



Working Paper Series

**Multidimensional well-being:
A Bayesian Networks approach**

Lidia Ceriani
Chiara Gigliarano

ECINEQ WP 2016 - 399

Multidimensional well-being: A Bayesian Networks approach

Lidia Ceriani[†]

The World Bank, U.S.A.

Chiara Gigliarano

Università degli Studi dell'Insubria, Italy

Abstract

In the multidimensional well-being literature, it has been long advocated that it is important to consider how the different well-being domains interact. Nevertheless, none of the existing approaches is useful to tackle this issue. In this paper, we show that the statistical technique of Bayesian Networks is an intuitive and powerful instrument that allows to graphically model the dependence structure among the different dimension of well-being. Moreover, Bayesian Networks can be used to understand the effectiveness of given interventions addressed to one or more dimensions, as well as to design more effective policies to reach the desired outcome. The new approach is illustrated with an empirical application based on data for a selection of Western and Eastern European countries.

Keywords: Multivariate analysis, directed acyclic graphs, probabilistic inference, well-being.

[†]Corresponding author.
chiara.gigliarano@uninsubria.it.

E-mails: (L.Ceriani) lceriani@worldbank.org; (C.Gigliarano)

1 Introduction

Starting from Amartya Sen seminal contribution (Sen, 1980, Sen, 1985), the economic literature has underlined the necessity of defining individual well-being as a multidimensional concept rather than relying only on income or consumption expenditure per capita. A general agreement about what aspects of someone's life matter for defining her well-being has been consolidating around the following dimensions: (i) *material living standards*, including level of income, consumption and wealth; (ii) *health*; (iii) *education*; (iv) *extent of personal activities*, including work; (v) *political voice and governance*; (vi) *participation to social activities and civil society*; (vii) *social connections and relationships*; (viii) *security*, of economic as well as physical nature; (ix) *satisfaction with life*, (x) *environment*, taking into account the present and discounting the future condition of it. See, among others, Stiglitz, Sen, and Fitoussi (2010); OECD (2015); Cnel and Istat (2015).

The current literature typically distinguishes between two different approaches for measuring multidimensional well-being. On the one hand, well-being can be monitored through a dashboard of several indicators. This approach has the advantages of highlighting changes over time in each specific well-being dimension or indicator, as well as of avoiding any loss of information. However, such a broad collection of indicators does not usually allow for a parsimonious representation of well-being of a society, thus making cross-country comparisons rather difficult. The dashboard approach has been applied, among others, in the OECD's Better Life Index (OECD, 2015), the Italian Equitable and Sustainable Well-being (Cnel and Istat, 2015), and the U.K.'s National Well-being Measure (Office for National Statistics, 2015). On the other hand, multidimensional well-being can be monitored using a composite index, which provides a synthetic analysis of well-being trends, by aggregating all the well-being indicators into a single number. This approach has the advantage of allowing for easy comparisons across countries, while the main drawbacks are the relevant loss of information due to the extreme synthesis, as well as the arbitrariness due to the different possible aggregation methods. See for instance OECD (2008) for a broad discussion on the composite indicators' methodology. Examples of composite indicators are the Human Development Index (HDI) proposed by the United Nation Development Programme (UNDP), the Canadian Index of Well-being (Canadian Index of Wellbeing, 2012), the Happy Planet Index (NEF, 2013), or multi-dimensional poverty or inequality indices (Alkire and Foster, 2011, Aaberge and Brandolini, 2015, Bosmans, Decancq, and Ooghe, 2015).

Beside these two opposite approaches, several other studies have been mainly interested in the subjective well-being, with the aim of explaining differences in life satisfaction both at country and at individual level, using econometric models (Clark, forthcoming). For example, starting from the well-known Easterlin paradox of the non-linear relationship between life satisfaction and income (see, e.g., Easterlin, 2001), several other scholars analyzed the role of economic well-being on life satisfaction. In particular,

Ferrer-i-Carbonell (2005) tested the importance of relative income on individual well-being, revealing that individuals are happier the larger their income is in comparison with the income of the reference group. Moreover, Pittau, Zelli, and Gelman (2010) investigated the role of economic variables in predicting regional disparities in life satisfaction of European Union citizens. Other studies analyzed the role of work status on life satisfaction, showing that individual unemployment has strong negative effects on subjective well-being; see, among others, Clark and Oswald (1994) and Winkelmann and Winkelmann (1998). The regression models proposed are, therefore, mainly interested in understanding the link between subjective well-being and some other objective dimensions.

Despite this rich set of diverse analytical methods, an important aspect of multi-dimensional well-being has been neglected: none of the current approaches is fitted to investigate the relationship between the different dimensions. The importance of analyzing how the different well-being dimension interact has been recently stressed by Stiglitz, Sen, and Fitoussi (2010): “[...] *it is critical to address questions about how developments in one domain of quality of life affect other domains*” (p.59) and, the authors continue, “*when designing policies in specific fields, impacts on indicators pertaining to different quality-of-life dimensions should be considered jointly, to address the interactions between dimensions*” (p.16).

The aim of this paper is to fill this gap in the literature by introducing a new approach which can offer an insight of the structure of dependence existing among the different dimensions of well-being: the statistical technique of Bayesian Networks. A Bayesian Network is a probabilistic graphical model that represents a set of variables and their probabilistic dependencies. It shares the transparency of a dashboard approach, since the different dimensions appear as distinct nodes in the graph; it allows for intuitive cross-country comparisons as in the composite index approach; it is also possible to identify dependent and independent variables, resembling econometric modeling. Novelty and explicatory power of this new approach lie in its ability to visualize and to analyze the dependency patterns among all the well-being dimensions, which can be further exploited to understand the effectiveness of given policies directed to one or more dimensions, as well as to design more effective interventions to reach the desired outcome. To the best of our knowledge, this paper constitutes the first attempt to apply Bayesian Networks to the analysis of multi-dimensional well-being.

The paper is organized as follows. In Section 2 we will discuss the theoretical framework supporting Bayesian Network analysis, using simplified examples to sketch out the main properties. Section 3 provides an empirical application based on *LiTS* (Life in Transition Survey) dataset for a sample of Western European countries and a selection of new European member states. Section 4 concludes and provides suggestions for further applications.

2 Theoretical background: Bayesian Networks

Bayesian Networks belong to the graphical model structures known in the literature as *Directed acyclic graph* (or *DAG*): a *DAG* represents conditional independence relations between variables (not necessarily causal interactions). The advantage of Bayesian Network approach is that it provides an intuitive visualization of the complex relationships among variables, in particular the non-linear relationship between dependent and independent variables and also among independent variables themselves. Moreover, it allows to analyze direct implications of policies and actions through prediction and diagnostics that can be understood by policy-makers and academics (see Anderson, Mackoy, Thompson, and Harrell, 2004).

Thanks to the growing availability of high dimensional datasets, and the promising link between graph theory and probability theory shed by Bayesian Network (Nagarajan, Scutari, and Lèbre, 2013), this approach has been applied in different disciplines. Among others, the range of application spans from on-line analytical processing (OLAP) performance enhancement (Margaritis, 2003) to medical service performance analysis (Acid, de Campos, Fernández-Luna, Rodríguez, Rodríguez, and Salcedo, 2004), gene expression analysis (Friedman, Linial, Nachman, and Peér, 2000), breast cancer prognosis and epidemiology (Nagl, Williams, and Williamson, 2008) and costumer satisfaction analysis (Salini and Kenett, 2009; Anderson, Mackoy, Thompson, and Harrell, 2004).

2.1 Definitions

A Bayesian Network is represented by a set $\{\mathcal{G}, \Theta\}$. The first element $\mathcal{G} = \{\underline{X}, A\}$ defines a *DAG*, where \underline{X} is the *node set* and A is the *arc set*. Each node $X_i \in \underline{X}$, $i = 1, \dots, n$, represents a random variable (in this paper we use node and random variable interchangeably), and each directed arc $a_{ij} \in A$ between nodes X_i and X_j , $i \neq j$, represents a probabilistic dependence between associated nodes. If the direction of arc a_{ij} goes from X_i to X_j , this means that values taken by X_j depend on the values taken by X_i . Usually, X_i is referred to as a *parent* of X_j and X_j is the *child* of X_i . Also, the set of nodes directly reachable from X_j are denoted as *descendants*, and the set of nodes from which X_j can be directly reached are denoted as *ancestors*. The acyclic property of this kind of graphs guarantees that a node cannot be simultaneously its ancestor and its descendent. In other words, feedback loops like $X \rightarrow Y \rightarrow X$ are not permitted. Moreover, the graph is directed because two-headed arrows depicting non-causal association (e.g., $X \leftrightarrow Y$) are not allowed. The second element of a Bayesian Network, Θ , denotes the set of parameters θ_i which quantifies the intensity of the dependence between a child-node X_i and its parent(s). The strength of a dependency is represented by the conditional probabilities that are attached to the cluster of parents-child nodes in the network.

