KYBERNETIKA — VOLUME *62* (2026), NUMBER 3, PAGES 400–426

ENHANCING BAYESIAN NETWORKS THROUGH LOGISTIC REGRESSION: A CASE STUDY OF CZECH SOCIETY DIVISIONS

Jiří Vomlel

One of the most frequently employed criteria for the structural learning of Bayesian networks (BNs) is the Bayesian Information Criterion (BIC). The objective is to identify a model that exhibits an optimal balance between the fit to the training data, as measured by the log-likelihood, and the model's complexity, as quantified by the number of model parameters. A significant challenge associated with this approach pertains to the exponential increase in the number of parameters required for conditional probability tables (CPTs) as the number of parents of each node increases. This phenomenon leads to a substantial growth in the complexity of the CPTs, which serve as the fundamental building blocks of BNs. However, there exist models of CPTs whose number of parameters grows linearly with the number of parents, and they often represent a better fit to data than general CPTs. In this paper, we examine models with CPTs, either in their general form or in the form corresponding to multinomial logistic regression (MLR) or ordinal logistic regression (OLR). We employ data from a sociological study entitled "Dividing Lines in Czech Society" to demonstrate the enhancement through the incorporation of MLR and OLR models as CPTs within the framework of BN structural learning.

Keywords: Bayesian networks, Bayesian information criterion, sociology, structural learning

Classification: 68T37, 62P25

1. INTRODUCTION

Bayesian networks [9, 10, 11] have found applications across diverse domains, with their interpretability making them particularly attractive for modeling in the social sciences. We argue that Bayesian networks can provide deeper insights into complex phenomena than conventional statistical techniques or standard machine learning approaches [16].

A Bayesian network (BN) consists of two components: a directed acyclic graph that encodes conditional independence relationships among variables and conditional probability tables (CPTs) that quantify these relationships for each variable given its parents.

DOI: 10.14736/kyb-2026-3-0400
Published by the Institute of Information Theory and Automation of the Czech Academy of Sciences.

This factorization offers significant computational advantages, as CPTs can be learned locally using data restricted to each node and its parents (the node's family). This local learning property makes BN training efficient for sparse graphs, though structural learning remains NP-hard in general. However, a fundamental limitation of BN models is the exponential growth of CPT parameters with the number of parent nodes. This exponential scaling not only increases computational complexity but also reduces parameter estimation reliability when training data are limited.

This paper addresses the scalability challenge by proposing structured alternatives to general CPTs that remain tractable even with many parents. Specifically, we investigate multinomial logistic regression (MLR) and ordinal logistic regression (OLR) as compact CPT representations. Strictly speaking, these compact representations are not tables but alternative parameterizations of conditional probability distributions. However, since MLR and OLR can always be represented using a CPT in Bayesian networks (as defined in formulas (13) and (16)), we will use the term CPT in a generalized sense to denote any representation of a conditional probability distribution. Using logistic regression models for CPTs is not a novel contribution, but to the best of our knowledge, the optional selection of OLR, MLR, and standard CPTs has not been proposed before. Our approach aligns with previous work [12, 14, 15] advocating for BNs with local structure in their CPTs, which can accommodate larger CPTs while maintaining low parameter counts, thereby improving model log-likelihood without overfitting. We empirically evaluate the viability of MLR and OLR as CPT approximations using real data from a comprehensive sociological survey.

The remainder of this paper is structured as follows. In Section 2, we present the standard BN structure learning approach, while in Section 3, we enhance it with the structured CPT representations. Section 4 introduces the survey data and applies the enhanced BN learning to it. The quality of MLR and OLR approximations for the learned CPTs is assessed. In Section 5, we provide a detailed analysis of the learned model, accompanied by a characterization of the phenomena represented by individual CPTs and their visualizations in the appendix. We conclude with a summary of our findings and directions for future research in Section 6.

2. BAYESIAN NETWORK MODELS

In this section, we review the standard approach to BN learning. More details can be found, e. g., in [3, 10]. Let $V = \{1, \ldots, d\}$ denote the index set for random variables $X_v, v \in V$, where each variable X_v takes values x_v from a finite state space $\mathcal{X}_v$. Consider a BN model that represents a joint probability distribution P, assigning a probability $P(\mathbf{x})$ to each realization $\mathbf{x} = (x_1, \ldots, x_d)$ of the multidimensional random variable $\mathbf{X} = (X_1, \ldots, X_d)$. The network structure is defined by a directed acyclic graph G, which induces a parent function $pa(v)$ that maps each node v to its set of parent nodes in G. Specifically, node u is a parent of node v if there exists a directed edge $u \to v$ in G. Throughout this work, we use nodes and their corresponding variables interchangeably when the meaning is clear from context.

Let $\mathbf{D}$ represent a dataset consisting of data vectors $\mathbf{x} = (x_1, \ldots, x_d)$, which are realizations of the random variables $\mathbf{X} = (X_1, \ldots, X_d)$. For any $A \subseteq V$, we define an $|A|$-dimensional variable $\mathbf{X}_A$ indexed by A. Its possible values $\mathbf{x}_A$ belong to $\mathcal{X}_A = \times_{a \in A} \mathcal{X}_a$.

For convenience, when $A = \{v\} \cup U$ with $U \subset V \setminus \{v\}$, we write $\mathbf{X}_{v,U}$ instead of $\mathbf{X}_{\{v\}\cup U}$. The likelihood of observing independent and identically distributed data $\mathbf{D}$ under the BN model P is:

$$
\begin{aligned}
L(\mathbf{D}|P) &= \prod_{\mathbf{x}\in\mathbf{D}} P(\mathbf{x}) \\
&= \prod_{\mathbf{x}\in\mathbf{D}} \prod_{v\in V} P(x_v|\mathbf{x}_{pa(v)}) \ , \qquad (1)
\end{aligned}
$$

where the second equality follows from the probability chain rule combined with conditional independencies represented in the Bayesian network [10, Section 3.2.3.2]. If $pa(v) = \emptyset$ then $P(x_v|\mathbf{x}_{pa(v)}) = P(x_v)$.

For any subset $A \subseteq V$, let $N_A : \mathcal{X}_A \to \mathbb{N}_0$ denote the count function that returns the number of occurrences of configuration $\mathbf{x}_A \in \mathcal{X}_A = \times_{a\in A}\mathcal{X}_a$ in dataset $\mathbf{D}$. We define $fa(v) = \{v\} \cup pa(v)$ as the family of node v. Using formula (1), the log-likelihood[1] can be decomposed as follows:

$$
LL(\mathbf{D}|P) = \log \prod_{\mathbf{x}\in\mathbf{D}} P(\mathbf{x}) = \sum_{\mathbf{x}\in\mathbf{D}} \sum_{v\in V} \log P(x_v|\mathbf{x}_{pa(v)}) \qquad (2)
$$

Equivalently to summing over all data vectors, we can sum over all possible data vectors $\mathbf{x} \in \mathcal{X}$ with their multiplicity defined by the count function N_V:

$$
LL(\mathbf{D}|P) = \sum_{v\in V} \sum_{\mathbf{x}\in\mathcal{X}} N_V(\mathbf{x}) \cdot \log P(x_v|\mathbf{x}_{pa(v)}) \qquad (3)
$$

By changing the order of the summation, we obtain:

$$
\begin{aligned}
LL(\mathbf{D}|P) &= \sum_{v\in V} LL_v(P|\mathbf{D}) \ , \text{ where} \qquad (4) \\
LL_v(P|\mathbf{D}) &= \sum_{\mathbf{x}_{fa(v)}} N_{fa(v)}(\mathbf{x}_{fa(v)}) \cdot \log P(x_v|\mathbf{x}_{pa(v)}) \ . \qquad (5)
\end{aligned}
$$

This decomposition reveals that the log-likelihood of a BN has a local structure: it can be expressed as a sum of local contributions $LL_v(\mathbf{D}|P)$ for each node $v \in V$, where each local term depends only on the family $fa(v)$ consisting of node v and its parents $pa(v)$. In this paper, for simplicity, we assume all variables $X_v, v = 1, \ldots, d$ have the same number k of values $x_{v,i}$ indexed by integers i from the set $\{1, 2, \ldots, k\}$. Note that the total number of free parameters of a general CPT associated with node v is

$$
C_v = (k-1) \cdot k^n \ , \qquad (6)
$$

where $n = |pa(v)|$.

When learning BN structures from data $\mathbf{D}$, naive log-likelihood maximization produces overly dense graphs that overfit the training data. Since the complete graph always achieves the maximum log-likelihood, effective structure learning requires scoring functions that penalize model complexity. We employ the Bayesian Information Criterion

[1] In this paper, we always use the natural logarithm, which we denote log.

(BIC) [13], which balances model fit against complexity by penalizing the number of parameters. BIC[2] combines the log-likelihood $LL(P|\mathbf{D})$ with a penalty term proportional to the parameter count $C(P)$ of the BN representing the distribution P:

$$BIC(\mathbf{D}|P) \quad = \quad LL(P|\mathbf{D}) \quad - \quad \frac{\log|\mathbf{D}|}{2} \cdot C(P) \ , \tag{7}$$

where $|\mathbf{D}|$ is the number of observations in the data D. The complexity penalty $C(P)$ aggregates the parameter counts across all CPTs:

$$C(P) \quad = \quad \sum_{v \in V} C_v \ . \tag{8}$$

Due to BIC's decomposability, this penalization applies locally to individual node distributions:

$$BIC_v(\mathbf{D}|P) \quad = \quad LL_v(P|\mathbf{D}) - \frac{\log|\mathbf{D}|}{2} \cdot C_v \ . \tag{9}$$

3. ENHANCED BAYESIAN NETWORK LEARNING

Next, we present two alternative representations of a CPT defining conditional probabilities $P(x_{v,j}|\mathbf{x}_{pa(v)})$ for all values $x_{v,j} \in \mathcal{X}_v$ of the variable X_v and all configurations $\mathbf{x}_{pa(v)}$ of the parents of node X_v. Recall that variable X_v has k values $x_{v,i}$ indexed by integers i from the set $\{1, 2, \ldots, k\}$. One state of variable X_v is considered the reference state; say it is $x_{v,1}$ for all variables X_v in the BN model.

