A Generalized Voting Game for Categorical Network Choices

Yueh Lin* Stefano Nasini*† Martine Labbé ‡

December 3, 2024

Abstract

This paper presents a game-theoretical framework for data classification and network discovery, focusing on pairwise influences in multivariate choices. The framework consists of two complementary games in which individuals, connected through a signed weighted graph, exhibit network similarity. A voting rule captures the influence of an individual's neighbors, categorized as attractive (friend-like) or repulsive (enemy-like), and encodes individuals' payoffs based on network similarity. We establish a duality between these two games, distinguishing between the endogeneity of choices (direct voting) and network information (inverse voting). While the latter has applications in network discovery, the direct voting game results in a data classification methodology that generalizes the K-nearest neighbors approach. Our theoretical results provide conditions for the existence of Nash equilibria and demonstrate the NP-completeness of their characterization. On the empirical side, we test our methodology with three applications and show the advantages (in terms of goodness-of-fit) of our game-theoretical framework in addressing both data classification and network discovery.

Keywords: Combinatorial games, Combinatorial optimization, Network influence, Categorical regression.

*IESEG School of Management, Univ. Lille, CNRS, UMR 9221 -LEM - Lille Economie Management, F-59000 Lille, France

†Corresponding author. Email: s.nasini@ieseg.fr.

‡Free University of Bruxelles, CP212, boulevard du Triomphe, 1050 Bruxelles, Belgium

1 Introduction

Pairwise (network) dependencies among categorical data are ubiquitous across various empirical contexts. In the social and economic sciences, empirical studies showcase the sizable impact of these dependencies, spanning from countries taking governmental actions that mirror the ones of their allies, to teenagers adopting behaviors that imitate the ones of their peers (Patacchini and Zenou 2012). Beyond the social and economic context, network dependencies have been frequently used to encode patterns of similarities in machine learning models. In this regard, the metaphorical notion of network has proven successful in capturing complex structural relationships within multidimensional data (Nasini et al. 2017, Lin et al. 2019).

In the statistical and operations research literature, distinct modeling perspectives have traditionally addressed network dependencies in categorical data. Stochastic simulation models, like the linear thresholds model (Granovetter 1978) and the independent cascade model (Goldenberg et al. 2001), paved the way for encoding pattern dependencies among interconnected individuals, through individual-by-individual propagation mechanisms. These approaches capture different forms of pairwise influences, with direct bearing on the determination of influence maximization decisions (Kempe et al. 2015, Wu and Küçükyavuz 2018). Alternatively, a distinct modeling viewpoint is offered by econometric approaches, exemplified by space-time autoregressive models (Cliff and Ord 1975, Stoffer 1986, Borovkova et al. 2008) and their generalization to categorical data (Cohen-Cole et al. 2018, Barigozzi and Brownlees 2018, Nasini and Martínez-de Albéniz 2020). These focus on estimating cross-sectional correlations and pattern similarities (mirroring proximity or relational structures), with a view to identifying social interactions from observed multivariate choices.

While the benefit of stochastic simulation models (in caparison to space-time autoregressive approaches) lies in their flexibility to capture non-linear dependencies, their methodologically accurate applicability is limited by the absence of a dedicated estimation theory, as well as by the lack of a theoretical micro-foundation. In fact, to the best of our knowledge, a small attention has been given to the game theoretic characterization of statistical dependencies, which would serve not only as a modeling micro-foundation for some machine learning approaches, but also as an operational methodology to tackle categorical network data in an integrated and flexible manner.[1]

This paper introduces two complementary games that involve a collection of individuals/players $\mathcal{I}$ connected through a signed weighted graph $\mathcal{G}_{+-} = \langle \mathcal{I}, \mathcal{E}^+, \mathcal{E}^-, W \rangle$ and a set of categorical features/choices $\mathcal{F}$ (e.g., political views, fashion preferences, or social attitudes), with each individual being associated to some of these features.[2] Links in the network are categorized as attractive

[1]See Gambella et al. (2021) for game theoretic models and optimization frameworks for commonly used machine learning approaches.

[2]Throughout the paper, the terms individual and player are used indistinctly, as well as the terms feature and choice.

(friend-like relationships $\mathcal{E}^+$) or repulsive (enemy-like relationships $\mathcal{E}^-$), with the weighting function $W : \mathcal{E}^+ \cup \mathcal{E}^- \to \mathbb{R}$ attaching a degree of strength to each relationship. In both games, payoffs are assigned based on a voting rule reflecting the influence of an individual's neighbors, so that individuals' features are positively influenced by attractive neighbors and negatively influenced by repulsive neighbors. This integrates the effects of attractive and repulsive neighbors into individuals' features, by treating each neighborhood as a voting pool that each individual examines prior to make a choice (adopt a categorical feature).

In social network literature, the phenomenon of pattern similarities between features of connected individuals is referred to as *assortativity* (McPherson et al. 2001). As Nasini et al. (2017) points out, addressing the assortativity issue raises an epistemological question regarding the direction of causality. Specifically, one could either hypothesize that individual features influence the formation of network linkages, or conversely, that the latter drive the emergence of similar features among connected individuals. To navigate this dichotomy, our game-theoretic framework establishes a form of duality between two games based on the endogeneity of choices (direct voting) versus network information (inverse voting). While the direct form assumes that $\mathcal{E}^+$, $\mathcal{E}^-$ and W are exogenous (with features being endogenously characterized by the equilibrium conditions), the inverse form relies upon the opposite design. We refer to these models as the Direct Voting Game for Categorical Network Choices (DVG-CNC) and the Inverse Voting Game for Categorical Network Choices (IVG-CNC), respectively.

The IVG-CNC supports network discovery and the estimation of cross-sectional dependencies from observed choices, in line with the work of Cohen-Cole et al. (2018), Barigozzi and Brownlees (2018) and Nasini and Martínez-de Albéniz (2020). Conversely, the DVG-CNC addresses the statistical classification. In this vein, unlike space-time autoregressive approaches, the DVG-CNC has the capability to discern various forms of non-linearity in multivariate choices while providing a theoretical micro-foundation for the presence of network auto-correlation between categorical data (cross-sectional dependencies). More generally, the DVG-CNC can tackle a spectrum of machine learning problems, extending beyond the scope of social influence and the identification of social interaction effects. By assuming a specific voting form for the aggregation rule that maps neighbors states into the update of an individual state, the DVG-CNC generalizes the K-nearest neighbors methodology (Lin et al. 2019, Lutu and Engelbrecht 2013), resulting in a normative framework for the statistical classification problem.[3]

Our theoretical contributions include deducing sufficient conditions for the existence of a Nash

[3]The main effort in the theoretical analysis of the K-nearest neighbors problem pivots on its convergence properties, when transformed to continuous time stochastic processes. In this vein, Bhattacharya and Mack (1987) show that the K-nearest neighbor estimator has a common limiting structure under the second-order smoothness conditions as the sample size tends to infinity. Likewise, Mack (1981) deriveS the rates of convergence for the bias and variance as well as asymptotic normality of the K-nearest neighbor estimator.

equilibrium, as well as a closed-form characterization of individual best responses. Forthermore, following the recent operations research literature on the numerical characterization of Nash equilibria in noncooperative games (Carvalho et al. 2022, 2017, Porter et al. 2008), we design mixed-integer linear programming (MILP) formulations to approach the equilibrium solutions of both DVG-CNC and IVG-CNC using state-of-the-art optimization solvers.[4] The need for these MILP formulations relates to one of our theoretical results, demonstrating the NP-completeness of our problem. This is in line with previous studies on the computational complexity of finding equilibrium solutions, which have been extensively discussed in the literature (Gilboa and Zemel 1989, Ben-Porath 1990, Koller and Megiddo 1992, Koller et al. 1996, Chu and Halpern 2001).

Our empirical analysis encompasses three applications: two pertaining to direct voting (in which the network is fixed, and choices are determined endogenously) and one for inverse voting (in which the network is endogenously determined while choices remain fixed). The first application delves into geopolitical alliances among 184 countries and allows for predicting their political stance on key issues, such as sustainability, LGBT legal protection, migration policy, and Palestine versus Israel recognition. Our out-of-sample analysis shows that the obtained Nash equilibrium solution of the DVG-CNC is a better fit than the social welfare solution for the observed national choices (as expected by the strategic behaviour of countries). This allows to predict all national choices related to the political stance on LGBT legal protection, and up to 83% national choices related to the political stance on CO2 emission reduction. For the second application, we use the DVG-CNC to impute missing data and highlight its competitive advantage over the KNN approach. Through our voting game, we are able to correctly impute up to 20% more data, in comparison to the KNN approach. For the third application, we apply the IVG-CNC to discover network influences among 164 secondary school students, unveiling network dependencies among students' choices.

The rest of the paper is organized as follows. Section 2 presents the baseline modeling framework for network voting. Section 3 studies the DVG-CNC, providing sufficient conditions for the existence of a Nash equilibrium and the closed-form characterization of players' best responses. Section 4 studies the IVG-CNC. Section 5 addresses the integer programming representation of both the DVG-CNC and the IVG-CNC models. The numerical and empirical analysis is provided in Section 6. Finally, Section 7 concludes the paper. Appendix A contains the mathematical proofs.

2 Baseline modeling framework for network voting

This section presents a baseline modeling framework encoding the pattern dependencies between choices of individuals connected in a network. The introduced definitions and notation are valid for

[4]On the algorithmic side, Dragotto and Scatamacchia (2023) employed integer programming techniques to compute Nash equilibria in some classes of noncooperative games based on a cutting-plane algorithm.

both DVG-CNC and IVG-CNC.

Definitions and notation. We consider a signed weighted graph $\mathcal{G}_{+-} = \langle \mathcal{I}, \mathcal{E}^+, \mathcal{E}^-, W \rangle$, where $\mathcal{I}$ is a collection of individuals/players (with $n = |\mathcal{I}|$), $\mathcal{E} = \mathcal{E}^+ \cup \mathcal{E}^- \subseteq \mathcal{I} \times \mathcal{I}$ is a collection of edges, and $W : \mathcal{I} \times \mathcal{I} \longrightarrow \mathbb{R}$ is a weight function that attaches to each pair of individuals a weight capturing the strength of the relationship. The weight on the edge (i, j) is denoted by $w_{i,j}$, with $w_{i,j} > 0$ iff $(i, j) \in \mathcal{E}$. We denote the neighborhood of $i \in \mathcal{I}$ as $\mathcal{E}_i = \{j \in \mathcal{I} : w_{i,j} > 0\}$, with $n_i = |\mathcal{E}_i|$.

Let $\mathcal{F}$ (with $F = |\mathcal{F}|$) be a set of features (e.g., political orientations, fashion preferences, etc.), and $\boldsymbol{x}_i \subseteq \mathcal{F}$ the subset of these features that are associated with the i-th individual.[5] We denote the feature profile as the n-tuple $\boldsymbol{x} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)$, and the feature profile of all players except i as the n_i-tuple $\boldsymbol{x}_{-i} = (\boldsymbol{x}_j, : j \in \mathcal{E}_i)$, so that $\boldsymbol{x} = \boldsymbol{x}_{-i} \cup \boldsymbol{x}_i$.[6] Additionally, we represent this feature profile by the indicator function $x : \mathcal{I} \times \mathcal{F} \to \{0, 1\}$, defined as $x(i, t) = 1$ if $t \in \boldsymbol{x}_i$, $x(i, t) = 0$ if $t \notin \boldsymbol{x}_i$.

To design a voting rule that mirrors a network dependency between individual features, the elements in $\boldsymbol{x}_i$ are assumed to be influenced by the ones of its neighbors $\boldsymbol{x}_{-i}$. Specifically, each individual can be positively or negatively sensitive to the choices of their neighbors, so that the choice of individual i must positively reflect the ones of individuals in $\mathcal{E}_i^+$ and negatively reflect the ones of individuals in $\mathcal{E}_i^-$. We introduce the binary indicators of attractive and repulsive neighbors:

$$\overline{\psi}_{i,j} = \begin{cases} 1 & \text{if } j \in \mathcal{E}_i^+, \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad \underline{\psi}_{i,j} = \begin{cases} 1 & \text{if } j \in \mathcal{E}_i^-, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Frequently, the notation $\overline{w}_{i,j} = w_{i,j}\overline{\psi}_{i,j}$ and $\underline{w}_{i,j} = w_{i,j}\underline{\psi}_{i,j}$ is adopted to quantify the specific weight/relevance that player i attaches to the choice of player j, for attractive and repulsive neighbors, respectively. When $j \notin \mathcal{E}_i$, then no wight is attached to the choice of j by player i, so that $\overline{w}_{i,j} = \underline{w}_{i,j} = 0$. For ease of notation, we write $\boldsymbol{w}_i$ for the vector $(\overline{w}_{i,1}, \dots, \overline{w}_{i,n}, \underline{w}_{i,1}, \dots, \underline{w}_{i,n})$ so long as no ambiguity arises.

A given feature is consistent with an individual's neighborhood if the total number of matching choices (mismatching choices for repulsive neighbors) are above a certain threshold $\theta_i \in \Theta_i \subseteq \{1, 2, \dots, d_i\}$, where $d_i = \sum_{j \in \mathcal{E}_i} w_{i,j}$. This threshold represents the player's sensitivity to the choices of its neighbors, whose value (i.e., number of votes) is computed by a neighborhood weighting sum:

$$\mathrm{WS}_{i,t} \;=\; \sum_{j \in \mathcal{E}_i^+ :\ t \in \boldsymbol{x}_j} w_{i,j} + \sum_{j \in \mathcal{E}_i^- :\ t \notin \boldsymbol{x}_j} w_{i,j} \;=\; \sum_{j \in \mathcal{I}} \left(\overline{w}_{i,j} x(i,t) + \underline{w}_{i,j}(1 - x(i,t)) \right).$$

This encodes the number of attractive and repulsive matches of the choices of neighbors of i, with the edge weight $w_{i,j}$ quantifying in absolute terms the influence that j has over i.

[5] See Cohen-Cole et al. (2018) and by Nasini and Martínez-de Albéniz (2020) for related multiple features designs.

[6] We use the set notation $\boldsymbol{x}_{-i} \cup \boldsymbol{x}_i$ for a tuple, referring to appending the set $\boldsymbol{x}_i$ at the i-th position of $\boldsymbol{x}_{-i}$.

Two complementary games. We establish a form of duality between the DVG-CNC model and the IVG-CNC model, both having a payoff structure accounting for the total number of matching choices from an individual's neighbors (or mismatching choices for repulsive neighbors). This is based on the endogeneity of features versus network information. The DVG-CNC assumes that $\mathcal{E}^+$, $\mathcal{E}^-$ and W are exogenous (with choices being endogenously characterized by the equilibrium conditions). Therefore, the DVG-CNC model allows predicting individual features selection $\boldsymbol{x}_1, \ldots,$ $\boldsymbol{x}_n$, from the observed network. The IVG-CNC represents the inverse problem of observing players choices over $\mathcal{F}$ and deducing a potential network of pairwise relationships (namely $\mathcal{E}^+$, $\mathcal{E}^-$ and W) that could have influenced these choices. Table 1 summarizes the distinction between endogenous variables (player strategies) versus exogenous parameters in DVG-CNC and IVG-CNC.

	DVG-CNC	IVG-CNC
Endogenous variables (player strategies)	$\boldsymbol{x}_i$ and θ_i	$\boldsymbol{w}_i$ and θ_i
Exogenous parameters	$\boldsymbol{w}_i$	$\boldsymbol{x}_i$

Table 1: Player strategy sets versus exogenous parameters.

Therefore, in the DVG-CNC model players choose features, whereas in the IVG-CNC model players choose network weights. Consistently, the neighborhood weighting sum is regarded as a function of neighbours' features selection in the DVG-CNC model, and denoted as $\mathrm{WS}_{i,t}(\boldsymbol{x}_{-i})$. Conversely, in the context of the IVG-CNC model, the neighborhood weighting sum is regarded as a function of the network weights decision, and denoted as $\mathrm{WS}_{i,t}(\boldsymbol{w}_i)$.

3 The DVG-CNC model

The DVG-CNC model is a simultaneous discrete-game, where n players take categorical decisions about which feature to select (which choices to make), where $\mathcal{E}^+$, $\mathcal{E}^-$ and W are assumed exogenous. We assume that individuals might have mandatory features, as well as forbidden features, in such a way that if $t \in \mathcal{F}$ is a mandatory feature for individual i, then $t \in \boldsymbol{x}_i$ (even when $\mathrm{WS}_{i,t}(\boldsymbol{x}_{-i}) < \theta_i$). Similarly, if t is a forbidden feature for individual i, then $t \notin \boldsymbol{x}_i$ (even when $\mathrm{WS}_{i,t}(\boldsymbol{x}_{-i}) \geq \theta_i$). We denote the set of mandatory and forbidden features as $\mathcal{P}_i^1$ (with $|\mathcal{P}_i^1| = P_i^1$) and $\mathcal{P}_i^0$ (with $|\mathcal{P}_i^0| = P_i^0$) respectively. The set $\mathcal{F}_i$ (with $|\mathcal{F}_i| = F_i$) contains the remaining features that are neither mandatory nor forbidden, so that the subsets of features $\mathcal{F}_i$, $\mathcal{P}_i^1$ and $\mathcal{P}_i^0$ constitute a partitioning of $\mathcal{F}$ (i.e., $\mathcal{F}_i \cup \mathcal{P}_i^1 \cup \mathcal{P}_i^0 = \mathcal{F}$ and $F_i + P_i^1 + P_i^0 = F$). For notational convenience, we also introduce $\mathcal{P}^+ = \{(i,t) \ : \ t \in \mathcal{P}_i^+, \ i \in \mathcal{I}\}$ and $\mathcal{P}^- = \{(i,t) \ : \ t \in \mathcal{P}_i^-, \ i \in \mathcal{I}\}$, and consider the following notation for the DVG-CNC.

Sets:

$$\begin{aligned}
&\tilde{\mathcal{X}}_i \subseteq \mathscr{P}(\mathcal{F}_i) &&: \text{ the set of non-fixed choices that player } i \text{ can take;}\\
&\Theta_i \subseteq \{1,\dots,d_i\} &&: \text{ the set of feasible voting thresholds of player } i;\\
&\mathcal{X}_i := \{\boldsymbol{x} \cup \mathcal{P}_i^1 \ : \ \boldsymbol{x} \in \tilde{\mathcal{X}}_i\} &&: \text{ the strategy set for player } i;\\
&\mathcal{X}_{-i} := \prod_{j \in \mathcal{E}_i^+ \cup \mathcal{E}_i^-} \mathcal{X}_j &&: \text{ the set of } i\text{-th neighbors strategies.}
\end{aligned}$$

Functions:

$$\begin{aligned}
&u_i : \tilde{\mathcal{X}}_i \times \Theta_i \times \mathcal{X}_{-i} \to \mathbb{R} && \text{payoff function of player } i \in \mathcal{I};\\
&v_t : \textstyle\prod_{j \neq i} \mathcal{X}_j \times \Theta_i \to \{0,1\} && \text{voting function of player } i \in \mathcal{I}.
\end{aligned}$$

It must be noted that as i is forced to select features in $\mathcal{P}_i^1$ and forbidden to select the ones in $\mathcal{P}_i^0$, its strategy set $\mathcal{X}_i \times \Theta_i$ can be fully characterized by its choice over the pure strategy set $\tilde{\mathcal{X}}_i \times \Theta_i$. Hence, $\boldsymbol{x}_i = \tilde{x}_i \cup \mathcal{P}_i^1 \in \mathcal{X}_i$. The payoff function u_i is specified by assuming that players wish to take decisions (select a subset of features in $\mathcal{F}$) based on the choices of their attractive neighbors (network assortativity) and repulsive neighbors (network dissortativity):

$$u_i((\tilde{x}_i,\theta_i),\boldsymbol{x}_{-i}) := \begin{cases} u_i^+((\tilde{x}_i,\theta_i),\boldsymbol{x}_{-i}) + u_i^-((\tilde{x}_i,\theta_i),\boldsymbol{x}_{-i}) + u_i^1(\theta_i,\boldsymbol{x}_{-i}) + u_i^0(\theta_i,\boldsymbol{x}_{-i}), & \text{if } n_i > 0 \\ F & \text{otherwise,} \end{cases}$$

where $u_i^+((\tilde{x}_i,\theta_i),\boldsymbol{x}_{-i}) = |\{t : v_t(\boldsymbol{x}_{-i},\theta_i) = 1, t \in \tilde{x}_i\}|$ is the payoff term accounting for free choices supported by neighours, $u_i^-((\tilde{x}_i,\theta_i),\boldsymbol{x}_{-i}) = |\{t : v_t(\boldsymbol{x}_{-i},\theta_i) = 0, t \in \mathcal{F} \setminus \tilde{x}_i\}|$ relates to free choices non-supported by neighours, $u_i^1(\theta_i,\boldsymbol{x}_{-i}) = |\{t : v_t(\boldsymbol{x}_{-i},\theta_i) = 1, t \in \mathcal{P}_i^1\}|$ corresponds to the matching of mandatory features, and $u_i^0(\theta_i,\boldsymbol{x}_{-i}) = |\{t : v_t(\boldsymbol{x}_{-i},\theta_i) = 0, t \in \mathcal{P}_i^0\}|$ corresponds to the matching of forbidden features, with

$$v_t(\boldsymbol{x}_{-i},\theta_i) = \begin{cases} 1 & \text{if} \quad \mathrm{WS}_{i,t}(\boldsymbol{x}_{-i}) \geq \theta_i \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Payoffs can also be expressed as the sum of the contributions of individual-features pairs (IF pairs):

$$u_i((\tilde{x}_i,\theta_i),\boldsymbol{x}_{-i}) = \sum_{t \in \mathcal{F}} u_{i,t}((x(i,t),\theta_i),\boldsymbol{x}_{-i}), \tag{3}$$

where $u_{i,t}((x(i,t),\theta_i),\boldsymbol{x}_{-i}) = \mathbf{1}\left(v_t(\boldsymbol{x}_{-i},\theta_i) = x(i,t)\right)$ and $\mathbf{1}\left(\right)$ denotes the indicator function. When $n_i = 0$ (disconnected node), a player cannot be influenced, so its utility is by convention set to F (as no influence implies complete freedom of choice).

3.1 Key properties and generality of the DVG-CNC model

While the main literature on network influence has focused on average rules to map neighbors states into the update of an individual state (Nasini and Martínez-de Albéniz 2020), few theoretical figures are available for the network propagation properties of voting aggregation rules.[7] In this subsection, we highlight some of their key properties and illustrate the generality of the DVG-CNC model.

Remark 1. *The payoff terms satisfy: (i) $u_i^1(\theta_i, \boldsymbol{x}_{-i}) \in [0, P_i^1]$ is piece-wise constant and non-increasing in θ_i; (ii) $u_i^0(\theta_i, \boldsymbol{x}_{-i}) \in [0, P_i^0]$ is piece-wise constant and non-decreasing in θ_i.*

As studied in Section 3.2, this remark fundamental to the existence conditions of a Nash equilibrium of the DVG-CNC. Properties (i) and (ii) in Remark 1 underscore that the existence of mandatory and forbidden features entails a trade-off in the player's choice of threshold level θ_i, as the increment/decrement of the payoff terms $u_i^0(\theta_i, \boldsymbol{x}_{-i})$ and $u_i^1(\theta_i, \boldsymbol{x}_{-i})$ depends on the neighborhood $\mathcal{E}_i$ and the neighbors' choices $\boldsymbol{x}_{-i}$. In line with the linear threshold model of Granovetter (1978), θ_i captures the susceptibility level of player i. However, unlike in Granovetter (1978), this threshold level is an endogenous variable in the DVG-CNC model (which can be fixed in case $|\Theta_i| = 1$).

This arrangement of components creates a multi-dimensional and constrained formulation, aligning the DVG-CNC with established influence propagation models rooted in voting rules. Specifically, our deterministic simultaneous game offers connections to well-known stochastic processes related to voting systems and network propagation, such as the voter model of Granovsky and Madras (1995). One desirable (and expected) property of this model is that payoff values (and best responses) must behave symmetrically with respect to any feature re-labeling (similar to the models proposed by Granovsky and Madras (1995) and Granovetter (1978)). This property is established hereafter.

