 28(7): 450-463.

Astbury B and Leeuw FL (2010) Unpacking black boxes: mechanisms and theory building in evaluation. *American Journal of Evaluation* 31: 363–81.

Bachtler J and Wren C (2007) Evaluation of European Union Cohesion policy: Research questions and policy challenges. *Regional Studies* 40(2): 143-153.

Befani B and Stedman-Bryce G (2017) Process Tracing and Bayesian Updating for impact evaluation. *Evaluation* 23(1): 42-60.

Ben-Gal I (2007) Bayesian Networks, In: Ruggeri F, Faltin F and Kenett R (Eds.), *Encyclopedia of Statistics in Quality and Reliability*, John Wiley & Sons.

Bozeman B and Gaughan M (2007) Impacts of grants and contracts on academic researchers' interactions with industry. *Research Policy* 36: 694–707.

Buckley AP (2016) Using Contribution Analysis to evaluate small & medium enterprise support policy. *Evaluation* 22(2): 129-148.

Busetti S and Dente B (2017) Using process tracing to evaluate unique events: the case of EXPO Milano 2015. *Evaluation* 23(3): 256-273.

Chen H-t (1997) Applying mixed methods under the framework of theory-driven evaluations. *New Directions for Evaluation* 1997(74): 61-72.

Cohen W and Levinthal D (1990) Absorptive capacity: A new perspective on learning and innovation. *Administration Science Quarterly* 35: 128-152.

Cunningham PN and Gök A (2012) *Compendium of Evidence on the Effectiveness of Innovation Policy Intervention*. NESTA/MIOIR: London/Manchester.

Daly R, Shen Q and Aitken S (2011) Learning Bayesian networks: approaches and issues. *The knowledge engineering review* 26.02: 99-157.

Damanpour F and Aravind D (2012) Managerial innovation: conceptions, processes, and antecedents, *Management and Organization Review* 8(2): 423-454.

de Jong JPJ and Freel M (2009) Absorptive capacity and the reach of collaboration in high technology small firms. *Research Policy* 39: 47–54.

Delahais T and Toulemonde J (2012) Applying contribution analysis: lessons from five years of practice. *Evaluation* 18: 281–93.

Ebrahim AN, Ahmed S and Taha Z (2010) SMEs; Virtual research and development (R&D) teams and new product development: A literature review. *International Journal of the Physical Sciences*, 5(7): 916-930.

European Commission (2015) Support to SMEs – Increasing Research and Innovation in SMEs and SME development, Third Intermediate Report. Work package 2, Ex post evaluation of Cohesion Policy programmes 2007-2013, focusing on the European Regional Development Fund (ERDF) and the Cohesion Fund (CF). Prepared by CSIL in partnership with CSES and ZEW. Contract 2014CE16BAT002. Available at: http://ec.europa.eu/regional_policy/sources/docgener/evaluation/pdf/expost2013/wp2_3rd_intermediate_report_1.pdf.pdf.

European Commission (2016) Support to SMEs – Increasing Research and Innovation in SMEs and SME development, Final Report. Work package 2, Ex post evaluation of Cohesion Policy programmes 2007-2013, focusing on the European Regional Development Fund (ERDF) and the Cohesion Fund (CF). Prepared by CSIL in partnership with CSES and ZEW. Contract 2014CE16BAT002. Available at:

http://ec.europa.eu/regional_policy/sources/docgener/evaluation/pdf/expost2013/wp2_final_en.pdf.

Florio M, Parteka A and Sirtori E (2017) The mechanisms of technological innovation in SMEs – a Bayesian Network Analysis of EU policy impact on Polish firms. CSIL working paper, latest version June 9th 2017. Available upon request.

Glymour C and Cooper G (1999) *Causation, computation and discovery*. Cambridge, MA: MIT/AAAI Press.

Hawkins AJ (2016) Realist evaluation and randomised controlled trials for testing program theory in complex social systems. *Evaluation* 22(3): 270-285.

Horny M (2014) *Bayesian Networks*. Technical Report No.5. Boston University. Available at: http://www.bu.edu/sph/files/2014/05/bayesian-networks-final.pdf

Jensen FV (1996) *An Introduction to Bayesian Networks*. Springer, ISBN 978-038-7915029.

Jensen FV and Nielsen TD (2007) *Bayesian Networks and Decision Graphs*. Second edition, Information Science and Statistics, Springer.

Jensen MB, Johnson B, Lorenz E and Lundvall BA (2007) Forms of knowledge and modes of innovation. *Research Policy* 36: 680-693.