3.1. Multinomial logistic regression

In the Multinomial Logistic Regression (MLR) model (see, e. g., [1]), a vector of parameters $\boldsymbol{\beta}_j$ is specified for each $j \in \{2, \ldots, k\}$. It has one value for each non-reference value of each parent and an additional parameter called intercept:

$$\boldsymbol{\beta}_j \quad = \quad (\beta_{0,j}, \beta_{1,j}, \ldots, \beta_{n(k-1),j}) \tag{10}$$

where n is the number of parents of X_v. The vector $\boldsymbol{x}_{pa(v)}$ of parent values is transformed into another vector $\boldsymbol{\delta}_{pa(v)}$ of the same length as $\boldsymbol{\beta}_j$, starting with the constant one and further containing one value for each dummy variable defined for each non-reference state r of each parent variable X_p:

$$\delta(X_p = x_{p,r}) \quad = \quad \begin{cases} 1 & \text{if } X_p = x_{p,r} \\ 0 & \text{otherwise.} \end{cases} \tag{11}$$

All values of vector $\boldsymbol{\beta}_1$ are defined to be zero, i. e., $\boldsymbol{\beta}_1 = \mathbf{0}$. The number of parameters in the MLR model is thus

$$C'_v \quad = \quad (k-1) \cdot (1 + n \cdot (k-1)) \ , \tag{12}$$

[2] We note that, contrary to the machine learning and BN research communities, BIC is usually defined as $-2 \cdot LL(P|\mathbf{D}) + \log|\mathbf{D}| \cdot C(P)$ in statistics.

where again $n = |pa(v)|$. Compare this to the general CPT, which requires $(k-1)k^n$ parameters, see formula (6) in Section 2.

Then, for $j \in \{1, \ldots, k\}$, the conditional probability of variable X_v taking the state $x_{v,j}$ is defined as:

$$P(X_v = x_{v,j} | \boldsymbol{x}_{pa(v)}) \quad = \quad \frac{\exp(\boldsymbol{\beta}_j^T \cdot \boldsymbol{\delta}_{pa(v)})}{\sum_{i=1}^k \ \exp(\boldsymbol{\beta}_i^T \cdot \boldsymbol{\delta}_{pa(v)})} \ = \ \sigma\left(\boldsymbol{\beta}_i^T \cdot \boldsymbol{\delta}_{pa(v)}\right) \ , \tag{13}$$

where σ is the softmax function.

The above formula can be derived from the more common log-odds parameterization as follows. For $j \in \{1, \ldots, k\}$

$$\begin{aligned}
\log \frac{P(X_v = x_{v,j} | \boldsymbol{x}_{pa(v)})}{P(X_v = 1 | \boldsymbol{\delta}_{pa(v)})} &= \boldsymbol{\beta}_i^T \cdot \boldsymbol{\delta}_{pa(v)} \\
\frac{P(X_v = x_{v,j} | \boldsymbol{x}_{pa(v))}}{P(X_v = 1 | \boldsymbol{x}_{pa(v)})} &= \exp(\boldsymbol{\beta}_i^T \cdot \boldsymbol{\delta}_{pa(v)}) \\
P(X_v = x_{v,j} | \boldsymbol{\delta}_{pa(v)}) &= P(X_v = 1 | \boldsymbol{x}_{pa(v)}) \cdot \exp(\boldsymbol{\beta}_i^T \cdot \boldsymbol{\delta}_{pa(v)}). \qquad (14)
\end{aligned}$$

By enforcing normalization $\sum_{j=1}^k P(X_v = x_{v,j} | \boldsymbol{\delta}_{pa(v)}) = 1$, we obtain the formula for the reference state probability:

$$\begin{aligned}
P(X_v = 1 | \boldsymbol{\delta}_{pa(v)}) \cdot \left(1 + \sum_{i=2}^k \exp(\boldsymbol{\beta}_i^T \cdot \boldsymbol{\delta}_{pa(v)}\right) &= 1 \\
P(X_v = 1 | \boldsymbol{\delta}_{pa(v)}) &= \frac{1}{1 + \sum_{i=2}^k \exp(\boldsymbol{\beta}_i^T \cdot \boldsymbol{\delta}_{pa(v)})}. \qquad (15)
\end{aligned}$$

Using the definition $\boldsymbol{\beta}_1 = \mathbf{0}$ and substituting (15) back to (14) gives (13).

Example 3.1. (MLR) Assume a MLR model of C with $n = 2$ parents, namely A and B, each having $k = 3$ states, which is the number of states we actually used in the real-world application discussed in Section 4. The corresponding CPT of this MLR model is fully specified by these parameter vectors:

$$\begin{aligned}
\boldsymbol{\beta}_2 &= (\beta_{0,2}, \beta_{1,2}, \beta_{2,2}, \beta_{3,2}, \beta_{4,2}) \\
\boldsymbol{\beta}_3 &= (\beta_{0,3}, \beta_{1,3}, \beta_{2,3}, \beta_{3,3}, \beta_{4,3}) \ .
\end{aligned}$$

This means that this MLR model requires $2 \cdot (1 + 2n) = 2 \cdot 5 = 10$ parameters. Since, by the MLR definition

$$\boldsymbol{\beta}_1 \quad = \quad (0, 0, 0, 0, 0)$$

we can compute each entry $(a, b, c) \in \{1, 2, 3\}^3$ of the general CPT as

$$\begin{aligned}
&P(C = c | A = a, B = b) \\
&= \frac{\exp\left(\beta_{0,c} + \beta_{1,c} \cdot \delta(a = 2) + \beta_{2,c} \cdot \delta(a = 3) + \beta_{3,c} \cdot \delta(b = 2) + \beta_{4,c} \cdot \delta(b = 3)\right)}{\sum_{i=1}^3 \exp\left(\beta_{0,i} + \beta_{1,i} \cdot \delta(a = 2) + \beta_{2,i} \cdot \delta(a = 3) + \beta_{3,i} \cdot \delta(b = 2) + \beta_{4,i} \cdot \delta(b = 3)\right)} \ .
\end{aligned}$$

3.2. Ordinal logistic regression

In Ordinal Logistic Regression (OLR) (see, e. g. [1]), the number of parameters is further reduced. Each parent of X_v has its own parameter, which is independent of the value $x_{v,j}$ of X_v. Together, these parameters form a vector of parameters $\boldsymbol{\alpha}$ describing the strength of the influence of the parents on X_v. A specific parameter ζ_j is associated with each state $x_{v,j}$ of X_v. The zeta parameters correspond to the thresholds or cut-points that separate the different ordered values of the variable X_v. The zeta parameters must be ordered ($\zeta_1 < \ldots < \zeta_{k-1}$), because the values of X_v are assumed to have a natural order. They essentially define how "spread out" or "compressed" the categories are on the underlying scale. A large gap between consecutive ζ_i, ζ_{i+1} suggests that those two values are well-separated, while a small gap suggests that they are harder to distinguish.

For $j \in \{1, \ldots, k-1\}$, the conditional probability of variable X_v taking a state higher than $x_{v,j}$ is defined as:

$$P(X_v > x_{v,j} | \boldsymbol{x}_{pa(v)}) = \frac{\exp(\zeta_j + \boldsymbol{\alpha}^T \cdot \boldsymbol{x}_{pa(v)})}{1 + \exp(\zeta_j + \boldsymbol{\alpha}^T \cdot \boldsymbol{x}_{pa(v)})} \ . \tag{16}$$

The number of parameters in this model is thus:

$$C''_v = (k-1) + n \ , \tag{17}$$

where again $n = |pa(v)|$.

Example 3.2. (OLR) Assume an OLR model of the same CPT as in Example 3.1, i. e., for C with $n = 2$ parents, namely A and B, each having $k = 3$ states, which is the number of states we actually used in the real-world application discussed in Section 4. The corresponding CPT of this OLR model is fully specified by the following four parameters:

$$\zeta_1, \zeta_2, \alpha_1, \alpha_2 \ ,$$

from which we can compute for $c \in \{1, 2\}$ and $(a, b) \in \{1, 2, 3\}^2$

$$P(C > c | A = a, B = b) = \frac{\exp(\zeta_c + \alpha_1 \cdot a + \alpha_2 \cdot b)}{1 + \exp(\zeta_c + \alpha_1 \cdot a + \alpha_2 \cdot b)} \ .$$

This, together with

$$\begin{aligned} P(C > 0 | A = a, B = b) &= 1 \\ P(C > 3 | A = a, B = b) &= 0 \end{aligned}$$

specifies each entry $(a, b, c) \in \{1, 2, 3\}^3$ of the general CPT as

$$P(C = c | A = a, B = b) = P(C > c - 1 | A = a, B = b) - P(C > c | A = a, B = b) \ .$$

A natural generalization of the OLR models is Generalized additive models (GAMs), in which the dependent variable depends on the values of its parents transformed via a function. This would definitely allow for more flexibility in learning CPTs and enable a better fit. Note that MLR models with a dummy variable for each non-reference state used in this paper offer flexibility comparable to GAMs. The only difference is that they do not assume the ordinality of the dependent variable, while the GAM generalization of OLR would assume ordinality. However, we leave the study of these generalizations as a topic for future research.

3.3. Bayesian network learning

Algorithm 1 outlines our two-stage BN learning procedure. The first stage applies standard structural learning to maximize BIC under the assumption of general CPTs. In the second stage, each learned CPT is evaluated for potential replacement with MLR or OLR representations that achieve higher BIC scores, with the best-performing alternative selected for the final model.

input : $\mathbf{D}$ – training dataset consisting of $|\mathbf{D}|$ complete data vectors
output: G – the structure of BN with CPTs being either standard CPTs, MLR, or OLR models maximizing local BIC score

```
w = log|D|/2 ;                                   /* the penalty weight */
G = arg max_G BIC(D|G) ;            /* learn G using a standard CPTs */
for v ∈ V do
    s_STD = BIC_v^STD(D|P) ;                   /* BIC of general CPT */
    s_MLR = BIC_v^MLR(D|P) ;              /* BIC of MLR model of CPT */
    s_OLR = BIC_v^OLR(D|P) ;                 /* BIC of OLR model CPT */
    Select CPT of m maximizing s_m.
end
```

$w = \frac{\log|\mathbf{D}|}{2}$; /* the penalty weight */
$G = \arg\max_G BIC(\mathbf{D}|G)$; /* learn G using a standard CPTs */
for $v \in V$ **do**
- $s_{STD} = BIC_v^{STD}(\mathbf{D}|P)$; /* BIC of general CPT */
- $s_{MLR} = BIC_v^{MLR}(\mathbf{D}|P)$; /* BIC of MLR model of CPT */
- $s_{OLR} = BIC_v^{OLR}(\mathbf{D}|P)$; /* BIC of OLR model CPT */
- Select CPT of m maximizing s_m.

end

Algorithm 1: BN learning with standard CPTs, MLR, or OLR models.