A Bayesian Network is based on the assumption that each variable is independent of its non-descendants given its parents, which is referred to in the literature as *Markovian*

property. As a consequence, the joint probability distribution of variables X_1, X_2, \dots, X_n can be factorized as follows:

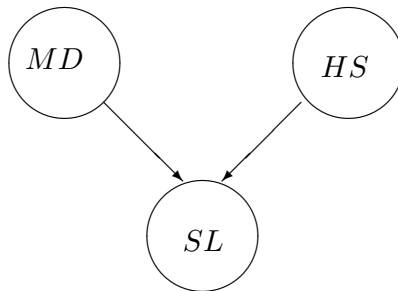
$$P(X_1, X_2, \dots, X_n) = \prod_{i=1}^n P(X_i | \Pi_i) \tag{1}$$

where Π_i is the set of *parents* of X_i and P is either the probability mass function or the probability density function when X_i is, respectively, discrete or continuous.

Building a Bayesian Network involves two steps: (i) graphically structuring the relations between variables, e.g. what depends on what, and (ii) estimating the parameters representing the strength of these dependences. Both steps can be a mixture of prior knowledge, e.g. temporal order and/or possible relationship restrictions, and data evidence, e.g. frequency tables. Learning the Bayesian Network, thus, means that given the training dataset and prior information (such as background knowledge or causal relationships), we need to estimate the network structure and the parameters of the joint probability density in the Bayesian Network.

Example 2.1. *As illustrative example, let us consider three of the well-being’s dimensions that we will consider later in the empirical application, namely Health status (HS), Material Deprivation (MD), and Satisfaction with life (SL). For simplicity, we consider all these variables as binary: $HS \in \{Good, Bad\}$; $MD \in \{True, False\}$; $SL \in \{True, False\}$.*

Assume that the qualitative relationships among these variables can be represented by the following DAG, whose arcs indicate which node affects/causes the others:



Satisfaction with Life (SL) is affected both by Health Status (HS) and by Material Deprivation (MD), while we assume there is no relationship between Health Status and Material Deprivation.

We have some prior information about the probability distributions of the two parent nodes, described in the following:

<i>MD</i>		<i>HS</i>	
<i>True</i> <i>False</i>	<i>Good</i> <i>Bad</i>	<i>Good</i> <i>Bad</i>	<i>Bad</i> <i>Good</i>
<i>0.1</i> <i>0.9</i>	<i>0.6</i> <i>0.4</i>	<i>0.6</i> <i>0.4</i>	<i>0.4</i> <i>0.6</i>

Hence, our prior information is that 10% of the sample is materially deprived, and 60% has good health status. We also have information about the conditional probabilities of Satisfaction with Life (SL) given its parent nodes:

		SL	
MD	HS	True	False
True	Good	0.4	0.6
True	Bad	0.2	0.8
False	Good	0.95	0.05
False	Bad	0.8	0.2

Given our priors, we can now compute the marginal probability of being satisfied with life using the law of total probability (where the capital letters T, F, G and B stands for, respectively, True, False, Good and Bad):

$$\begin{aligned}
 P(SL = T) &= P(SL = T|MD = T, HS = G) * P(MD = T) * P(HS = G) \quad (2) \\
 &+ P(SL = T|MD = T, HS = B) * P(MD = T) * P(HS = B) \\
 &+ P(SL = T|MD = F, HS = G) * P(MD = F) * P(HS = G) \\
 &+ P(SL = T|MD = F, HS = B) * P(MD = F) * P(HS = B) \\
 &= 0.4 * 0.1 * 0.6 + 0.2 * 0.1 * 0.4 + 0.95 * 0.9 * 0.6 + 0.8 * 0.9 * 0.4 \\
 &= 0.833
 \end{aligned}$$

From (2), it is straightforward to compute the probability of not being satisfied with life: $P(SL = F) = 1 - P(SL = T) = 0.167$.

2.2 Algorithms

Three approaches have been employed to evaluate potential Bayesian networks: (i) independence-testing algorithms (Cooper, 1997; Spirtes, Glymour, and Scheines, 1993); (ii) score-based algorithms (Cooper and Herskovits, 1992; Heckerman, Geiger, and Chickering, 1995); and (iii) a hybrid approach (Anderson, Mackoy, Thompson, and Harrell, 2004). We briefly discuss the three approaches in turn.

Independence-testing is a two-step procedure. In the first step, for each pair of nodes, independence and conditional-independence are tested, for example using Pearson-chi square test for discrete variables, and the Pearson's linear correlation for continuous variables. A rejected independence relationship implies a dependency linkage, and these linkages are assembled to provide the graphical structure. In the second step, conditional probabilities are estimated *via* maximum likelihood approach from sample data.

Score-based algorithms are heuristic optimization methods used to select the best network with respect to a goodness-of-fit score:

$$\arg \max_{\mathcal{G} \in \mathcal{G}^n} s(\mathcal{G}, D)$$

where $s(\cdot, \cdot)$ is a decomposable score function which evaluates the merit of \mathcal{G} with respect to the data D and \mathcal{G}^n is the set containing all DAGs with n nodes. The most common score-based algorithm found in the literature is the *Hill Climbing* which, starting from the empty network, chooses at each step to link the nodes providing the greatest improvement in the score function, one arc at the time. The algorithm then stops when there is no local improvement in the score function. Four score functions are prevalent in the literature: the Logarithm of the Dirichlet posterior density (or Logarithm of the K2-score), the Akaike Information Criterion, the Bayesian Information Criterion (or Schwarz Information Criterion) and the multinomial log-likelihood (see Scutari, 2010).¹

Finally, the hybrid approach uses conditional independence tests to reduce the dimension of the network, by confirming or rejecting hypotheses on the relations between nodes, and it uses the scores to find the optimal network and to estimate the conditional probability distributions. For an application of this third approach we refer to Anderson, Mackoy, Thompson, and Harrell (2004) and references therein.

2.3 Forcing the Relationships

Bayesian Network analysis allows also to include prior information on the relationship between nodes such as some temporal order among variables and some relationship restrictions. This is done by including *blacklist* and *whitelist* arguments. The former guarantees that a set of arcs will be missing from the Bayesian network, while the latter guarantees that it will be present. For example, if the arc from A to B is whitelisted, then it is guaranteed to be present in the graph in the specified direction. On the contrary, if the same arc is blacklisted, it means that the directed arc from A to B will not be present in the network, while the opposite arc from B to A may or may not appear in the graph. Any arc blacklisted in both directions is never present in the graph.

2.4 The strength of the dependence

For each arc it is also possible to determine its strength, establishing whether this arc is supported strongly enough by the data and can therefore be considered as significant. If the conditional independence test approach is used, then arc strengths are computed as the p-value of the test. If, instead, the network scores approach is considered, the arc strength is defined as the increase/decrease of the network score caused by the removal of the arc. Removing arcs with negative strength will decrease the network score,

¹In particular, the logarithm of the Dirichlet posterior density score assumes that the Bayesian network (BN) is a stochastic variable. Prior knowledge of BN is combined with sample data to estimate a posterior probability distribution for BN. Usually, conjugate prior and posterior probabilities are chosen, such as the family of Dirichlet distributions for multinomial sampling. Bayes theorem is used to update the Dirichlet prior distribution of BN, given sample data, to a Dirichlet posterior distribution of BN. Hence, the logarithm of the posterior probability is used as the Bayesian network score, which will always be negative since it is the logarithm of a probability; thus the smallest negative score among a group of BNs indicates the most probable BN.

while removing arcs with positive strength will increase the network score. Graphical representations of the arc strength usually plot positive-strength arcs as dashed lines. Also, arcs may be graphically grouped, by dividing the range of the strength coefficients into intervals. The interval each strength coefficient falls into determines the line width of the corresponding arc in the plot (see Scutari, 2010).

2.5 Probabilistic Inference

One of the most distinctive features of a Bayesian Network is the ability to represent how the configuration responds to policies and actions (Cowell, Dawid, Lauritzen, and Spiegelhalter, 1999; Pearl and Russel, 2003; Jensen, 2001). The direct implications of a change may affect the distribution of ancestors and descendants of the policy-targeted node. The Bayesian Network approach can be used also to analyze direct implications of these interventions both on ancestors (referred as *backward network inference*, or *diagnostic reasoning*) and descendants (*forward network inference*, or *predictive reasoning*) and therefore it is a useful tool both to choose the most effective policy and to assess the impact of a given policy or combination of policies. In the first case (from the target to input nodes), we fix a level of target variable to certainty, e.g. everybody is satisfied with her own life, and then, the probability distribution of inputs, e.g. poverty, are revised. In the second case (from input to target nodes) we fix a level of input variable to certainty, e.g. everybody is not poor, and then, probability distribution of targets, e.g. satisfaction with life, are updated. Examples 2.2 and 2.3 illustrates how probabilistic inference works.