Proposition 1. *Let us define the labeling symmetry as $u_{i,t}((x(i,t), \theta_i), \boldsymbol{x}_{-i}) = u_{i,t}((1 - x(i,t), d_i - \theta_i), \mathsf{C}_t(\boldsymbol{x}_{-i}))$, for all $\theta_i \in \Theta_i$ and for all $\boldsymbol{x}_i = \tilde{x}_i \cup \mathcal{P}_i^1 \in \mathcal{X}_i$, where the notation $\mathsf{C}_t(\boldsymbol{x}_{-i})$ refers to the n_i-tuple obtained from $\boldsymbol{x}_{-i}$ by relabeling each $x(j,t)$ as $1 - x(j,t)$, for all $j \in \mathcal{E}_i$. We claim that all instances of the DVG-CNC model for which $\theta_i \in \Theta_i$ and $(d_i - \theta_i) \in \Theta_i$ satisfy the labeling symmetry.*

An immediate consequence of this labeling symmetry is that the majority rule $\Theta_i = \{d_i/2\}$ enforces a special kind of labeling invariance, for which $u_{i,t}((x(i,t), \theta_i), \boldsymbol{x}_{-i}) = u_{i,t}((1-x(i,t), \theta_i), \mathsf{C}_t(\boldsymbol{x}_{-i}))$, for all $\boldsymbol{x}_i = \tilde{x}_i \cup \mathcal{P}_i^1 \in \mathcal{X}_i$.[8] Similarly, this majority rule specification of the DVG-CNC allows presenting it as a generalization of the KNN problem (Lin et al. 2019, Lutu and Engelbrecht 2013), as detailed in the following remark.

Remark 2 (Reduction to KNN)**.** *Let us define a collection of n data points $\mathcal{W} = \{(L_1, x_1), \ldots, (Z_n, x_n)\}$, where $Z_i \in \mathbb{R}^q$ is a q-dimensional vector encoding exogenous covariates for individual $i \in \mathcal{I}$, and x_i is the corresponding binary choice. Let us define a distance metric $D : \mathbb{R}^q \times \mathbb{R}^q \to \mathbb{R}$ and a total*

[7] Focusing on a specific social simulation model, some statistical consequences of using a voting aggregation rule to encode neighbor influence have been studied by Keuschnigg and Ganser (2017) using an agent based algorithm.

[8] As emerging from the theoretical analysis in the next section, a Nash equilibrium solution satisfying a majority rule labeling invariance might not exist.

order $\succ_i$, for each $i \in \mathcal{I}$, satisfying $j \succ_i j'$ iff $D(Z_i, Z_j) < D(Z_i, Z_{j'})$, for all $j,\ j' \in \mathcal{I}$. A directed weighted graph $\mathcal{G} = \langle \mathcal{I}, \mathcal{E}, W\rangle$ is obtained from these pairwise distances:

$$\begin{cases} (i,j) \in \mathcal{E}_i^+ & \text{if there exists } K \text{ distinct individuals } \mathcal{J}_i \subset \mathcal{I} \text{ such that } j \succ_i j' \text{ for all } j' \in \mathcal{J}_i. \\ (i,j) \notin \mathcal{E}_i^+ & \text{otherwise,} \end{cases}$$

and $\mathcal{E}_i^- = \emptyset$, for all $i \in \mathcal{I}$. Further, $w_{i,j} = 1$ for all $(i,j) \in \mathcal{E}_i^+$ and $w_{i,j} = 0$ for all $(i,j) \notin \mathcal{E}_i^+$. As a result, the voting rule for a DVG-CNC representation of a KNN problem is

$$v_t(\boldsymbol{x}_{-i}, \theta_i) = \begin{cases} 1 & \text{if } WS_{i,t}(\boldsymbol{x}_{-i}) \geq \dfrac{K}{2} \\ 0 & \text{otherwise.} \end{cases}, \qquad \text{where} \quad WS_{i,t}(\boldsymbol{x}_{-i}) = \sum_{j \in \mathcal{E}_i^+} x(i,t).$$

Remark 2 reveals the most important level of generality of the DVG-CNC model, establishing the KNN problem as one its particular specifications. This allows not only to establish a game theoretic framework to the statistical classification problem, but also to address the determination of the optimal K using the endogeneity of the thresholds θ_i (for each $i \in \mathcal{I}$). As already mentioned, θ_i is an endogenous choice in the DVG-CNC, so that specific definitions of the set Θ_i can be adopted when using the DVG-CNC for data classification purposes.

In the same vein, the sets of mandatory and forbidden features $\mathcal{P}_i^1$ and $\mathcal{P}_i^0$ play an important role in the applicability of the DVG-CNC game to the problem of missing data imputation, as numerically studied in Subsection 6.2. In this case, each element $(Z_i, x_i) \in \mathcal{W}$ is associated with a partitioning between observed choices (corresponding to mandatory and forbidden features $\mathcal{P}_i^1$ and $\mathcal{P}_i^0$) and missing choices (corresponding to the set $\mathcal{F}_i$). Consequently, missing data constitute the endogenous choices for the DVG-CNC game, which are determined either by a social welfare criterion or by a Nash equilibrium criterion, as studied in the next session.

3.2 The existence of equilibrium in the DVG-CNC model

This section examines sufficient conditions for the existence of a Nash equilibrium of the DVG-CNC model and provides a closed-form characterization of players best responses.

3.2.1 From mandatory and forbidden features to DVG-CNC supportability

As previously mentioned, the existence of mandatory and forbidden features (based on the sets $\mathcal{P}_i^1$ and $\mathcal{P}_i^0$, respectively) entails a trade-off in players' choice of their threshold levels, which is critical in the characterization of equilibrium solutions (see Remark 1). An algebraic formulation of this trade-off is a necessary step to establish our theoretical result. Hence, for technical convenience, we represent $\mathcal{P}_i^1$ and $\mathcal{P}_i^0$ by the prerequisite function $p_0 : \text{dom}(p_0) \to \{0,1\}$, which is defined as $p_0(i,t) = 1$ if $t \in \mathcal{P}_i^1$ and $p_0(i,t) = 0$ if $t \in \mathcal{P}_i^0$, where

$$\text{dom}(p_0) = \mathcal{P}^+ \cup \mathcal{P}^- = \bigcup_{i \in \mathcal{I}} \left(\{i\} \times \left(\mathcal{P}_i^1 \cup \mathcal{P}_i^0\right)\right) \subseteq \mathcal{I} \times \mathcal{T}.$$

This subsection investigates the interplay of mandatory and forbidden features with the network topology of attractive and repulsive neighbors by introducing five new notions: *prerequisiteness*, *satisfiability*, *supportedness*, *supportability* and *extension*.

Definition 1 (Prerequisiteness). *A Boolean-valued function* $p : \mathrm{dom}(p) \to \{0, 1\}$ *is called a prerequisite for a defined voting game iff dom*$(p_0) \subseteq \mathrm{dom}(p) \subseteq \mathcal{I} \times \mathcal{F}$. *Further, if* p *and* q *are two prerequisites, we say that* q *extends* p *if* $\mathrm{dom}(p) \subset \mathrm{dom}(q)$ *and* $\forall (i,t) \in \mathrm{dom}(p),\ p(i,t) = q(i,t)$.

Definition 2 (Satisfiability). *Given a choice profile* $\boldsymbol{x}$*, we say that* $\boldsymbol{x}$ *satisfies a prerequisite* p *if* $x|_{\mathrm{dom}(p)}$ *(the choice indicator function* x *restrained to the domain of* p*) is pointwise equal to* p.

The notion of prerequisiteness in Definition 1 generalizes the idea of mandatory and forbidden feature selection by considering the inclusion of extra IF pairs in addition to the ones already in $\mathrm{dom}(p_0)$. This is achieved by the extension of a prerequisite function, which consists of enlarging the collection of IF pairs starting from the baseline prerequisite p_0. Definition 2 formalizes the consistency between individual choices and a given prerequisite. Hence, the notion of satisfiability of a strategy profile $\boldsymbol{x}$ with respect to a prerequisite p implies that for any IF pair $(i,t) \in \mathrm{dom}(p)$, if $p(i,t) = 1$, then $x(i,t) = 1$ and if $p(i,t) = 0$, then $x(i,t) = 0$.

Definition 3 (Supportedness and supportability).

(i) Let $s \in \{0,1\}$. *An IF pair* (i,t) *is supported by* $\langle \boldsymbol{x}_{-i}, s \rangle$*, iff*

$$\begin{cases} \mathrm{WS}_{i,t}(\boldsymbol{x}_{-i}) > 0 & \textit{in case } s = 1 \textit{ (positive supportedness)}, \\ \mathrm{WS}_{i,t}(\boldsymbol{x}_{-i}) < d_i & \textit{in case } s = 0 \textit{ (negative supportedness)}. \end{cases}$$

(ii) Given a prereaquisite p *and* $s \in \{0,1\}$*, an IF pair* (i,t) *is supportable by* $\langle p, s \rangle$*, iff there exists* $\boldsymbol{x}_{-i}$ *satisfying* p*, such that* (i,t) *is supported by* $\langle \boldsymbol{x}_{-i}, s \rangle$.

Definition 4 (One-step supportability and extension).

(i) Let p *be a prerequisite and* $s \in \{0,1\}$. *An IF pair* (i,t) *is one-step supportable by* $\langle p, s \rangle$ *iff there exists* $j \in \mathcal{E}_i : (j,t) \in \mathrm{dom}(p)$*, such that*

$$\begin{cases} \textit{if } p(j,t) = 1,\ \textit{then } t \in \boldsymbol{x}_j \textit{ and if } p(j,t) = 0 \textit{ then } t \notin \boldsymbol{x}_j & \textit{in case } s = 1, \\ \textit{if } p(j,t) = 1,\ \textit{then } t \notin \boldsymbol{x}_j \textit{ and if } p(j,t) = 0 \textit{ then } t \in \boldsymbol{x}_j. & \textit{in case } s = 0. \end{cases}$$

(ii) Let p *and* q *be two prerequisites. We say that* q *is a one-step extension of* p *iff for all* $(i,t) \in \mathrm{dom}(q) \setminus \mathrm{dom}(p)$*,* (i,t) *is one-step supportable by* $\langle p, s \rangle$*, for some* $s \in \{0,1\}$*, and*

$$\begin{cases} \textit{if } (i,t) \textit{ is one-step supportable by } \langle p, 1 \rangle \textit{ then } q(i,t) = 1, \\ \textit{if } (i,t) \textit{ is one-step supportable by } \langle p, 0 \rangle \textit{ then } q(i,t) = 0. \end{cases}$$

The idea of supportedness (point (i) of Definition 3) formalizes the possibility of obtaining a voting result by adjusting the value of θ_i. Hence, if (i,t) is supported by $\langle \boldsymbol{x}_{-i}, s\rangle$, then i must have a neighbor whose t-th choice is consistent with it. Formally stated, there exists $j \in \mathcal{E}_i$, such that

$$\begin{cases} \text{if } j \in \mathcal{E}_i^+, \text{ then } t \in \boldsymbol{x}_j \text{ and if } j \in \mathcal{E}_i^- \text{ then } t \notin \boldsymbol{x}_j & \text{in case } s = 1 \text{ (positive supportedness)}, \\ \text{if } j \in \mathcal{E}_i^+, \text{ then } t \notin \boldsymbol{x}_j \text{ and if } j \in \mathcal{E}_i^- \text{ then } t \in \boldsymbol{x}_j. & \text{in case } s = 0 \text{ (negative supportedness)}. \end{cases}$$

Similarly, the supportability (point (ii) of Definition 3) designates the possibility of being supported, so that (i,t) being supportable entails node i having a neighbor whose t-th choice is either not restricted by p or consistently restricted by p, where the consistency relates to the value of s. The latter is a binary indicator encoding the orientation (positive or negative) of this consistency.

Generally, supportedness and supportability are not exclusive in s, so that (i,t) might be supported both by $\langle \boldsymbol{x}_{-i}, 1\rangle$ (i.e., positive support) and by $\langle \boldsymbol{x}_{-i}, 0\rangle$ (i.e., negative support) (see Example 2 in Appendix C). The general statements "(i,t) being supported by $\boldsymbol{x}_{-i}$" and "(i,t) being supportable by p" are used throughout this section to describe an IF pair that is supported (supportable for the second statement) either positively or negatively by $\boldsymbol{x}_{-i}$ (p for the second statement). Further, the concept of one-step supportability, as outlined in point (i) of Definition 4, formalizes the idea of a node having neighbors whose choices are consistently restricted by p. In the same vein, the one-step extension q, introduced at point (ii) of Definition 4, entails that each newly appended IF pairs $(i,t) \in \text{dom}(q) \setminus \text{dom}(p)$ is connected to individuals that are already in $\text{dom}(p)$ and whose t-th choice is supportable by p.

Remark 3 (Transitivity)**.** *Notice that if an IF pair (i,t) is one-step supportable by $\langle p, s\rangle$, and q is an extension of p, then (i,t) is also one-step supportable by $\langle q, s\rangle$.*[9] *Conversely, if an IF pair (i,t) is supportable by $\langle p, s\rangle$, and q is an extension of p, then (i,t) is might not be supportable by $\langle q, s\rangle$.*

The introduced notions of satisfiability, supportedness, supportability, and extension refer to single IF pairs (i,t). Example 2 in Appendix C clarifies these notions on an illustrative four-nodes network. The next definition outlines uniform supportability with respect to a family of IF pairs, which will be used in the next subsection to characterize individual best responses and sufficient conditions for the existence of a Nash equilibrium. To do so, we formalize the effect of attractive and repulsive neighbors on the orientation of the support, by defining the *voting orientation mapping σ* that attaches to each IF pair a Boolean value.

Definition 5. *Let $\mathcal{D} \subseteq \mathcal{I} \times \mathcal{T}$ be a family of IF pairs, p a prerequisite function, and $\sigma : \mathcal{D} \longrightarrow \{0,1\}$ a voting orientation mapping. We say that $\mathcal{D}$ is uniformly supportable by $\langle p, \sigma\rangle$, iff there exists $\boldsymbol{x}$ satisfying p (Definition 2), such that for all $(i,t) \in \mathcal{D}$, (i,t) is supported by $\langle \boldsymbol{x}_{-i}, \sigma(i,t)\rangle$.*

[9]This comes from the fact that there exists a neighbor $j \in \mathcal{E}_i$ whose t-th choice is in p, namely $(j,t) \in \text{dom}(p)$. Further, given that q is an extension of p, it also contains the IF pair (j,t) with the same value.

3.2.2 Best responses of the DVG-CNC model

This subsection provides different characterizations of the uniform supportability and shows that this notion encodes the interplay of mandatory and forbidden features with the network of attractive and repulsive neighbors. This is needed to determine players' best responses. Lemma 1 and Theorem 2 below establish sufficient conditions for the uniform supportability of a collection of IF pairs.

Lemma 1. *Let $\mathcal{D}$ be a set of IF pairs, p be a prerequisite, and σ be a voting orientation in $\mathcal{D}$. If for each $(i,t) \in \mathcal{D}$, (i,t) is one-step supportable by $\langle p, \sigma_{i,t} \rangle$, then $\mathcal{D}$ is uniformly supportable by $\langle p, \sigma \rangle$.*

Theorem 1. *Let $(i,t) \in \mathcal{I} \times \mathcal{T}$ be an IF pair and p be a prerequisite. We denote as $\tilde{i} = (i_0, i_1, \ldots, i_k)$ an arbitrary path in $\mathcal{G}_{+-}$ (with $i_0 = i$) and as $\phi(\tilde{i}) = (\phi_1(\tilde{i}), \ldots, \phi_k(\tilde{i})) \in \{0,1\}^k$ the corresponding indicator of the attractiveness and repulsiveness of each element in $\tilde{i}$ (i.e., $\phi_\ell(\tilde{i}) = \overline{\psi}_{i_{\ell-1}, i_\ell}$). If there exists $\tilde{i}$ with $(i_k, t) \in dom(p)$, then (i,t) is supportable by $\langle p, s \rangle$, where $s = 1$ iff $1 - p(i_k, t) + \sum_{l=1}^{k}(1 - \phi_l(\tilde{i}))$ is even, and $s = 0$, otherwise.*

Theorem 2. *Let $\mathcal{D} \subseteq \mathcal{I} \times \mathcal{T}$ be a family of IF pairs. Assume that there exists a voting orientation mapping σ and a sequence of prerequisites $P_K = (p_0, p_1, \ldots, p_K)$, satisfying the following conditions:*

(i) For all $k = 2, \ldots, K$, p_k is a one-step extension of p_{k-1}.

(ii) $\mathcal{D} \subseteq \mathrm{dom}(p_K)$.

(iii) For all $(i,t) \in \mathcal{D}$, we have $p_K(i,t) = \sigma_{i,t}$.

We claim that $\mathcal{D}$ is uniformly supportable by $\langle p_0, \sigma \rangle$.

Corollary 1. *Let $\mathcal{D} = \{(i^1, t^1), \ldots, (i^D, t^D)\} \subseteq \mathcal{I} \times \mathcal{T}$ be a family of IF pairs, with $D = |\mathcal{D}|$, and p be a prerequisite. Let $\tilde{i}^d$ and $\tilde{\Sigma}^d$ for $d = 1, \ldots, D$ be a family of paths in $\mathcal{G}_{+-}$ and a family of voting orientations, respectively:*

$$\tilde{i}^d = (i_1^d, \ldots, i_{k_d}^d) \in \mathcal{I}^{k_d}, \quad \text{and} \quad \tilde{\Sigma}^d = (\Sigma_1^d \ldots, \Sigma_{k_d}^d) \in \{0,1\}^{k_d}, \quad \forall\, d = 1, \ldots, D.$$

Assume that for all $d = 1, \ldots, D$, the following conditions are satisfied :

(i) $i_1^d = i^d$ and $p(i_{k_d}^d, t) \in dom(p)$;

(ii) $\forall\, q \leq k_d$ and $q' \leq k_{d'}$, if $i_q^d = i_{q'}^{d'}$ and $t^d = t^{d'}$, then $\Sigma_q^d = \Sigma_{q'}^{d'}$;

(iii-a) $\forall\, q : 2 \leq q \leq k_d$, if i_q^d is an attractive neighbor of i_{q-1}^d (i.e., $\overline{\psi}_{i_q^d, i_{q-1}^d} = 1$), then $\Sigma_q^d = \Sigma_{q-1}^d$;

(iii-b) $\forall\, q : 2 \leq q \leq k_d$, if i_q^d is a repulsive neighbor of i_{q-1}^d (i.e., $\underline{\psi}_{i_q^d, i_{q-1}^d} = 1$), then $\Sigma_q^d = 1 - \Sigma_{q-1}^d$.

We claim that $\mathcal{D}$ is uniformly supportable by $\langle p, \sigma \rangle$, where $\sigma(i^d, t^d) = \Sigma_1^d$, for all $d = 1, \ldots, D$.

By Theorem 1, if a directed path between i and i_k exists in $\mathcal{G}_{+-}$ and $p(i_k, t) = 1$, an even number of enemies is needed to support (i, t) positively, whereas if $p(i_k, t) = 0$, an odd number of enemies is needed to support (i, t) negatively. The same result can be attained as a consequence of Theorem 2, by letting $\mathcal{D} = \{(i, t)\}$ be a singleton set of IF pairs. If there exists a sequence of one-step extending prerequisites $(p_0, p_1, \ldots, p_k)$, with each extension consisting of one individual at a time (namely, $\mathrm{dom}(p_l) \setminus \mathrm{dom}(p_{l-1}) = \{i_l\}$ for some $i_l \in \mathcal{I}$, for $l = 1, \ldots, k$), then $\mathcal{D}$ is uniformly supportable by p with proper orientation. Accordingly, a sequence of one-step extensions generalizes the notion of directed path and allows for the assessment of the uniform supportability of IF pairs. Example 3 in Appendix C clarifies the idea of one-step extensions on an illustrative four-nodes network.

Lemma 2. *If an IF pair (i, t) is positively (negatively) supported by $\boldsymbol{x}_{-i}$, then there exists θ_i such that $v_t(\boldsymbol{x}_{-i}, \theta_i) = 1$ ($v_t(\boldsymbol{x}_{-i}, \theta_i) = 0$). Furthermore, θ_i can be chosen as $\min \Theta_i$ (as $\max \Theta_i$).*

Theorem 3. *We have the following result.*

(i) If $\mathcal{D}^+ = \bigcup_{i \in \mathcal{I}} \left(\{i\} \times \mathcal{P}_i^1\right)$ is uniformly supportable by $\langle p_0, \sigma\rangle$ with $\sigma_{i,t} = 1$, for all $(i, t) \in \mathcal{D}^+$, there exists choice profile $\boldsymbol{x}$ satisfying p_0, such that for all $i \in \mathcal{I}$, and $\theta_i \leq \theta_i^{1,*}(\boldsymbol{x}_{-i})$, we have

$$u_i^1(\theta_i, \boldsymbol{x}_{-i}) = P_i^1, \quad \text{where} \quad \theta_i^{1,*}(\boldsymbol{x}_{-i}) := \begin{cases} \max\limits_{\theta_i \in [0, d_i]} & \theta_i \\ s.t. & v_t(\boldsymbol{x}_{-i}, \theta_i) = 1, \quad t \in \mathcal{P}_i^1. \end{cases} \tag{4}$$

(ii) If $\mathcal{D}^- = \bigcup_{i \in \mathcal{I}} \left(\{i\} \times \mathcal{P}_i^0\right)$ is uniformly supportable by $\langle p_0, \sigma\rangle$ with $\sigma_{i,t} = 0$, for all $(i, t) \in \mathcal{D}^-$, there exists a choice profile x satisfying p_0, such that for all $i \in \mathcal{I}$, and $\theta_i \geq \theta_i^{0,*}(\boldsymbol{x}_{-i})$, we have

$$u_i^0(\theta_i, \boldsymbol{x}_{-i}) = P_i^0, \ \text{where} \ \theta_i^{0,*}(\boldsymbol{x}_{-i}) := \begin{cases} \min\limits_{\theta_i \in [0,1]} & \theta_i \\ s.t. & v_t(\boldsymbol{x}_{-i}, \theta_i) = 0 \quad \forall\ t \in \mathcal{P}_i^0. \end{cases} \tag{5}$$

By Theorem 3, if the set of mandatory and forbidden IF pairs is such that $\mathcal{D}^+$ is uniformly supportable by $\langle p_0, \sigma\rangle$ with $\sigma_{i,t} = 1$, for all $(i, t) \in \mathcal{D}^+$, and $\mathcal{D}^-$ is uniformly supportable by $\langle p_0, \sigma\rangle$ with $\sigma_{i,t} = 0$, for all $(i, t) \in \mathcal{D}^-$, then $u_i^0(\theta_i, \boldsymbol{x}_{-i})$ and $u_i^1(\theta_i, \boldsymbol{x}_{-i})$ attain their maximum values P_i^0 and P_i^1, respectively, in line with Remark 1. Further, Theorem 2 offers a sufficient condition for the uniform supportability of the set of mandatory and forbidden IF pairs.

The shorter notations $\theta_i^{1,*}$ and $\theta_i^{0,*}$ shall be used in place of $\theta_i^{1,*}(\boldsymbol{x}_{-i})$ and $\theta_i^{0,*}(\boldsymbol{x}_{-i})$, respectively, when the context allows.

Theorem 4. *If the set of mandatory and forbidden IF pairs are uniformely supportable, then for all $i \in \mathcal{I}$, we have the following properties:*

(i) $\overline{u}_i(\boldsymbol{x}_{-i}) = \max\limits_{(\tilde{x}_i, \theta_i) \in \mathscr{P}(\mathcal{F}_i) \times \Theta_i} u_i((\tilde{x}_i, \theta_i), \boldsymbol{x}_{-i}) = F \quad$ iff $\quad \theta_i^{0,*} \leq \theta_i^{1,*}$,

(ii) $\overline{u}_i(\boldsymbol{x}_{-i}) = \max_{(\tilde{x}_i,\theta_i)\in\mathscr{P}(\mathcal{F}_i)\times\Theta_i} u_i((\tilde{x}_i,\theta_i),\boldsymbol{x}_{-i}) \geq F_i + \max\left\{P_i^0 + u_i^1(\underline{\theta}_i^{0,*},\boldsymbol{x}_{-i}), P_i^1 + u_i^0(\overline{\theta}_i^{1,*},\boldsymbol{x}_{-i})\right\}.$

Therefore, building on Theorems 3 and 4, we can characterize players best responses and their relations with $\theta_i^{0,*}$, $\theta_i^{1,*}$, P_i^0, and P_i^1. It results that $\theta_i^{0,*}$ and $\theta_i^{1,*}$ contains all relevant information about the prerequisite structure, concerning the set of mandatory and forbidden IF pairs and the pairwise links between players in $\mathcal{G}_{+-}$.