When computing the BIC score, the penalty term considers the smaller number of parameters required for MLR and OLR. The BIC and AIC are believed to be useful criteria for avoiding overfitting, which log-likelihood maximization alone cannot guarantee. This means that, even if the log-likelihood drops when using MLR or OLR, we could still end up with a higher BIC score if the penalty is reduced sufficiently, as this is the intended behavior.

While the approach of Algorithm 1 is computationally efficient, it could be enhanced by integrating CPT type selection directly into the structural learning process. Such integration might yield models with superior BIC values by simultaneously optimizing network structure and CPT representations. However, implementing this approach would require developing efficient search space methods, likely incorporating parent set pruning techniques similar to those proposed by de Campos [4] for general CPTs. Given the non-trivial nature of this extension, we defer it to future research.

4. APPLICATION TO SURVEY DATA

4.1. The survey, dividing lines in Czech society

From 26th August to 4th September 2022, the STEM Institute for Empirical Research surveyed a total of 1661 Czech citizens (over the age of 18). The structure of the respondents followed the structure of the general Czech population according to the criteria of gender, age, education, size of place of residence, and region. The respondents provided their opinions on 30 key topics of public life ($A_1, \ldots, A_{30}$) and 14 assessments of their own life situation ($D_1, \ldots, D_{14}$). See Appendices A and B for details. The study served as an analytical basis for the publication [2].

Responses used the Likert scale $\{1, 2, 3, 4, 5, 6, 7\}$. For modeling based on BNs, we transformed the scale to $\{-1, 0, +1\}$ so that the values from $\{1, 2, 3\}$ were mapped to -1, the value 4 to 0, and values from $\{5, 6, 7\}$ to $+1$. This means that the number of states $k = 3$. To support this simplification, we present probability distributions for six typical variables from this survey. We can see that the most frequent answers are either at an extreme position, i. e., 1 or 7, or at the value exactly in the middle, i. e., the value 4. See Figure 1. Actually, there was no variable in the data that would have the most probable value different from 1, 4, or 7.

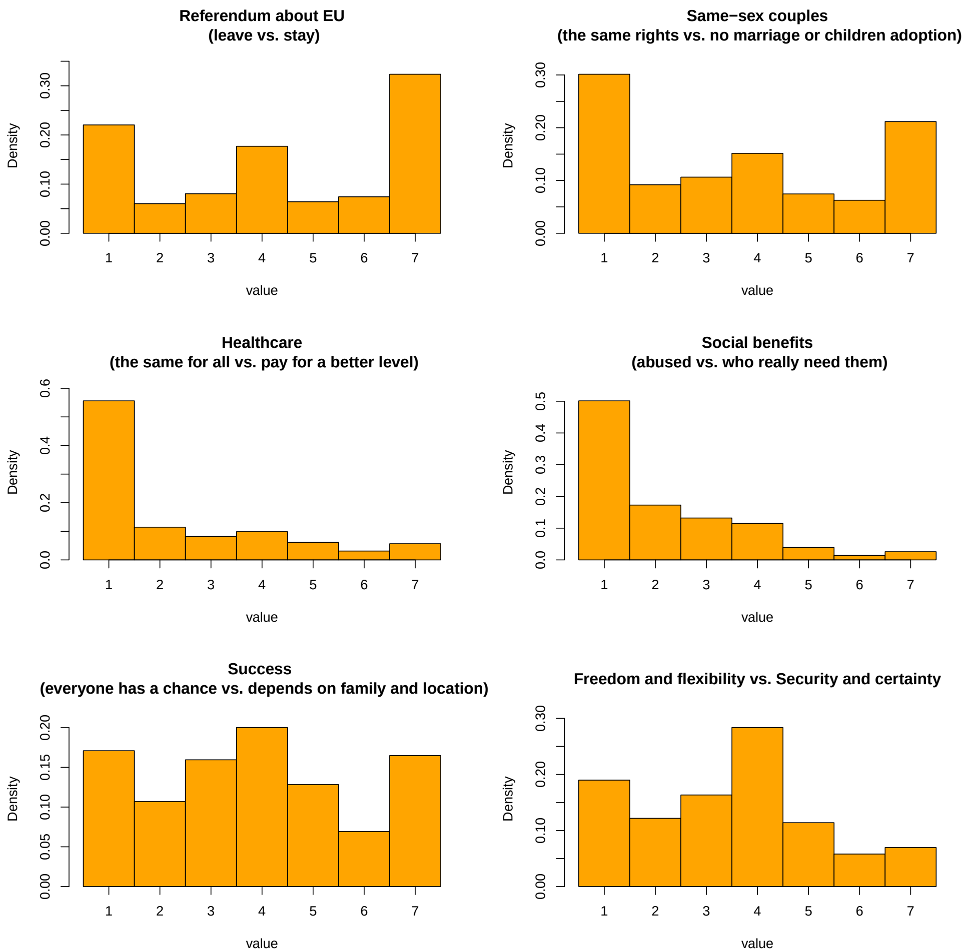


Fig. 1. Probability distributions of six typical variables.

The dataset also included demographic data on respondents and how they voted in the last general election (in 2021), but to keep the models clear and legible, we have not considered these variables in this paper. While omitting demographic variables means we cannot make causal claims about opinion relationships, unobserved confounders may create spurious associations; some observed correlations may result from shared demographic influences. However, our conditional independence analysis reveals meaningful structure: which opinions are directly related versus those connected only through others. This captures how beliefs cohere within the opinion space itself. While demographics influence individual belief combinations, controlling for a rich set of opinions can indirectly capture some of this variation, and the conditional structure we identify reflects substantive constraints in belief systems that warrant investigation in their own right. Note that the i.i.d. assumption used in Section 2 for the likelihood definition refers to the observations (respondents) being independent and identically distributed. Unobserved confounders do not necessarily violate this; each respondent can still be an independent draw from the same population.

4.2. Pairwise association analysis

To obtain a quick and simple overview of the studied problem, we start our analysis by examining pairwise associations between the studied variables. Note that we have transformed the original variable domains into a domain consisting of only three states, $\{-1, 0 + 1\}$. We do not assume that variables with transformed domains are ordinal. There are several arguments to support this. Generally, the neutral state 0 cannot be interpreted as a state between -1 and $+1$, since it often has its own typical characteristics. Additionally, the general CPTs and multinomial logistic regression models do not assume ordinality. We will see later in this section that they represent a large majority of selected representations, as they provide a better fit than ordinal logistic regression for most model variables. For these reasons, we decided to use Cramér's V, which is a measure of association between nominal variables. Its values range between 0 (no association) and $+1$ (complete association). To counteract the problem of false positives, which is common when multiple comparisons are performed, we used Holm's correction method [8]. In Figure 2, we present the upper diagonal of the matrix of Cramér's V values for all pairs of variables. The larger and darker the bullet, the stronger the association between the corresponding variables. The statistically insignificant values are denoted by crosses.

4.3. Bayesian network maximizing BIC

Figure 3 depicts the structure of the BN learned through BIC maximization. Nodes are color-coded by domain: yellow nodes represent opinions on key aspects of public life, while green nodes correspond to assessments of personal life circumstances. Node positioning reflects the strength of statistical associations, determined by Cramér's V correlation coefficients, with more strongly associated variables placed in closer proximity. The layout was generated using the Fruchterman-Reingold force-directed algorithm [7], which used the matrix of Cramér's V values presented in Figure 2 as its input and was implemented in the R package qgraph [6].

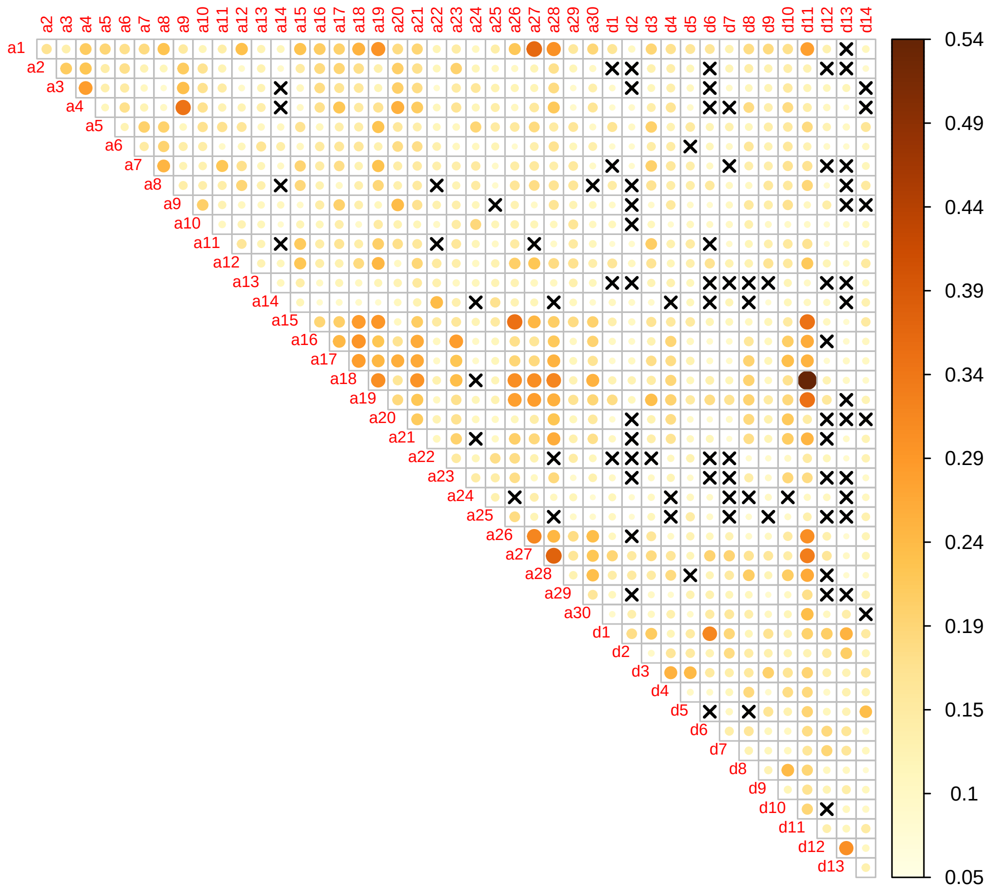


Fig. 2. Upper diagonal of the matrix of Cramér's V values for all pairs of variables. Crosses denote insignificant values.