Example 2.2 (Example 2.1 (continued) - Diagnostic Reasoning). *In the diagnostic reasoning we move from the effects to the cause, in the opposite direction to the network arcs. For example, we know that an individual is satisfied with life, and we update our beliefs on the probability (i) of being materially deprived ($MD=T$) and (ii) of having a bad health ($HS=B$), using the Bayes formula.*

i) The probability of being materially deprived given that the individual is satisfied with life is:

$$\begin{aligned}
 P(MD = T|SL = T) &= \frac{P(SL = T|MD = T) * P(MD = T)}{P(SL = T)} & (3) \\
 &= \frac{[P(SL = T|MD = T, HS = B) * P(HS = B)] * P(MD = T)}{P(SL = T)} \\
 &+ \frac{[P(SL = T|MD = T, HS = G) * P(HS = G)] * P(MD = T)}{P(SL = T)} \\
 &= \frac{[0.4 * 0.6 + 0.2 * 0.4] * 0.1}{0.833} = 0.023
 \end{aligned}$$

which is much smaller than the prior probability of being materially deprived, $P(MD = T) = 0.1$. Also, $P(MD = F|SL = T) = 1 - 0.023 = 0.977$.

ii) *The probability of having bad health status given that the individual is satisfied with life is:*

$$\begin{aligned}
 P(HS = B|SL = T) &= \frac{P(SL = T|HS = B) * P(HS = B)}{P(SL = T)} & (4) \\
 &= \frac{[P(SL = T|MD = T, HS = B) * P(MD = T)] * P(HS = B)}{P(SL = T)} \\
 &+ \frac{[P(SL = T|MD = F, HS = B) * P(MD = F)] * P(HS = B)}{P(SL = T)} \\
 &= \frac{[0.2 * 0.1 + 0.8 * 0.9] * 0.4}{0.833} = 0.355,
 \end{aligned}$$

which is slightly smaller than the prior probability of having bad health, $P(HS = B) = 0.4$.

Example 2.3 (Example 2.1 (continued) - Predictive Reasoning). *In the predictive reasoning, we move from the cause to the effects, following the direction of the network arcs. In this example, we know that an individual has bad health ($HS=B$), and we update consequently our belief about the probability of being satisfied with life, as follows:*

$$\begin{aligned}
 P(SL = T|HS = B) &= P(SL = T|HS = B, MD = T) * P(MD = T) & (5) \\
 &+ P(SL = T|HS = B, MD = F) * P(MD = F) \\
 &= 0.2 * 0.1 + 0.8 * 0.9 \\
 &= 0.296,
 \end{aligned}$$

which is much smaller than the prior probability of being satisfied with life, $P(SL = T) = 0.833$.

3 Application to well-being analysis

Bayesian Network is here used to model the non-linear relationships among the several dimensions of well-being, and to understand their branching patterns. We also take advantage of this methodology to conduct probabilistic inference for prediction and diagnostics, which could be understood not only by academics but also by policy-makers.

3.1 Data Description

We use the second wave of the Life in Transition Survey (LITS II), which has been conducted in late 2010 by the European Bank for Development and Reconstruction jointly with the World Bank. The survey covers almost 39,000 households in 34 countries, mainly in the former communist East, but also in five western European countries (France, Germany, United Kingdom, Italy, Sweden), used as benchmark advanced-market economies. It has the scope of assessing governance, public service delivery, public, economic and social attitudes, well-being, the impact of economic and political

change, and the impact of the 2007-2008 crisis.

The richness of the dataset allows to include all the key well-being dimensions suggested by literature as discussed in the introduction, with the exception of *environment*: (i) *material living standards*, (ii) *health*, (iii) *education*; (iv) *personal activities and work*, (v) *political voice and governance*, (vi) *participation to social activities and civil society*, (vii) *social connections and relationships*, (viii) *security of physical as well as economic nature*. To this list, we add the dimension (ix) *personal satisfaction*. Since many of the dimensions listed above may be correlated with the age of respondents, and the different demographic structure of the societies may influence the overall results, we also include *age*, and *household size* as a control variables. In fact, when analyzing well-being indicators, it is essential to control for some individual and household characteristics. For instance, when measuring poverty, we take into account the size and composition of the household by means of equivalence scales. Also, individual variables may be correlated to the age of the respondent. For instance, health status is likely to be worse for elderly individuals, being student is most likely linked to belonging to a younger cohort, and satisfaction with life may vary at different phases in the life of an individual (see, for instance, Steptoe, Deaton, and Stone, 2015). The complete list of variables is presented in Table 1. The unit of analysis is the household head. We apply the proposed Bayesian Network analysis to understand the relations between the different well-being dimensions, comparing the set of Western European countries (henceforth, WestEU) with a group of selected new European Member States, namely, Bulgaria, Hungary, Romania, Croatia, Slovakia (henceforth, EastEU).

[Table 1 about here]

Figure 1 provides a basic set of descriptive statistics for the variables involved in the analysis for each country of the WestEU and EastEU groups, as well as the groups' averages. Overall, the figures show homogeneity among the countries belonging to the same group, in particular for countries belonging to the WestEU group, and heterogeneity between the two groups of countries. Almost 70 percent of households heads living in WestEU are satisfied or strongly satisfied with their lives, as opposed to less than 30 percent in the selected New Member States. Nevertheless, the Slovak Republic looks very similar to Italy, with about 50 percent households heads feeling satisfied or strongly satisfied with their lives. A similar pattern is observed for the variable satisfaction with finances and satisfaction with job: the former is more similarly distributed across countries in the same group, while for the latter, Croatia and the Slovak Republic are more similar to the WestEU block than the other three countries in the EastEU group. Almost 80 percent of individuals in the EastEU, on average, do not feel safe in their countries, as opposed to less than 50 percent in WestEU, the only exception being Croatia, where the distribution of the variable safety is similar to Great Britain and the WestEU average. Evidently, the five selected New Member States have been more severely hit by the 2007/2008 global crises: the share of individuals who experienced at least one dimen-

sion of material deprivation in EastEU is 70 percent as opposed to 43 percent in WestEU.

[Figure 1 about here]

As an exploratory analysis we have also performed a logistic regression model aimed at understanding the relations between satisfaction with life, as target variable, and the other well-being dimensions, as explanatory variables. To reduce the multicollinearity problem among variables belonging to the same well-being domain, we have used the first principal component for material living standards and security. Age and household size are included as control variables. The dependent variable is a dummy variable indicating whether the household head is satisfied with his life or not. Table 2 illustrates the results of the analysis conducted separately for WestEU and EastEU countries, as well as for the pooled sample. In both groups of countries we find that the worse are material living standards, the lower is the probability to be satisfied with life. Health status is positively and significantly correlated with satisfaction with life in both blocks of countries. Education is a significant correlate of life satisfaction only in WestEU. Concerning work status, the unemployed are less likely of being satisfied with life than employed individuals in WestEU, while retirees living in EastEU are more likely of being satisfied with life than employed individuals. Commonly to both groups, a person who feels safe in his neighborhood, and who believes that in his country there are freedom of speech, peace and stability, is more likely to be satisfied with his life. Also, being satisfied with job and with the financial situation increases the probability of being satisfied with one's whole life. Overall, as already detected in the summary statistics, people living in the EastEU countries feel less satisfied with their life than those living in the WestEU.

[Table 2 about here]

3.2 Results

For each of the two groups of countries we have performed Hill-Climbing algorithm for learning the Bayesian Network structure, based on three different score methods: (i) the Logarithm of the K2-score (henceforth, K2), (ii) the Akaike Information Criterion (AIC), and (iii) the Bayesian Information Criterion (BIC). We have used the R package *bnlearn*, implemented by Scutari (2010). We have also forced a set of arcs not to be included in the graph, by means of a blacklist, in order to avoid the control variables age and household size to be considered dependent of the well-being dimensions. Also, we have excluded the directed arcs from status in employment (WORK) and poverty (M_POOR) to education (EDU) and from satisfaction with life (SA_LIFE) to health (HEALTH). Moreover, we have whitelisted the directed arc from EDU to WORK, as our prior is that level of education should affect status in employment.

Figures 2 and 3 compare the graphs obtained from each of the three score methods (K2, AIC and BIC) for the WestEU and the EastEU groups, respectively. The arcs highlighted in bold are the ones common to both methods, while the thinner arcs represent

the ones not in common following a pairwise comparison. Looking at the six pairwise comparisons, we find that for both groups of countries (WestEU and EastEU) the K2 and the AIC score methods propose very similar networks, with very few arcs not in common, while the BIC score method differs the most from the other two.

[Figure 2 about here]

[Figure 3 about here]

In order to select the best Bayesian Network for each group of countries, we have evaluated the output of the three score methods by means of a goodness-of-fit analysis based on the Bayesian Information Criterion (the smaller the better).² Table 3 shows that for the WestEU countries the best Bayesian Network is the one based on the AIC score, while for the EastEU countries the best Bayesian Network is based on the K2 score.

[Table 3 about here]

This is also confirmed by the results provided by a cross-validation procedure aimed to assess the accuracy of the networks (see Table 4). In particular, the last two rows of Table 4 compare the overall networks in terms of prediction error, and confirm that for the WestEU group the smallest prediction error (18.13%) is provided by the AIC score method, while for the EastEU group the K2 method allows for a prediction error equal to 19.78%. Overall, being the prediction error below 20%, we are confident about the goodness of fit of the Bayesian Networks for both groups of countries. Table 4 also illustrates the prediction error related to some selected target variables, in particular expected economic security, economic security relative to the past, satisfaction with life and satisfaction with finances. For all these target variables the classification can be considered good, as the prediction error is always smaller than 20%.