3.2.3 Nash equilibrium of the DVG-CNC model

A pure strategy profile $\{(\tilde{x}_i^*,\theta_i^*)\ :\ i\in\mathcal{I}\}$ is an equilibrium point of the DVG-CNC iff it satisfies: $u_i((\tilde{x}_i^*,\theta_i^*),\tilde{\boldsymbol{x}}_{-i}^*) \geq u((\tilde{x}_i,\theta_i),\tilde{\boldsymbol{x}}_{-i}^*)$, for each $(\tilde{x}_i,\theta_i)\in\tilde{\mathcal{X}}_i\times\Theta_i,\ i\in\mathcal{I}$. A pure strategy profile $\{(\tilde{x}_i^W,\theta_i^W)\ :\ i\in\mathcal{I}\}$ is a social welfare point iff it solves the following problem

$$\overline{\overline{u}}(\mathcal{G}_{+-}) = \max \sum_{i\in\mathcal{I}} u_i((\tilde{x}_i,\theta_i),\tilde{\boldsymbol{x}}_{-i}) \ \text{ subj. to } (\tilde{x}_i,\theta_i)\in\tilde{\mathcal{X}}_i\times\Theta_i\ \forall i\in\mathcal{I}. \tag{6}$$

By construction $\overline{\overline{u}}(\mathcal{G}_{+-}) \leq nF$, and the case $\overline{\overline{u}}(\mathcal{G}_{+-}) = nF$ corresponds to a strategy profile where each player's choice satisfies the voting rule.

Theorem 5. *We have the following properties:*

(1.) If $\overline{\overline{u}} = nF$, then any social welfare solution is an equilibrium point.

(2.) If for each $i\in\mathcal{I}$ and each neighbor's strategy $\boldsymbol{x}_{-i}\in\mathcal{X}_{-i}$, one has $\theta_i^{0,*}(\boldsymbol{x}_{-i}) > \theta_i^{1,*}(\boldsymbol{x}_{-i})$, then any equilibrium point (if it exists) is a welfare solution.

(3.) If for each $i\in\mathcal{I}$ and each neighbor's strategy $\boldsymbol{x}_{-i}\in\mathcal{X}_{-i}$, one has $\theta_i^{0,*}(\boldsymbol{x}_{-i}) \leq \theta_i^{1,*}(\boldsymbol{x}_{-i})$ and $\overline{\overline{u}} < nF$, then the voting game has no equilibrium point.

Theorem 6. *If $\mathcal{G}_{+-}$ is a negative cycle and $\mathcal{P}_i^1 = \mathcal{P}_i^0 = \emptyset$, for all $i\in\mathcal{I}$, then the voting game admits no equilibrium point and $\overline{\overline{u}}(\mathcal{G}_{+-}) < Fn$.*

Theorem 5 showcases how the prerequisite structure, which is encoded in $\theta_i^{0,*}(\boldsymbol{x}_{-i})$ and $\theta_i^{1,*}(\boldsymbol{x}_{-i})$, constitutes the fundamental determinant for the relationship between the social welfare solution and the Nash equilibrium solution. Theorem 6 establishes that the latter might not exist under specific network topologies, even in the absence of mandatory and forbidden IF pairs.

For the sake of clarity, the theoretical notions examined in Theorems 3, 4, 5 and 6 are illustrated in the Example 4 in Appendix C using a three-node network.

Building upon an extensive body of literature examining the computational intricacies of equilibrium solutions (Gilboa and Zemel 1989, Ben-Porath 1990, Koller and Megiddo 1992, Koller et al. 1996, Chu and Halpern 2001), the subsequent proposition establishes the NP-completeness of the problem of finding a Nash equilibrium of the game.

Proposition 2 (NP-completeness)**.** *Let* R-DVG-CNC *be the recognition version of the proposed DVG-CNC model, which asks whether there exists a Nash equilibrium such that all utility functions are larger than or equal to a given threshold. There exists at least one specification of* $\mathcal{I}$, $\mathcal{E}^+$, $\mathcal{E}^-$, $\mathcal{P}^+$, $\mathcal{P}^-$, W, *and* Θ *for which* R-DVG-CNC *is NP-complete.*

4 The IVG-CNC model

The IVG-CNC model represents the inverse problem of observing players choices over $\mathcal{F}$ and deducing a potential network of pairwise relationships that could have influenced these choices, based on the voting mechanism studied so far. To motivate this inverse problem, it must be noticed that, when presented with a set of noisy observations of individual choices $\mathcal{V} = \{x_{i,t}\}_{i,t}$, the quest is the characterization of the individual threshold levels $\{\theta_i\}_{i\in\mathcal{I}}$, as well as the structures of attractive and repulsive neighborhoods $\mathcal{E}_i^+ i \in \mathcal{I}$ and $\mathcal{E}_i^- i \in \mathcal{I}$.

Specifically, in the IVG-CNC, players decide about how to partition their neighborhoods into attractive and repulsive neighbors, as well as about the specific weight attached to each of them. Therefore, the strategy space of the IVG-CNC is defined as:

$$\mathcal{Q}_i = \{(\boldsymbol{w}_i, \theta_i) \in \mathcal{W}_i \times \Theta_i \ : \ 0 \leq \theta_i \leq ||\boldsymbol{w}_i||_1 = 1\},$$

where $\Theta_i \subseteq \mathbb{R}$ is the set of feasible voting thresholds for player i, and

$$\mathcal{W}_i = \{\boldsymbol{w}_i = (\overline{w}_{i,1}, \dots, \overline{w}_{i,n}, \underline{w}_{i,1}, \dots, \underline{w}_{i,n}) \in \mathbb{R}_+^{2n} \ : \ \overline{w}_{i,j}\underline{w}_{i,j} = 0, \ \forall j \in \mathcal{E}_i, \overline{w}_{i,j} = \underline{w}_{i,j} = 0, \ \forall j \notin \mathcal{E}_i\}.$$

Here, we recall that $\overline{w}_{i,j}$ and $\underline{w}_{i,j}$ quantify the specific weight/relevance that player i attaches to the choice of player j, for attractive and repulsive neighbors, respectively. When $j \notin \mathcal{E}_i$, then no wight is attached to the choice of j by player i, so that $\overline{w}_{i,j} = \underline{w}_{i,j} = 0$. Additionally, we require that at most one of the two have positive values, as a neighbor cannot be both attractive and repulsive at the same time. Mirroring the DVG-CNC formulation, we re-define the domains of the weighted sum, voting, and payoff functions. Therefore, we have $\text{WS}_{i,t} : \mathcal{W}_i \to \mathbb{R}$, $v_{i,t} : \mathcal{W}_i \times \Theta_i \to \{0,1\}$, $u_{i,t} : \mathcal{W}_i \times \Theta_i \to \mathbb{Z}_{\geq 0}$ and $u_i : \mathcal{W}_i \times \Theta_i \to \mathbb{Z}_{\geq 0}$, respectively.

We notice that the IVG-CNC has no strategic interaction (as players payoffs are independent), so the game is separable and equilibrium solutions are social welfare solutions. These can be obtained by maximizing the sum of u_i's:

$$u_i^* = \max_{(w_i,\theta_i)_{i\in\mathcal{I}}} \sum_{t\in\mathcal{F}} u_i(w_i, \theta_i),$$

The following propositions reproduce properties of the DVG-CNC model in the IVG-CNC model, highlighting strict theoretical relationship between them.

Proposition 3 (Labeling symmetry)**.** *Let* (Ins) *and* (Ins′) *be instances of the IVG-CNC, parametrized by* $(\mathcal{I}, \mathcal{E}, \mathcal{F}, X, \Theta)$ *and* $(\mathcal{I}, \mathcal{E}, \mathcal{F}, \mathtt{C}_t(X), \Theta)$*, respectively, where the notation* $\mathtt{C}_t(X)$ *is obtained from* X *by replacing* $x_{1,t}, \ldots, x_{n,t}$ *with* $1 - x_{1,t}, \ldots, 1 - x_{n,t}$*. The following labeling symmetry property holds:*

$$u_{i,t}^{(\text{Ins})}(\boldsymbol{w}_i, \theta_i) = u_{i,t}^{(\text{Ins}')}(\boldsymbol{w}_i, 1 - \theta_i), \qquad \text{for any } \boldsymbol{w}_i \in \mathcal{W}_i, \text{ and some } \theta_i \in \Theta_i.$$

Proposition 4 (Best responses)**.** *We have the following sufficient conditions:*

(i) If there exists $t_1 \neq t_2 \in \mathcal{F}$*, such that* $x_{i,t_1} \neq x_{i,t_2}$ *and for all* $j \in \mathcal{E}_i$ $x_{j,t_1} = x_{j,t_2}$*, then* $u_i^* < F$*.*

(ii) If $\min_t \{ WS_{i,t}(\boldsymbol{w}_i) \mid x(i,t) = 1\} > \max_t \{ WS_{i,t}(\boldsymbol{w}_i) \mid x(i,t) = 0\}$*, for all* $\boldsymbol{w}_i \in \mathbb{R}_+^{2n}$*, then* $u_i^* < F$*.*

(iii) If there exists $\boldsymbol{w}_i$*, such that* $\min_{t:x(i,t)=1} WS_{i,t}(\boldsymbol{w}_i) > \max_{t:x(i,t)=0} WS_{i,t}(\boldsymbol{w}_i)$*, then* $u_i^* = F$*.*

Proposition 3 is an exact reflection of Proposition 1. Conversely, while Theorem 4 and Proposition 4 aim at identifying sufficient conditions for the best response value to attain F, they reveal that the IVG-CNC presents a simpler combinatorial structure compared to the DVG-CNC.

5 Integer programming representation

While the presented game provides a normative framework for statistical classification (DVG-CNC) and cross-sectional dependency estimation (IVG-CNC), for the resulting methodology to be operational, its solutions must be numerically computed. Designing numerical approaches to characterize Nash equilibria poses important challenges in algorithmic game theory and optimization. Adding to the recent literature exploring the computation of Nash equilibria via integer programming (Dragotto and Scatamacchia 2023, Carvalho et al. 2022, 2017, Porter et al. 2008), this section provides an integer programming representation of the DVG-CNC and IVG-CNC models to approach their social welfare solutions and Nash equilibria in a computationally tractable manner.

5.1 DVG-CNC formulation

Let us introduce the variable $\eta_{i,t}$ to encode the agreement between the t-th choice of player i and its corresponding voting outcome. The social welfare problem (6) can be approached by the following

MILP problem:

$$\begin{cases} \min\limits_{\boldsymbol{\eta},\, \boldsymbol{x},\, \boldsymbol{\omega},\, \boldsymbol{\theta}} \displaystyle\sum_{i\in\mathcal{I}}\sum_{t\in\mathcal{F}}(1-\eta_{i,t}) & & \text{(7a)}\\ \text{s. to } \ (1-\eta_{i,t}) \geq x_{i,t} - \omega_{i,t} & & \text{(7b)}\\ (1-\eta_{i,t}) \geq \omega_{i,t} - x_{i,t} & & \text{(7c)}\\ \theta_i - 1 \geq \left(\displaystyle\sum_{j\in\mathcal{E}_i^+} w_{i,j}x_{j,t} + \sum_{j\in\mathcal{E}_i^-} w_{i,j}(1-x_{j,t})\right) - \omega_{i,t}d_i & i\in\mathcal{I}, d_i>0,\ t\in\mathcal{F} & \text{(7d)}\\ \theta_i \leq \left(\displaystyle\sum_{j\in\mathcal{E}_i^+} w_{i,j}x_{j,t} + \sum_{j\in\mathcal{E}_i^-} w_{i,j}(1-x_{j,t})\right) + (1-\omega_{i,t})d_i & i\in\mathcal{I}, d_i>0,\ t\in\mathcal{F} & \text{(7e)}\\ \theta_i \in \Theta_i,\ x_{i,t}\in\{0,1\},\ \omega_{i,t}\in\{0,1\}. & & \text{(7f)} \end{cases}$$

Problem (7) involves $2nF$ binary variables, $n(F+1)$ continuous variables, and $4nF$ inequality constraints, where variables $\omega_{i,t}$ play the role of the voting result. Hence $\omega_{i,t}=0$, if $\mathrm{WS}_{i,t}(\boldsymbol{x}_{-i}) < \theta_i$, and $\omega_{i,t}=1$, if $\mathrm{WS}_{i,t}(\boldsymbol{x}_{-i}) \geq \theta_i$. Notably, based on constraints (7d) and (7e), if a player $i\in\mathcal{I}$ has no neighbors (namely, $d_i=0$), then such a player has undetermined choices (i.e., the value of the objective function is invariant with respect to the its choices).

Based on Theorem 5, if $\overline{\overline{u}} = nF$, then any social welfare solution $\{(\tilde{x}_i^W, \theta_i^W) \ : \ i\in\mathcal{I}\}$ is an equilibrium point. Therefore, when problem (7) attains this value, the obtained solution is an equilibrium point. In the opposite case, a linearized version of inequality (ii) in Theorem 4 can serve as a necessary condition for approaching an equilibrium point by solving problem (7). This is done by including the continuous variables $\theta_i^{0,*}, \theta_i^{1,*}$, and s_i, as well as the binary variables q_i^0 and q_i^1, for each $i\in\mathcal{I}$, obtaining the following variant of problem (7):

$$\begin{cases} \min & \displaystyle\sum_{i\in\mathcal{I}}\sum_{t\in\mathcal{F}}(1-\eta_{i,t}) + F\sum_{i\in\mathcal{I}} s_i \\ \text{subject to} & \text{constraints (7b)-(7f) and constraints (9a)-(9k),} \\ & q_i^0,\ q_i^1 \ \in\ \{0,1\}, \quad \theta_i^{0,*},\ \theta_i^{1,*} \ \in\ \{1,\ldots,d_i\} \qquad i\in\mathcal{I}. \end{cases} \tag{8}$$

The objective function of (8) quantifies the total number of agreements between the t-th choice of player i and its corresponding voting outcome, so that

$$\eta_{i,t} = x_{i,t}\mathbf{1}\left(\sum_{j\in\mathcal{E}_i^+ : t\in\tilde{\boldsymbol{x}}_j} w_{i,j} + \sum_{j\in\mathcal{E}_i^- : t\notin\tilde{\boldsymbol{x}}_j} w_{i,j} \geq \theta_i\right).$$

It also contains the penalty term $F\sum_{i\in\mathcal{I}} s_i$, whose value is equal to zero iff the conditions of Theorem 2 hold, where s_i is characterized by in system of constraints (9a)-(9k):

$$
\begin{cases}
\sum_{t\in\mathcal{F}} \eta_{i,t} = F_i + \sum_{t\in\mathcal{P}_i^1} \omega_{i,t} + \sum_{t\in\mathcal{P}_i^0} (1-\omega_{i,t}) & i\in\mathcal{I}, \quad (9a)\\
\sum_{t\in\mathcal{P}_i^1} \omega_{i,t} + \sum_{t\in\mathcal{P}_i^0} (1-\omega_{i,t}) + s_i \geq \max\{P_i^0,\ P_i^1\} & i\in\mathcal{I}, \quad (9b)\\
\theta_i^{0,*} - 1 \geq \sum_{j\in\mathcal{E}_i^+} w_{i,j}x_{j,t} + \sum_{j\in\mathcal{E}_i^-} w_{i,j}(1-x_{j,t}) & t\in\mathcal{P}_i^0,\ i\in\mathcal{I}, \quad (9c)\\
\theta_i^{1,*} \leq \sum_{j\in\mathcal{E}_i^+} w_{i,j}x_{j,t} + \sum_{j\in\mathcal{E}_i^-} w_{i,j}(1-x_{j,t}) & t\in\mathcal{P}_i^1,\ i\in\mathcal{I} \quad (9d)\\
P_i^0 - \sum_{t\in\mathcal{P}_i^0} (1-\omega_{i,t}) \leq P_i^0 q_i^0 & i\in\mathcal{I}, \quad (9e)\\
P_i^0 - \sum_{t\in\mathcal{P}_i^0} (1-\omega_{i,t}) \geq q_i^0 & i\in\mathcal{I}\ :\ P_i^0\geq 1, \quad (9f)\\
P_i^1 - \sum_{t\in\mathcal{P}_i^1} \omega_{i,t} \leq P_i^1 q_i^1 & i\in\mathcal{I}, \quad (9g)\\
P_i^1 - \sum_{t\in\mathcal{P}_i^1} \omega_{i,t} \geq q_i^1 & i\in\mathcal{I}\ :\ P_i^1\geq 1, \quad (9h)\\
d_i q_i^0 \geq \theta_i^{0,*} - \theta_i & i\in\mathcal{I}, \quad (9i)\\
d_i q_i^1 \geq \theta_i - \theta_i^{1,*} & i\in\mathcal{I}, \quad (9j)\\
\sum_{t\in\mathcal{F}} \eta_{i,t} \leq F - q_i^0 - q_i^1 & i\in\mathcal{I} \quad (9k)
\end{cases}
$$

Visibly, s_i acts as a slack variable to relax constraints (9b), which formulates together with constraints (9a), a weak version of inequality (ii) in Theorem (4). Further, since $\overline{u}_i(\boldsymbol{x}_{-i}) = \sum_{t\in\mathcal{F}} \overline{\eta}_{i,t}$, constraints (9a) establish that players choices must match the voting results of all free features $\mathcal{F}_i$ (so that best response values must be equal to F_i) plus the additional matching of mandatory and forbidden features $\mathcal{P}_i^0$ and $\mathcal{P}_i^1$ (whose value must be at least as large as $\max\{P_i^0,\ P_i^1\}$). The threshold level θ_i is characterized by enforcing a feasible range of $\theta_i^{1,*}$ and $\theta_i^{1,*}$ in constraints (9c) and (9d). It is worth noticing that if $s_i > 0$ then (9b) is binding, so that constraints (9a) and (9b) combined imply a maximum number of correctly classified choices:

$$
\sum_{t\in\mathcal{F}} \eta_{i,t} \leq F - s_i.
$$

Conversely, when $s_i = 0$ inequality (ii) in Theorem 4 holds. Then, variable $\theta_i^{0,*}$ quantifies the smallest threshold level for which all fixed features $\mathcal{P}_i^0$ are correctly matched by the player selection of θ_i. Similarly, $\theta_i^{1,*}$ is the largest threshold level for which all fixed features $\mathcal{P}_i^1$ are correctly matched by the player selection of θ_i. Hence, if $\theta_i^{0,*} \leq \theta_i^{1,*}$, player i can choose a strategy $(\tilde{x}_i, \theta_i)$, for which

$u_i((\tilde{x}_i, \theta_i), \boldsymbol{x}_{-i}) = F_i + P_i^0 + P_i^1 = F$. Constraints, (9e)-(9j) make use of binary variables q_i^0 and q_i^1 to characterize the best response of player i when $\theta_i^{0,*} > \theta_i^{1,*}$. Hence, if $\theta_i < \theta_i^{0,*}$ then $q_i^0 = 1$ and the number of zero voting results $\sum_{t\in\mathcal{P}_i^0}(1-\omega_{i,t})$ must be strictly less than P_i^0. An analogous reasoning is valid for the case when $\theta_i > \theta_i^{1,*}$, as enforced in constraints (9f), (9h) and (9j). Table 2 formalizes the bounds on θ_i induced by constraints (9e)-(9j).

q_i^0	q_i^1	θ_i range
0	0	$\max\{\mathrm{WS}_{i,t}(\boldsymbol{x}_{-i}) \ : \ t \in \mathcal{P}_{i,t}^0\} + 1 \leq \theta_i \leq \min\{\mathrm{WS}_{i,t}(\boldsymbol{x}_{-i}) \ : \ t \in \mathcal{P}_{i,t}^1\}$
0	1	$\max\{\mathrm{WS}_{i,t}(\boldsymbol{x}_{-i}) \ : \ t \in \mathcal{P}_{i,t}^0\} + 1 \leq \theta_i$
1	0	$\theta_i \leq \min\{\mathrm{WS}_{i,t}(\boldsymbol{x}_{-i}) \ : \ t \in \mathcal{P}_{i,t}^1\}$
1	1	θ_i free

Table 2: Lower and upper bounds on θ_i induced by constraints (9e)-(9j).

Finally, constraint (9k) establishes that the best response value $\overline{u}_i(\boldsymbol{x}_{-i})$ must be less or equal than $F-1$ if either $\theta_i^{0,*} > \theta_i^{1,*}$ or $\theta_i > \theta_i^{1,*}$. Additional valid inequalities can be appended to enforce bounds on the excess quantity s_i, based on whether $\theta_i^{0,*} > \theta_i^{1,*}$ or $\theta_i > \theta_i^{1,*}$:

$$\begin{cases} s_i \leq q_i^0 P_i^0 + q_i^1 P_i^1 & i \in \mathcal{I}, \quad (10a) \\ s_i \geq q_i^0 P_i^0 - \sum_{t\in\mathcal{P}_i^0}(1-\omega_{i,t}) & i \in \mathcal{I}, \quad (10b) \\ s_i \geq q_i^1 P_i^1 - \sum_{t\in\mathcal{P}_i^1}\omega_{i,t} & i \in \mathcal{I}. \quad (10c) \end{cases}$$

Notably, constraints (9a)-(9k) and (10a)-(10c) are necessary but not sufficient conditions to guarantee that the optimal solution of (8) is an equilibrium point. However, as numerically shown in Section 6, they represent a computationally tractable linear programming representation of the non-linear inequalities in Theorem 2, allowing this formulation to approach equilibrium solutions in all analyzed instances with limited computational time.[10]

5.2 IVG-CNC formulation

While in Subsection 5.1 we presented a MILP formulation of the DVG-CNC, this subsection adopts an analogous procedure to approach the IVG-CNC. We also use the decision variable $\eta_{i,t}$ to encode the agreement between the t-th choice of player i and its corresponding voting outcome.

Notably, to construct this MILP formulation, an alternative approach to capture the complementary nature of w_i in the IVG-CNC strategy space $\mathcal{Q}_i$ is to introduce binary variables $\overline{\psi}_{i,j}$ and

[10]See Glover and Woolsey (1974) and Sherali and Adams (1998) for linear programming representations of systems of non-linear constraints.

$\underline{\psi}_{i,j}$ which encodes whether $\overline{w}_{i,j} > 0$ and $\overline{w}_{i,j} > 0$ respectively. We obtain the following MILP formulation of the IVG-CNC model (in line with problem (8) for the DVG-CNC model):

$$\begin{cases} \max_{\boldsymbol{\eta}} \sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{F}} \eta_{i,t} & \text{(11a)} \\ \text{s. to } \sum_{j \in \mathcal{I}} (\overline{w}_{i,j}\ x_{j,t} + \underline{w}_{i,j}(1 - x_{j,t})) \ \leq \ \theta_i - 1 + M_i(1 - \eta_{i,t}), \quad t \in \mathcal{F} : x_{i,t} = 0 & \text{(11b)} \\ \sum_{j \in \mathcal{E}_i} w_{i,j}(\overline{w}_{i,j}\ x_{j,t} + \underline{w}_{i,j}(1 - x_{j,t})) \ \geq \ \theta_i - M_i(1 - \eta_{i,t}), \quad t \in \mathcal{F} : x_{i,t} = 1 & \text{(11c)} \\ \overline{\psi}_{i,j} + \underline{\psi}_{i,j} \leq 1, \quad \forall j \in \mathcal{I}, & \text{(11d)} \\ \overline{w}_{i,j} \leq \overline{\psi}_{i,j} & \text{(11e)} \\ 1 \leq \theta_i \leq \sum_{j \in \mathcal{I}} (\underline{w}_{i,j} + \overline{w}_{i,j}) & \text{(11f)} \\ (\overline{\psi}, \underline{\psi}) \in \Psi, & \text{(11g)} \end{cases}$$

where $M_i = \sum_{j \in \mathcal{I}} w_{i,j}$ and Ψ is a set of feasible network dependency configurations, encoding a-priory information about attractive and repulsive neighbors (our empirical application in the next section will consider constraints involving the network density and the nodal degrees, among others).