For the network structure learned through BIC maximization with a standard CPT. We evaluated whether multinomial logistic regression (MLR) or ordinal logistic regression (OLR) representations provided a superior fit for each CPT. When an alternative representation achieved better BIC scores, we replaced the original CPT with the optimal alternative, as detailed in Algorithm 1. In Table 1, we compare the number of parameters of three different CPT types as a function of the number of parents. Table 2 summarizes the types of CPTs in the final learned model using the notation: $\bullet$ for standard CPTs, $\star$ for multinomial logistic regression, and $\circ$ for ordinal logistic regression.

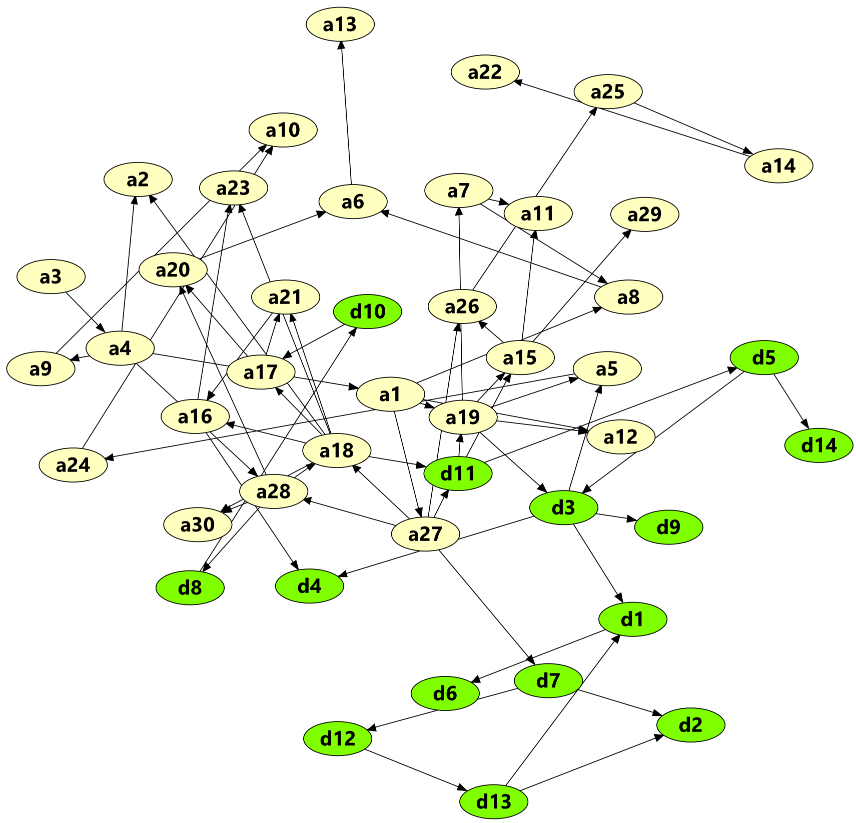


Fig. 3. The structure of the BN model learned by maximization of BIC. The yellow nodes correspond to the key topics of public life and the green nodes to assessments of own life situation.

The distribution comprises 19 standard CPTs, 22 MLR representations, and 3 OLR representations. It is worth mentioning that all 17 standard CPTs correspond either to variables with no parents (A_3) or to those with just one parent. This means that if a variable has more than one parent, it is always better to represent it with an OLR or MLR model for this dataset. Interestingly, two CPTs of variables with just one parent are better represented by OLR models.

Notably, multinomial logistic regression emerges as the most prevalent representation, which is expected given its balance between expressiveness and compactness—less restrictive than OLR yet more parameter-efficient than general CPTs. The high frequency of MLR adoption, even for nodes with only two parents, suggests substantial potential for parameter reduction in networks with higher-degree nodes, where the computational savings scale rapidly with the parent set size.

model type	general CPT	MLR	OLR
no. of parents			
0	2	2	2
1	6	6	3
2	18	10	4
3	54	14	5

Tab. 1. Comparison of the number of parameters.

$P(A_3)$	$\bullet$

$P(A_1\|A_4)$	$\bullet$
$P(A_4\|A_3)$	$\bullet$
$P(A_7\|A_{19})$	$\bullet$
$P(A_9\|A_4)$	$\bullet$
$P(A_{13}\|A_6)$	$\bullet$
$P(A_{14}\|A_{25})$	$\bullet$
$P(A_{22}\|A_{14})$	$\bullet$
$P(A_{24}\|A_5)$	$\bullet$
$P(A_{25}\|A_{26})$	$\bullet$
$P(A_{27}\|A_1)$	$\bullet$
$P(A_{29}\|A_{15})$	$\bullet$
$P(D_5\|D_{11})$	$\bullet$
$P(D_6\|D_1)$	$\bullet$
$P(D_7\|A_{27})$	$\circ$
$P(D_8\|A_{28})$	$\circ$
$P(D_9\|D_3)$	$\bullet$
$P(D_{10}\|D_8)$	$\bullet$
$P(D_{12}\|D_7)$	$\bullet$
$P(D_{13}\|D_{12})$	$\bullet$
$P(D_{14}\|D_5)$	$\bullet$

$P(A_2\|A_4, A_{18})$	$\star$
$P(A_5\|A_{19}, D_3)$	$\star$
$P(A_6\|A_8, A_{20})$	$\star$
$P(A_8\|A_1, A_7)$	$\star$
$P(A_{10}\|A_9, A_{24})$	$\star$
$P(A_{11}\|A_7, A_{15})$	$\star$
$P(A_{12}\|A_1, A_{19})$	$\star$
$P(A_{15}\|A_{19}, D_{11})$	$\star$
$P(A_{16}\|A_{18}, A_{21})$	$\star$
$P(A_{17}\|A_{18}, D_{10})$	$\star$
$P(A_{18}\|A_{27}, A_{28})$	$\star$
$P(A_{19}\|A_1, D_{11})$	$\star$
$P(A_{20}\|A_{17}, A_{28})$	$\circ$
$P(A_{21}\|A_{17}, A_{18})$	$\star$
$P(A_{23}\|A_{16}, A_{18})$	$\star$
$P(A_{26}\|A_{15}, A_{27})$	$\star$
$P(A_{28}\|A_4, A_{27})$	$\star$
$P(A_{30}\|A_{18}, A_{28})$	$\star$
$P(D_1\|D_3, D_{13})$	$\star$
$P(D_2\|D_7, D_{13})$	$\star$
$P(D_3\|A_{19}, D_5)$	$\star$
$P(D_4\|A_{16}, D_3)$	$\star$
$P(D_{11}\|A_{18}, A_{27})$	$\star$

Tab. 2. CPTs of the learned model and their types
($\bullet$ for a standard CPT, $\star$ for a MLR, and $\circ$ for an OLR).

Figure 4 compares BIC values, the number of model parameters, log-likelihood, and prediction accuracy for the proposed approach and standard BNs. Prediction accuracy was evaluated using ten-fold cross-validation: models trained on nine folds were tested by randomly selecting evidence variables and predicting the remaining variables against ground truth values. This effectively means that for each relative ratio of variables with evidence, each respondent was tested exactly once, which indicates that in the bottom right plot of Figure 4, we present values that represent averages over all 1661 respondents and all predicted variables. We tested the relative ratio of variables with evidence to

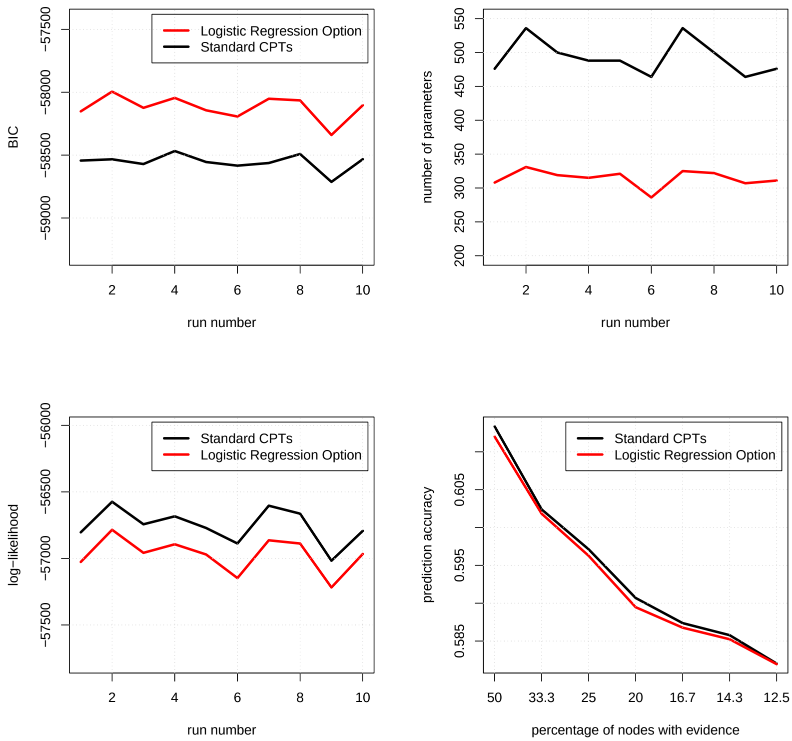


Fig. 4. Comparison of BIC values (top left), number of model parameters (top right), log-likelihood (bottom left), and prediction accuracy (bottom right) for the BN model with optional MLR or OLR and the standard BN model. Note that the last plot uses the logarithmic scale for the horizontal axis.

be $\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{4}$, $\frac{1}{5}$, $\frac{1}{6}$, $\frac{1}{7}$, and $\frac{1}{8}$, as we believe these may correspond to realistic scenarios in the model applications. The accuracy is the relative ratio of variables whose most likely predicted value equals the variable's value in the testing data. The accuracy ranges between 0.61 for 50% of nodes with evidence and 0.58 for 12.5% of nodes with evidence. The results show that adaptive selection between general CPT and logistic regression representations achieves consistently better BIC scores[3] while maintaining comparable predictive performance with about 60% parameters, confirming that structured CPT representations reduce model complexity without sacrificing accuracy.