[Table 4 about here]

Henceforth, the analysis will be based on the best model for each group of countries, namely, AIC score based for WestEU and K2 score based for the EastEU. Figure 4 shows the selected networks for both sets of countries, underlying the differences and similarities between the two diagrams. Thick lines indicate arcs common to both groups.

[Figure 4 about here]

²Note that the Bayesian Information Criterion is here used both for constructing the Bayesian Network (Bayesian Network score method) and to evaluate the goodness of fit of the overall Bayesian Network (Network evaluation).

Looking at the Bayesian Network in Figure 4 we see that Age influences health, work, education, social activities and expected economic security. Moreover, in the EastEU age affects also participation to social activities, while in WestEU it affects also how the household copes with the crisis. Education commonly affects status in employment and material deprivation. In WestEU education also parents social participation, being poor, economic security relative to the past, and perception of safety. Work influences in both groups of countries health and satisfaction with job. As with the case of education, WestEU presents more complex interrelations, and status in employment affects also material deprivation, economic security relative to the past, and satisfaction with finances. The objective dimensions of well-being are in both graphs located in higher nodes with respect to subjective dimensions: as we might reasonably expect, objective circumstances influence individuals' subjective perceptions of their well-being. It is worth noting that the analysis returns satisfaction with life as an output node, being located at the end of the network. Interestingly, in both sets of countries, voice and governance is parented by life satisfaction. Contrarily to our priors, life satisfaction is directly affected by other satisfaction variables, but not by subjective health status or material well being variables.

Figure 5 shows the arcs' strengths for the selected networks of the two sets of countries. The strength is measured by the score gain/loss which would be caused by the arc's removal. In Figure 5 the thicker the lines, the higher the loss that would be suffered by the model by removing the arc. A dotted line, on the other hand, represents an arc which could be removed to gain a better overall score for the graph (the underlined strength scores are reported in Table 9 and Table 10).

[Figure 5 about here]

[Table 9 and 10 about here]

In both models, the stronger arc is the one linking satisfaction with finances to satisfaction with life, followed by the arc from economic security to satisfaction with finances. In WestEU, status in employment is strongly linked to satisfaction with job. In EastEU, material deprivation is strongly linked to economic security. These findings support our prior beliefs that subjective variables are strongly interlinked, as well as that objectives dimensions influence subjective variables.

3.3 Probabilistic Inference

As discussed in the theoretical section, a remarkable property of Bayesian Networks is the capacity to show the effects of a change in a given policy-targeted node on the distribution of its parents and children. Assessing the effect on ancestors' nodes is defined backward network inference, or diagnostic reasoning; assessing the effects on descendants' nodes is defined forward network inference, or predictive reasoning.

Let us assume that the policy-maker would like to maximize the number of individuals who are strongly satisfied with their life, and needs to assess which policy, or combination of policies, to implement in order to reach this target. Using diagnostic reasoning, we can evaluate the distribution of each parent node resulting from setting satisfaction with life to *very satisfied* for 100 percent of individuals. Table 5 and Table 6 show the results for a selected number of parent nodes (input), in, respectively, WestEU countries and EastEU countries. In the sample of WestEU countries, reaching the goal of 100 percent of individuals fully satisfied with their life calls, in particular, for labor market policies and the implementation of safety nets to ease the effect of crisis pathways affecting households. As shown in Table 5, in fact, the share of employed individuals should increase from 58 to 65 percent, and the share of unemployed individuals should decrease by 3 percentage points, from 5 to 2 percent. Moreover, the share of individuals not affected by the crises should increase from 46 to 60 percent. The implementation of education policies targeted to improve access to tertiary education should also be considered, as the share of individuals with at least upper secondary education should increase by 4 percentage points. Improving the share of individuals with upper secondary and tertiary education, and the share of employed individuals is even more crucial in the case of our sample of New Member States, as shown in Table 6. In those countries, also decreasing the share of individuals experiencing material deprivation has a key role to reach the target of 100 percent of individuals with strong life satisfaction.

[Table 5 about here]

[Table 6 about here]

Bayesian Networks may be also interrogated to assess the efficacy of one or a combination of policies on the target node, through predictive reasoning. Table 7 and Table 8 show the effects of selected policies on the distribution of the target node satisfaction with life in WestEU and EastEU, respectively. We first assume to implement a policy aimed at completely eliminating poverty (intervention a) in the tables). In both groups of countries such a policy has a moderate effect on the target node, as the percentage of individuals satisfied or strongly satisfied increases only by 1 point in WestEU and by 2 points in EastEU. A further policy simulated to both groups assumes that 100 percent of the population has tertiary education (intervention c) in the tables). This intervention, however, increases the share of persons satisfied with their life only by 1%. On the other hand, policies aimed at improving the material living standards (intervention b) in the tables) have stronger impacts. In particular, if all households living in WestEU countries had zero crisis pathways, then the percentage of satisfied people would increase by 11 points. Moreover, if all heads of the households living in EastEU countries had no material deprivation, then the percentage of satisfied people would be 6 points higher.

[Table 7 about here]

[Table 8 about here]

4 Concluding Remarks

This paper is the first attempt to graphically model the dependence structure among the different dimensions of well-being. We showed that the Bayesian Network approach is an intuitive and powerful instrument both from a theoretical and a policy-oriented point of view.

Bayesian Networks address the long advocated need to take into account the correlation among dimensions of well-being (see, for instance, Stiglitz, Sen, and Fitoussi, 2010), which has not been solved yet in the multidimensional well-being literature. By visualizing all directional relationships among different variables, scholars and policy-makers can immediately see the map of dependencies and interactions among all dimensions, and whether the relationship between two domains of well-being is direct or mediated through some other domains.

Bayesian Networks can also be used as a tool to understand the impact of simulated policies not only on target variables, but on all the covariates involved in the analysis. Diagnostic reasoning allows to set target levels for input variables in order to reach the maximum desirable level of well-being, which is a helpful preliminary analysis to decide on the set of policies to implement in a country, as well as on how to allocate the budget among the different domains. Predictive reasoning allows to analyze the impact of a specific policy or a combination of policies on the target variable, which is helpful to discriminate the best course of action among conflicting programs.

Future analysis should explore the possibility of including some information contained in the Bayesian Network in existing approaches to multivariate well-being. For instance, dependence patterns among different variables could suggest interactions among covariates in econometric models. Also, information about the strength of the different arcs might be useful to create a set of weights for composite indices based on the actual importance of the different dimensions in determining overall well-being (see Decancq and Lugo, 2013 for a review of existing weighting methods).

References

- AABERGE, R., AND A. BRANDOLINI (2015): “Multidimensional Poverty and Inequality,” in *Handbook of Income Distribution*, ed. by A. B. Atkinson, and F. Bourguignon, pp. 141 – 216. Elsevier.
- ACID, S., L. M. DE CAMPOS, J. M. FERNÁNDEZ-LUNA, S. RODRÍGUEZ, J. M. RODRÍGUEZ, AND J. L. SALCEDO (2004): “A comparison of learning algorithms for Bayesian networks: a case study based on data from an emergency medical service,” *Artificial Intelligence in Medicine*, 30, 215–232.
- ALKIRE, S., AND J. E. FOSTER (2011): “Counting and Multidimensional Poverty Measurement,” *Journal of Public Economics*, 95(7-8).
- ANDERSON, R. D., R. D. MACKOY, V. B. THOMPSON, AND G. HARRELL (2004): “A Bayesian Network Estimation of the Service-Profit Chain for Transport Service Satisfaction,” *Decision Sciences*, 34(4), 665–689.
- BOSMANS, K., K. DECANCQ, AND E. OOGHE (2015): “What do normative indices of multidimensional inequality really measure?,” *Journal of Public Economics*, 130, 94 – 104.
- CANADIAN INDEX OF WELLBEING (2012): *How are Canadians Really Doing?* Canadian Index of Wellbeing and University of Waterloo. The 2012 CIW Report, Waterloo, ON.
- CLARK, A. (forthcoming): “SWB as a Measure of Individual Well-Being,” in *Oxford Handbook of Well-Being and Public Policy*, ed. by M. Adler, and M. Fleurbaey, Oxford. Oxford University Press.
- CLARK, A. E., AND A. OSWALD (1994): “Unhappiness and Unemployment,” *Economic Journal*, 104, 648–659.
- CNEL, AND ISTAT (2015): *Report on Equitable and Sustainable Wellbeing (BES 2014)*. Istat, Rome.
- COOPER, G. (1997): “A simple constraint-based algorithm for efficiently mining observational databases for causal relationships,” *Data Mining and Knowledge Discovery*, 1, 203–224.
- COOPER, G. F., AND E. HERSKOVITS (1992): “A Bayesian method for the induction of probabilistic networks from data,” *Machine Learning*, 9(4), 309–347.
- COWELL, R. G., A. P. DAWID, S. L. LAURITZEN, AND D. J. SPIEGELHALTER (1999): *Probabilistic Networks and Expert Systems*. Springer-Verlag.
- DECANCQ, K., AND M. A. LUGO (2013): “Weights in Multidimensional Indices of Wellbeing: An Overview,” *Econometric Reviews*, 32(1), 7–34.