Solving problem (11b)-(11g) results in the formation of a signed weighted graph $\mathcal{G}_{+-}$ from the observation of $x_{i,t}$. This translates the IVG-CNC into a network discovery approach for the estimation of cross-sectional dependency in categorical data.

6 Empirical and numerical analysis

With a view to providing an empirical grounding for our game-theoretical framework, this section presents three applications: two pertaining to the DVG-CNC (in which the network is fixed, and choices are determined endogenously) and one for the IVG-CNC (in which the network is endogenously determined while choices remain fixed). Hereafter, each of these three applications is treated in a dedicated subsection.

6.1 Country-by-country influences under geopolitical alignments

For the first application, we consider geopolitical alignments among 183 countries: $\mathcal{I} = \{$Afganistan, Albania, ..., Zambia, Zimbabwue$\}$.[11] We study how the combined influences of attractive and repulsive neighbors shape their political orientation on $m = 5$ internationally relevant issues. Categorical features correspond to these issues: $\mathcal{F} = \{$LGBT legal protection, CO2 reduction policy, gender equality legislation, immigration policy, Palestine versus Israel recognition$\}$.

[11] From the original 193 UN countries, nine countries have been dropped for the lack of complete information with respect different required measures of the whole data architecture. These nine countries are San Marino, Andorra, Marshall Islands, Liechtenstein, Nauru, Palau, the Federated States of Micronesia, South Sudan and Monaco.

We construct the data architecture required to calibrate $\mathcal{I}$, $\mathcal{E}^+$, $\mathcal{E}^-$, $\mathcal{P}^+$, $\mathcal{P}^-$, and W in the DVG-CNC model using eight publicly available data sources: (i) UN general assembly resolutions data from January 1st 2022 to June 1st 2023, (ii) monetary, military, and bilateral trade agreements, from January 1st 2022 to June 1st 2023, (iii) country-by-country bilateral imports from 2017 to 2023, (iv) GDP data from the World Bank, (v) Environmental Performance Index (EPI), (vi) LGBT equality index, and (vii) Migrant Integration Policy index. A detailed description of these data sources is reported in the dedicated online Appendix D.

Attractive and repulsive links $\mathcal{E}^+$ and $\mathcal{E}^-$. Using the data sources (i)-(iv), we define $L_1(i,j)$ and $L_2(i,j)$ the number of co-votes and contra-votes in 91 UN resolutions, respectively, between a pair of countries (i,j). Two countries co-vote iff either they both vote "yes" or "no" on one of the 91 topics. In the second case, two countries contra-vote iff if their votes are opposite. Similarly, we define $L_3(i,j)$ as the number of shared monetary, military, or bilateral trade agreements, and we denote their bilateral export and import from WTO data with $L_4(i,j)$ and $L_5(i,j)$, respectively. The criteria for two countries (i,j) being regarded as neighbors are the following:

- their co-votes in UN resolutions are larger than an upper quantile ($L_1(i,j) > \overline{q}_1$) for attractive neighbors, or smaller than a lower quantile ($L_1(i,j) < \underline{q}_1$) for repulsive neighbors;
- their contra-votes in UN resolutions are smaller than a lower quantile ($L_2(i,j) < \underline{q}_2$) for attractive neighbors, or larger than an upper quantile ($L_2(i,j) > \overline{q}_2$) for repulsive neighbors;
- they share at least one of the considered alliances ($L_3(i,j) \geq 1$) for attractive neighbors, or they don't share any alliances ($L_3(i,j) = 0$) for repulsive neighbors;
- their bilateral export is larger than a specified quantile ($L_4(i,j) > q_4$) for attractive neighbors, or smaller than a lower quantile ($L_4(i,j) < \underline{q}_4$) for repulsive neighbors;
- their bilateral import is larger than a specified quantile ($L_5(i,j) > q_5$) for attractive neighbors, or smaller than a lower quantile ($L_5(i,j) < \underline{q}_5$) for repulsive neighbors.

Building on these measurements, attractive links $\mathcal{E}^+$ are obtained by setting the condition $[\mathbf{1}(L_1 > \overline{q}_1) + \mathbf{1}(L_2 < \underline{q}_2) + \mathbf{1}(L_3 \geq 1) + \mathbf{1}(L_4 > \overline{q}_4) + \mathbf{1}(L_5 > \overline{q}_5)] \geq 3$; repulsive links $\mathcal{E}^-$ are obtained by setting the condition $[\mathbf{1}(L_1 < \underline{q}_1) + \mathbf{1}(L_2 > \overline{q}_2) + \mathbf{1}(L_3 = 0) + \mathbf{1}(L_4 < \underline{q}_4) + \mathbf{1}(L_5 < \underline{q}_5)] \geq 3$. In other words, two countries are regarded as neighbors iff at least three out of five criteria are satisfied. Hence, different network structures can be constructed by tuning the quantile levels. Table 3 reports summary statistics for $\mathcal{E}^+$ and $\mathcal{E}^-$, based on two quantile levels.

The repulsive links appear to be more numerous in comparison to the attractive links (with double density). This gap remains invariant when generating the network based on the 80% percentile and the 90%. However, in the latter case, the attractive network appears to be disconnected with 10

	Above 80-percentile		Above 90-percentile	
	attractive links	repulsive links	attractive links	repulsive links
singletons	0	0	10	0
edges	2437	5653	1280	2551
degr. max	143 (THA)	180 (STP)	112 (SGP)	174 (TUV)
degr. min	1 (TLS)	13 (CIV)	0 (CAF)	6 (MDV)

Table 3: Networks of the attractive and repulsive neighbors. The countries corresponding to the trigrams are: THA for Thailand, STP for Sao Tome and Principe, SGP for Singapore, TUV for Tuvalu, TLS for Timor-Leste, CIV for Cote d'Ivoire, CAF for the Central African Republic, and MDV for the Maldives.

singletons (namely, 10 countries having no attractive neighbors to be influenced by). See Figure 5 in Appendix D for the network plots corresponding to these generated networks.

The weight function W. To empirically calibrate edge weights W in the country-by-country geopolitical alignments, we use GDP data from the World Bank. The working assumption that we make is that the degree of influence that player j has on its neighbor i is increasing in the ratio between the GDP of j and the one of i. The idea is that big countries are more influential than smaller countries due to the size of their economies. Similarly, this influence must be decreasing with respect to their geographical distance (i.e., far apart countries are less exposed to reciprocal influence). Accordingly, we define W as follows:

$$w_{i,j} = \rho^{\mathbf{1}\left((i,j)\in\mathcal{E}_i^-\right)} \left(1 + \frac{\text{GDP}_j}{\text{GDP}_i}\right)^{d_{i,j}}, \text{ where } d_{i,j} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are contiguous,} \\ \dfrac{\Delta_{\max} - \Delta_{i,j}}{\Delta_{\max}} & \text{otherwise,} \end{cases} \tag{12}$$

with $\Delta_{i,j}$ being the geographical distance between the capital of country i and the one of country j; $\Delta_{\max}$ is the largest distance in the data set; $\rho \in (0, 1]$ is an exogenous parameter that mediates the impacts of attractive versus repulsive neighbors. Note that the effect of the geographical distance on $w_{i,j}$ is decreasing for non geographically adjacent countries. Conversely, when two countries share a common border, the distance is set at zero by convention. See Figure 6 in Appendix D for the distribution of weights among pairs of countries.

The prerequisite sets $\mathcal{P}^+$ and $\mathcal{P}^-$. The collection of features $\mathcal{F}$ represents national choices under the geopolitical influence. To generate the prerequisite sets $\mathcal{P}_t^+$ and $\mathcal{P}_t^-$, for each $t \in \mathcal{F}$, we use observable country-level indexes, and retrieve the top and bottom countries for each index. The prerequisite elements for the first categorical feature (LGBT legal protection) are obtained by including all countries in which homosexuality is sanctioned by the death penalty in $\mathcal{P}_1^-$. Consistently, we include all countries in which same-sex marriage existed for more than 15 years and where child adoption is allowed for homosexual couples in $\mathcal{P}_1^+$ (see Appendix D for an extensive list of these countries). The second categorical feature (CO2 reduction policy) is constructed based on the EPI 2022. The top and bottom 15 countries are included in $\mathcal{P}^+$ and $\mathcal{P}^-$, respectively. The third categorical feature is constructed based on the migrant integration policy index, which contains measurements regarding access to nationality, anti-discrimination laws, family reunion, labor market mobility, among others. Countries that are above average in at least three categories and below average in at most two categories are considered to be in $\mathcal{P}^+$, those below average in at least three categories and above average in at most two categories are considered $\mathcal{P}^-$. The last categorical feature is constructed based on their legal recognition of Israel and Palestine: the countries that recognize the sovereignity of Palestine are considered in $\mathcal{P}^+$; those who recognize Israel instead are put in $\mathcal{P}^-$. This is summarized in Table 4.

Categorical feature	P_t^+	P_t^-	
$t = 1$: LGBT legal protection	30	37	Criminalization of homosexuality, legal marrage/adoption
$t = 2$: CO2 reduction policy	10	10	10% percentile of the EPI
$t = 3$: Immigration policy	17	21	Customized immigration rights index (see description)
$t = 4$: Palestine vs Israel recognition	26	22	Palestine recognition versus Israel recognition

Table 4: Number of countries in the prerequisite set for each categorical feature.

Numerical tests. Building on the characterization of $\mathcal{I}$, $\mathcal{E}^+$, $\mathcal{E}^-$, W, $\mathcal{P}^+$, and $\mathcal{P}^-$, we apply the direct voting model and approach its social welfare solutions and Nash equilibrium solutions by problems (7) and (8), respectively. This aims to analyze how the combined influences of attractive and repulsive neighbors shape national political orientations based on the aforementioned four features.

The computational experiment consists of 24 instances obtained by the cross-combination of two network structures (one constructed by the 80% percentile and the other by the 90% percentile), three levels of ρ (namely, $\rho \in \{0.1, 0.5, 1\}$), and four specifications of the Θ_i sets. The latter is constructed as $\Theta_i = \{(1+\lfloor(1-\kappa)(d_i-1)/2\rfloor), \ldots, (d_i - \lceil(1-\kappa)(d_i-1)/2\rceil)\}$, where κ is an exogenous parameters in $[0, 1]$ whose role is to shrink the range of feasible threshold levels. In the conputational experiment, we set $\kappa \in \{0.7, 0.8, 0.9, 1\}$. The extensive results covering the 48 instances, i.e., 24 instances for the social welfare problem (7) plus 24 instances for the Nash equilibrium problem (8), are reported in Table D.3 (Appendix D).

We report hereafter a goodness-of-fit analysis, where we analyze the difference between the country choices predicted by solving problems (7) and (8) and the ones observed in the data set, for features $t = 1$, $t = 2$ and $t = 3$. As for the last item, no direct observation (beyond the one already contained in the prerequisite) can be adopted to establish our comparison.[12] As far as feature $t = 1$

[12]The information contained in $\mathcal{P}^+$ and $\mathcal{P}^-$ is correctly predicted by construction of the DVG-CNC parametrization.

is concerned, our comparison assesses the percentage of countries opposing LGBT legal protection (i.e., the 60 countries in which homosexuality is illegal) that are predicted to take choice $x(i, 1) = 0$ by our model. From these 60 countries we only take into account the 23 that are not fixed by the prerequisite when measuring the goodness of fit. As far as feature $t = 2$ (CO2 reduction policy) is concerned, our comparison focuses on assessing the percentage of *eco-friendly* and *eco-unfriendly* countries (i.e., the 72 countries that are beyond the top and bottom 80% percentiles of the EPI index). From these 72 countries we only take into account the 52 that are not fixed by the prerequisite when measuring the goodness of fit. Finally, for feature $t = 3$ (immigration policy), the observed political stance on immigration of the whole list of 133 countries is compared with our prediction. We have one instance for the social welfare solution and one for the Nash equilibrium solution maximizing the fit between observed and predicted choices. The social welfare instance corresponds to $\kappa = 1.0$, percentile = 80%, $\rho = 1$. For the Nash equilibrium, the selected instance corresponds to $\kappa = 1.0$, percentile = 80%, $\rho = 0.5$. Table 5 summarizes these results. Table D.4 (Appendix D) reports the goodness-of-fit comparison for the whole collection of 48 instances.

	$\kappa = 1.0$, Perc. = 80%, $\rho = 0.1$ (SW)	$\kappa = 0.8$, Perc = 80%, $\rho = 0.5$ (NE)
Objective function	1.00	1.00
Correct classifications $t = 1$	1.00 (pval = 1.19e-07)	0.95 (pval = 2.86e-06)
Correct classifications $t = 2$	0.83 (pval = 1.02e-06)	0.83 (pval = 1.02e-06)
Correct classifications $t = 3$	0.88 (pval = 1.90e-20)	0.83 (pval = 4.21e-15)

Table 5: Goodness-of-fit matching proportions. The cells contain the value of the objective functions and the proportion of matching points for two instances. Into parenthesis, we include the p-value of the result compared to binary random variables, which is associated with the probability of a similar (or better) result by a random guess.

Visibly, both instances result in equivalent total payoffs. However, the social welfare solution for the instance parametrized by $\kappa = 1.0$, percentile = 80%, and $\rho = 0.1$ gives rise to a higher empirical fit for feature $t = 1$ (LGBT legal protection) and feature $t = 3$ (Immigration policy). The corresponding p-values supports the sizable predictive power of the DVG-CNC solution when applied to forecast national choices about internationally relevant issues. Note that by exploring the overall collection of 48 instances from Tables D.3, we note that all analyzed Nash equilibrium solutions are social welfare solutions, although among the latter there are some that are not Nash equilibrium. Therefore, constraints (9a)-(9k) and (10a)-(10c) remove social welfare solutions.

6.2 Imputation of missing data by DVG-CNC

In this subsection, we use the DVG-CNC to impute missing data, building on its reduction to a KNN problem (Remark 2), where the directed weighted graph $\mathcal{G}_{+-} = \langle \mathcal{I}, \mathcal{E}^+, \mathcal{E}^-, W \rangle$ is constructed by selecting the K-nearest data points for each individual. Following the notation in Remark 2, let us define n data points $\mathcal{W} = \{(L_1, x_1), \ldots, (Z_n, x_n)\}$, where $Z_i \in \mathbb{R}^q$ is a q-dimensional vector encoding exogenous covariates for individual $i \in \mathcal{I}$, and x_i is the corresponding binary choice, and consider a distance metric $D : \mathbb{R}^q \times \mathbb{R}^q \to \mathbb{R}$.

To test the applicability of the DVG-CNC to missing data imputation, we intentionally remove data point from $\mathcal{W}$, based on the following scheme: (i) pick a random subset $\mathcal{I}_0$ (with $n_0 = |\mathcal{I}_0|$) from $\mathcal{I}$ and mark $\mathcal{W}_0 = \{(L_i, x_i) \ : \ i \in \mathcal{I}_0\}$ as missing points; (ii) for each $i \in \mathcal{I}_0$, mark as missing all $j \in \mathcal{I}$, such that $D(Z_i, Z_j) < r$. Hence, the constructed set of missing data is created in such a way that they are clustered into n_0 groups of neighbor points:

$$\mathcal{W}_M = \{(L_j, x_j) \in \mathcal{W} \ : \ D(Z_i, Z_j) < r, \ i \in \mathcal{I}_0, \ j \in \mathcal{I}\}.$$

We test the applicability of the DVG-CNC to predict $\mathcal{W}_M$, based on the following parametrization of $\mathcal{E}^+$ and $\mathcal{E}^-$: for all $i \in \mathcal{I}$, we have that $(i,j) \in \mathcal{E}_i^+$ iff $j \in \mathcal{J}_i^+$ and $(i,j) \in \mathcal{E}_i^-$ iff $j \in \mathcal{J}_i^-$, where $\mathcal{J}_i^+$ and $\mathcal{J}_i^-$ are the respective closest and farthest neighbors of i, with respect to the distance D:

$$\begin{cases} \mathcal{J}_i^+ & = \{j_1, \ldots, j_K \neq i : D(Z_i, Z_{j_1}) \leq \ldots \leq D(Z_i, Z_{j_K}); \ \forall j \in \mathcal{I} \setminus \mathcal{J}_i^+, \ D(Z_i, Z_{j_K}) \leq D(Z_i, Z_j)\} \\ \mathcal{J}_i^- & = \{j_1, \ldots, j_K \neq i : D(Z_i, Z_{j_1}) \geq \ldots \geq D(Z_i, Z_{j_K}); \ \forall j \in \mathcal{I} \setminus \mathcal{J}_i^-, \ D(Z_i, Z_{j_K}) \leq D(Z_i, Z_j)\}. \end{cases}$$

By construction $(i,i) \notin \mathcal{E}^+$ and $(i,i) \notin \mathcal{E}^-$. Further, to complete the characterization of the directed weighted graph $\mathcal{G}_{+-}$, we numerically compare two alternative parametrizations of the weight function W: uniform weights and distance-based weights. The uniform weight setting is such that $w_{i,j} = 1$ for all $(i,j) \in \mathcal{E}_i$. On the other hand, the distance-based weights for the attractive and repulsive neighbors are respectively defined as

$$w_{i,j} = \begin{cases} \lceil M - D(i,j) + 1 \rceil & \text{if } D(i,j) \leq M \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad w_{i,j} = \begin{cases} \lceil D(i,j) - M + 1 \rceil & \text{if } D(i,j) > M \\ 0 & \text{otherwise.} \end{cases}$$

where $M = Me[D(i,j) : j \in \mathcal{I}, j \neq i]$ is the median of $D(i,1), \ldots D(i,i-1), D(i,i+1) \ldots, D(i,n)$.

To provide a comprehensive performance benchmark for the DVG-CNC, we utilized publicly available datasets from the UCI machine learning repository (Bache and Lichman 2013), which are commonly used for reporting and comparing the performance of various classification methods (Bertsimas and Dunn 2017). The datasets we are using to illustrate the missing data imputation are the monk dataset, the balance scale data set, the tic-tac-toe dataset and the car evaluation dataset.[13] Table 6 reports the proportions of correctly imputed missing points.

The columns "in prerequisit" refers to parametrizations in which the neighborhood of each individual can only contain individuals whose choices are fixed by the prerequisit. Conversely, the

[13]The monk dataset contains 2 classes (for the response variable), 432 observations and 6 binary covariates. The balance scale dataset contains 3 classes (for the response variable), 625 observations and 4 covariates. The car evaluation dataset contains 4 classes (for the response variable), 1728 observations and 6 binary covariates. More details are provided in Siegler (1994).

Table 6: Proportions of correctly imputed missing points across various datasets using different imputation methods and different inclusion of points in the weight function W The results are compared with KNN imputation. Datasets include monk1, monk2, monk3, balance scale, car evaluation, and tic-tac-toe.

Data set	Distance-based weight				Uniform weight				KNN
	Nash equilibrium (problem (8))		social welfare (problem (7))		Nash equilibrium (problem (8))		social welfare (problem (7))		
	In test	In prerequisite	In test	In prerequisite	In test	In prerequisite	In test	In prerequisite	
monk1	0.87	0.90	0.87	0.96	0.82	0.90	0.83	0.90	0.91
	K=9, $\theta = 35$	K=11, $\theta = 22$	K=9, $\theta = 35$	K=11, $\theta = 22$	K=13, $\theta = 13$	K=11, $\theta = 6$	K=9, $\theta = 10$	K=7, $\theta = 4$	K=13
monk2	0.69	0.74	0.69	0.70	0.68	0.71	0.70	0.71	0.70
	K=5, $\theta = 35$	K=5, $\theta = 11$	K=5, $\theta = 35$	K=7, $\theta = 19$	K=3, $\theta = 6$	K=5, $\theta = 3$	K=7, $\theta = 8$	K=7, $\theta = 4$	K=11
monk3	0.91	0.84	0.90	0.93	0.84	0.81	0.92	0.95	0.89
	K=9, $\theta = 22$	K=3, $\theta = 10$	K=9, $\theta = 22$	K=17, $\theta = 35$	K=9, $\theta = 5$	K=3, $\theta = 2$	K=9, $\theta = 5$	K=21, $\theta = 11$	K=9
balance scale	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	1
	K=7, $\theta = 3$	K=7, $\theta = 6$	K=7, $\theta = 3$	K=7, $\theta = 6$	K=7, $\theta = 13$	K=7, $\theta = 22$	K=7, $\theta = 13$	K=7, $\theta = 22$	K=7
car evaluation	0.75	0.75	0.75	0.75	0.76	0.74	0.75	0.75	0.99
	K=7, $\theta = 2$	K=7, $\theta = 4$	K=7, $\theta = 2$	K=7, $\theta = 4$	K=7, $\theta = 2$	K=7, $\theta = 30$	K=7, $\theta = 8$	K=7, $\theta = 30$	K=3
tic-tac-toe	0.99	0.95	0.99	0.95	0.99	0.95	0.99	0.95	0.99
	K=7, $\theta = 2$	K=7, $\theta = 2$	K=7, $\theta = 2$	K=7, $\theta = 3$	K=7, $\theta = 8$	K=7, $\theta = 22$	K=7, $\theta = 9$	K=7, $\theta = 15$	K=3

columns “in test” refers to parametrizations in which the neighborhoods can also contain individuals whose choices are not fixed. The latter case correspond to the possibility of using the imputed missing data (endogenous choices) to impute other missing data.

Mostly, the social welfare criterion (problem (7)) results in better missing data imputation, in comparison to the Nash equilibrium solution (problem (8)). This is related to the fact that no strategic choices are to be invoked when trying to explain patterns of similarities between neighbor points in a missing data imputation context. The only two cases in which the Nash equilibrium solution provides better prediction is in the monk2 data set, with 74% of correctly imputed missing. Further, “in prerequisit” is usually better than “in test”, except for very few cases, such as the previously mentioned case of the monk2 dataset. When it comes to the weight function, there is no clear pattern: distance-based weight is as good as uniform weight.

Overall, the DVG-CNC seems to be able to impute missing data in a way that is either better or as good as KNN, in the majority of the analysed instances.

6.3 Student-by-student influences in cannabis consumption

This section presents a real world application of the IVG-CNC to the consumption of three addictive substances (that we index as $j = 1$, for tobacco, $j = 2$ for alcohol, and $j = 3$ for cannabis) in a cohort of $n = 160$ students over their second, third and fourth year at a secondary school in Glasgow (Scotland). The goal is to discover attractive and repulsive peer influences that are capable of explaining the observed consumption behaviour. For this analysis we use a publicly available dataset, previously used in the influence propagation literature (Tchouya et al. 2023).

The dataset contains the consumption levels of addictive substances: $\tilde{m}(1) = 5$ for alcohol, $\tilde{m}(2) = 3$ for tobacco, $\tilde{m}(3) = 4$ for cannabis. Contextually, alcohol use is coded as 1 (non), 2 (once or twice a year), 3 (once a month), 4 (once a week) and 5 (more than once a week); tobacco use is coded as 1 (non), 2 (occasional) and 3 (regular, i.e. more than once per week); cannabis use is coded as 1 (non), 2 (tried once), 3 (occasional) and 4 (regular). For the purpose of constructing the features for our model, the substance consumption is transformed into cumulative representation, which are encoded by the following binary choices:

$$x(i, 3L(t-1)+(j-1)(\tilde{m}(j)-1)+q) = \begin{cases} 1 & \text{if consumption of substance } j \text{ is above } q+1 \text{ in period } t, \\ 0 & \text{otherwise,} \end{cases}$$

where $L = \tilde{m}(1) + \tilde{m}(2) + \tilde{m}(3) - 3 = 9$. For instance, $x(i, 1) = 1$ if player i has consumed alcohol at least twice a year in the first period, $x(i, 5) = 1$ if player i has consumed tobacco at least twice a year in the second period. Overall, after this enumeration of periods, substances, and consumption levels, we obtain $F = 27$ features among $n = 160$ individuals. Out of these 160 students, 129 of them are recorded in all 3 years of observation; the other are removed from the data set. Finally, the dataset also contains a network of friendship between the students during the three observed periods, that we use to define the network constraints $(\overline{\psi}, \underline{\psi}) \in \Psi$ in problem (11b)-(11g).