[3] It should be noted that, due to the algorithm design, the BIC value cannot be worse.

5. MODEL ANALYSIS

The importance of variables can be measured by their centrality in the graph of the BN. A good measure of centrality for a variable in BNs is the size of its Markov Blanket [16]. The Markov Blanket of a variable X in a BN is the set that contains the parents of X, the children of X, and the other parents of all children of X. An important property of the Markov Blanket is that all other variables in the BN model are independent of X given the states of the nodes in the Markov Blanket. In Table 3, we present the nine variables with the largest Markov blankets. We use abbreviated names of variables; see the full names in Appendixes A and B.

	short name	size
A_{18}	Referendum about EU	11
A_{19}	Trend after 1989	8
A_{27}	Refugees from Ukraine	8
D_3	Trust in future	8
A_1	Migration	7
A_4	Support to Roma people	7
A_{15}	Geopolitics	7
A_{28}	State treatment of refugees	7
D_{11}	Benefits of EU	6

Tab. 3. Markov Blanket sizes.

The MLR models from the learned BN are presented in Tables 5 and 7. In the last column, we provide a keyword for the phenomenon that the corresponding probability distribution may represent, and in the footnotes, we provide more details about each phenomenon. Although logical implications are often used; care should be taken when interpreting them. For example, the causal interpretation based on the direction of the implication need not be true. Furthermore, the relationships are stochastic and not deterministic; i. e., they attempt to reflect the most common relationship between the variables of each CPT, which does not need to hold and typically does not hold for all participants of the survey. The visualizations of the corresponding CPTs can be found in Appendix C. They can be used to provide a deeper insight into the corresponding phenomenon.

Note the large number of various patterns of the CPTs. These patterns provide a good justification for the preference of MLR over OLR models, since quite a number of these patterns cannot be represented by a CPT that assumes the ordinality of variables; therefore, an OLR model would not work well for them. Also, as we mentioned in Subsection 4.2, there are variables for which the neutral state 0 cannot be interpreted as a state between -1 and $+1$, since it often has its own typical characteristics.

Three OLR models from the learned BN are presented in Table 4. In the last column, we provide a keyword for the phenomenon that the corresponding probability distribution may represent. The conditional probability distribution $P(A_{20}|A_{17}, A_{28})$ is specified by four parameters of the OLR model: two thresholds $\zeta_1 = 1.284$ (deciding between $A_{20} = -1$ and $A_{20} = 0$), $\zeta_2 = 2.764$ (deciding between $A_{20} = 0$ and $A_{20} = 1$), and two coefficients $\alpha_1 = 0.891$ (for A_{17}) and $\alpha_2 = 0.669$ (for A_{28}). We can see that

A_{17} is more influential than A_{28}. The conditional probability distribution $P(D_7|A_{27})$ is specified by three parameters: $\zeta_1 = 0.543$, $\zeta_2 = 1.600$, and $\alpha_1 = 0.600$. $P(D_8|A_{28})$ is also specified by three parameters: $\zeta_1 = -0.293$, $\zeta_2 = 1.567$, and $\alpha_1 = 0.619$.

The remaining CPTs have the general form, and they are presented in Table 8. Note that there is never more than one parent. The CPTs are visualized in Appendix D.

The learned model reveals several interconnected psychological and political patterns in Czech society, centering on core attitudes towards security, belonging, and institutional trust.

CPT	dependent var.	explanatory variables	phenomenon
$P(A_{20}\|A_{17}, A_{28})$	Politicians	Partnership with West, State treat. of refugees	elite accountability attribution[4]
$P(D_7\|A_{27})$	National Pride	Refugees from Ukr.	vicarious moral elevation[5]
$P(D_8\|A_{28})$	State aid	State treat. of refugees	state failure perception[6]

Tab. 4. OLR models.

CPT	dependent var.	explanatory variables	phenomenon
$P(A_2\|A_4, A_{18})$	Minorities	Supp. to Roma p., Referendum about EU	cultural protectionism[7]
$P(A_5\|A_{19}, D_3)$	Meritocracy	Trend after 1989, Trust in future	systemic pessimism[8]
$P(A_6\|A_8, A_{20})$	Healthcare	State paternalism, Politicians	apolitical individualism[9]
$P(A_8\|A_1, A_7)$	State paternalism	Migration, Inequalities	fear-driven paternalism[10]
$P(A_{10}\|A_9, A_{24})$	Basic needs	Social benefits, Women Oppor.	inequality awareness[11]
$P(A_{11}\|A_7, A_{15})$	Influence of the rich	Inequalities, Geopolitics	neutrality cascade effect[12]

Tab. 5. MLR models (part one).

[4]Whether problem is external (Western contempt, $A_{17} = -1$) or internal (refugee handling, $A_{28} = -1$) politicians get blamed and distrusted.

[5]Feeling proud of nation ($D_7 = -1$) because of its moral choices ($A_{27} = -1$).

[6]State failure with the refugees from Ukr. ($A_{28} = -1$) relates to its failure to help me ($D_8 = -1$).

[7]Support for Roma people is balanced ($A_4 = 1$) and Czechia should stay in EU ($A_{18} = 1$) implies the right for minorities to preserve their own customs and traditions ($A_2 = 1$).

[8]Czech society not moving in the correct direction ($A_{19} \geq 0$) and Distrust in future ($D_3 = 1$) indicate that wealth or social class are important for success ($A_5 = 1$).

[9]Individual responsibility ($A_8 = -1$) and a neutral trust in politicians ($A_2 = 0$) imply support for partially paid healthcare ($A_6 = 1$).

[10]State reliance preference ($A_8 = 1$) can be activated by either cultural-nationalist concerns ($A_1 = 1$) or by social-democratic concerns ($A_7 = 1$).

[11]Trust in the social benefits system ($A_9 = 1$) and awareness of structural barriers ($A_{24} \geq 0$) implies we shouldn't require people to 'earn' basic needs ($A_{10} = 1$).

[12]Generally, it is believed that the big economic players in Czechia have too much influence on how the state is run. However, neutrality on related issues ($A_7 = 0, A_{15} = 0$) produces neutrality on the target issue ($A_{11} = 0$).

CPT	dependent var.	explanatory variables	phenomenon
$P(A_{12}\|A_1, A_{19})$	Democracy	Migration, Trend after 1989	authoritarian nostalgia[13]
$P(A_{15}\|A_{19}, D_{11})$	Geopolitics	Trend after 1989, Benefits of EU	dissatisfied neutralism[14]
$P(A_{16}\|A_{18}, A_{21})$	Globalization	Referendum about EU, General Referenda	anti-globalization axis[15]
$P(A_{17}\|A_{18}, D_{10})$	Partnership with West	Referendum about EU, Feeling heard	feeling of disrespect[16]
$P(A_{18}\|A_{27}, A_{28})$	Referendum about EU	Refugees from Ukr., State treat. of refugees	humanitarian Europeanism[17]
$P(A_{19}\|A_1, D_{11})$	Trend after 1989	Migration, Benefits of EU	post-1989 regret syndrome[18]
$P(A_{21}\|A_{17}, A_{18})$	General Referenda	Partnership with West, Referendum about EU	liberal democratic core[19]
$P(A_{23}\|A_{16}, A_{18})$	Climate change measures	Globalization, Referendum about EU	liberal ecological modernization[20]
$P(A_{26}\|A_{15}, A_{27})$	War in Ukraine	Geopolitics, Refugees from Ukr.	Russophile rationalization[21]
$P(A_{28}\|A_4, A_{27})$	State treatment of refugees	Supp. to Roma p., Refugees from Ukr.	justice validates performance[22]
$P(A_{30}\|A_{18}, A_{28})$	Vaccination	Referendum about EU, State treat. of refugees	institutional trust cascade[23]

Tab. 6. MLR models (part two).

[13] External threat (migration, $A_1 = 1$) plus internal failure (wrong direction since 1989, $A_{19} \geq -1$) implies authoritarian solution (strong leader, $A_{12} = 1$).

[14] Western path failure: wrong direction after 1989 ($A_{19} \geq 0$) and EU not beneficial personally ($D_{11} \geq 0$) leads to the solution: be a bridge between both West and East ($A_{15} = 1$).

[15] Two different routes to the same conclusion: Leave the EU nationalist route ($A_{18} = -1$) and Direct democracy populist route ($A_{21} = -1$) leads to the conclusion that globalization is a threat ($A_{16} = -1$).

[16] The feeling of disrespect, whether from outside (from the West, $A_{18} = -1$) or inside (D_{10}=-1), gives the impression of Western arrogance.

[17] Support for refugees (morally, $A_{27} = -1$, pragmatically, $A_{28} = 1$) relates to EU support ($A_{18} = 1$).

[18] Either cultural threat (migration, $A_1 = 1$) or economic disappointment (EU not beneficial, $D_{11} = 1$) leads to rejecting the entire post-communist trajectory $A_{19} = 1$.

[19] Comfort with both international equality ($A_{17} = 1$) and EU membership ($A_{18} = 1$) produces trust in domestic representative democracy (A_{21}).

[20] Economic openness ($A_{16} = 1$), political integration ($A_{18} = 1$), and environmental action ($A_{23} = 1$) are mutually reinforcing.

[21] Adopting narratives that align with Russian interests through: Opposing Ukrainian refugee support ($A_{27} = 1$), Advocating Czech neutrality ($A_{15} = 1$), Accepting Russian historical narrative about NATO expansion as provocation ($A_{26} = 1$).

[22] Believing in fair treatment ($A_4 = 1$) or moral rightness ($A_{27} = -1$) leads to perceiving state competence ($A_28 = 1$).

[23] Trust in supranational institutions (EU, $A_{18} = 1$) or national institutions (state refugee handling, $A_{28} = 1$) implies trust in public health institutions (vaccination, $A_{30} = 1$).

CPT	dependent var.	explanatory variables	phenomenon
$P(D_1 \mid D_3, D_{13})$	Perception of success	Trust in future, Loneliness	depressive cognitive triad[24]
$P(D_2 \mid D_7, D_{13})$	Trust in other people	Pride in being Czech, Loneliness	compound belonging[25]
$P(D_3 \mid A_{19}, D_5)$	Trust in future	Trend after 1989, Changes vs. traditions	validated worldview optimism[26]
$P(D_4 \mid A_{16}, D_3)$	Children's lives	Globalization, Trust in future	optimistic prosperity projection[27]
$P(D_{11} \mid A_{18}, A_{27})$	Benefits of EU	Referendum about EU, Refugees from Ukr.	self-interest as moral cover[28]

Tab. 7. MLR models (part three).