- EASTERLIN, R. A. (2001): “Income and happiness: towards a unified theory,” *The Economic Journal*, 111, 465–484.
- FERRER-I-CARBONELL, A. (2005): “Income and well-being: an empirical analysis of the comparison income effect,” *Journal of Public Economics*, 89, 997–1019.
- FRIEDMAN, N., M. LINIAL, I. NACHMAN, AND D. PEÉR (2000): “Using Bayesian Networks to Analyze Expression Data,” *Journal of Computational Biology*, 7(3-4), 601–620.
- HECKERMAN, D., D. GEIGER, AND D. M. CHICKERING (1995): “Learning Bayesian Networks: The Combination of Knowledge and Statistical Data,” *Machine Learning*, 20, 197–243.
- JENSEN, F. V. (2001): *Bayesian Networks and Decision Graphs*. Springer-Verlag.
- MARGARITIS, D. (2003): “Learning Bayesian network model structure from data,” Ph.D. thesis, Carnegie-Mellon University, School of Computer Science, Pittsburgh.
- NAGARAJAN, R., M. SCUTARI, AND S. LÈBRE (2013): *Bayesian Networks in R with Applications in Systems Biology*. Springer.
- NAGL, S., M. WILLIAMS, AND J. WILLIAMSON (2008): “Objective Bayesian Nets for Systems Modelling and Prognosis in Breast Cancer,” in *Innovations in Bayesian Networks*, vol. SCI 156. Springer-Verlag.
- NEF (2013): “The Happy Planet Index: 2012 Report. A global index of sustainable well-being,” .
- OECD (2008): *Handbook on constructing composite indicators: Methodology and User guide*. OECD publishing, Paris.
- (2015): *How’s Life? 2015: Measuring Well-being*. OECD Publishing, Paris.
- OFFICE FOR NATIONAL STATISTICS (2015): “Measuring National Wellbeing: Personal Well-being in the UK, 2014 to 2015,” .
- PEARL, J., AND S. RUSSEL (2003): “Bayesian networks,” in *Handbook of Brain Theory and Neural Networks*, ed. by M. Arbib, pp. 157–160. MIT Press.
- PITTAU, M. G., R. ZELLI, AND A. GELMAN (2010): “Economic Disparities and Life Satisfaction in European Regions,” *Social Indicators Research*, 96, 339–361.
- SALINI, S., AND R. S. KENETT (2009): “Bayesian networks of customer satisfaction survey data,” *Journal of Applied Statistics*, 36(11), 1177–1189.
- SCUTARI, M. (2010): “Learning Bayesian Networks with the bnlearn R Package,” *Journal of Statistical Software*, 35(2).

- SEN, A. K. (1980): “Equality of What?,” *The Tanner Lecture on Human Values*, I, 197–220.
- (1985): *Commodities and capabilities*. North-Holland, Amsterdam.
- SPIRITES, P., C. GLYMOUR, AND R. SCHEINES (1993): *Causation, Prediction, and Search*. Springer Verlag, New York.
- STEPTOE, A., A. DEATON, AND A. A. STONE (2015): “Subjective wellbeing, health, and ageing,” *The Lancet*, 385(9968), 640–648.
- STIGLITZ, J. E., A. SEN, AND J.-P. FITOUSSI (2010): “Report by the Commission on the Measurement of Economic Performance and Social Progress,” Discussion paper, Institut National de la Statistique at des études économiques.
- WINKELMANN, L., AND R. WINKELMANN (1998): “Why Are the Unemployed So Unhappy? Evidence from Panel Data,” *Economica*, 65(257), 1–15.

Appendix: Tables and Figures

Table 1: Well-being Dimensions and their definitions

Well-being dimension	Lits variables and definition	label
Material living standards	Poverty. Dummy equal to 1 if the monthly expenditure of the household (q2.22-q2.24) plus household savings (q2.24) is less than 60 percent of the median household's	M_POOR
	Material deprivation (q2.25), equal to the number of material dimensions not owned by the households (car, mobile, computer, internet), from 0 to 4	M_MMD
	Number of crisis pathways (q804) (reduced consumption of staple foods, luxury goods, alcohol and tobacco, reduced vacations, postponed or skipped visits to the doctor after falling ill, stopped buying regular medications, delayed payments on utilities (gas, water, electric), delayed or defaulted on a loan installment, cut TV / phone / internet service, had utilities cut)	M_CRISIS
Health	Self assessed health status (q7.04): 1-Very Bad, 2-Bad, 3-Medium, 4-Good, 5-Very Good	HEALTH
Education	Highest level of education attained (q5.15): 1-No education, 2-Primary, 3-Lower Secondary, 4-Upper Secondary, 5- Tertiary	EDU
Personal activities and work	Status in employment (q5.01 - q5.22 - q5.24): 1-Employed, 2-Unemployed, 3-Student, 4-Housewife, 5-Retired, 6-Disabled, 7-Other	WORK
Political voice and governance	Voice and Governance (q3.12): To what extent do you agree that [Free and fair elections, Law and Order, Freedom of speech, Peace and stability, A press that is independent from the government, A courts system that treats all citizens equally] exists in your country (1-Strongly disagree, 2-Disagree, 3-Neither disagree nor agree, 4-Agree, 5-Strongly agree), Average score of the seven indicators	POL
Participation to social activities and civil society	Civil Participation: Dummy equal to 1 if the individual voted in the most recent elections (q3.19), and/or is currently a member of a political party (q7.12), and/or belongs to some voluntary organizations (q7.13)	CIV
Social connections and relationships	Social Connection: How often do you meet up with relatives who are not living with you (q3.24) or friends (q3.25) (5-On most days, 4 Once or twice a week, 3 Once or twice a month, 2 Less often than once a month, 1 Never), highest score for the two variables	SOC

Security, Physic Nature:	Safety (q3.23): Suppose you lost your (purse/wallet) containing your address details, and it was found in the street by someone living in this neighborhood. How likely is it that it would be returned to you with nothing missing? (Recoded to 1-Not at all likely, 2- Not very likely, 3-Quite likely, 4-Very likely)	S.SAFETY
Security, Economic Nature:	Economic Security: Please imagine a ten-step ladder where on the bottom, the first step, stand the poorest 10% people in our country, and on the highest step, the tenth, stand the richest 10% of people in our country. On which step of the ten is your household today? (q2.27). Recoded from 1 (first 2 steps in the 10-step ladder) to 5 (steps 9 and 10 in the 10-step ladder).	S.TODAY
	Economic Security relative to the past: Please imagine a ten-step ladder where on the bottom, the first step, stand the poorest 10% people in our country, and on the highest step, the tenth, stand the richest 10% of people in our country. On which step of the ten is your household today? (q2.27) On which step was your household 4 years ago? (q2.28). Dummy equal to 1 if the step referred to today's situation is higher than the step referred to the past	S.PAST
	Expected economic security: Please imagine a ten-step ladder where on the bottom, the first step, stand the poorest 10% people in our country, and on the highest step, the tenth, stand the richest 10% of people in our country. On which step of the ten is your household today? (q2.27). Where on the ladder do you believe your household will be 4 years from now? (q2.29). Dummy equal to 1 if the step referred to future situation is higher than the step referred to today's	S.FUTURE
Personal Satisfac- tion:	Satisfaction with Life: All things considered, I am satisfied with my life now (1- strongly disagree, 2-Disagree, 3-Neither disagree nor agree, 4- Agree, 5- Strongly Agree (q3.01, e)	SA.LIFE
	Satisfaction with Job: All things considered, I am satisfied with my job as a whole (1- strongly disagree, 2-Disagree, 3-Neither disagree nor agree, 4- Agree, 5- Strongly Agree (q3.01, j)	SA.JOB
	Satisfaction with Finances: All things considered, I am satisfied with my financial situation as a whole (1- strongly disagree, 2-Disagree, 3-Neither disagree nor agree, 4- Agree, 5- Strongly Agree (q3.01, k)	SA.FIN

Table 2: Logistic Regression on Satisfaction with Life

	WestEU	EastEU	Overall
Material Living Standards ^a	-0.104*	-0.230***	-0.155***
Self Assessed Health Status	0.284***	-0.155***	0.288***
Level of Education	0.092**	0.031	0.069**
Status in Employment			
Employed (omitted category)			
Unemployed	-0.408**	-0.069	-0.236*
Retired	0.021	0.363**	0.203*
Out of the Labor Force	-0.050	0.235*	0.079
Voice and Governance			
Strongly Disagree (omitted category)			
Disagree	0.625	0.540*	0.552*
Neither disagree nor agree	0.863*	0.628*	0.672**
Agree	1.270**	0.642*	0.905***
Strongly Agree	1.374***	0.832**	1.054***
Civil Participation	0.090	0.228*	0.180*
Social Connections			
Never (omitted category)			
Less than once a month	-0.573	0.166	-0.224
Once or twice a month	-0.380	0.361	-0.034
Once or twice a week	-0.294	0.507	0.073
On most days	-0.246	0.369	0.037
Security ^b	0.085*	0.292***	0.190***
Satisfaction with job	0.554***	0.654***	0.603***
Satisfaction with finances	0.893***	0.800***	0.838***
Country Group			
WestEU (omitted category)	-	-	
East EU	-	-	-0.777***
Constant	-6.152***	-7.678***	-6.415***
Pseudo R2	0.273	0.257	0.337
Observation	5503	5162	10665

Note: Age and Household size included as controls in each model (not shown). Asterisks denote p-values: * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$.