Tchouya et al. (2023) applied an influential-imitator approach to this data set, assuming imitation patterns, where the choices of popular children (influential) represent the leading effects to determine the habits of others. In our application, we compare a collection of different modeling specifications of the IVG-CNC, by appending network constraints to problem (11), as described below:

- fixed attractive neighbors (FAN): $\overline{\Phi}_{i,j} = \mathcal{E}_{i,j}$, for all $i, j, \in \mathcal{I}$,

- bounded attractive neighbors (BAN): $\overline{\Phi}_{i,j} \leq \mathcal{E}_{i,j}$, for all $i, j, \in \mathcal{I}$,

- attractive larger than repulsive (ALR): $\sum_{j\in\mathcal{I}} \overline{\Phi}_{i,j} \geq \sum_{j\in\mathcal{I}} \underline{\Phi}_{i,j}$, for all $i \in \mathcal{I}$,

- fixed degree (FD): $\sum_{j\in\mathcal{I}} \overline{\Phi}_{i,j} = \sum_{j\in\mathcal{I}} \mathcal{E}_{i,j}$, for all $i \in \mathcal{I}$.

Table 7 reports the optimal value of (11), as well as its difference with respect to the observed links, for different specification of these network constraints.

	attractive neighbors	repulsive neighbors	degree-type	Optimal	$\frac{\sum \lvert\overline{\Phi}_{i,j} - \mathcal{E}_{i,j}\rvert}{\sum \overline{\Phi}_{i,j}}$
(cfg 1)	free	free	free	3227	6.7%
(cfg 2)	free	free	FD	2363	94%
(cfg 3)	FAN	free	FD	2344	0%
(cfg 4)	BAN	free	FD	2351	52%
(cfg 5)	BAN	free	free	2880	3%
(cfg 6)	BAN	ALR	FD	2351	28%
(cfg 7)	BAN	ALR	free	2498	14%

Table 7: The configurations with different neighbor strategies and their degree type. The optimal value and the error rate relative to friendship and relative to the network.

The highest within-sample fit (in terms of predicted choices) is obtained when attractive neighbors are fixed (with objective function value of 2344). However, letting all connections free in problem (11b)-(11g) allows obtaining a network that is similar to the observed one (6.7% of differences) and a much higher amount of correctly classified choices (3227). In other words, the network topology that maximizes choices' prediction under the voting influence assumption of the IVG-CNC model closely resembles the observed friendship network in the data set, validating the underplaying empirical story of pairwise influences in categorical choices.

7 Conclusion

In this paper, we introduced a novel game-theoretical framework designed to study the interplay of pairwise (network) influences in multivariate choices, by incorporating nuanced dependencies between individuals connected in a signed weighted graph. The model considers both attractive and repulsive influences, enabling a flexible representation of the diverse nature the statistical dependence. The payoff functions capture the social behavior of individuals to match (mismatch for repulsive neighbors) the similarity between their choices and the ones of their neighbors.

We establish a duality between the DVG-CNC (data classification) and the IVG-CNC (network discovery), grounded in the endogeneity of choices versus network information. We showed that, by the inclusion of mandatory ($\mathcal{P}_i^1$) and forbidden ($\mathcal{P}_i^0$) choices, the DVG-CNC model results in a natural generalization of the KNN method for missing data imputation. In this case, each individual observation is associated to a partitioning between observed choices (corresponding to $\mathcal{P}_i^1$ and $\mathcal{P}_i^0$) and missing choices (corresponding to $\mathcal{F}_i$). This generalization not only provides a theoretical bridge between game theory and statistical classification but also offers insights into the optimal choice for the number of neighbors K in the KNN approach.

Both games are reformulated as mixed-integer linear programs and applied to three different problems: two for the DVG-CNC (with a fixed network and endogenous choices) and one for the IVG-CNC (with an endogenous network and fixed choices). In the first application, we analyze geopolitical alliances among 184 countries, predicting their stances on key issues. The second application demonstrates the proposed approach's superiority in imputing missing data compared to the KNN method. In the third application, the inverse voting game unveils network influences among 164 secondary school students based on substance consumption.

This flexibility of our game-theoretical framework to tackle diverse empirical contexts opens avenues for further research. An important modeling aftermath of the DVG-CNC is its applicability to the analysis of power indexes, measuring the ability of players to influence voting outcomes (Leech 2002, 2003). In fact, when players are linked on a network carrying pairwise influence channels, the decisional power of a player is not solely driven by the weight an organization can acknowledge to it (for instance, the International Monetary Fund and World Bank give each member country a number of votes based on its financial contribution), but also on the influence that its decision has on others. From this viewpoint, our equilibrium solution can be used to design power indexes in voting games. Further, conditions for the uniqueness of a Nash equilibrium remains to be investigated. Also, specialized combinatorial and decomposition algorithms can be designed to tackle large instances of problem (8) (for the direct form) and problem (11) (for the inverse form).

References

Bache, K. and Lichman, M.: 2013, Uci machine learning repository.

Barigozzi, M. and Brownlees, C.: 2018, NETS: Network estimation for time series, *Journal of Applied Econometrics* **34**, 347–364.

Ben-Porath, E.: 1990, The complexity of computing a best response automaton in repeated games with mixed strategies, *Games and Economic Behavior* **2**(1), 1–12.

Bertsimas, D. and Dunn, J.: 2017, Optimal classification trees, *Machine Learning* **106**(7), 1039–1082.

Bhattacharya, P. and Mack, Y.: 1987, Weak convergence of k-nn density and regression estimators with varying k and applications, *The Annals of Statistics* pp. 976–994.

Borovkova, S., Lopuhaä, H. P. and Ruchjana, B. N.: 2008, Consistency and asymptotic normality of least squares estimators in generalized star models, *Statistica Neerlandica* **62**(4), 482–508.

Carvalho, M., Lodi, A. and Pedroso, J. P.: 2022, Computing equilibria for integer programming games, *European Journal of Operational Research* **303**(3), 1057–1070.

Carvalho, M., Lodi, A., Pedroso, J. P. and Viana, A.: 2017, Nash equilibria in the two-player kidney exchange game, *Mathematical Programming* **161**, 389–417.

Chu, F. and Halpern, J.: 2001, On the np-completeness of finding an optimal strategy in games with common payoffs, *International Journal of Game Theory* **30**, 99–106.

Cliff, A. and Ord, J.: 1975, Model building and the analysis of spatial pattern in human geography, *Journal of the Royal Statistical Society: Series B (Methodological)* **37**(3), 297–328.

Cohen-Cole, E., Liu, X. and Zenou, Y.: 2018, Multivariate choices and identification of social interactions, *Journal of Applied Econometrics* **33**(2), 165–178.

Dragotto, G. and Scatamacchia, R.: 2023, The zero regrets algorithm: Optimizing over pure nash equilibria via integer programming, *INFORMS Journal on Computing* .

Gambella, C., Ghaddar, B. and Naoum-Sawaya, J.: 2021, Optimization problems for machine learning: A survey, *European Journal of Operational Research* **290**(3), 807–828.

Gilboa, I. and Zemel, E.: 1989, Nash and correlated equilibria: Some complexity considerations, *Games and Economic Behavior* **1**(1), 80–93.

Glover, F. and Woolsey, E.: 1974, Converting the 0-1 polynomial programming problem to a 0-1 linear program, *Operations Research* **22**(1), 180–182.

Goldenberg, J., Libai, B. and Muller, E.: 2001, Talk of the network: A complex systems look at the underlying process of word-of-mouth, *Marketing Letters* **12**(3), 211–223.

Granovetter, M.: 1978, Threshold models of collective behavior, *American Journal of Sociology* pp. 1420–1443.

Granovsky, B. L. and Madras, N.: 1995, The noisy voter model, *Stochastic Processes and their applications* **55**(1), 23–43.

Kempe, D., Kleinberg, J. and Éva Tardos: 2015, Maximizing the spread of influence through a social network, *Theory of Computing* **11**(4), 105–147.

Keuschnigg, M. and Ganser, C.: 2017, Crowd wisdom relies on agents' ability in small groups with a voting aggregation rule, *Management Science* **63**(3), 818–828.

Koller, D. and Megiddo, N.: 1992, The complexity of two-person zero-sum games in extensive form, *Games and Economic Behavior* **4**(4), 528–552.

Koller, D., Megiddo, N. and Von Stengel, B.: 1996, Efficient computation of equilibria for extensive two-person games, *Games and Economic Behavior* **14**(2), 247–259.

Leech, D.: 2002, Voting power in the governance of the international monetary fund, *Annals of Operations Research* **109**, 375–397.

Leech, D.: 2003, Computing power indices for large voting games, *Management Science* **49**(6), 831–837.

Lin, Y., Nelson, B. L. and Pei, L.: 2019, Virtual statistics in simulation via k nearest neighbors, *INFORMS Journal on Computing* **31**(3), 576–592.

Lutu, P. E. and Engelbrecht, A. P.: 2013, Base model combination algorithm for resolving tied predictions for k-nearest neighbor ova ensemble models, *INFORMS Journal on Computing* **25**(3), 517–526.

Mack, Y.-P.: 1981, Local properties of k-nn regression estimates, *SIAM Journal on Algebraic Discrete Methods* **2**(3), 311–323.

McPherson, M., Smith-Lovin, L. and Cook, J. M.: 2001, Birds of a feather: Homophily in social networks, *Annual review of sociology* pp. 415–444.

Nasini, S. and Martínez-de Albéniz, V.: 2020, Pairwise influences in dynamic choice: network-based model and application, *Journal of Applied Statistics* pp. 1–34.

Nasini, S., Martínez-de Albéniz, V. and Dehdarirad, T.: 2017, Conditionally exponential random models for individual properties and network structures: Method and application, *Social Networks* **48**, 202–212.

Patacchini, E. and Zenou, Y.: 2012, Juvenile delinquency and conformism, *The Journal of Law, Economics, & Organization* **28**(1), 1–31.

Porter, R., Nudelman, E. and Shoham, Y.: 2008, Simple search methods for finding a nash equilibrium, *Games and Economic Behavior* **63**(2), 642–662.

Sherali, H. D. and Adams, W. P.: 1998, Reformulation-linearization techniques for discrete optimization problems, *Handbook of Combinatorial Optimization: Volume1–3* pp. 479–532.

Siegler, R.: 1994, Balance Scale, UCI Machine Learning Repository.

Stoffer, D. S.: 1986, Estimation and identification of space-time armax models in the presence of missing data, *Journal of the American Statistical Association* **81**(395), 762–772.

Tchouya, R. T., Nasini, S. and Dabo-Niang, S.: 2023, An estimation approach for the influential-imitator diffusion, *Computers & Operations Research* p. 106315.

Wu, H.-H. and Küçükyavuz, S.: 2018, A two-stage stochastic programming approach for influence maximization in social networks, *Computational Optimization and Applications* **69**, 563–595.

Appendices: Supplementary Material

A : Mathematical proofs

Proposition 1

Proof. By definition of $u_{i,t}((x(i,t),\theta_i),\boldsymbol{x}_{-i})$ in (3), we have:

$$\begin{aligned}
u_{i,t}((x(i,t),\theta_i),\boldsymbol{x}_{-i}) &= |\mathbf{1}\,(\mathrm{WS}_{i,t}(\boldsymbol{x}_{-i}) \geq \theta_i) - x(i,t)| \\
&= |\mathbf{1}\left(\mathrm{WS}_{i,t}(\mathsf{C}_t(\boldsymbol{x}_{-i})) < d_i - \theta_i\right) - x(i,t)| \\
&= |\mathbf{1}\left(\mathrm{WS}_{i,t}(\mathsf{C}_t(\boldsymbol{x}_{-i})) \geq d_i - \theta_i\right) - (1 - x(i,t))| \\
&= u_{i,t}((1 - x(i,t), d_i - \theta_i), \mathsf{C}_t(\boldsymbol{x}_{-i})).
\end{aligned}$$

□

Lemma 1

Proof. It is sufficient to note that if (i,t) is one-step supportable by $\langle p, \sigma_{i,t}\rangle$, then (by Denifition 4), i has a neighbor $j \in \mathcal{E}_i$ in the prerequisite p (namely, $(j,t) \in \mathrm{dom}(p)$), such that

$$\begin{cases}
\text{if } p(j,t) = 1 \text{ then } t \in \boldsymbol{x}_j \text{ and if } p(j,t) = 0 \text{ then } t \notin \boldsymbol{x}_j & \text{in case } \sigma_{i,t} = 1, \\
\text{if } p(j,t) = 1 \text{ then } t \notin \boldsymbol{x}_j \text{ and if } p(j,t) = 0 \text{ then } t \in \boldsymbol{x}_j. & \text{in case } \sigma_{i,t} = 0.
\end{cases}$$

Let x be an arbitrary choice profile satisfying p. For the same j, we have that

$$\begin{cases}
\text{if } x(j,t) = 1 \text{ then } t \in \boldsymbol{x}_j \text{ and if } x(j,t) = 0 \text{ then } t \notin \boldsymbol{x}_j & \text{in case } \sigma_{i,t} = 1, \\
\text{if } x(j,t) = 1 \text{ then } t \notin \boldsymbol{x}_j \text{ and if } x(j,t) = 0 \text{ then } t \in \boldsymbol{x}_j. & \text{in case } \sigma_{i,t} = 0.
\end{cases}$$

This implies that x supports $(i,t) \in \mathcal{D}$, with the sign $\sigma_{i,t}$. Hence, by Definition 5, $\mathcal{D}$ is uniformly supported by $\langle p, \sigma\rangle$. □

Theorem 1

Proof. We proceed by induction. For $k = 1$, we have $\tilde{i} = (i_0, i_1)$ and $p(i_1,t) = \phi_1(\tilde{i})$. We have four possibilities

$$\begin{aligned}
&\text{case 1:} \quad \phi_{i_1}(\tilde{i}) = 1,\ p(i_1,t) = 1 \\
&\text{case 2:} \quad \phi_{i_1}(\tilde{i}) = 1,\ p(i_1,t) = 0 \\
&\text{case 3:} \quad \phi_{i_1}(\tilde{i}) = 0,\ p(i_1,t) = 1 \\
&\text{case 4:} \quad \phi_{i_1}(\tilde{i}) = 0,\ p(i_1,t) = 0
\end{aligned}$$

Cases 1 and 2 entail that (i,t) is positively supportable by p, whereas cases 3 and 4 entail that (i,t) is negatively supportable by p.

Suppose the proposition is true for k, meaning that the existence of a path $\tilde{i}$ of length $k+1$, connecting $i_0 = i$ and i_k, with $(i_k, t) \in \text{dom}(p)$ implies that (i, t) is supportable by $\langle p, s \rangle$.[14] We show that if there exist a path $\tilde{i}'$ of length $k+2$ connecting $i = i_0$ and i'_{k+1}, with $(i'_{k+1}, t) \in \text{dom}(p)$, then (i, t) is supportable by $\langle p, s \rangle$, where

$$s = \begin{cases} 1 & \text{if } 1 - p(i_{k+1}, t) + \sum_{l=1}^{k+1}(1 - \phi_l(\tilde{i})) \text{ is even.} \\ 0 & \text{otherwise.} \end{cases}$$

To do that, let $\tilde{i}'' = (i''_0, i''_1 \ldots, i''_k, \; i''_{k+1})$ be such a path and $\phi(\tilde{i}'')$ be corresponding indicator of attractiveness. By focusing our attention to the subsequence $(i''_1, \cdots, i''_{k+1})$, we note that since $(i''_{k+1}, t) \in \text{dom}(p)$, the existence of any path of length $k+1$ from any arbitrary node i''_{k+1} entails (by the induction hypothesis) the supportability of that node by $\langle p, s \rangle$. Therefore, by the induction hypothesis, (i''_1, t) is supportable by $\langle p, s' \rangle$ where

$$s' = \begin{cases} 1 & \text{if } 1 - p(i''_{k+1}, t) + \sum_{l=2}^{k+1}(1 - \phi_l(\tilde{i}'')) \text{ is even.} \\ 0 & \text{otherwise.} \end{cases}$$

Let x satisfy p, such that $x(i''_1, t) = 1$, for $s' = 1$ (resp. $x(i''_1, t) = 0$, for $s' = 0$). We have four possibilities

$$\begin{aligned} &\text{case 1:} \quad \phi_{i''_1}(\tilde{i}) = 1, \; x(i''_1, t) = 1 \\ &\text{case 2:} \quad \phi_{i''_1}(\tilde{i}) = 1, \; x(i''_1, t) = 0 \\ &\text{case 3:} \quad \phi_{i''_1}(\tilde{i}) = 0, \; x(i''_1, t) = 1 \\ &\text{case 4:} \quad \phi_{i''_1}(\tilde{i}) = 0, \; x(i''_1, t) = 0 \end{aligned}$$

Cases 1 and 4 entail that (i, t) is positively supported by x, whereas cases 2 and 3 entail that (i, t) is negatively supported by x. Case 1 and case 4 both imply that

$$1 - p(i''_{k+1}, t) + \sum_{l=2}^{k+1}(1 - \phi_l(\tilde{i}'')) \text{ is even}$$

and case 2 and case 4 both imply that

$$1 - p(i''_{k+1}, t) + \sum_{l=2}^{k+1}(1 - \phi_l(\tilde{i}'')) \text{ is odd}$$

Which proves the proposition for the case $k+2$.

□

Theorem 2

[14] Note that the induction hypothesis is the truthness of the implication.

Proof. To prove Theorem 2, we prove the following stronger statement (called Stat(K)), whose truthiness implies theorem 2 (by Lemma 1):

Stat(K): If (i) (ii) (iii) are satisfied, then for all $(i,t) \in \mathcal{D}$, (i,t) is one-step supportable by $\langle p_K, \sigma \rangle$.

We proceed with mathematical induction. For $K = 1$, it is true by Definition 4. In fact, since $P_1 = (p_0, p_1)$ and since p_1 is a one-step extension of p_0, the notion of one-step extension in Definition 4 establishes that for all $(i,t) \in \text{dom}(p_1) \setminus \text{dom}(p_0)$, (i,t) is one-step supportable by $\langle p_0, \sigma_{i,t} \rangle$. In particular, this is true for all $(i,t) \in \mathcal{D}$.

Assume Stat(K) to be true (for an arbitrary K), and let's consider Stat($K+1$). Assume there exists a sequence $P_{K+1} = (p_0, \ldots, p_K, p_{K+1})$ satisfying:

(i) For all $k = 2, \ldots, K+1$, p_k is a one-step extension of p_{k-1}.

(ii) $\mathcal{D} \subseteq \text{dom}(p_{K+1})$.

(iii) For all $(i,t) \in \mathcal{D}$, we have $p_{K+1}(i,t) = \sigma_{i,t}$.

Let's consider a restricted collection of IF pairs $\tilde{\mathcal{D}} := \mathcal{D} \cap \text{dom}(p_k)$ and the corresponding restricted sign function $\tilde{\sigma} = \sigma|_{\tilde{\mathcal{D}}}$. We have:

(i. a) For all $k = 2, \ldots, K$, p_k is a one-step extension of p_{k-1}.

(ii. a) $\tilde{\mathcal{D}} \subseteq \text{dom}(p_K)$.

(iii. a) For all $(i,t) \in \tilde{\mathcal{D}}$, we have $p_K(i,t) = \tilde{\sigma}_{i,t}$.

Note that (i) implies $(i.a)$. Next, p_{K+1} being a one-step extension of p_K together with (ii) imply $(ii.a)$. Concerning (iii), since p_{K+1} is a one-step extension of p_K, for all $(i,t) \in \tilde{\mathcal{D}}$, $p_{K+1}(i,t) = p_K(i,t)$, thus $\tilde{\sigma}_{i,t} = p_{K+1}(i,t) = p_K(i,t)$.

By the induction hypothesis Stat(K), we conclude that for all $(i,t) \in \tilde{\mathcal{D}}$, (i,t) is one-step supportable by $\langle p_K, \tilde{\sigma}_{i,t} \rangle$. Consequently, (i,t) is one-step supportable by $\langle p_{K+1}, \tilde{\sigma}_{i,t} \rangle$.

Since p_{K+1} is a one-step extension of p_K, the incremental part $\partial\mathcal{D} := \mathcal{D} \backslash \tilde{\mathcal{D}}$ is one-step supportable by $\langle p_K, \sigma \rangle$. Consequently, for all $(i,t) \in \partial\mathcal{D}$, (i,t) is one-step supportable by $\langle p_{K+1}, \tilde{\sigma}_{i,t} \rangle$, which finishes the proof.

□

Corollary 1

Proof. Consider the family of paths and orientations $(\tilde{i}^1, \ldots, \tilde{i}^D)$ and $(\tilde{\Sigma}^1, \ldots, \tilde{\Sigma}^D)$, satisfying (i) -(iii-b). Let us define an ordering g of $\{(d,q) : d = 1, \ldots, D,\ q = 1, \ldots, k_d\}$:

$$g(d,q) = (d-1)k_d + (k_d - q + 1),$$

so as to follow a counter lexicographic ordering

$$((1,k_1),(1,k_1-1),\dots,(1,1),(2,k_2),(2,k_2-1),\dots,(D,k_D),\dots,(D,1)).$$

Let's construct a sequence of prerequisite, which we is attached to the collection of extending paths :

$$P_K = \Big(p_0, \underbrace{p_1, p_2, \dots, p_{k_1}}_{\tilde{i}^1}, \underbrace{p_{k_1+1}, p_{k_1+2}, \dots, p_{k1+k2}}_{\tilde{i}^2}, \dots, \underbrace{p_{g(D,k_D)}, p_{g(D,k_D-1)}, \dots, p_{g(D,k_1)}}_{\tilde{i}^D}\Big), \tag{A.1}$$

where the prerequisite $p_{g(d,q)}$ is associated to the q-th element of the d-th path. Specifically, we define $p_{g(d,q)}$ recursively from $p_{g(d,q)-1}$, in the following way:

$$p_{g(d,q)} \; : \; \mathrm{dom}(p_{g(d,q)-1}) \cup \{(i^d_q, t^d)\} \longrightarrow \{0,1\}, \tag{A.2}$$

and

$$p_{g(d,q)}(i',t') = \begin{cases} \tilde{\Sigma}^d_q & \text{for } i' = i^d_q \text{ and } t' = t^d, \\ p_{g(d,q)-1} & \text{otherwise.} \end{cases} \tag{A.3}$$

The rest of this proof consists of two steps:

1. We first show that the family of paths and orientations $(\tilde{i}^1,\dots,\tilde{i}^D)$ and $(\tilde{\Sigma}^1,\dots,\tilde{\Sigma}^D)$ is uniquely associated to a sequence of extending paths (A.1). To do that, we note that $p_{g(d,q)}$ extends $p_{g(d,q)-1}$. In fact, by (A.2) and (A.3), if $(i^d_q,t^d) \notin \mathrm{dom}(p_{g(d,q)-1})$ then $\mathrm{dom}(p_{g(d,q)}) \setminus \mathrm{dom}(p_{g(d,q)-1}) = \{(i^d_q,t^d)\}$ (i.e., $p_{g(d,q)}$ extends the domain of $p_{g(d,q)-1}$ by a single element, without changing anything already defined). Conversely, if $(i^d_q,t^d) \in \mathrm{dom}(p_{g(d,q)-1})$, by (A.3), $p_{g(d,q)}(i',t') = p_{g(d,q)-1}(i',t')$, for all $(i',t') \in \mathrm{dom}(p_{g(d,q)})$.

By condition (ii) in Corollary 1, if there is $d' \le d$ and $q' \le k_{d'}$ for which $i^d_q = i^{d'}_{q'}$ and $t^d = t^{d'}$ then $\Sigma^d_q = \Sigma^{d'}_{q'}$. Let us define

$$(\delta,\nu) = \min\{d' \; : \; i^d_q = i^{d'}_{q'}, \; t^d = t^{d'}\},$$

so that $g(\delta,\nu) < g(d,q)$. By (A.3), we have

$$\Sigma^\delta_\nu = p_{g(\delta,\nu)}(i^\delta_\nu, t^\delta) = p_{g(\delta,\nu)}(i^d_q, t^d).$$

Since $p_{g(d,q)-1}$ is an extension of $p_{g(\delta,\nu)}$, we know that $p_{g(d,q)-1}(i^d_q,t^d) = \Sigma^\delta_\nu$, which also equals Σ^d_q. Hence, if $(i^d_q,t^d) \in \mathrm{dom}(p_{g(d,q)-1})$, then both $\mathrm{dom}(p_{g(d,q)-1}) = \mathrm{dom}(p_{g(d,q)})$ and $p_{g(d,q)-1} = p_{g(d,q)}$, elementwise.

2. In the second step, we show that the sequence of one-step extending paths (A.1) satisfies conditions (i)-(iii) of Theorem 2.