5.1. Personal wellbeing and social integration

At an individual level, psychological distress manifests in predictable patterns: people who experience loneliness and fear about the future also report feeling unsuccessful in life, a state that occurs alongside financial strain (the success-debt linkage $P(D_6|D_1)$ and the depressive cognitive triad $P(D_1|D_3, D_{13})$). Conversely, individuals who perceive themselves as socially connected, both through practical support networks and social support ($P(D_{13}|D_{12})$), also express national pride and a generalized sense of trust in others (social belonging, $P(D_{12}|D_7)$ and compound belonging, $P(D_2|D_7, D_{13})$). A coherent syndrome emerges from personal orientation towards change versus stability: those who value traditions also prefer security over flexibility (stability preference, $P(D_{14}|D_5)$).

5.2. Institutional trust cascades

Trust in institutions operates as a cascading phenomenon. Those who have personally experienced state failure (e. g., lack of help when needed) also perceive systemic marginalization (e. g., feeling unheard), creating a coherent pattern of institutional alienation ($P(D_{10}|D_8)$). This extends to perceptions of refugee policy ($P(D_8|A_{28})$), attitudes towards vaccination ($P(A_{30}|A_{18}, A_{28})$), and ultimately political trust, where both external grievances (e. g. Western disrespect) and internal grievances (e. g. policy failures) converge to create distrust of politicians (elite accountability attribution $P(A_{20}|A_{17}, A_{28})$, institutional trust cascade, $P(A_{30}|A_{18}, A_{28})$). Conversely, positive institutional experiences create self-reinforcing trust.

[24] Emotional distress (fear, $D_3 = 1$ and loneliness, $D_{13} = -1$) leads to negative self-evaluation about life success ($D_1 = 1$).

[25] Feeling connected both nationally (Czech pride, $D_7 = -1$) AND personally (social network $D_{13} = 1$) leads to generalized social trust ($D_2 = 1$).

[26] Positive historical assessment ($A_{19} = -1$) and personal openness to change ($D_5 = -1$) implies future optimism ($D_3 = -1$).

[27] Global opportunity perception ($A_{16} = 1$) and personal optimism ($D_3 = -1$) leads to next generation advancement expectation ($D_4 = 1$).

[28] People believe their opposition to outgroups - EU exit preference ($A_{18} = -1$) and refugees ($A_{27} = 1$) - is based on competition for resources ("they harm my interests", $D_{11} = 1$) rather than prejudice.

CPT	dependent var.	explanatory var.	phenomenon
$P(A_1 \mid A_4)$	Migration	Supp. to Roma p.	minority threat multiplication[29]
$P(A_3)$	Cond. of Roma p.		
$P(A_4 \mid A_3)$	Supp. to Roma p.	Cond. of Roma p.	deservingness denial pattern[30]
$P(A_7 \mid A_{19})$	Inequalities	Trend after 1989	post-communist disillusionment[31]
$P(A_9 \mid A_4)$	Social benefits	Supp. to Roma p.	welfare chauvinism escalation[32]
$P(A_{13} \mid A_6)$	Taxes	Healthcare	equality principle generalization[33]
$P(A_{14} \mid A_{25})$	Envir. protection	Same-sex couples	centrist coherence[34]
$P(A_{22} \mid A_{14})$	Climate change threat	Envir. protection	cognitive consistency on climate[35]
$P(A_{24} \mid A_5)$	Women	Meritocracy	structural ineq. generalization[36]
$P(A_{25} \mid A_{26})$	Same-sex couples	War in Ukraine	traditionalist geopol. alignment[37]
$P(A_{27} \mid A_1)$	Refugees from Ukr.	Migration	threat frame dominance[38]
$P(A_{29} \mid A_{15})$	Education	Geopolitics	alternative worldview coherence[39]
$P(D_5 \mid D_{11})$	Changes vs. traditions	Benefits of EU	stability-seeking euroscepticism[40]
$P(D_6 \mid D_1)$	Savings vs. debts	Percept. of success	success-debt linkage[41]
$P(D_9 \mid D_3)$	Responsibility	Trust in future	fearful disengagement[42]
$P(D_{10} \mid D_8)$	Feeling heard	State aid	institutional alienation[43]
$P(D_{12} \mid D_7)$	Care from others	National Pride	social belonging[44]
$P(D_{13} \mid D_{12})$	Loneliness	Care from others	social support[45]
$P(D_{14} \mid D_5)$	Flexibility vs. security	Changes. vs. trad.	stability preference[46]

Tab. 8. General CPTs.

[29] Resentment about Roma support ($A_4 = -1$) strengthens antimigration attitudes ($A_1 = 1$).

[30] To justify resentment about support ($A_4 = -1$), deny conditions that would make support deserved ($A_3 = -1$).

[31] People who believe 1989 path was wrong ($A_{19} = 1$) actively seek or amplify evidence supporting that inequalities are dangerous ($A_7 = 1$).

[32] Targeted resentment (against Roma, $A_9 = -1$) expands into generalized welfare skepticism ($A_4 = -1$).

[33] Egalitarian values in one domain (healthcare, $A_6 = -1$) amplify in another (taxation, $A_{13} = -1$).

[34] Neutral/moderate positions cluster together across issues, $A_{14} = 0, A_{25} = 0$.

[35] Believing problem is serious ($A_{22} = -1$) requires believing action should be prioritized ($A_{14} = -1$).

[36] Believing in class-based inequality ($A_5 = 1$) strengthens recognition of gender-based inequality ($A_{24} = 1$).

[37] Sympathy for Russia's geopolitical position ($A_{26} = 1$) linked to sympathy for its traditional values stance ($A_{25} = 1$).

[38] Viewing migration as threat ($A_1 = 1$) overrides other considerations (solidarity, geopolitics, humanitarian values).

[39] Rejecting mainstream geopolitics (Western alignment, $A_{15} = 1$) correlates with rejecting mainstream knowledge sources (formal education, $A_{29} = 1$).

[40] EU seen as source of unwanted change/uncertainty.

[41] Abstract self-evaluation (lacking success, $D_1 = 1$) co-occurs with a concrete material circumstance (financial struggle, $D_6 = 1$).

[42] Fearfulness ($D_3 = 1$) is associated with lack of responsibility $D_9 = 1$.

[43] Personal experience of state failure ($D_8 = -1$) and perceived systemic marginalization ($D_{10} = -1$) are both dimensions of feeling abandoned by institutions.

[44] Being integrated into or connected with one's community at both symbolic ($D_7 = -1$) and practical levels ($D_{12} = -1$).

[45] Practical support (people who care for you when ill) co-occurs with emotional support (people to spend time with and confide in).

[46] Valuing traditions/certainties ($D_5 = 1$) co-occurs with preferring security over flexibility ($D_{14} = 1$) - both reflecting an underlying preference for predictability and established order over novelty and openness.

5.3. Cultural protectionism and threat perception

Resentment towards specific minorities (such as the Roma community) can escalate into broader anti-immigration attitudes (known as minority threat multiplication, $P(A_1|A_4)$ and threat frame dominance, $P(A_{27}|A_1)$), which can then lead to generalized skepticism towards welfare (known as welfare chauvinism escalation, $P(A_9|A_4)$). These cultural concerns lead to demands for state paternalism (fear-driven paternalism, $P(A_8|A_1, A_7)$) and authoritarian solutions, especially when combined with the perception of societal failure since 1989 (authoritarian nostalgia, $P(A_{12}|A_1, A_{19})$).

5.4. Post-1989 disillusionment syndrome

A significant pattern emerges centering on disappointment with the post-communist trajectory. Those who believe that Czech society has moved in the wrong direction since 1989 perceive dangerous inequalities (post-communist disillusionment, $P(A_7|A_{19})$), doubt meritocracy (systemic pessimism, $P(A_5|A_{19}, D_3)$), and lose faith in the future (inverse validated worldview optimism, $P(D_3|A_{19}, D_5)$). This disillusionment creates two distinct response pathways: some advocate geopolitical neutrality between East and West (dissatisfied neutralism, $P(A_{15}|A_{19}, D_{11})$), while others embrace Russian geopolitical narratives (Russophile rationalization, $P(A_{26}|A_{15}, A_{27})$).

5.5. Liberal-conservative coherence

The data reveal internally consistent worldview packages. The liberal package integrates support for the EU, a humanitarian approach to refugee policy (humanitarian Europeanism, $P(A_{18}|A_{27}, A_{28})$), trust in representative democracy (the liberal democratic core, $P(A_{21}|A_{17}, A_{18})$), climate action (liberal ecological modernization, $P(A_{23}|A_{16}, A_{18})$) and recognition of structural inequalities across multiple domains, whether economic, gender-based, or racial (generalization of structural inequality, $P(A_{24}|A_5)$, awareness of inequality, $P(A_{10}|A_9, A_{24})$). Traditionalist positions also demonstrate coherence: Euroscepticism is connected to a preference for stability (stability-seeking euroscepticism, $P(D_5|D_{11})$), a rejection of mainstream knowledge institutions (alternative worldview coherence, $P(A_{29}|A_{15})$), and an alignment with Russian geopolitical and social positions (traditionalist geopolitical alignment, $P(A_{25}|A_{26})$).

5.6. Motivated reasoning patterns

Several findings suggest self-serving cognition. For example, people may deny that the Roma face disadvantaged conditions while simultaneously resenting the support given to them (the 'deservingness denial' pattern, $P(A_4|A_3)$). Those who oppose EU membership and refugees tend to frame their position in terms of economic self-interest rather than prejudice (using self-interest as a moral cover, $P(D_{11}|A_{18}, A_{27})$). Conversely, supporting refugee policy can produce national pride through vicarious moral elevation ($P(D_7|A_{27})$). People perceive state competence when outcomes align with their moral preferences (justice validates performance, $P(A_{28}|A_4, A_{27})$).

5.7. Summary

Overall, the Czech population appears to be divided into clusters with coherent worldviews, organised around fundamental orientations such as openness versus security-seeking, institutional trust versus alienation, and acceptance versus rejection of the post-1989 trajectory. These orientations cascade across political attitudes, personal psychology and social integration in predictable, mutually reinforcing patterns.