^a First Principal Component of a principal component analysis including the following variables: poverty, material deprivation and number of crisis pathways.

^b First Principal Component of a principal component analysis including the following variables: economic security, economic security relative to the past, expected economic security

Table 3: Network Evaluation based on the Bayesian Information Criterion

	Bayesian Network score method		
	<i>K2</i>	AIC	BIC
WestEU	-105460	-106588	-103649
EastEU	-110969	-108291	-105993

Note: the Bayesian Information Criterion is here used both for constructing the Bayesian Network (Bayesian Network score method) and to evaluate the goodness of fit of the overall Bayesian Network (Network evaluation).

Table 4: Network Validation for selected response variables (Prediction Error, %)

Response Variable	Bayesian Network score method		
	<i>k2</i>	AIC	BIC
Expected economic security (WestEU)	18.26	18.20	18.51
Expected economic security (EastEU)	19.90	19.80	20.21
Economic security relative to the past (WestEU)	18.25	18.19	18.52
Economic security relative to the past (EastEU)	19.90	19.81	20.21
Satisfaction with life (WestEU)	18.26	18.19	18.52
Satisfaction with life (EastEU)	19.90	19.81	20.22
Satisfaction with finances (WestEU)	18.26	18.19	18.52
Satisfaction with finances (EastEU)	19.91	19.82	20.21
Average (WestEU)	18.20	18.13	18.38
Average (EastEU)	19.78	19.85	20.09

Table 5: Choice of policy interventions (backward probabilistic assessment), WestEU

			Baseline	Diagnostic
TARGET	Life Satisfaction	Strongly satisfied	16	100
	Poverty	0	84	87
	Education	No Education	6	6
		Primary	5	5
		Lower Secondary	24	22
		Upper Secondary	26	27
		Tertiary	38	41
	Status in Employment	Disabled	3	1
		Employed	58	65
		Housewife	5	3
INPUT		Other	3	3
		Retired	24	24
		Student	2	1
	Unemployed	5	2	
	Number of Crisis Pathways	0	46	60
		1	22	19
		2	18	12
		3	8	4
		4	4	2
		5	2	1
		6	1	0

Table 6: Choice of policy interventions (backward probabilistic assessment), EastEU

			Baseline	Diagnostic
TARGET	Life Satisfaction	Strongly satisfied	5	100
		No Education	4	3
		Primary	14	10
	Education	Lower Secondary	20	19
		Upper Secondary	32	36
		Tertiary	31	35
		Disabled	2	1
		Employed	48	57
		Housewife	5	5
INPUT	Status in Employment	Other	5	4
		Retired	29	26
		Student	3	2
		Unemployed	7	8
		0	37	47
		1	16	17
	Material Deprivation	2	14	13
		3	19	15
		4	15	10

Table 7: Assessment of the impact of policy interventions on the target variable (forward probabilistic assessment), WestEU

			Baseline	Interventions		
				a	b	c
INPUTS	Poverty	0	84	100		
	# of Crisis Pathways	0	46		100	
	Education	Tertiary	39			100
TARGET	Life Satisfaction	Strongly dissatisfied	4	3	1	3
		Dissatisfied	12	12	7	11
		Neither Satisfied nor dissatisfied	15	14	11	15
		Satisfied	54	54	59	53
		Strongly satisfied	16	17	22	18

Table 8: Assessment of the impact of policy interventions on the target variable (forward probabilistic assessment), EastEU

			Baseline	Interventions		
				a	b	c
INPUTS	Poverty	0	84	100		
	Material Deprivation	0	37		100	
	Education	Tertiary	31			100
TARGET	Life Satisfaction	Strongly dissatisfied	15	15	12	14
		Dissatisfied	23	22	21	22
		Neither Satisfied nor dissatisfied	29	29	29	28
		Satisfied	28	29	32	29
		Strongly satisfied	5	6	7	5

Table 9: WestEU, scores for Bayesian Network k2

from	to	strength
age	WORK	-1793.02
SA_FIN	SA_LIFE	-926.84
age	hhsiz	-577.65
SA_FIN	M_CRISIS	-555.76
S_TODAY	SA_FIN	-515.05
WORK	SA_JOB	-418.97
SA_FIN	SA_JOB	-395.87
hhsiz	M_POOR	-333.16
age	EDU	-282.64
age	S_FUTURE	-229.62
WORK	M_MD	-198.59
M_POOR	M_MD	-189.38
SA_LIFE	POL	-167.96
M_MD	S_TODAY	-155.34
S_PAST	S_TODAY	-145.45
WORK	SA_FIN	-145.04
S_FUTURE	S_TODAY	-136.72
WORK	HEALTH	-126.84
EDU	M_MD	-115.97
WORK	S_PAST	-94.00
SA_JOB	SA_LIFE	-87.48
EDU	WORK	-86.60
SA_FIN	S_SAFETY	-57.69
S_SAFETY	CIV	-57.20
age	HEALTH	-52.14
age	SOC	-47.68
S_TODAY	CIV	-43.18
age	M_CRISIS	-40.15
EDU	S_SAFETY	-36.81
S_PAST	S_FUTURE	-33.23
M_POOR	S_FUTURE	-29.49
EDU	M_POOR	-26.68
EDU	S_PAST	-25.64
CIV	POL	-7.22
EDU	SOC	-5.47

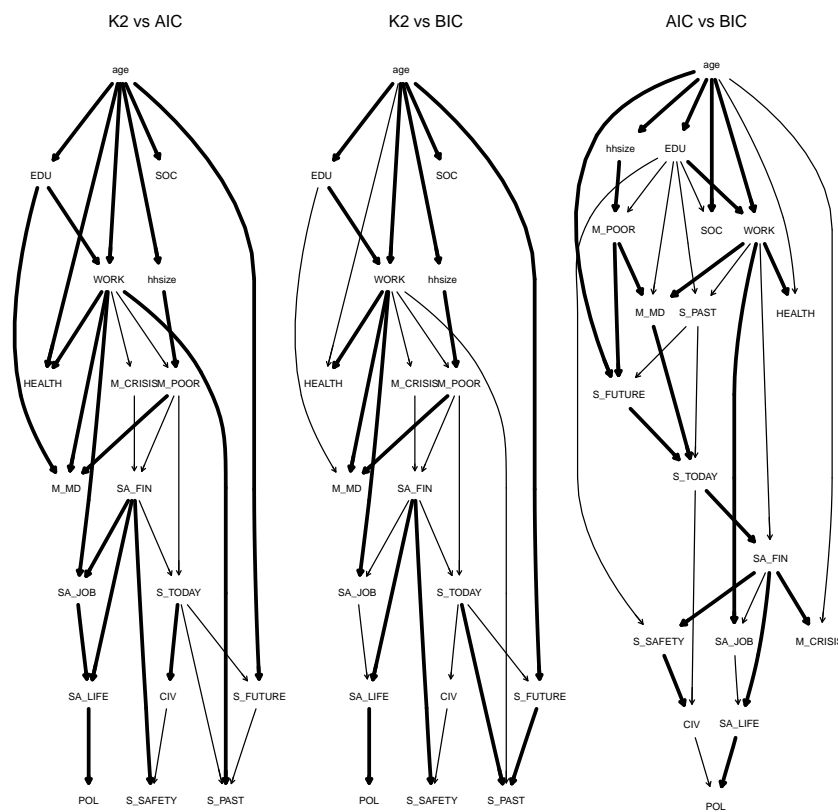
Table 10: EastEU, scores for Bayesian Network k2

from	to	strength
age	WORK	-1529.44
SA_FIN	SA_LIFE	-878.56
S_TODAY	SA_FIN	-576.38
age	hhsiz	-484.75
M_MD	S_TODAY	-436.50
SA_FIN	SA_JOB	-388.41
age	EDU	-307.18
SA_FIN	M_CRISIS	-288.89
WORK	SA_JOB	-173.93
S_TODAY	S_PAST	-170.17
age	HEALTH	-154.50
age	M_MD	-154.29
hhsiz	M_MD	-153.75
M_MD	M_POOR	-151.56
EDU	M_MD	-150.83
SA_LIFE	POL	-114.84
SA_JOB	SA_LIFE	-113.43
WORK	HEALTH	-100.61
SA_JOB	S_SAFETY	-82.99
age	SOC	-52.03
age	S_FUTURE	-49.92
S_FUTURE	S_TODAY	-44.61
hhsiz	M_POOR	-44.61
S_FUTURE	S_PAST	-44.40
age	CIV	-32.07
CIV	S_SAFETY	-6.70
EDU	WORK	22.63