- **Condition (i) of Theorem 2.** Let's show that for each $d = 1, \dots, D$ and $q = 1, \dots, k_d$, $p_{g(d,q)}$ is a one-step extension of $p_{g(d,q)-1}$. If $q = k_d$, we know that

$$(i^d_{k_d}, t^d) \in \text{dom}(p),$$

so that $(i^d_q, t^d) \in \text{dom}(p) \subseteq \text{dom}(p_{g(d,q)-1})$. Therefore, $(i^d_{k_d}, t^d)$ is supportable by $p_{g(d,q)-1}$, so that by Definition 4, $p_{g(d,q)}$ is a one-step-extension of $p_{g(d,q)-1}$.

Next, if $q < k_d$, we can show that (i^d_{q+1}, t^d) is one-step supportable by $\langle p_{g(i^d_{q+1},t^d)-1}, \Sigma^d_q \rangle$. Let x be a choice profile satisfying $p_{g(d,q+1)}$ and such that $x(i^d_q, t^d) = \Sigma^d_q$. Using conditions (iii-a) and (iii-b) from Corollary A.2, if i^d_{q+1} is an attractive neighbor of i^d_q, then $p^d_{q+1}(i^d_{q+1}, t^d) = \Sigma^d_{q+1} = \Sigma^d_q = x(i^d_q, t^d)$. If i^d_{q+1} is a repulsive neighbor of i^d_q, then $p^d_{q+1}(i^d_{q+1}, t^d) = \Sigma^d_{q+1} = 1 - \Sigma^d_q = 1 - x(i^d_q, t^d)$. Hence, (i^d_{q+1}, t^d) is one-step supportable by $\langle p_{g(i^d_{q+1},t^d)-1}, \Sigma^d_q \rangle$.

- **Condition (ii) of Theorem 2.** By construction, the first node of each path i^d_1 is such that $(i^d_1, t^d) \in \mathcal{D}$. Since $(i^d_q, t^d) \in \text{dom}(p_{g(d,q)})$ and P_K is a sequence of extending prerequisites, we have $\text{dom}(p) \subseteq \text{dom}(p_{g(1,k_1)}) \subseteq \cdots \text{dom}(p_{g(D,1)})$. Combining above points, we have that for all

$$(i^d_1, t^d) \in \text{dom}(p_{g(D,1)}) = \text{dom}(p_K), \quad \text{for all } d = 1, \dots, D.$$

- **Condition (iii) of Theorem 2.** For each $d = 1, \dots, D$, $q = 1, \dots, k_d$, $(i^d_1, t^d) \in \mathcal{D}$ and $p_K(i^d_1, t^d) = p^d_1(i^d_1, t^d) = \Sigma^d_1 = \sigma(i^d, t^d)$ by definition of σ.

□

Lemma 2

Proof. Since (i, t) is supported positively by $\boldsymbol{x}_{-i}$, we know that $\text{WS}_{i,t}(\boldsymbol{x}_{-i}) \geq \min_{j \in \mathcal{E}_i} w_{i,j}$. By choosing $\theta_i = \min_{j \in \mathcal{E}_i} w_{i,j}$, we have $\text{WS}_{i,t}(\boldsymbol{x}_{-i}) \geq \theta_i$. Thus $v_t(\boldsymbol{x}_{-i}, \theta_i) = 1$. Similar argument can be made for the negative supportedness.

□

Theorem 3

Proof. As a consequence of Lemma 2, we have that

(1.) if for all $t \in \mathcal{P}^1_i$, (i, t) is positively supported by $\boldsymbol{x}_{-i}$, then $\max_{\theta_i} u^1_i(\theta_i, \boldsymbol{x}_{-i}) = |\mathcal{P}^1_i|$;

(2.) if for all $t \in \mathcal{P}^0_i$, (i, t) is negatively supported by $\boldsymbol{x}_{-i}$, then $\max_{\theta_i} u^0_i(\theta_i, \boldsymbol{x}_{-i}) = |\mathcal{P}^0_i|$.

In fact, by letting $\theta_i = \min_{j\in\mathcal{E}_i} w_{i,j}$, we have $v_t(\boldsymbol{x}_{-i}, \theta_i) = 1$, for all $t \in \mathcal{P}_i^1$ (Lemma 2). Since (i,t) is positively supported by $\boldsymbol{x}_{-i}$, then x satisfies p_0. Therefore, for all $t \in \mathcal{P}_i^1$, $x(i,t) = 1$. By construction of u_i^1, we have

$$\forall\ i \in \mathcal{I},\ \exists\ \boldsymbol{x}_{-i},\ \text{ such that } u_i^1(\min_{j\in\mathcal{E}_i} w_{i,j}, \boldsymbol{x}_{-i}) = \sum_{t\in\mathcal{P}_i^1} \mathbf{1}\left(x(i,t) = v_t(\boldsymbol{x}_{-i}, \theta_i)\right) = |\mathcal{P}_i^1|. \tag{A.4}$$

This proves (1.). An analogous argument is valid for (2.).

Note that (A.4) does not entail $\boldsymbol{x}_{-i} \cup \boldsymbol{x}_i = \boldsymbol{x}_{-j} \cup \boldsymbol{x}_j$, for all $i,j \in \mathcal{I}$. To obtain (4) (which entails $\boldsymbol{x}_{-i} \cup \boldsymbol{x}_i = \boldsymbol{x}_{-j} \cup \boldsymbol{x}_j$, for all $i,j \in \mathcal{I}$), we requires the uniform supportability of $\mathcal{D}^+$ (Definition 5).

Let $\boldsymbol{x}$ be a choice profile satisfying p_0 such that for all $(i,t) \in \mathcal{D}^+$, (i,t) is supported by $\langle \boldsymbol{x}_{-i}, 1\rangle$, where $\boldsymbol{x} = \boldsymbol{x}_{-i} \cup (\mathcal{P}_i^1 \cup \tilde{x}_i)$, for all $i \in \mathcal{I}$. Using (A.4), we know that and setting $\theta_i = 1$, for all $i \in \mathcal{I}$, we have $v_t(\boldsymbol{x}_{-i}, \theta_i) = 1$ for every $(i,t) \in \mathcal{D}^+$, or using the definition of $\mathcal{D}^+$, for all $t \in \mathcal{P}_i^1$ and for all $i \in \mathcal{I}$. Therefore,

$$u_i^1(\theta_i, \boldsymbol{x}_{-i}) = |\{t : v_t(\boldsymbol{x}_{-i}, \theta_i) = 1, t \in \mathcal{P}_i^1\}| = P_i^1.$$

The proof for (ii) is analogous. □

Theorem 5

Proof. Based on Theorem 3, $u_i^1(\theta_i, \boldsymbol{x}_{-i})$ and $u_i^0(\theta_i, \boldsymbol{x}_{-i})$ attain their highest values P_i^1 and P_i^0 in $[0, \theta_i^{1,*}]$ and $[\theta_i^{0,*}, 1]$, respectively. Hence, if $\theta_i^{0,*}(\boldsymbol{x}_{-i}) \leq \theta_i^{1,*}(\boldsymbol{x}_{-i})$, for any $\tilde{\theta}_i \in [\theta_i^{0,*}, \theta_i^{1,*}]$,

$$\max_{\tilde{x}_i \in \mathscr{P}(\mathcal{F}_i)} u_i((\tilde{x}_i, \tilde{\theta}_i), \boldsymbol{x}_{-i}) = H_i^0 + H_i^1 + \max_{\tilde{x}_i \in \mathscr{P}(\mathcal{F}_i)} \Big(u_i^+((\tilde{x}_i, \tilde{\theta}_i), \boldsymbol{x}_{-i}) + u_i^-((\tilde{x}_i, \tilde{\theta}_i), \boldsymbol{x}_{-i}) \Big)$$

Let us construct a strategy $\boldsymbol{s}_i'$, as follows:

$$\forall\ t \in \mathcal{F}_i,\ t \in \boldsymbol{s}_i' \iff \sum_{j\in\mathcal{E}_i^+ : t\in\boldsymbol{x}_j} w_{i,j} + \sum_{j\in\mathcal{E}_i^- : t\notin\boldsymbol{x}_j} w_{i,j} \geq \tilde{\theta}_i \sum_{j\in\mathcal{E}_i} w_{i,j}.$$

We have

$$\max_{\tilde{x}_i \in \mathscr{P}(\mathcal{F}_i)} \Big(u_i^+((\tilde{x}_i, \tilde{\theta}_i), \boldsymbol{x}_{-i}) + u_i^-((\tilde{x}_i, \tilde{\theta}_i), \boldsymbol{x}_{-i}) \Big) \geq u_i^+((\tilde{x}_i', \tilde{\theta}_i), \boldsymbol{x}_{-i}) + u_i^-((\tilde{x}_i', \tilde{\theta}_i), \boldsymbol{x}_{-i}) = F_i.$$

Since $\overline{u}_i(\boldsymbol{x}_{-i}) \leq F$, then $\overline{u}_i(\boldsymbol{x}_{-i}) = F_i + H_i^0 + H_i^1 = F$. This proves (i).

As for (ii), notice that a lower bound of $\overline{u}_i$ can be obtained by fixing θ to some fixed values. Let $\underline{\theta}_i^{0,*}$ and $\overline{\theta}_i^{1,*}$ be the bounds from lemma 1, we have

$$u_i((\tilde{x}_i, \underline{\theta}_i^{0,*}), \boldsymbol{x}_{-i}) = u_i^+((\tilde{x}_i, \underline{\theta}_i^{0,*}), \boldsymbol{x}_{-i}) + u_i^-((\tilde{x}_i, \underline{\theta}_i^{0,*}), \boldsymbol{x}_{-i}) + u_i^1(\underline{\theta}_i^{0,*}, \boldsymbol{x}_{-i}) + \underbrace{u_i^0(\underline{\theta}_i^{0,*}, \boldsymbol{x}_{-i})}_{=P_i^0}$$

Notice that $u_i^1(\underline{\theta}_i^{0,*}, \boldsymbol{x}_{-i})$ doesn't depend on $\tilde{x}_i$, we can choose $\boldsymbol{x}_i'$ same as previously to make sure the sum of first two terms equals F_i.

$$\overline{u}_i(\boldsymbol{x}_{-i}) = \max_{(\tilde{x}_i, \theta_i) \in \mathscr{P}(\mathcal{F}_i) \times \Theta_i} u_i((\tilde{x}_i, \theta_i), \boldsymbol{x}_{-i}) \geq \max_{\tilde{x}_i \in \mathscr{P}(\mathcal{F}_i)} u_i((\tilde{x}_i, \underline{\theta}_i^{0,*}), \boldsymbol{x}_{-i}) \geq F_i + P_i^0 + u_i^1(\underline{\theta}_i^{0,*}, \boldsymbol{x}_{-i})$$

Similarly, by evaluating θ_i at $\overline{\theta}_i^{1,*}$, we obtain $\overline{u}_i(\boldsymbol{x}_{-i}) \geq F_i + P_i^1 + u_i^0(\overline{\theta}_i^{1,*}, \boldsymbol{x}_{-i})$.

□

Proposition 2

Proof. To prove the NP-completeness of R-GVG-CNC we establish an equivalence with respect to the 3-SAT problem, which consists in determining the satisfiability of a formula in conjunctive normal form, where each clause is limited to at most three literals.[15]

To establish this equivalence, we introduce the following notation. Let us consider a collection of m clauses and L literals $y_1, \ldots, y_L$. We define $\iota : \{1, \ldots, m\} \times \{1, 2, 3\} \to \{1, \ldots, L\} \times \{0, 1\}$ and characterize a Boolean expression (in disjunctive normal form) as

$$\mathrm{BE}(y_1, \ldots, y_L) = \bigwedge_{i=1}^{m} \left(\bigvee_{j=1}^{3} (y_{\iota_0(i,j)})^{\iota_1(i,j)} (1 - y_{\iota_0(i,j)})^{1-\iota_1(i,j)} \right). \tag{A.5}$$

To establish an equivalence between this Boolean expression and a signed graph $\mathcal{G} = \langle \mathcal{I}, \mathcal{E}^+, \mathcal{E}^- \rangle$, we define

$$\begin{aligned}
\mathcal{I} &= \{c_1, \ldots, c_m, v_{1,1}, v_{1,2}, v_{1,3}, v_{2,1}, v_{2,2}, v_{2,3}, \ldots v_{m,1}, v_{m,2}, v_{m,3}, l_1, \ldots, l_L, a_0, a_1\}, \\
\mathcal{E}^+ &= \{(c_i, v_{i,j}) \ : \ i = 1, \ldots, m, \ j = 1, \ldots, 3\} \\
&\quad \cup \left\{(v_{i,j}, l_{\iota_0(i,j)}) \ : \ i = 1, \ldots, m, \ j = 1, \ldots, 3, \iota_1(i,j) = 1\right\} \\
&\quad \cup \{(l_k, a_0) \ : \ k = 1, \ldots, L\}, \\
\mathcal{E}^- &= \left\{(v_{i,j}, l_{\iota_0(i,j)}) \ : \ i = 1, \ldots, m, \ j = 1, \ldots, 3, \iota_1(i,j) = 0\right\} \\
&\quad \cup \{(l_k, a_1) \ : \ k = 1, \ldots, L\},
\end{aligned} \tag{A.6}$$

This is elucidated in the following illustrative example.

Example 1. *To visualize this equivalence between disjunctive normal forms with three literals per clause and signed graphs, we consider a small example:* $(y_1 \vee y_2 \vee y_3) \wedge (\neg y_1 \vee y_2 \vee y_4) \wedge (\neg y_3 \vee y_4 \vee y_5)$. *The equivalence between this Boolean expression and a signed graph* $\mathcal{G} = \langle \mathcal{I}, \mathcal{E}^+, \mathcal{E}^- \rangle$ *is established by*

[15] Note that any conjunctive normal form with clause having less than three literals (e. g., $(y_1 \vee y_2)$) can be equivalently expressed as conjunctive normal form with clause having exactly three literals. This is obtained by introducing dummy literals ε and replicating the clause by affirming and negating ε $((y_1 \vee y_2 \vee \varepsilon) \wedge (y_1 \vee y_2 \vee \neg\varepsilon))$.

defining $\mathcal{I}$, $\mathcal{E}^+$, $\mathcal{E}^-$ as follows:

$$\mathcal{I} = \{c_1, c_2, c_3, v_{1,1}, v_{1,2}, v_{1,3}, v_{2,1}, v_{2,2}, v_{2,3}, v_{3,1}, v_{3,2}, v_{3,3}, l_1, l_2, l_3, l_4, l_5, a_0, a_1\},$$

$$\begin{aligned}\mathcal{E}^+ &= \{(c_i, v_{i,j}) \;:\; i = 1, \ldots, m,\; j = 1, \ldots, 3\} \\ &\quad \cup \big\{(v_{i,j}, l_{\iota_0(i,j)}) \;:\; i = 1, \ldots, m,\; j = 1, \ldots, 3, \iota_1(i,j) = 1\big\} \\ &\quad \cup \{(l_k, a_0) \;:\; k = 1, \ldots, L\},\end{aligned}$$

$$\begin{aligned}\mathcal{E}^- &= \big\{(v_{i,j}, l_{\iota_0(i,j)}) \;:\; i = 1, \ldots, m,\; j = 1, \ldots, 3, \iota_1(i,j) = 0\big\} \\ &\quad \cup \{(l_k, a_1) \;:\; k = 1, \ldots, L\}.\end{aligned}$$

Figure 1 illustrates the corresponding signed graphs.

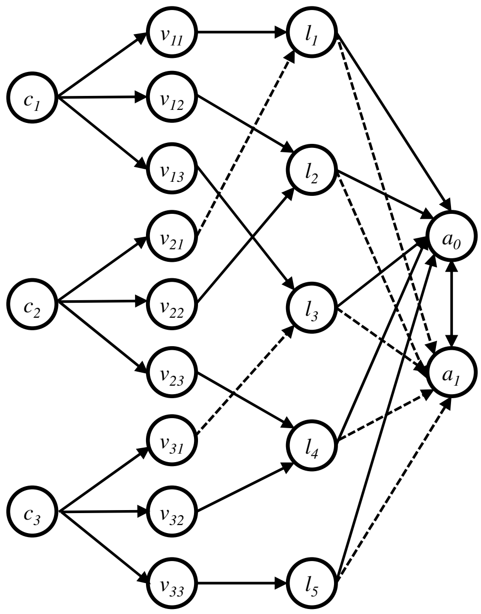


Figure 1: The constructed network for the formula $(y_1 \vee y_2 \vee y_3) \wedge (\neg y_1 \vee y_2 \vee y_4) \wedge (\neg y_3 \vee y_4 \vee y_5)$.

Visibly, this example reveals that the equivalent signed graph (A.6) has a four-partite structure, where the first layer refers to the m clauses, the third layer refers to the L literals, and the second layer consists of placeholders for the literals inside the clauses (namely, in which place of each clause a literal appear). Finally, the fourth layer contains two auxiliary nodes whose purpose will be clarified next in the proof.

Next, for each disjunctive normal form with three literals per clause (encoded in a signed graph (A.6)), we construct an instance of GVG-CNC($\mathcal{I}$, $\mathcal{E}^+$, $\mathcal{E}^-$, $\mathcal{P}^+$, $\mathcal{P}^-$, W, Θ) with one item ($\mathcal{F} = \{t\}$)

and $4m + L + 2$ individuals, by defining

$$
\begin{aligned}
&\mathcal{P}^+ &&= \{(c_i, t) \ : \ \forall i = 1, \ldots, m\} \cup \{(a_0, t), (a_1, t)\} && \\
&\mathcal{P}^- &&= \emptyset && \\
&W(i, j) &&= 1 && \forall \ (i, j) \in \mathcal{E}^+ \cup \mathcal{E}^- \\
&\Theta_{c_i} &&= \{1\} && \\
&\Theta_{v_{i,j}} &&= \{1\} && \forall i = 1, \ldots m, \ \forall j = 1, 2, 3 \\
&\Theta_{l_k} &&= \{1, 2\} && \forall k = 1, \ldots, L \\
&\Theta_{a_0} = \Theta_{a_1} &&= \{1\}. &&
\end{aligned}
$$

We now proceed to prove the equivalence of the 3-SAT problem and the recognition version of GVG-CNC problem, by distinguishing the two implications.

(3-SAT) implies GVG-CNC. If 3-SAT has a solution, let's construct a solution to the GVG-CNC problem which has a total utility greater or equal to Fn. Let $b_1, \ldots, b_L \in \{\text{True}, \text{False}\} \equiv \{1, 0\}$ be the values for the literals $y_1, \ldots, y_L$, such that $\text{BE}(b_1, \ldots, b_L) = 1$. We construct the following voting profiles $x_i \subseteq \mathcal{F}$ for each $i \in \mathcal{I}$:

$$
\begin{aligned}
&(\boldsymbol{x}_{c_i}, \theta_{c_i}) &&= (\{t\}, 1) && \forall i = 1, , \ldots, m \\
&(\boldsymbol{x}_{v_{i,j}}, \theta_{v_{i,j}}) &&= (\{t\}, 1) && \forall i = 1, \ldots, m \ \forall j = 1, 2, 3 \ \text{ if } b_{\iota_0(i,j)} = \iota_1(i, j) \\
&(\boldsymbol{x}_{v_{i,j}}, \theta_{v_{i,j}}) &&= (\emptyset, 1) && \forall i = 1, \ldots, m \ \forall j = 1, 2, 3 \ \text{ if } b_{\iota_0(i,j)} \neq \iota_1(i, j) \\
&(x_{l_k}, \theta_{l_k}) &&= (\{t\}, 1) && \forall k = 1, \ldots, L, \ \text{ if } b_k = \text{True} \\
&(\boldsymbol{x}_{l_k}, \theta_{l_k}) &&= (\emptyset, 2) && \forall k = 1, \ldots, L, \ \text{ if } b_k = \text{False} \\
&(\boldsymbol{x}_{a_0}, \theta_{a_0}) = (x_{a_1}, \theta_{a_1}) &&= (\{t\}, 1), &&
\end{aligned}
$$

where $\boldsymbol{x}_i$ is the subset of $\mathcal{F}$ denoting the collection of features selected by individual i.

We can show that if $\text{BE}(b_1, \ldots, b_L) = 1$, then the utility function $u_i((\tilde{x}_i, \theta_i), \boldsymbol{x}_{-i}) = F = 1$, for all $i \in \mathcal{I}$. Since we have a single item, this boils down to $v_t(\boldsymbol{x}_{-i}, \theta_i) = x(i, t)$, where x is the associated choice profile function. Therefore, using the weighted sum we can further reformulate it as follows:

$$
x(i, t) = 1 \Longleftrightarrow \sum_{j \in \mathcal{E}_i^+} x_{j,t} + \sum_{j \in \mathcal{E}_i^-} (1 - x_{j,t}) \geq \theta_i \quad \text{for } i \in \mathcal{I} \text{ and } t \in \mathcal{F}, \tag{A.7}
$$

where a_0 is only influenced by a_1 (more precisely $\{a_1\} = \mathcal{E}_{a_0} = \mathcal{E}_{a_0}^+$). Since $x(a_1, t) = 1$ and $\theta_{a_0} = 1$ are assigned by construction, (A.7) holds for a_0, so that $u_{a_0}((\boldsymbol{x}_{a_0}, \theta_{a_0}), \boldsymbol{x}_{-a_0}) = 1$. The same argument is valid for a_1, which implies that the fourth layer of the four-partite graph attains the highest utility.

Focusing on the third layer, for each $k = 1, \ldots, L$, the node l_k is influenced attractively by a_0 and repulsively by a_1. More formally, $\mathcal{E}_{l_k}^+ = \{a_0\}$ and $\mathcal{E}_{l_k}^- = \{a_1\}$. Since its weighted sum is always 1, by varying the threshold level $\theta_i = 1$ or 2, the voting result would suggest either 1 or 0 accordingly, which matches the assigned value for y_{l_k}.

Focusing on the second layer, for each $i = 1, \ldots, m$, $j = 1, 2, 3$, the node $v_{i,j}$ is only influenced by $l_{\iota_0(i,j)}$. Hence, we have $\{l_{\iota_0(i,j)}\} = \mathcal{E}^+_{v_{i,j}}$ when $\iota_1(i,j) = 1$ and $\{l_{\iota_0(i,j)}\} = \mathcal{E}^-_{v_{i,j}}$ when $\iota_1(i,j) = 0$. Inspecting whether weighted sum being greater than the threshold yields:

$$\begin{aligned} x(l_{\iota_0(i,j)}, t)\iota_1(i,j) + (1 - x(l_{\iota_0(i,j)}), t)(1 - \iota_1(i,j)) \geq \underbrace{\theta_{v_{i,j}}}_{=1} &\iff x(l_{\iota_0(i,j)}, t) = \iota_1(i,j) \\ &\iff b_{\iota(i,j)} = \iota_1(i,j) \\ &\iff x(v_{i,j}, t) = 1. \end{aligned}$$

This implies the utility associated to $v_{i,j}$ is maximized.