Kenett R and Salini S (2011b) *Modern analysis of customer surveys: with applications using R*. Vol. 117. John Wiley & Sons.

Kenett RS and Salini S (2011a) Modern Analysis of Customer Surveys: comparison of models and integrated analysis, with discussion. *Applied Stochastic Models in Business and Industry* 2: 465–475.

Kennet RS (2012) *Applications of Bayesian Networks*. Available at SSRN: http://ssrn.com/abstract=2172713 or http://dx.doi.org/10.2139/ssrn.2172713

Lam A (2005) *Organizational innovation*. In J. Fagerberg, D. Mowery, R. Nelson (Eds.), The Oxford Handbook of Innovation, Oxford University Press, Oxford (2005), pp. 115-148.

Lipnack J and Stamps J (2000) *Why The Way to Work. Virtual teams: People Working across Boundaries with Technology*. Second Edition ed. John Wiley & Sons, New York

Mackenzie M and Blamey A (2005) The Practice and the Theory. Lessons from the Application of a Theories of Change Approach. *Evaluation* 11(2): 151-168.

Majocchi A, Dalla Valle L and D'Angelo A (2015) Internationalisation, cultural distance and country characteristics: a Bayesian analysis of SMEs financial performance. *Journal of Business Economics and Management*, 16 (2): 307-324, DOI: 10.3846/16111699.2012.720600Mayne J (2001) Addressing attribution through contribution analysis: using performance measures sensibly. *Canadian Journal of Program Evaluation* 16(1): 1-24.

Mowery DC (1983) The relationship between intrafirm and contractual forms of industrial research in American manufacturing. 1900-1940. *Explorations in Economic History* 20: 351-374.

Murphy K (2001) *An introduction to graphical models*. Technical report, University of California, Berkeley, May 2001. Available at: https://www.cs.ubc.ca/~murphyk/Papers/intro_gm.pdf

Nadkarni S and Shenoy PP (2001) A Bayesian network approach to making inferences in causal maps. *European Journal of Operational Research* 128(3): 479-498.

Parrilli MD, Fitjar R, Rodriguez-Pose A (2016) *Business innovation modes: a review from a country perspective*. M.D. Parrilli, R. Fitjar, A. Rodriguez-Pose (Eds.), Innovation Drivers and Regional Innovation Strategies, Routledge, London and New York.

Pawson R and Tilley N (1997) *Realistic Evaluation*. Thousand Oaks, CA: SAGE.

Pearl J (2000) Causal inference without counterfactuals: comment. *Journal of the American Statistical Association*: 428-431.

Peng H and Ding C (2003) Structure search and stability enhancement of Bayesian networks. *Proceedings of the Third IEEE International Conference on Data Mining* (ICDM'03) 0-7695-1978-4/03.

Ranmuthugala G, Cunningham FC, Plumb JJ, Long J, Georgiou A, Westbrook JI, Braithwaite J (2011) A realist evaluation of the role of communities of practice in changing healthcare practice. *Implementation Science* 6(49), doi: 10.1186/1748-5908-6-49.

Schmitt K and Beach D (2015) The contribution of process tracing to theory-based evaluations of complex aid instruments. *Evaluation* 21(4): 429-447.

Sirtori E, Vallino E and Vignetti S (2017) Testing intervention theory using Bayesin Network Analysis: evidence from a pilot exercise on SMEs support. In Pokorrski J, Popis Z, Wyszynska T and Hermann-Pawlowska K (Eds) *Theory-based evaluation in complex environments*, PARP, ISBN 978-83-7633-334-2 Available at https://en.parp.gov.pl/images/PARP_publications/pdf/20017_theory-based-evaluation.pdf.

Spirtes P (2001) An anytime algorithm for causal inference. in *The Presence of Latent Variables and Selection Bias in Computation, Causation and Discovery*, MIT press.

Stame N (2004) Theory-based evaluation and varieties of complexity. *Evaluation* 10(1): 58-76.

Weiss CH (1995) Nothing as practical as good theory: Exploring theory-based evaluation for comprehensive community initiatives for children and families. In J. Connell, A. Kubisch, L. B. Schorr, & C. H. Weiss (Eds.), *New approaches to evaluating community*.

Weiss CH (1997) How can theory-based evaluation make greater headway. *Evaluation Review* 21: 501-24.

White H (2009) Theory-based impact evaluation: principles and practice, *Journal of Development Effectiveness* 1(3): 271-284.

Williamson J (2005) *Bayesian Networks and Causality: Philosophical and Computational Foundations*. Oxford University Press, Oxford.