	East EU					West EU					Average	
	Bulgaria	Croatia	Hungary	Romania	Slovakia	Average	France	Germany	Great Britain	Italy		Sweden
Household Size	2.42	2.69	2.14	2.44	2.92	2.45	2.36	2.01	2.03	2.64	1.96	2.27
Age	3.65	4.59	5.27	6.28	12.84	6.23	2.97	3.27	3.93	3.61	1.62	3.37
18-24	13.23	16.48	12.32	18.39	28.77	17.30	11.91	11.84	12.19	12.04	6.01	11.84
25-34	16.43	17.98	14.36	18.84	23.44	17.97	20.81	20.72	14.84	26.61	17.31	20.55
35-44	17.46	16.78	12.74	14.43	21.78	15.58	20.28	24.05	12.65	20.43	15.12	19.49
45-54	16.35	15.83	17.47	14.30	9.94	14.93	17.89	13.44	10.46	13.39	18.62	13.87
55-64	32.88	28.35	37.83	27.76	3.23	28.00	26.14	26.70	45.94	23.93	38.91	30.88
65+	19.31	19.86	13.99	21.17	10.72	18.15	12.08	16.87	20.32	12.35	22.02	15.89
Poor	28.92	44.30	25.37	21.74	60.95	29.89	66.42	57.59	50.62	53.48	73.60	57.55
Material Deprivation	14.24	9.70	14.39	17.26	10.96	11.34	10.96	11.34	13.24	11.48	18.10	12.01
Deprived in 0 dimensions	15.26	17.30	11.49	12.07	9.05	12.58	12.48	14.42	12.97	23.24	3.24	15.18
Deprived in 1 dimension	24.45	13.99	24.96	23.57	8.95	21.56	6.73	10.14	14.30	8.77	4.18	9.85
Deprived in 2 dimensions	17.13	14.71	23.78	21.92	1.85	18.71	3.41	6.50	8.87	3.04	0.88	5.41
Deprived in 3 dimensions	6.02	24.97	16.66	19.79	50.50	20.79	41.46	57.12	56.87	12.83	76.82	44.84
Deprived in 4 dimensions	1	14.83	25.45	26.11	18.42	19.07	23.62	20.12	17.99	34.44	12.18	23.20
Number of crisis pathways	2	28.17	21.61	24.88	23.63	17.40	23.63	21.47	13.69	31.01	6.52	18.85
	3	29.11	14.64	15.53	20.21	8.04	8.52	5.50	6.65	13.61	2.65	8.04
	4	14.69	10.14	10.55	11.93	3.77	3.31	2.34	3.40	4.67	1.14	3.25
	5	7.17	3.20	6.27	6.28	1.23	1.62	1.16	1.49	3.44	0.69	1.82
	6					5.60						
Self Assessed Health Status	2.98	4.27	5.18	5.06	0.56	4.19	1.20	1.06	1.38	0.47	0.83	1.02
Very Bad	13.82	11.29	16.49	20.95	4.26	16.24	4.11	5.65	9.38	4.01	6.61	5.86
Bad	36.04	31.13	39.93	30.48	23.29	32.57	24.19	27.03	24.98	30.10	16.48	26.22
Medium	36.01	34.84	31.30	33.80	47.59	35.24	54.11	51.27	42.98	48.36	45.41	49.06
Good	11.15	18.46	7.10	9.71	24.29	11.77	16.39	15.00	21.27	17.06	30.66	17.83
Very Good	0.83	7.85	0.97	7.76	2.58	4.71	15.53	2.28	9.59	1.48	6.90	6.69
Highest Level of Education Attained	1.4	10.32	30.01	22.71	5.11	18.25	9.64	1.70	2.28	14.65	2.18	6.35
None or incomplete	18.04	9.84	25.41	19.92	24.49	20.36	14.10	34.26	32.02	29.34	10.11	27.52
Primary	22.45	49.98	15.98	23.70	45.13	26.57	11.20	38.51	22.04	41.89	10.14	28.71
Lower Secondary	54.55	22.01	27.64	25.92	22.70	30.10	49.52	23.25	34.09	12.64	70.66	30.73
Upper Secondary	44.16	43.81	40.42	41.54	71.90	45.25	56.93	58.43	41.01	53.92	60.78	53.22
Tertiary	8.15	8.11	4.47	7.60	6.34	6.96	3.60	5.62	4.58	6.13	1.81	4.92
Employed	37.46	29.59	46.24	36.86	8.17	35.09	31.48	29.40	43.55	21.44	31.55	31.39
Unemployed	10.23	18.49	8.87	14.00	13.59	12.70	7.98	6.56	10.86	18.51	5.86	10.46
Retired	1	1.70	4.48	5.44	4.40	4.40	0.31	1.08	0.44	1.39	0.18	0.81
Out of the Labour Force	2	28.77	24.69	15.79	28.97	7.13	7.13	2.69	5.20	15.86	0.61	7.01
	3	54.41	48.82	37.64	47.41	49.30	34.72	17.63	32.04	39.90	9.20	29.04
	4	13.96	17.83	32.62	15.58	35.49	41.70	54.23	62.56	31.86	46.00	48.97
	5	1.15	4.18	8.52	2.27	3.43	16.15	24.37	9.76	10.99	44.01	17.17
Civil Participation	82.82	89.76	78.88	87.22	72.59	83.44	87.97	90.24	87.53	88.71	89.46	89.20
Social Connections	0.23	0.14	0.47	0.20	0.19	0.25	0.40	0.26	0.30	0.09	0.00	0.25
Never	6.98	9.72	10.14	15.78	2.51	11.25	7.33	5.67	3.70	2.28	2.29	4.69
Less than once a month	16.60	19.44	22.12	20.53	18.07	19.78	25.42	21.47	12.54	6.34	19.27	17.45
Once or twice a week	41.08	42.92	40.90	35.12	45.19	39.03	44.95	50.08	50.79	43.34	52.69	47.60
On most days	35.11	27.78	26.37	28.57	34.03	29.68	21.91	22.51	32.56	44.95	25.75	29.61
Safety	36.02	16.49	44.33	54.08	38.87	44.27	8.69	14.16	15.70	29.13	8.03	16.46
Not at all safe	33.98	25.44	42.23	28.63	39.25	33.14	24.10	43.70	23.34	35.08	26.37	32.51
Not very safe	20.86	34.32	10.62	13.20	17.53	16.21	36.70	29.21	45.51	26.10	45.30	34.87
Quite safe	9.13	23.75	2.81	4.09	4.35	6.38	28.51	12.93	15.45	9.11	20.30	16.16
Very safe	2.11	2.56	2.21	2.28	2.89	2.33	2.62	2.79	2.50	2.89	3.02	2.72
Economic security	8.88	9.94	7.34	14.60	24.96	12.94	21.37	17.73	16.71	13.17	29.70	17.73
Economic security relative to the past	37.30	30.99	26.84	20.09	31.04	26.35	22.15	17.82	27.76	14.44	27.26	20.63
Expected economic security	11.68	12.17	27.96	21.08	2.39	18.17	3.74	2.38	2.67	6.14	0.93	3.50
Satisfaction with life	28.63	14.85	19.68	34.37	15.21	26.65	10.83	8.10	14.55	17.37	2.98	11.98
Strongly Dissatisfied	26.46	26.69	33.55	26.68	30.29	28.45	13.92	13.59	10.38	25.78	6.39	15.34
Dissatisfied	31.09	36.83	14.23	15.81	44.61	22.90	58.20	55.71	58.68	45.14	53.85	54.48
Neither satisfied nor dissatisfied	2.15	9.46	4.57	2.07	7.51	3.84	13.31	20.22	13.71	5.58	35.64	14.70
Satisfied	1.85	2.40	9.67	8.99	2.40	6.70	1.87	0.60	1.44	3.53	0.88	1.72
Strongly Dissatisfied	5.47	2.68	8.16	9.11	10.08	7.88	6.36	2.91	4.78	7.65	3.34	5.12
Dissatisfied	8.67	8.52	62.60	64.22	24.54	45.87	6.65	8.14	4.82	14.90	4.06	8.40
Neither satisfied nor dissatisfied	75.46	79.08	12.03	15.38	55.93	34.22	72.07	68.73	81.81	66.61	70.86	72.03
Satisfied	8.55	7.32	7.55	2.31	7.04	5.32	13.05	19.62	7.14	7.31	20.86	12.73
Strongly Satisfied	19.96	28.97	35.89	28.77	10.97	26.90	8.32	6.66	8.86	11.89	4.60	8.58
Satisfaction with finances	44.84	25.82	24.82	32.00	26.06	31.37	24.22	13.22	24.90	23.31	10.25	19.66
Strongly Dissatisfied	20.28	23.01	27.49	24.57	35.64	25.56	15.80	17.43	10.73	26.65	17.41	17.41
Dissatisfied	11.80	18.18	8.91	12.63	23.42	13.36	47.51	48.50	48.80	35.29	49.00	45.46
Neither satisfied nor dissatisfied	3.12	4.02	2.89	2.12	3.91	2.80	7.10	14.19	6.71	2.86	23.40	8.89
Satisfied												
Strongly Satisfied												