6. CONCLUSION AND FUTURE WORK

In this paper, we have proposed an extension to BN learning that incorporates structured CPTs representations based on multinomial logistic regression (MLR) and ordinal logistic regression (OLR) as alternatives to standard general CPTs. Through empirical evaluation using real data from a comprehensive opinion survey, we have demonstrated that these structured representations frequently achieve superior model fit compared to general CPTs when assessed using the Bayesian Information Criterion (BIC). The proposed approach yields BNs with substantially fewer parameters while maintaining comparable predictive performance, effectively addressing the curse of dimensionality that affects traditional CPT parameterization.

Our model analysis has some limitations. In the model, we omitted all demographic variables; some of them may explain certain relationships as a common explanation or a common cause for variables connected by an edge. The relationships are simplified to the most relevant variables; the less relevant ones might be included if an integrated structural learning algorithm were used.

Our future work will focus on developing an integrated structural learning algorithm that simultaneously optimizes both the network structure and the CPT representations, rather than treating these as separate sequential steps. Additionally, the framework can be extended to incorporate other types of local CPT structures, including Noisy-MIN and Noisy-MAX models [5], as well as Noisy-Threshold models [17] and Generalized additive models (GAMs), which are a natural generalization of the OLR models. Such extensions would provide practitioners with a comprehensive toolkit for learning parsimonious yet expressive BN models from complex real-world data.

ACKNOWLEDGEMENT

I would like to thank the reviewers for their constructive comments, which have helped improve the readability and clarity of the paper.

This work was supported by the European Regional Development Fund project ”Beyond Security: Role of Conflict in Resilience-Building” (reg. no.: CZ.02.01.01/00/22_008/0004595).

(Received July 15, 2025)

REFERENCES

[1] A. Agresti: Categorical Data Analysis. Third edition. John Wiley and Sons, Hoboken, NJ 2013.

[2] M. Buchtík: Různá vyprávění o jedné společnosti. Friedrich–Ebert–Stiftung and Masarykova demokratická akademie, Praha 2023.

[3] J. Cussens: Bayesian network learning with cutting planes. In: Proc. 27th Conference on Uncertainty in Artificial Intelligence, (F. Cozman and A. Pfeffer, eds.), AUAI Press, Corvallis 2011, pp. 153–160. DOI:10.5771/9780810873834-153

[4] C. P. de Campos, M. Scanagatta, G. Corani, and M. Zaffalon: Entropy-based pruning for learning Bayesian networks using BIC. Artificial Intelligence *260* (2018), 42–50. DOI:10.1016/j.artint.2018.04.002

[5] F. J. Díez and S. F. Galán: An efficient factorization for the noisy MAX. Int. J. Intell. Systems *18* (2003), 165–177. DOI:10.1002/int.10080

[6] S. Epskamp, A. O. J. Cramer, L. J. Waldorp, V. D. Schmittmann, and D. Borsboom: Network visualizations of relationships in psychometric data. J. Statist. Software *48* (2012), 4, 1–18. `http://www.jstatsoft.org/v48/i04/`

[7] T. M. J. Fruchterman and E. M. Reingold: Graph drawing by force-directed placement. Software: Practice Experience *21* (1991), 11, 1129–1164. DOI:10.1002/spe.4380211102

[8] S. Holm: A simple sequentially rejective multiple test procedure. Scand. J. Statist. *6* (1979), 65–70. `http://www.jstor.org/stable/4615733`

[9] F. V. Jensen and T. D. Nielsen: Bayesian Networks and Decision Graphs. Second edition. Information Science and Statistics, Springer New York, NY 2007. DOI:10.1007/978-0-387-68282-2

[10] D. Koller and N. Friedman: Probabilistic Graphical Models: Principles and Techniques. The MIT Press, 2009.

[11] J. Pearl: Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann Publishers Inc., San Francisco 1988.

[12] F. Rijmen: Bayesian networks with a logistic regression model for the conditional probabilities. Int. J. Approx. Reasoning *48* (2008), 2, 659–666. In memory of Philippe Smets 1938–2005. DOI:10.1016/j.ijar.2008.01.001

[13] G. Schwarz: Estimating the dimension of a model. Ann. Statist. *6* (1978), 2, 461–464. `http://www.jstor.org/stable/2958889`

[14] C. Sharma, Z. A. Liao, J. Cussens, and P. van Beek: A score-and-search approach to learning Bayesian networks with noisy-or relations. In: Pro. 10th International Conference on Probabilistic Graphical Models (PGM 2020). Proc. Machine Learning Research *138* (2020), pp. 413–42. `https://proceedings.mlr.press/v138/sharma20a.html`

[15] J. Vomlel, V. Kratochvíl,and F. Kratochvíl: Structural learning of mixed noisy-OR Bayesian networks. Int. J. Approx. Reasoning *161* (2023), 108990. DOI:10.1016/j.ijar.2023.108990

[16] J. Vomlel, A. Kuběna, M. Šmíd, and J. Weinerova: Uncovering relationships using bayesian networks: A case study on conspiracy theories. In: Proc. 12th International Conference on Probabilistic Graphical Models. Proc. Machine Learning Research (J. Kwisthout and S. Renooij, eds.), PMLR *246* (2024), pp. 470–485. `https://proceedings.mlr.press/v246/vomlel24a.html`

[17] J. Vomlel and P. Tichavský: Probabilistic inference with noisy-threshold models based on a CP tensor decomposition. Int. J. Approx. Reasoning *55* (2014), 4, 1072–1092. DOI:10.1016/j.ijar.2013.12.002

Jiří Vomlel, The Czech Academy of Sciences, Institute of Information Theory and Automation, Pod Vodárenskou věží 4, 182 08 Praha 8 and Institute of Information Studies and Librarianship, Faculty of Arts, Charles University, Praha. Czech Republic.
e-mail: vomlel@utia.cas.cz

A. ATTITUDES

	name	value −1 value +1
A_1	Migration	Migration to Czechia enriches our society. Migration to Czechia is a threat to our society.
A_2	Minorities	Members of minorities in Czechia should adapt to Czech customs and traditions. Members of minorities in Czechia have the right to preserve their own customs and traditions.
A_3	Cond. of Roma people	The Roma people in Czechia have the same conditions as the majority of the population. The Roma people in Czechia have worse conditions than the majority of the population.
A_4	Supp. to Roma people	The Roma people in Czechia receive undeservedly more support than others. The Roma people in Czechia receive the same support as others who need it.
A_5	Meritocracy	Everyone in Czechia has a chance to achieve professional success and wealth through hard work and effort. Success in Czechia depends more on where and what family you are born into than on your own efforts.
A_6	Healthcare	The level of healthcare should be the same for all. There should be a way to pay for a better level of care in the healthcare system.
A_7	Inequalities	Economic and social inequalities are not a big problem in Czechia. Economic and social inequalities in Czechia are dangerously large.
A_8	State paternalism	Everybody should take care of himself, the state should take care of only the most basic provision for the Czech population. The state should take more responsibility for the care of the Czech population.
A_9	Social benefits	Social benefits in Czechia are often abused and given to people who do not deserve them. Social benefits in Czechia are mostly received only by those who really need them.
A_{10}	Basic needs	Everything a person receives must be deserved and earned. The fulfillment of basic needs should not depend on whether a person earns it or works for it.
A_{11}	Influence of the rich	The big economic players in Czechia are under sufficient democratic control. The big economic players in Czechia have too much influence on how the state is run.
A_{12}	Democracy	Representational democracy is best, even when decisions are compromises that do not always correspond to what I want. Czechia needs a strong leader to push through what is right regardless of differing attitudes.
A_{13}	Taxes	The more income and wealth people have, the higher the taxes they should pay. Taxes should be the same for everyone regardless of income and wealth.
A_{14}	Envir. protection	Environmental protection should take precedence over economic growth. Economic growth should take precedence over environmental protection.
A_{15}	Geopolitics	Czechia should be an integral part of Western Europe. Czechia should be a neutral bridge between the West and the East.