Finally, focusing on the first layer, we have that for each $i = 1, \ldots, m$, the node c_i is influenced by exactly three nodes $v_{i,1}$, $v_{i,2}$, and $v_{i,3}$. Inspecting whether the weighted sum being greater than the threshold yields:

$$\begin{aligned} x(v_{i,1}, t) + x(v_{i,2}, t) + x(v_{i,3}, t) \geq \underbrace{\theta_{c_i}}_{=1} &\iff (x(v_{i,1}, t) = 1) \vee (x(v_{i,2}, t) = 1) \vee (x(v_{i,3}, t) = 1) \\ &\iff \bigvee_{j=1}^{3} (x(v_{i,j}, t) = 1) \\ &\iff \bigvee_{j=1}^{3} (b_{\iota_0(i,j)} = \iota_1(i,j)) \\ &\iff \bigvee_{j=1}^{3} (b_{\iota_0(i,j)})^{\iota_1(i,j)} (1 - b_{\iota_0(i,j)})^{1-\iota_1(i,j)} \\ &\iff \text{True} \iff x(c_i, t) = 1. \end{aligned}$$

(DVG-CNC) implies (3-SAT). Conversely, if there exists a solution to the GVG-CNC problem with total utility greater or equal to Fn, let's show that 3-SAT has a solution. To do so, let $\boldsymbol{x}_i$ and θ_i be the choices corresponding to this maximal utility level, for $i \in \mathcal{I}$. We proceed to show that the Boolean expression (A.5) is satisfied by the following evaluation of literals:

$$y_k = b_k = \begin{cases} \text{True} & \text{if } x_{l_k} = \{t\} \\ \text{False} & \text{otherwise .} \end{cases} \tag{A.8}$$

In fact, the maximal utility in DVG-CNC is attained when every player utility is equal to one. Firstly by focusing on the first layer, $u_{c_i}((\boldsymbol{x}_{c_i}, \theta_{c_i}), (\boldsymbol{x}_{-c_i}) = 1$ entails

$$x(c_i, t) = 1 \iff x(v_{i,1}, t) + x(v_{i,2}, t) + x(v_{i,3}, t) \geq \underbrace{\theta_{c_i}}_{=1}, \quad \text{i=1, \ldots, m} .$$

Similarly, for the second layer, the utility of players $v_{i,j}$ are maximized, by noticing that the only neighbor of $v_{i,j}$ is $l_{\iota_0(i,j)}$, with the attractiveness determined by $\iota_1(i,j)$. Formally, this can be expressed as

$$\begin{aligned} x(v_{i,j},t)=1 \quad &\iff x(l_{\iota_0(i,j)},t)\iota_1(i,j)+(1-x(l_{\iota_0(i,j)}),t)(1-\iota_1(i,j))\geq 1 \\ &\iff x(l_{\iota_0(i,j)},t)=\iota_1(i,j). \end{aligned}$$

Therefore, for every $i=1,\ldots,m$,

$$\begin{aligned} \bigvee_{j=1}^{3}(b_{\iota_0(i,j)})^{\iota_1(i,j)}(1-b_{\iota_0(i,j)})^{1-\iota_1(i,j)} &= \bigvee_{j=1}^{3}(b_{\iota_0(i,j)}=\iota_1(i,j)) \\ &= \bigvee_{j=1}^{3}(x(l_{\iota_0(i,j)},t)=\iota_1(i,j)) \\ &= \bigvee_{j=1}^{3}(x(v_{i,j},t)=1) \\ &= (x(v_{i,1},t)=1)\vee(x(v_{i,2},t)=1)\vee(x(v_{i,3},t)=1) \\ &= (x(v_{i,1},t)+x(v_{i,2},t)+x(v_{i,3},t)\geq \underbrace{1}_{=\theta_{c_i}}) \\ &= (x(c_i,t)=1)=\text{True} \end{aligned}$$

It follows that the Boolean expression (A.5) is satisfied by (A.8).

□

Theorem 5

Proof. Based on Lemma 4, if $\overline{\overline{u}}=nF$, then any social welfare solution $\{(\tilde{x}_i^W,\theta_i^W)\ :\ i\in\mathcal{I}\}$ constitutes a best respond strategy profile. Hence, $\{(\tilde{x}_i^W,\theta_i^W)\ :\ i\in\mathcal{I}\}$ is an equilibrium point for $\mathcal{G}$ (no player can improve its payoff by unilaterally departing from it). This proves the statement (1.).

Next, let $\{(\tilde{x}_i^E,\theta_i^E)\ :\ i\in\mathcal{I}\}$ be an equilibrium point for which

$$\sum_{i\in\mathcal{I}} u_i((\tilde{x}_i^E,\theta_i^E),\tilde{\boldsymbol{x}}_{-i}^E)<nF.$$

This implies that there must exist $i'\in\mathcal{I}$ for which

$$u_{i'}((\tilde{x}_{i'}^E,\theta_{i'}^E),\tilde{\boldsymbol{x}}_{-i'}^E)<F.$$

Based on Lemma 4, this implies $\theta_{i'}^{0,*}(\boldsymbol{x}_{-i})>\theta_{i'}^{1,*}(\boldsymbol{x}_{-i})$. Hence, if for all $i\in\mathcal{I}$, $\theta_i^{0,*}(\boldsymbol{x}_{-i}^E)\leq\theta_i^{1,*}(\boldsymbol{x}_{-i}^E)$, then

$$u_i((\tilde{x}_i^E,\theta_i^E),\tilde{\boldsymbol{x}}_{-i'}^E)=F,$$

so that any equilibrium point (if it exists) is a welfare solution.

□

Theorem 6

Proof. By way of contradiction, let $(\boldsymbol{x}_i)_{i=1\dots n}$ be an equilibrium point and let $(x_{i,t})_{i=1\dots n, t\in\mathcal{F}}$ be the corresponding binary representation encoding of whether $t \in \boldsymbol{x}_i$. Since $\mathcal{G}$ is a negative cycle, by reordering the nodes, we assume that node i is only connected to the node $i+1$. Also, θ_i is constrained to be 1. Define

$$e_i = \begin{cases} 1 & \text{if } i+1 \in \mathcal{E}_i^+ \\ -1 & \text{if } i+1 \in \mathcal{E}_i^- \end{cases}$$

Since $\mathcal{G}$ is a negative cycle, $\prod_{i=1}^n e_i = -1$. Further, by Theorem (4), since $P_i^0 = P_i^1 = 0$, then $1 = \theta_i^{0,*}(\boldsymbol{x}_{-i}) \leq \theta_i^{1,*}(\boldsymbol{x}_{-i}) = d_i = 1$, then

$$\max_{(\tilde{x}_i,\theta_i)\in\mathscr{P}(\mathcal{F}_i)\times\Theta_i} u_i((\tilde{x}_i,\theta_i),\boldsymbol{x}_{-i}) = F.$$

This implies

$$\forall t \in \mathcal{F}\ , t \in \boldsymbol{x}_i \iff v_t(\boldsymbol{x}_{-i},\theta_i) = 1$$

By definition of voting,

$$\begin{aligned} x_{i,t} = 1 \quad &\iff \quad v_t(\boldsymbol{x}_{-i},\theta_i) = 1 \\ &\iff \quad \sum_{j\in\mathcal{E}_i^+} x_{j,t} + \sum_{j\in\mathcal{E}_i^-} (1-x_{j,t}) \geq 1 \\ &\iff \quad x_{j,t} = \mathbf{1}\left(j \in \mathcal{E}_i^+\right), \text{ where } j = i+1. \end{aligned}$$

Therefore,

$$x(i+1,t) = v_t(\boldsymbol{x}_{-(i+1)},\theta_{i+1}) = \begin{cases} x_{i,t} & \text{if } e_i = 1 \\ 1-x_{i,t} & \text{if } e_i = -1 \end{cases}$$

Or equivalently $(-1)^{x_{i+1,t}} = (-1)^{x_{i,t}}\ e_i$, which leads to a contradiction by taking product over i.

□

Proposition 3

Proof. Let $\mathrm{WS}_{i,t}^{(\mathrm{Ins}')}(\boldsymbol{w}_i)$ and $\mathrm{WS}_{i,t}^{(\mathrm{Ins}')}(\boldsymbol{w}_i)$ be the weighted sum associated to X and $\mathtt{C}_t(X)$, respectively. We have the following equivalences:

$$\begin{aligned} u_{i,t}^{(\mathrm{Ins})}(\boldsymbol{w}_i,\theta_i) = 1 &\iff v_{i,t}^{(\mathrm{Ins})}(\boldsymbol{w}_i,\theta_i) = x_{i,t} \\ &\iff \mathrm{WS}_{i,t}^{(\mathrm{Ins})}(\boldsymbol{w}_i) \geq \theta_i \\ &\iff 1 - \sum_{j=1}^n \left(\overline{w}_{i,j}x_{j,t} + \underline{w}_{i,j}(1-x_{j,t})\right) \leq 1-\theta_i \\ &\iff \mathrm{WS}_{i,t}^{(\mathrm{Ins}')}(\boldsymbol{w}_i) \leq 1-\theta_i \\ &\iff v_{i,t}^{(\mathrm{Ins}')}(\boldsymbol{w}_i, 1-\theta_i) = 1-x_{i,t} \\ &\iff u_{i,t}^{(\mathrm{Ins}')}(\boldsymbol{w}_i, 1-\theta_i) = 1. \end{aligned}$$

To achieve full equivalence, we need the inequality to be strict. This can be achieved by replacing θ_i with $\theta_i - \varepsilon$ where $\varepsilon > 0$ is a small enough positive number. We just need to show that such $\varepsilon > 0$ exists. This is established from the following property:

$$\exists\, \varepsilon > 0 \quad \forall t \in \mathcal{F} \quad \mathrm{WS}_{i,t}^{(\textsc{Ins})}(\boldsymbol{w}_i) \geq \theta_i \Longleftrightarrow \mathrm{WS}_{i,t}^{(\textsc{Ins})}(\boldsymbol{w}_i) \geq \theta_i - \varepsilon.$$

In fact, we can pick ε small enough such that $\theta_i - \varepsilon > \max_{t\,:\,x_{i,t}=0} \mathrm{WS}_{i,t}^{(\textsc{Ins})}(\boldsymbol{w}_i)$. Therefore by replacing with a θ_i that avoid equality in every item,

$$\mathrm{WS}_{i,t}^{(\textsc{Ins}')}(\boldsymbol{w}_i) \leq 1 - \theta_i \iff \mathrm{WS}_{i,t}^{(\textsc{Ins}')}(\boldsymbol{w}_i) < 1 - \theta_i.$$

□

Proposition 4

Proof. Consider the weighted sum of the two items,

$$\mathrm{WS}_{i,t_1}(w_i) = \sum_{j=1}^{n} \left(\overline{w}_{i,j} x_{j,t_1} + \underline{w}_{i,j}(1 - x_{j,t_1})\right) = \sum_{j=1}^{n} \left(\overline{w}_{i,j} x_{j,t_2} + \underline{w}_{i,j}(1 - x_{j,t_2})\right) = \mathrm{WS}_{i,t_2}(w_i).$$

Since $\mathrm{WS}_{i,t_1}(w_i) = \mathrm{WS}_{i,t_2}(w_i)$, then $v_{i,t_1}(w_i, \theta_i) = v_{i,t_2}(w_i, \theta_i)$. To lighten the notation, we write v_{i,t_s} in place of $v_{i,t_s}(w_i, \theta_i)$ in the rest of this proof, when the input argument (w_i, θ_i) is clear.We know that either $v_{i,t_1} = x_{i,t_1}$ or $v_{i,t_2} = x_{i,t_2}$ (but not both). Hence, for any $(w_i, \theta_i) \in \mathcal{Q}_i$, we have

$$u_i(w_i, \theta_i) = \sum_{t \in \mathcal{F}} u_{i,t} \leq F - 2 + u_{i,t_1} + u_{i,t_2} = F - 2 + \mathbf{1}\left(v_{i,t_1} = x_{i,t_1}\right) + \mathbf{1}\left(v_{i,t_2} = x_{i,t_2}\right) = F - 1 < F.$$

To prove (iii), suppose that the inequality statement holds, define $\theta_i = \min \left\{\mathrm{WS}_{i,t}(\boldsymbol{w}_i) \mid t : x(i,t) = 1\right\}$ for each $t \in \mathcal{F}$

$$\begin{aligned} u_{i,t}(\boldsymbol{w}_i, \theta_i) = 1 \quad &\iff v_t(\boldsymbol{w}_i, \theta_i) = x(i,t) \\ &\iff \mathbf{1}\left(\mathrm{WS}_{i,t}(\boldsymbol{w}_i) \geq \theta_i\right) = x(i,t). \end{aligned}$$

The last term holds for each t by definition of our θ_i here. □

B : Auxiliary technical notes

Optimality bounds for $\theta_i^{0,*}$ and $\theta_i^{1,*}$

A closed-form sufficient condition for the best responds to attain the highest value F, is provided by the following bounds:

$$\begin{aligned}
\max_{\tilde{\boldsymbol{x}}_{-i}} \theta_i^{1,*}(\boldsymbol{x}_{-i}) &\geq \min_{t \in \mathcal{P}_i^1} \left(\sum_{j \in \mathcal{E}_i^+} \mathbf{1}\left(t \in \mathcal{P}_j^1\right) w_{i,j} + \sum_{j \in \mathcal{E}_i^-} \mathbf{1}\left(t \in \mathcal{P}_j^0\right) w_{i,j} \right) = \underline{\theta}_i^{1,*}, \\
\min_{\tilde{\boldsymbol{x}}_{-i}} \theta_i^{0,*}(\boldsymbol{x}_{-i}) &\leq \left(1 + \max_{t \in \mathcal{P}_i^0} \left(\sum_{j \in \mathcal{E}_i^+} \mathbf{1}\left(t \in \mathcal{F} \setminus \mathcal{P}_j^0\right) w_{i,j} + \sum_{j \in \mathcal{E}_i^-} \mathbf{1}\left(t \in \mathcal{F} \setminus \mathcal{P}_j^1\right) w_{i,j} \right) \right) = \overline{\theta}_i^{0,*}.
\end{aligned}$$

The detailed steps to determine this bonds are given below.

$$\begin{aligned}
\theta_i^{1,*}(\boldsymbol{x}_{-i}) &:= \begin{cases} \max\limits_{\theta_i \in [d_i]} & \theta_i \\ \text{s.t.} & v_t(\boldsymbol{x}_{-i}, \theta_i) = 1 \ \forall t \in \mathcal{P}_i^1 \end{cases} \\
&= \begin{cases} \max\limits_{\theta_i \in [d_i]} & \theta_i \\ \text{s.t.} & \min\limits_{t \in \mathcal{P}_i^1} v_t(\boldsymbol{x}_{-i}, \theta_i) = 1 \end{cases} \\
&= \begin{cases} \max\limits_{\theta \in [d_i]} & \theta_i \\ \text{s.t.} & \min\limits_{t \in \mathcal{P}_i^1} \left(\sum_{j \in \mathcal{E}_i^+} \mathbf{1}\left(t \in \boldsymbol{x}_j\right) w_{i,j} + \sum_{j \in \mathcal{E}_i^-} \mathbf{1}\left(t \in \mathcal{F} \setminus \boldsymbol{x}_j\right) w_{i,j} \right) \geq \theta_i \end{cases} \\
&= \min_{t \in \mathcal{P}_i^1} \left(\sum_{j \in \mathcal{E}_i^+} \mathbf{1}\left(t \in \boldsymbol{x}_j\right) w_{i,j} + \sum_{j \in \mathcal{E}_i^-} \mathbf{1}\left(t \in \mathcal{F} \setminus \boldsymbol{x}_j\right) w_{i,j} \right) \\
&= \min_{t \in \mathcal{P}_i^1} \left(\sum_{j \in \mathcal{E}_i^+} \mathbf{1}\left(t \in \tilde{\boldsymbol{x}}_j\right) w_{i,j} + \mathbf{1}\left(t \in \mathcal{P}_j^1\right) w_{i,j} + \sum_{j \in \mathcal{E}_i^-} \mathbf{1}\left(t \in \mathcal{F}_i \setminus \tilde{\boldsymbol{x}}_j\right) w_{i,j} + \mathbf{1}\left(t \in \mathcal{P}_j^0\right) w_{i,j} \right)
\end{aligned}$$

By properly choosing $\tilde{\boldsymbol{x}}_j$ for $j \in \mathcal{F}_i$ to zero out the first and the third term, we have

$$\max_{\tilde{\boldsymbol{x}}_{-i}} \theta_i^{1,*}(\boldsymbol{x}_{-i}) \geq \min_{t \in \mathcal{P}_i^1} \left(\sum_{j \in \mathcal{E}_i^+} \mathbf{1}\left(t \in \mathcal{P}_j^1\right) w_{i,j} + \sum_{j \in \mathcal{E}_i^-} \mathbf{1}\left(t \in \mathcal{P}_j^0\right) w_{i,j} \right)$$

and

$$\theta_i^{0,*}(\boldsymbol{x}_{-i}) := \begin{cases} \min\limits_{\theta_i \in [d_i]} & \theta_i \\ \text{s.t.} & v_t(\boldsymbol{x}_{-i}, \theta_i) = 0 \ \forall \ t \in \mathcal{P}_i^0 \end{cases}$$

$$= \begin{cases} \min\limits_{\theta \in [d_i]} & \theta_i \\ \text{s.t.} & \max\limits_{t \in \mathcal{P}_i^0} \left(\sum\limits_{j \in \mathcal{E}_i^+} \mathbf{1}\left(t \in \boldsymbol{x}_j\right) w_{i,j} + \sum\limits_{j \in \mathcal{E}_i^-} \mathbf{1}\left(t \in \mathcal{F} \setminus \boldsymbol{x}_j\right) w_{i,j} \right) \leq \theta_i - 1 \end{cases}$$

$$\leq \left(1 + \max_{t \in \mathcal{P}_i^0} \left(\sum_{j \in \mathcal{E}_i^+} \mathbf{1}\left(t \in \mathcal{F} \setminus \mathcal{P}_j^0\right) w_{i,j} + \sum_{j \in \mathcal{E}_i^-} \mathbf{1}\left(t \in \mathcal{F} \setminus \mathcal{P}_j^1\right) w_{i,j} \right)\right).$$

Doubled-instance equivalence of the IVG-CNC

Let IVG-CNC^R be a relaxed version of IVG-CNC, obtained by removing the complementary constraints $\overline{w}_{i,j}\underline{w}_{i,j} = 0$ from $\mathcal{W}_i$. Hence, the strategy space of IVG-CNC^R is associated to the following weights set:

$$\mathcal{W}_i^R = \{(\overline{w}_{i,1}, \ldots, \overline{w}_{i,n}, \underline{w}_{i,1}, \ldots, \underline{w}_{i,n}) \in \mathbb{R}_+^{2n} \ : \ \overline{w}_{i,j} = \underline{w}_{i,j} = 0, \ \forall j : (i,j) \notin \mathcal{E}\}.$$

Similarly, let IVG-CNC$^{FO}(\mathcal{I}, \mathcal{E}, \mathcal{F}, X, \Theta)$ be the friend-only re-formulation of IVG-CNC, with strategy set associated to the following weights set:

$$\mathcal{W}_i^{FO} = \{(\overline{w}_{i,1}, \ldots, \overline{w}_{i,n}, 0, \ldots, 0) \in \mathbb{R}_+^{2n} \ : \ \overline{w}_{i,j} = 0, \ \forall j : (i,j) \notin \mathcal{E}\}.$$

Therefore, $\mathcal{W}_i^{FO} \subset \mathcal{W}_i \subset \mathcal{W}_i^R$.

Proposition 5 (Doubled-instance equivalence)**.** *Let* (Ins) *and* (Ins$'$) *be instances of* IVG-CNC *and* IVG-CNC^{FO}*, parametrized by* $(\mathcal{I}, \mathcal{E}, \mathcal{F}, X, \Theta)$ *and* $(\mathcal{I}^D, \mathcal{E}^D, \mathcal{F}, X^D, \Theta)$*, respectively, where* $\mathcal{I}^D$*,* $\mathcal{E}^D$ *and* X^D *refers to a doubled-instance obtained by doubling the nodes and assign the corresponding doubled node with flipped features:*

$$\mathcal{I}^D = \{i_1, i_2, \ldots, i_n, i_1^D, \ldots, i_n^D\}, \quad \mathcal{E}^D = \mathcal{E} \cup \{(i, j^D) \ : \ \forall (i,j) \in \mathcal{E}\}, \quad X^D = \begin{pmatrix} X \\ \hline 1 - X \end{pmatrix} \in \mathbb{R}^{2n \times F}.$$

We have

$$\max_{(\boldsymbol{w}_i, \theta_i) \in \mathcal{W}_i^R \times \Theta_i} u_i^{(\text{Ins})}(\boldsymbol{w}_i, \theta_i) = \max_{(\boldsymbol{w'}_i, \theta_i) \in \mathcal{W}_i^{FO} \times \Theta_i} u_i^{(\text{Ins}')}(\boldsymbol{w'}_i, \theta_i), \quad \textit{for any } i \in \mathcal{I}. \tag{B.1}$$

Proof. We prove (B.1) in two parts. We first show that the LHS is less or equal than the RHS, and then the other way around.

($\leq$) We will show that for any feasible solution $(\boldsymbol{w}_i, \theta_i) \in \mathcal{W}_i^R \times \Theta_i$ of (INS), we can construct a feasible solution $(\boldsymbol{w}_i', \theta_i) \in \mathcal{W}_i^{FO} \times \Theta_i$ of (INS$'$) with the same utility.

$$u_i^{(\text{INS})}(\boldsymbol{w}_i, \theta_i) = u_i^{(\text{INS}')}(\boldsymbol{w'}_i, \theta_i) \leq \max_{(\boldsymbol{w'}_i, \theta_i) \in \mathcal{W}_i^{FO} \times \Theta_i} u_i^{(\text{INS}')}(\boldsymbol{w'}_i, \theta_i)$$

Since this inequality is true for any feasible $(\boldsymbol{w}_i, \theta_i)$, we have

$$\max_{(\boldsymbol{w}_i, \theta_i) \in \mathcal{W}_i^R \times \Theta_i} u_i^{(\text{INS})}(\boldsymbol{w}_i, \theta_i) \leq \max_{(\boldsymbol{w'}_i, \theta_i) \in \mathcal{W}_i^{FO} \times \Theta_i} u_i^{(\text{INS}')}(\boldsymbol{w'}_i, \theta_i).$$

Before we proceed to the construction, notice that if their weighted sums are equal then their utilities are equal. In fact, since $x_{i,t}$ and θ_i are equal in both formulation by definition, the utility depends only on the weighted sum. More specifically,

$$u_{i,t}^{(\text{INS})}(\boldsymbol{w}_i, \theta_i) = \mathbf{1}\left(v_{i,t}^{(\text{INS})}(\boldsymbol{w}_i, \theta_i) = x_{i,t}\right) = \mathbf{1}\left(\mathbf{1}\left(\text{WS}_{i,t}^{(\text{INS})}(\boldsymbol{w}_i, \theta_i) \geq \theta_i\right) = x_{i,t}\right).$$

To construct $\boldsymbol{w}_i'$, we make use of the fact that $\mathcal{W}_i^{FO}$ is of dimension $2|\mathcal{I}^{\mathcal{D}}| = 4|\mathcal{I}| = 4n$ with the last $2n$ dimensions filled with zero. We define a feasible solution $\boldsymbol{w'}_i$ as $\boldsymbol{w}_i$ followed by $2n$ zeros:

$$\boldsymbol{w'}_i = (\overline{w'}_{i,1}, \ldots, \overline{w'}_{i,n}, \overline{w'}_{i,n+1}, \ldots, \overline{w'}_{i,2n}, \underbrace{0, \ldots, 0}_{2n \text{ zeros}}) := (\overline{w}_{i,1}, \ldots, \overline{w}_{i,n}, \underline{w}_{i,1}, \ldots, \underline{w}_{i,n}, 0, \ldots, 0)$$

In other words, $\overline{w'}_{i,j} = \begin{cases} \overline{w}_{i,j} & \text{if } 1 \leq j \leq n \\ \underline{w}_{i,j-n} & \text{if } n+1 \leq j \leq 2n \end{cases}$.