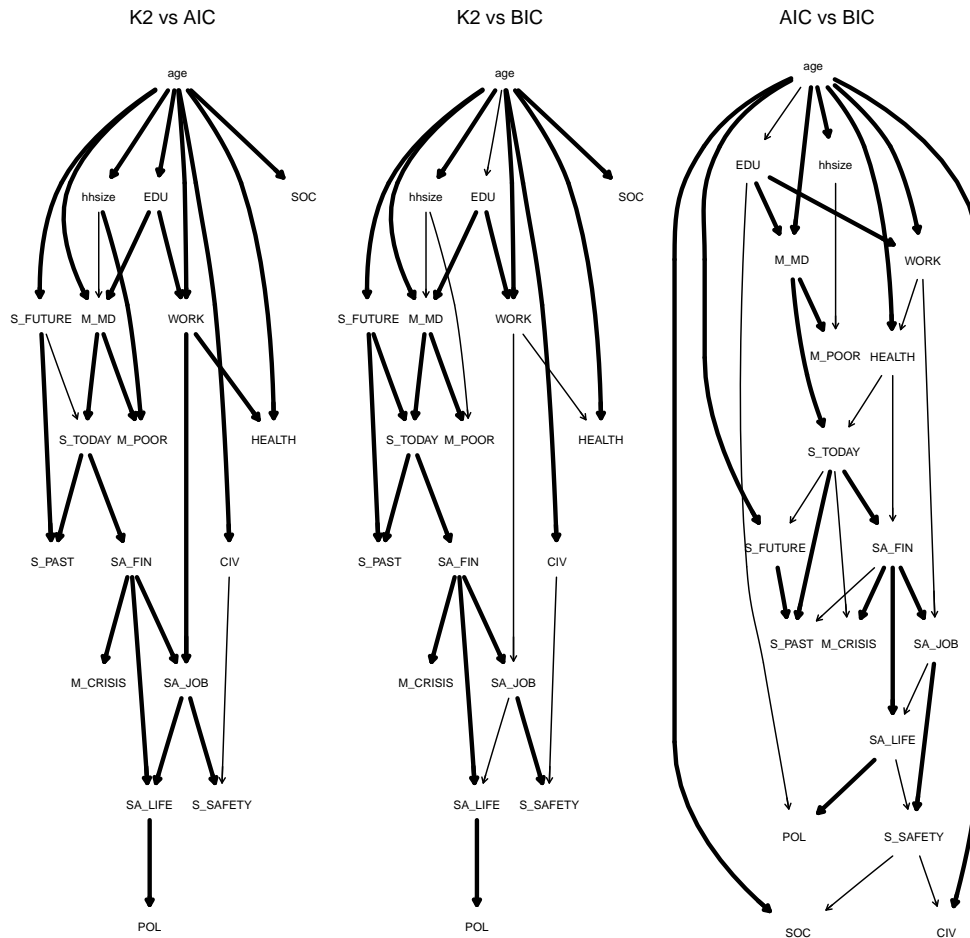
Figure 1: Summary Statistics

Figure 2: Bayesian Network, comparing K2 score with the AIC score (left), K2 score with the BIC score (center) and AIC score with BIC score (right), WestEU countries



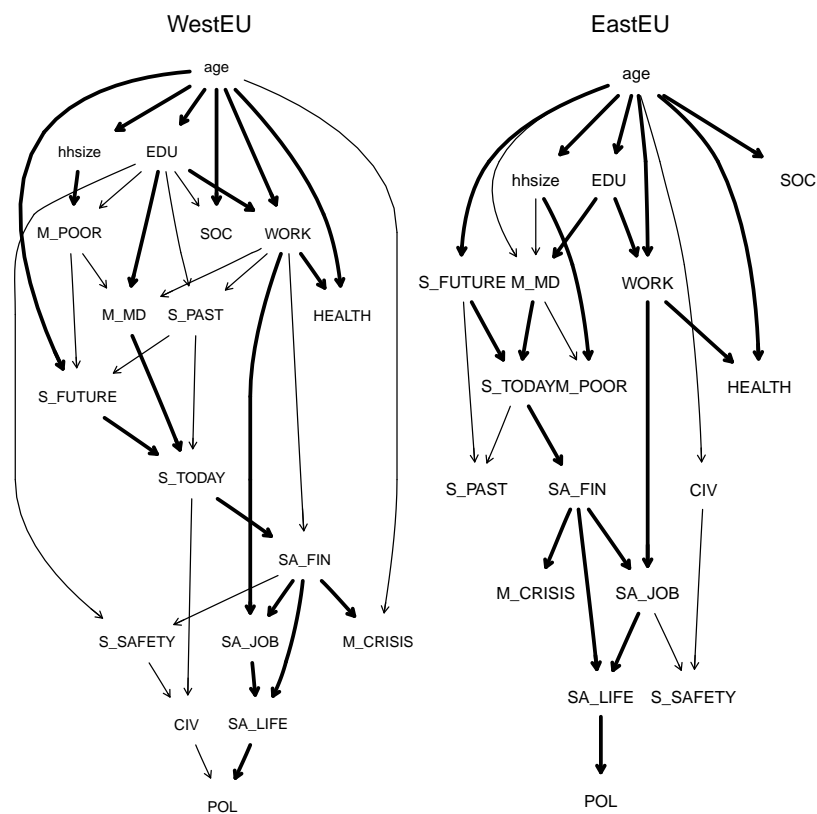
Note: Thick lines indicate arcs common to the graphs defined by both score functions, the thin lines arcs are present only in the graph defined by the first score function.

Figure 3: Bayesian Network, comparing K2 score with the AIC score (left), K2 score with the BIC score (center) and AIC score with BIC score (right), EastEU countries



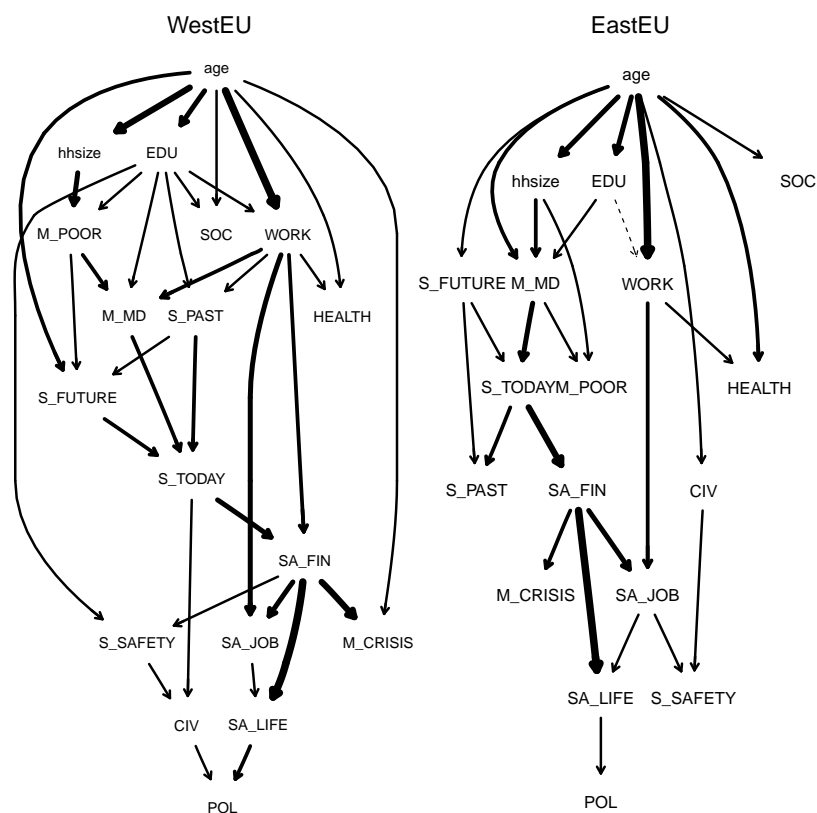
Note: Thick lines indicate arcs common to the graphs defined by both score functions, the thin lines arcs are present only in the graph defined by the first score function.

Figure 4: Bayesian Network, comparing WestEU countries (left, AIC score) with EastEU countries (right, K2 score).



Note: Thick lines indicate arcs common to both groups of countries, the thin lines arcs are present only in the graph of the group of countries defined above.

Figure 5: Bayesian Network, comparing arcs' strength in WestEU countries (left, AIC score) with EastEU countries (left, K2 score).



Note: Thicker lines correspond to arcs which, if removed, would cause a higher loss to the model. Dotted lines, on the other hand, could be removed to improve the model overall score.