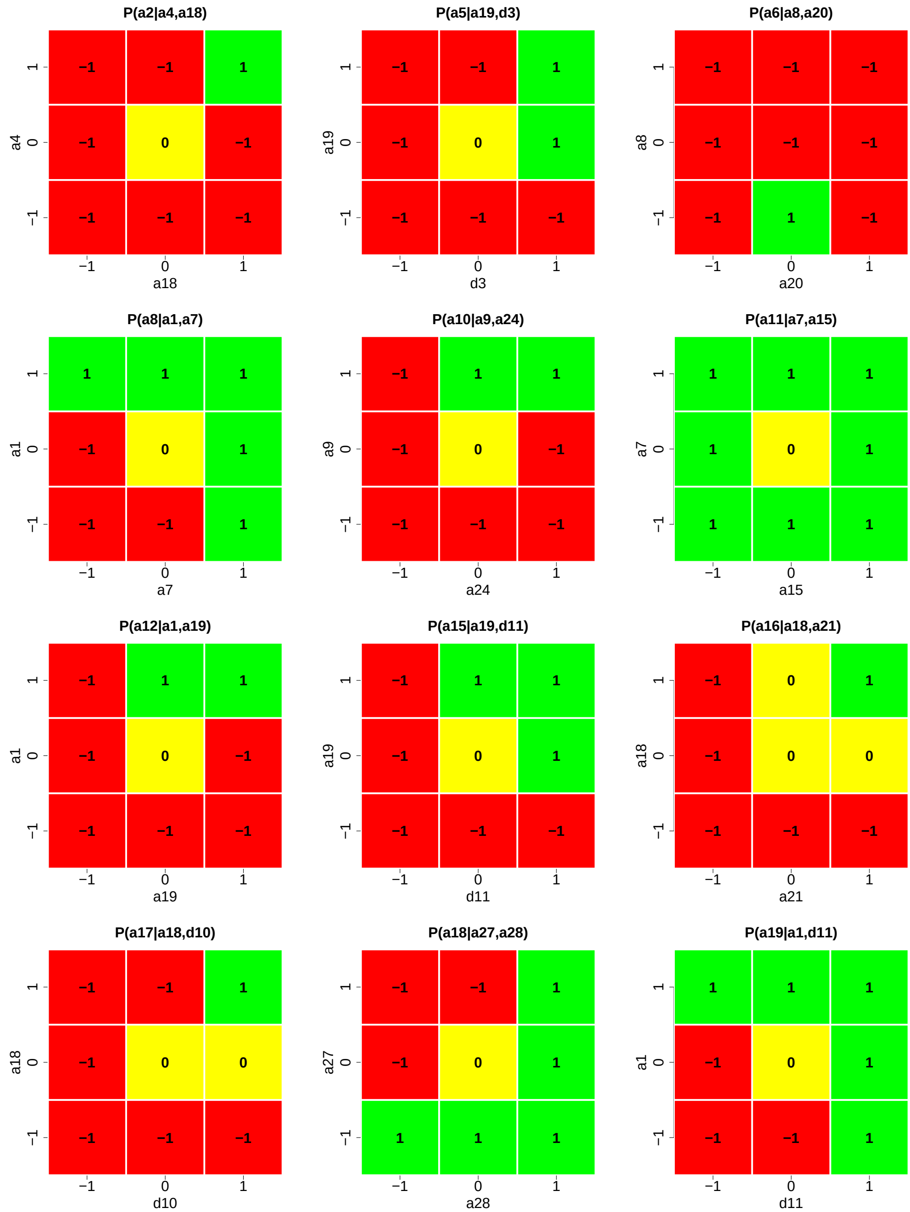
	name	value -1 value $+1$
A_{16}	Globalization	Globalization threatens the prosperity of Czechia. Globalisation is above all an opportunity for prosperity for Czechia.
A_{17}	Partnership with West	The countries of the West look down on us. We are equal partners for Western countries.
A_{18}	Referendum about EU	If a referendum were held in Czechia today, I would vote to leave the EU. If a referendum were held in Czechia today, I would vote to stay in the EU.
A_{19}	Trend after 1989	Czech society is generally moving in the correct direction after 1989. Czech society is generally moving in the wrong direction after 1989.
A_{20}	Politicians	Czech politicians are primarily concerned with their own benefit and cannot be trusted. Czech politicians act mainly for the benefit of society and are trustworthy.
A_{21}	General Referenda	People in referenda, not elected politicians, should make the most important decisions in Czechia. It is fine that elected politicians decide important social issues in Czechia.
A_{22}	Climate change threat	Man-made climate change is a major threat to our future. Climate change is not happening, so it does not threaten us.
A_{23}	Climate change measures	Measures against climate change are ill-conceived and threaten our economy. Action on climate change is an opportunity to improve our lives.
A_{24}	Women	Women in Czechia today have the same opportunities and possibilities as men. Women in Czechia today have worse opportunities and possibilities than men.
A_{25}	Same-sex couples	Same-sex couples in Czechia should have the same rights as others, including marriage and child adoption. Same-sex couples in Czechia should not be able to marry or adopt children.
A_{26}	War in Ukraine	Russia is the clear cause of the war in Ukraine. NATO caused the war in Ukraine in the first place by provoking Russia.
A_{27}	Refugees from Ukr.	It is right that Czechia has accepted refugees from Ukraine. Czechia should not have accepted refugees from Ukraine.
A_{28}	State treat. of refugees	The Czech state has failed to deal with the refugees from Ukraine. The Czech state succeeded in dealing with the refugees from Ukraine.
A_{29}	Education	Quality education is the key to understanding today's world. Life experience is more important for understanding today's world than formal education.
A_{30}	Vaccination	Vaccination against Covid-19 is dangerous. Vaccination against Covid-19 is an important step in overcoming the pandemic.

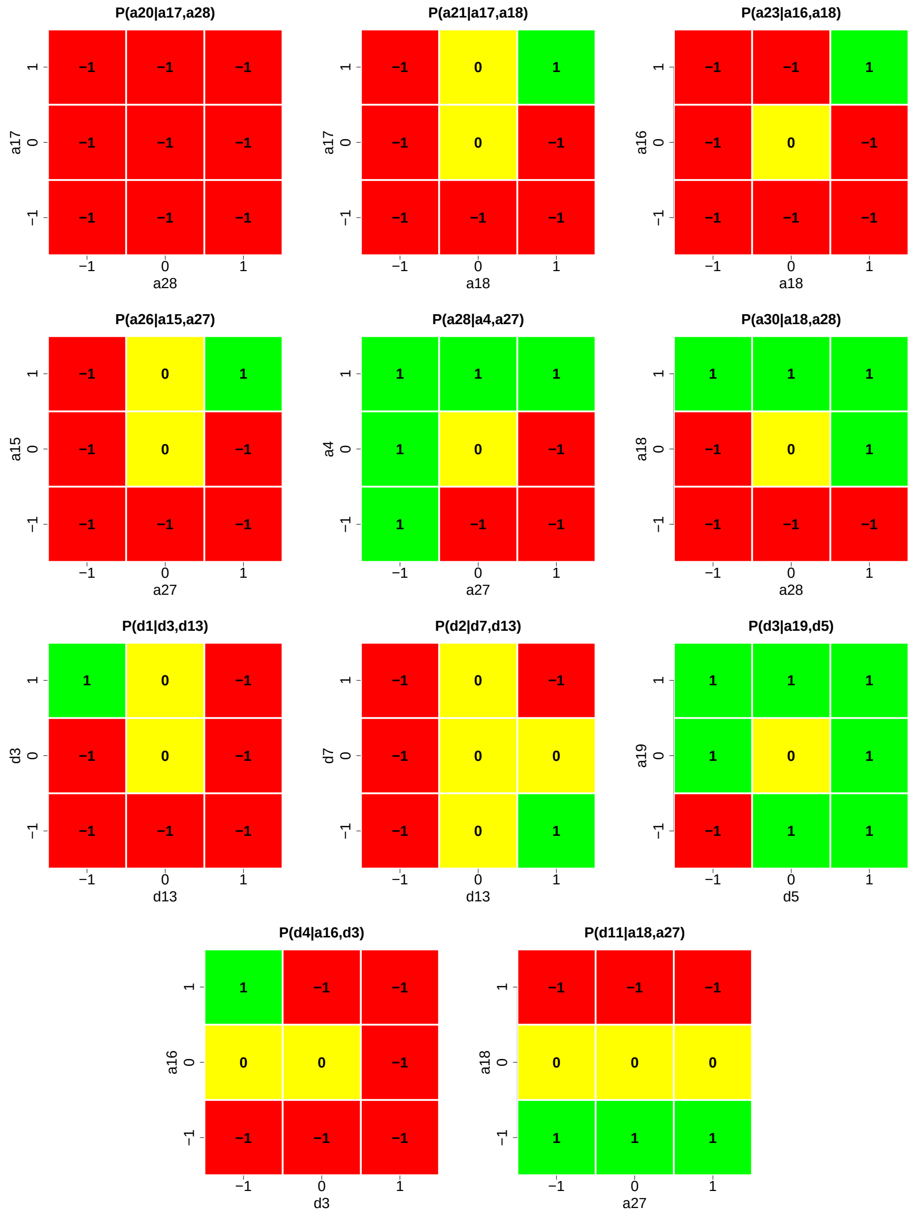
B. PERSONAL QUESTIONS

	name	value −1 value +1
D_1	Perception of success	I consider myself a successful person. I am not as successful in life as I would like to be.
D_2	Trust in other people	I usually don't trust the people around me, even those I know personally. I usually trust the people around me, even those I don't know.
D_3	Trust in future	The future fills me with optimism. The future makes me fearful.
D_4	Children's lives	I feel that my children will do worse than me. I feel my children will do better than me.
D_5	Changes vs. traditions	I am a person who welcomes change, I am open to change. I am a person for whom traditions and certainties are important.
D_6	Savings vs. debts	I am good at saving money over the long term. I usually don't make ends meet and sometimes I have to borrow money.
D_7	Pride in being Czech	I am proud to belong to the Czech nation. I am ashamed of belonging to the Czech nation.
D_8	State aid	When I needed help from the state, it did not come. When I needed help, I got help from the state.
D_9	Responsibility	I feel responsible for the current political state of Czechia. I do not feel responsible for the current political state of Czechia.
D_{10}	Feeling heard	The voices and interests of people like me are forgotten in political debates in Czechia. The voices and interests of people like me are often heard in political debates in Czechia.
D_{11}	Benefits of EU	EU membership is good for people like me. EU membership is not beneficial for people like me.
D_{12}	Care from others	I have many people around me who will look after me when I am ill or need something. I have no one to look after me when I am ill or need something.
D_{13}	Loneliness	I often feel lonely. I have enough people around me to spend time with and talk about what I am experiencing.
D_{14}	Flexibility vs. security	I prefer freedom and flexibility, although it also brings uncertainty. I prefer certainty and calculability, even if it brings less freedom and flexibility.

C. CONDITIONAL PROBABILITY TABLES WITH TWO PARENTS

In each CPT the most probable value of the dependent variable for the corresponding combination of parents' values is displayed. The red color corresponds to the value -1, yellow to the neutral value 0, and green to the value $+1$. Note that there is one CPT ($P(A_{20}|A_{17}, A_{28})$) where the most probable value of the dependent variable is the same for the corresponding combination of parents' values. The probabilities of $A_{20} = -1$ differ for different combinations of parents' values but are always higher than the probabilities of the other two states. Interestingly, this is the CPT of the only OLR model with two parents.


P(a2|a4,a18)
a4
a18
P(a5|a19,d3)
a19
d3
P(a6|a8,a20)
a8
a20
P(a8|a1,a7)
a1
a7
P(a10|a9,a24)
a9
a24
P(a11|a7,a15)
a7
a15
P(a12|a1,a19)
a1
a19
P(a15|a19,d11)
a19
d11
P(a16|a18,a21)
a18
a21
P(a17|a18,d10)
a18
d10
P(a18|a27,a28)
a27
a28
P(a19|a1,d11)
a1
d11


P(a20|a17,a28)
a17
a28
P(a21|a17,a18)
a17
a18
P(a23|a16,a18)
a16
a18
P(a26|a15,a27)
a15
a27
P(a28|a4,a27)
a4
a27
P(a30|a18,a28)
a18
a28
P(d1|d3,d13)
d3
d13
P(d2|d7,d13)
d7
d13
P(d3|a19,d5)
a19
d5
P(d4|a16,d3)
a16
d3
P(d11|a18,a27)
a18
a27

D. CONDITIONAL PROBABILITY TABLES WITH ONE PARENT

Note that there exist CPTs where some values are never the most probable. However, it does not mean that their probabilities are constant. Actually, there is always a trend, as described in the footnotes of Tables 5, 7, 4, and 8. Similarly, as in Section C, the red color corresponds to the value -1, yellow to the neutral value 0, and green to the value $+1$.