The weighted sum can be calculated as follows

$$\begin{aligned} \text{WS}_{i,t}^{(\text{INS})}(\boldsymbol{w}_i) &= \sum_{j=1}^{n} \overline{w}_{i,j} x_{j,t} + \underline{w}_{i,j}(1 - x_{j,t}) \\ &= \sum_{j=1}^{n} \overline{w}_{i,j} x_{j,t} + \underline{w}_{i,j} x_{j+n,t} \\ &= \sum_{j=1}^{n} \overline{w'}_{i,j} x_{j,t} + \overline{w'}_{i,j+n} x_{j+n,t} \\ &= \sum_{j=1}^{2n} \overline{w'}_{i,j} x_{j,t} \\ &= \sum_{j=1}^{2n} \overline{w'}_{i,j} x_{j,t} + \underbrace{\underline{w'}_{i,j}}_{=0}(1 - x_{j,t}) \\ &= \text{WS}_{i,t}^{(\text{INS}')}(\boldsymbol{w'}_i) \end{aligned}$$

($\geq$) Inversely, we construct $\boldsymbol{w}_i$ given $\boldsymbol{w'}_i \in \mathcal{W}_i^{FO}$, by simply picking its first $2n$ entries. More specifically,

$$\boldsymbol{w}_i = (\overline{w}_{i,1}, \ldots, \overline{w}_{i,n}, \underline{w}_{i,1}, \ldots, \underline{w}_{i,n}) := (\overline{w'}_{i,1}, \ldots, \overline{w'}_{i,n}, \overline{w'}_{i,n+1}, \ldots, \overline{w'}_{i,2n}).$$

Since this definition coincides with the previous part of the proof, all previous equations still hold. We conclude the equality of the weighted sum

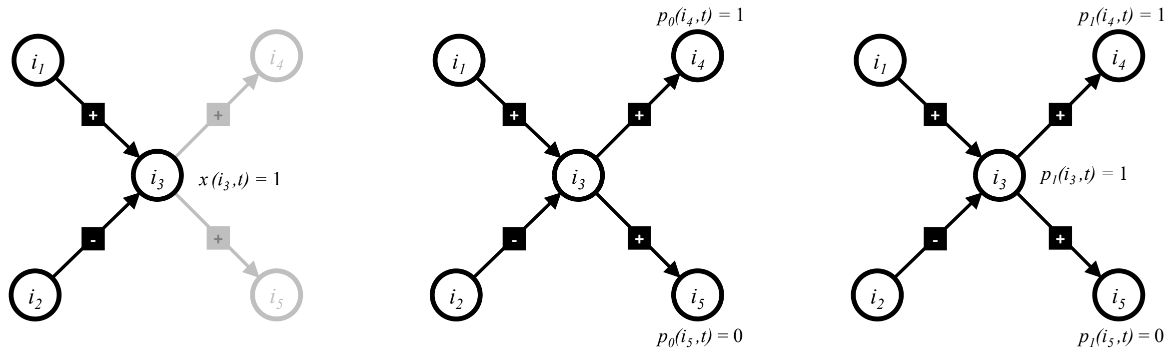
$$\text{WS}_{i,t}^{(\text{INS})}(\boldsymbol{w}_i) = \text{WS}_{i,t}^{(\text{INS}')}(\boldsymbol{w'}_i).$$

□

The idea of Proposition 5 is to replace each repulsive link $\underline{w}_{i,j}$ with a new attractive neighbor j^D with flipped choices, which implies doubling the number of nodes in IVG-CNC. The weight associated to this new neighbor is $\overline{w}_{i,j^D} = \underline{w}_{i,j}$. By this doubled-instance, the claim in Proposition 5 is that focusing only on the nodes in $\mathcal{I} \subset \mathcal{I}_D$ from IVG-CNC^{FO} (and ignore doubled nodes $i_1^D, \ldots, i_n^D$'s utilities), then both (Ins) and (Ins$'$) have the same global optimal solution.

C : Illustrative examples

Example 2 (Five-node graph). *Let us consider a case in which* $\mathcal{I} = \{i_1, i_2, i_3, i_4, i_5\}$ *and* $\mathcal{E} = \mathcal{E}^+ \cup \mathcal{E}^- = \{(i_1, i_3), (i_2, i_3), (i_3, i_4), (i_3, i_5)\}$*, as illustrated in Figure 2. This corresponds to the following lists of attractive and repulsive neighbors:* $\mathcal{E}^+_{i_1} = \{(i_1, i_3)\}$*,* $\mathcal{E}^+_{i_2} = \emptyset$*,* $\mathcal{E}^+_3 = \{(i_3, i_4), (i_3, i_5)\}$*,* $\mathcal{E}^+_{i_4} = \emptyset$*,* $\mathcal{E}^+_{i_5} = \emptyset$*, and* $\mathcal{E}^-_{i_1} = \emptyset$*,* $\mathcal{E}^-_{i_2} = \{(i_2, i_3)\}$*,* $\mathcal{E}^-_{i_3} = \emptyset$*,* $\mathcal{E}^+_{i_4} = \emptyset$*,* $\mathcal{E}^+_{i_5} = \emptyset$*.*



(a) Supported IF pair (b) Supportable IF pair (c) One-step supportable IF pair

Figure 2: Five-node graph with attached choice profiles and prerequisites. A directed (i, j) edge with sign $+$ means that j is an attractive neighbor of i, and one with sign $-$ means that j is a repulsive neighbor of i.

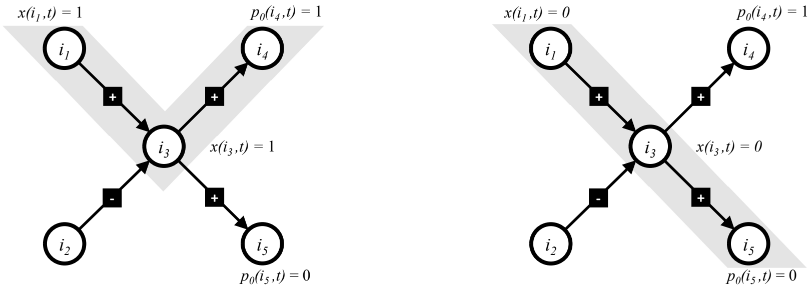
In panel (a) the IF pair (i_1, t) *is positively supported by* $\boldsymbol{x}_{-i_1}$*, as* i_1 *has an attracted neighbor (i.e.,* i_3*) that includes* t *in its selected features (i.e.,* $x(i_3, t) = 1$*). Likewise,* i_2 *is negatively supported by* $\boldsymbol{x}_{-i_2}$*, as* i_2 *has a repulsive neighbor (i.e.,* i_3*) that includes* t *in its selected features. Note that, given the presented graph structure, the only relevant information to assess whether* (i_1, t) *and* (i_2, t) *are supported is the choice of* i_3*. For this reason,* i_4 *and* i_5 *appears as shadow objects in panel (a).*

Panel (b) shows a prerequisite p_0 *that requires* i_4 *to select* t *and* i_5 *not to select it, while leaving everyone else free to choose. Therein, the IF pair* (i_1, t) *is supportable both positively and negatively by* p_0*, as there exists* $\boldsymbol{x}_{-i_1}$ *(which reduces to a feasible choice of* i_3 *in this specific graph structure) such that* (i_1, t) *is supported by* $\boldsymbol{x}_{-i_1}$*. The feasible choice* $x(i_3, t) = 1$ *supports* (i_1, t) *positively, while another feasible choice* $x(i_3, t) = 0$ *supports it negatively. Conversely,* (i_1, t) *is not one-step supportable, as none of the neighbors of* i_1 *is in* p_0*.*

Finally, panel (c) reports p_1*, a possible extension of* p_0*, in such a way that the IF pair* (i_1, t) *becomes positively one-step supportable by* p_1*. However, it stops being negatively supportable by* p_1 *(and consequently, not negatively one-step supportable).*

Example 3 (Supportability by paths and one-step extensions). *Given the same graph as in Example 2, panel (a) of Figure 3 shows the existence of a path (highlighted in the gray-shadowed region) connecting* i_1 *to an individual whose* t*-th choice belongs to the prerequisite, i.e.,* $(i_4, t) \in dom(p)$*.*

Since $1 - p(i_4, t) + (1 - \overline{\psi}_{i_1,i_3}) + (1 - \overline{\psi}_{i_3,i_4}) = 0$ *(even), then* (i_1, t) *is positively supportable by* p_0*. Similarly, panel (b) shows a different path which allows establishing the supportability of*

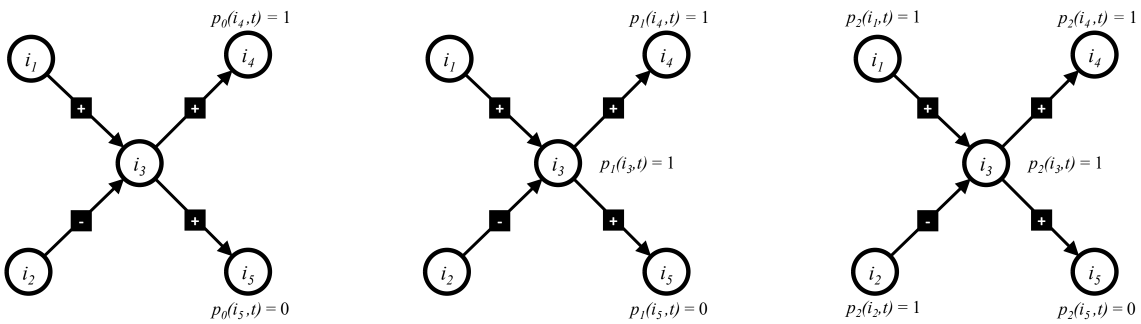


(a) Supported IF pair. (b) Supportable IF pair.

Figure 3: Five-node graph with attached choice profiles and prerequisites.

(i_1, t) *by* $\langle p, 1\rangle$. *This second path is* (i_1, i_3, i_5) *and its corresponding sign is* $(\overline{\psi}_{i_1,i_3}, \overline{\psi}_{i_3,i_5})$. *Since,* $1 - p(i_5, t) + (1 - \overline{\psi}_{i_1,i_3}) + (1 - \overline{\psi}_{i_3,i_5}) = 1$ *(odd), then* (i_1, t) *is negatively supportable by* p.

Based on Theorem 2, the supportability of (i_1, t) *can be achieved by the sequence of one-step extensions:* $P_2 = (p_0, p_1, p_2)$, *as shown in Figure 4. We have: dom*$(p_0) = \{(i_4, t), (i_5, t)\}$, *dom*$(p_1) = \{(i_4, t), (i_5, t), (i_3, t)\}$, *and dom*$(p_2) = \{(i_4, t), (i_5, t), (i_3, t), (i_2, t)\}$. *Hence,* p_1 *is a one-step extension of* p_0, *because* $\{(i_3, t)\} =$ *dom*$(p_1) \setminus$ *dom*(p_0) *is one step supportable by* p_0. *Similarly,* p_2 *is also a one-step extension of* p_1 *because all elements of* $\{(i_1, t), (i_2, t)\} =$ *dom*$(p_2) \setminus$ *dom*(p_1) *are one-step supportable by* p_1. *For* i_1, $p_2(i_1, t) = 1$ *since it has an attractive neighbor* i_3 *that chooses* 1. *For* i_2, $p_2(i_2, t) = 0$ *since it has a repulsive neighbor* i_3 *that chooses* 1.



(a) prerequisite p_0 (b) p_1 a one-step extension of p_0 (c) p_2 a one-step extension of p_2

Figure 4: Five-node graph with attached choice profiles and prerequisites.

Since at the end of the extension, $p_2(i_1, t) = 1$ *and* $p_2(i_2, t) = 0$, *we conclude that* (i_1, t) *and* (i_2, t) *are uniformly supportable by* $\langle p_0, \sigma\rangle$, *where* $\sigma(i_1, t) = 1$ *and* $\sigma(i_2, t) = 0$.[16]

Example 4 (Illustrative three-node example)**.** *Consider a three-node fully connected signed weighted graph* $\mathcal{G}_{+-} = \langle\{i_1, i_2, i_3\}, \mathcal{E}^+, \mathcal{E}^-, W\rangle$, *where* $\mathcal{E}^+_{i_1} = \{i_2\}$, $\mathcal{E}^-_{i_1} = \{i_3\}$, $\mathcal{E}^+_{i_2} = \{i_3\}$, $\mathcal{E}^-_{i_2} = \{i_1\}$, $\mathcal{E}^+_{i_3} = \{i_2\}$, $\mathcal{E}^-_{i_3} = \{i_1\}$, *so that* $\mathcal{E} = \{(i_1, i_2), (i_1, i_3), (i_2, i_3), (i_2, i_1), (i_3, i_1), (i_3, i_2)\}$. *We define*

[16]It is noteworthy to mention that we could have considered an alternative sequence of one-step extensions: $P'_2 = (p_0, p'_1, p'_2)$, where $p'_1(i_3, t) = 0$, $p'_2(i_1, t) = 0$, $p'_2(i_2, t) = 1$. In such a case, (i_1, t) and (i_2, t) would have been uniformly supportable by $\langle p_0, \sigma\rangle$, where $\sigma(i_1, t) = 0$ and $\sigma(i_2, t) = 1$.

a DVG-CNC with $\mathcal{F} = \{t_1, t_2\}$, and prerequisites $\mathcal{P}^+ = \{(i_3, t_1), (i_1, t_2)\}$ $\mathcal{P}^- = \{(i_2, t_1)\}$. To characterize players' best responses, we compute the weighted sums for all IF pairs:

$$\begin{cases} WS_{i_1,t_1}(\boldsymbol{x}_{-i_1}) & = \sum_{j\in\{i_2\}} x_{j,t_1} + \sum_{j\in\{i_3\}} (1 - x_{j,t_1}) = x_{i_2,t_1} + (1 - x_{i_3,t_1}) \\ WS_{i_1,t_2}(\boldsymbol{x}_{-i_1}) & = \sum_{j\in\{i_2\}} x_{j,t_1} + \sum_{j\in\{i_3\}} (1 - x_{j,t_1}) = x_{i_2,t_1} + (1 - x_{i_3,t_1}) \\ WS_{i_2,t_1}(\boldsymbol{x}_{-i_2}) & = \sum_{j\in\{i_3\}} x_{j,t_1} + \sum_{j\in\{i_1\}} (1 - x_{j,t_1}) = x_{i_3,t_1} + (1 - x_{i_1,t_1}) \\ WS_{i_2,t_2}(\boldsymbol{x}_{-i_2}) & = \sum_{j\in\{i_3\}} x_{j,t_1} + \sum_{j\in\{i_1\}} (1 - x_{j,t_1}) = x_{i_3,t_1} + (1 - x_{i_1,t_1}) \\ WS_{i_3,t_1}(\boldsymbol{x}_{-i_3}) & = \sum_{j\in\{i_2\}} x_{j,t_1} + \sum_{j\in\{i_1\}} (1 - x_{j,t_1}) = x_{i_2,t_1} + (1 - x_{i_1,t_1}) \\ WS_{i_3,t_2}(\boldsymbol{x}_{-i_3}) & = \sum_{j\in\{i_2\}} x_{j,t_1} + \sum_{j\in\{i_1\}} (1 - x_{j,t_1}) = x_{i_2,t_1} + (1 - x_{i_1,t_1}). \end{cases}$$

Fixing the prerequisite for (i_3, t_1), (i_1, t_2), (i_2, t_1), the weighed sums are reported in Table C.1.

$\mathrm{WS}_{i,t}(\boldsymbol{x}_i)$	i_1	i_2	i_3
t_1	$0 + (1 - 1)$	$1 + (1 - x_{i_1,t_1})$	$0 + (1 - x_{i_1,t_1})$
t_2	$x_{i_2,t_2} + (1 - x_{i_3,t_2})$	$x_{i_3,t_2} + (1 - 1)$	$x_{i_2,t_2} + (1 - 1)$

Table C.1: Weighed sum values of the six IF pairs (i_1, t_1), (i_2, t_1), (i_3, t_1), (i_1, t_2), (i_2, t_2), (i_3, t_2).

Using these weighed sum values, the thresholds $\theta_i^{0,*}$ and $\theta_i^{1,*}$ in Theorem 3 are the following:

$$\begin{cases} \theta_{i_1}^{0,*} & = \min_{\theta_{i_1}\in\{1,2\}} \theta_{i_1}(\ s.t.\ v_t(\boldsymbol{x}_{i_1}, \theta_{i_1}) = 0\ , \forall t \in \mathcal{P}_{i_1}^0) = 1 \\ \theta_{i_1}^{1,*} & = \max_{\theta_{i_1}\in\{1,2\}} \theta_{i_1}(\ s.t.\ v_t(\boldsymbol{x}_{i_1}, \theta_{i_1}) = 1\ , \forall t \in \mathcal{P}_{i_1}^1) \\ & = \max(1,\ x_{i_2,t_2} + (1 - x_{i_3,t_2}) \\ \theta_{i_2}^{0,*} & = \min_{\theta_{i_2}\in\{1,2\}} \theta_{i_2}(\ s.t.\ v_t(\boldsymbol{x}_{i_2}, \theta_{i_2}) = 0\ , \forall t \in \mathcal{P}_{i_2}^0) \\ & = \min(2,\ x_{i_3,t_1} + (1 - x_{i_1,t_1}) + 1) \\ \theta_{i_2}^{1,*} & = \max_{\theta_{i_2}\in\{1,2\}} \theta_{i_2}(\ s.t.\ v_t(\boldsymbol{x}_{i_2}, \theta_{i_2}) = 1\ , \forall t \in \mathcal{P}_{i_2}^1) = 2 \\ \theta_{i_3}^{0,*} & = \min_{\theta_{i_3}\in\{1,2\}} \theta_{i_3}(\ s.t.\ x_{i_2,t_1} + (1 - x_{i_1,t_1}) < \theta_{i_3}) \\ & = \min(2,\ x_{i_2,t_1} + (1 - x_{i_1,t_1}) + 1) \\ \theta_{i_3}^{1,*} & = \max_{\theta_{i_3}\in\{1,2\}} \theta_{i_3}(\ s.t.\ v_t(\boldsymbol{x}_{i_3}, \theta_{i_3}) = 1\ , \forall t \in \mathcal{P}_{i_3}^1) = 2 \end{cases}$$

These closed-form expressions for $\theta_i^{0,*}$ and $\theta_i^{1,*}$ highlight players sensitivity to the choices of their neighbors. Using these quantities, we obtain the following:

i_1*: To maximize i_1's utility, its optimal choice is $x(i_1, t_1) = 0$ as $WS_{i_1,t_1} = 0$ and $x(i_1, t_2) = 1$,* *since* $(i_1, t_2) \in \mathcal{P}^+$*. Therefore, we have*

$$\overline{u}_{i_1}(\boldsymbol{x}_{-i_1}) = \begin{cases} 2 & \text{if } x_{i_2,t_2} + (1 - x_{i_3,t_2}) \geq \theta_{i_1}, \\ 1 & \text{otherwise.} \end{cases}$$

In other words, i_1 *attains its highest score iff it is positively supported by* $\boldsymbol{x}_{-i_1}$*.*

i_2*: To maximize* i_2*'s utility, its optimal choice is* $x(i_2, t_1) = 1$ *as required by prerequisite, which forces its weighted sum* $WS_{i_3,t_1}(\boldsymbol{x}_{-i_2}) = 1 + (1 - x_{i_1,t_1})$ *to be at least* 1 *when* $\theta_{i_2} = 1$ *; or at least* 2 *when* $\theta_{i_2} = 2$*, in which case* $x_{i_1,t_1} = 0$ *is necessary. As for* t_2 *we have* $x(i_2, t_2) = 0$ *or* 1*, depending on* $WS_{i_2,t_2}(\boldsymbol{x}_{-i_2}) = x_{i_3,t_2}$*.Therefore, we have*

$$\overline{u}_{i_2}(\boldsymbol{x}_{-i_2}) = \begin{cases} 2 & \text{if } 1 + (1 - x_{i_1,t_1}) \geq \theta_{i_2}, \\ 1 & \text{otherwise.} \end{cases}$$

i_3*: To maximize* i_3*'s utility, its optimal choice is* $x(i_3, t_1) = 0$ *as required by prerequisite, which forces its weighted sum* $WS_{i_2,t_1}(\boldsymbol{x}_{-i_2}) = 0 + (1 - x_{i_1,t_1})$ *to be* 0 *when* $\theta_{i_3} = 1$ *in which case* $x_{i_1,t_1} = 0$ *is necessary.; or less or equal to* 1 *when* $\theta_{i_3} = 2$*, As for* t_2 *we have* $x(i_3, t_2) = 0$ *or* 1*, depending on* $WS_{i_2,t_2}(\boldsymbol{x}_{-i_2}) = x_{i_2,t_2}$*. Therefore, we have*

$$\overline{u}_{i_3}(\boldsymbol{x}_{-i_3}) = \begin{cases} 2 & \text{if } 1 - x_{i_1,t_1} \geq \theta_{i_3}, \\ 1 & \text{otherwise.} \end{cases}$$

Since $\theta_{i_1}, \theta_{i_2}, \theta_{i_3} \geq 1$*, it can be show that it does not exists* x_{i_1,t_1}*,* x_{i_2,t_2}*, and* x_{i_3,t_2}*, for which*

$$\begin{cases} x_{i_2,t_2} + (1 - x_{i_3,t_2}) & \geq 1 \\ 1 + (1 - x_{i_1,t_1}) & \geq 1 \\ 1 - x_{i_1,t_1} & \geq 1 \end{cases}$$

This is equivalent to case (ii) of Theorem 4, as the inequality $\theta_i^{0,*}(\boldsymbol{x}_{-i}) \leq \theta_i^{1,*}(\boldsymbol{x}_{-i})$ *does not hold for all* $i \in \mathcal{I}$*.*

Given the presented three-node fully connected signed weighted graph, there are four equilibria:

$$\begin{aligned} x_{i_1,t_1} &= 0, \quad x_{i_2,t_2} = 0, \quad x_{i_3,t_2} = 0, \\ x_{i_1,t_1} &= 0, \quad x_{i_2,t_2} = 1, \quad x_{i_3,t_2} = 0, \\ x_{i_1,t_1} &= 0, \quad x_{i_2,t_2} = 1, \quad x_{i_3,t_2} = 1, \\ x_{i_1,t_1} &= 0, \quad x_{i_2,t_2} = 0, \quad x_{i_3,t_2} = 1, \end{aligned}$$

all of them corresponding to the case $\overline{\overline{u}}(\mathcal{G}) < Fn$*.*

D : In-depth analysis of the country-by-country influences under geopolitical alignments (Subsection 6.1)

This appendix contains supplementary material to enrich the analysis of the country-by-country influences under geopolitical alignments presented in Subsection 6.1.

Attractive and repulsive links $\mathcal{E}^+$ and $\mathcal{E}^-$. The attractive and repulsive links $\mathcal{E}^+$ and $\mathcal{E}^-$ are constructed using the following data sources:

(i) UN general assembly resolutions data from January 1st 2022 to June 1st 2023;[17] This data set contains links to voting records for the general assembly resolutions in the United Nation digital library, which returns all adopted resolutions with or without vote, as well as a record for the vote of each country (either "yes", or "no", or "abstention", or "non-voting");

(ii) monetary, military, and bilateral trade agreements, from January 1st 2022 to June 1st 2023;

(iii) country-by-country bilateral imports from 2017 to 2023;[18] This data set contains statistical indicators related to WTO issues. Available time series cover merchandise trade and trade in services statistics (annual, quarterly and monthly), market access indicators (bound, applied and preferential tariffs), non-tariff information as well as other indicators.

We collected the voting results from the digital library of united nation votes. The selected topics are the dated from 1 January 2022 to 1 June 2023. There contains three types of resolutions : (i) resolutions where all of the 193 member countries are required to vote, (ii) resolutions where only some countries are required to vote, (iii) resolutions adopted without voting. Only case (i) is taken into account for the construction of this data set. Hence, out of four hundreds resolutions, only 91 topics are kept. For each topic, there can be four possible decisions : either "yes", or "no", or "abstention", or "non-voting".

To read the HTML file, filter and export the required information from the websites into spreadsheets, we use the Python library `request` and `BeautifulSoup`.[19]

We gathered data on the number of co-votes and contra-votes between countries. High numbers of co-votes indicate similar voting behaviors in the UN general assembly, suggesting an attractive relationship according to our model. Conversely, numerous contra-votes suggest a repulsive relationship. Specifically, we analyzed the distribution of co-votes and calculated percentiles to establish thresholds for determining whether the number of co-votes is relatively high or low.

---

[17] `https://research.un.org/en/docs/ga/quick/regular/77`

[18] `https://stats.wto.org/`

[19] The Python code used to reproduce this HTML reading, filtering and exporting is provided in the following Github directory: `https://github.com/StefanoNasini/UN-General-Assembly-2022-and-2023-voting`.

To define $L_3(i,j)$ we consider the following list of monetary and military alliances, as well as bilateral trade agreements, from January 1st 2022 to June 1st 2023:

The weight function W. The weight function W is constructed using the GDP data from World Bank.[20]

The prerequisite sets $\mathcal{P}^+$ and $\mathcal{P}^-$.

(v) Environmental Performance Index (EPI);[21] This data set provides summary indexes of the state of sustainability around the world as well as multiple environmental indicators, such as climate change performance, environmental health, and ecosystem vitality.

(vi) We use the LGBT equality index website[22] to determine the countries in which homosexuality is sanctioned by death penalty,[23] as well as the countries in which same-sex marriage existed for more than 15 years and children adoption is allowed to homosexual couples.[24]

(vii) Migrant integration policy index. [25] This data set contains measurements regarding the access to nationality, the anti-discrimination laws, the family reunion, the labor market mobility, among others.

The processed data required to construct L_1, L_2, L_3, L_4 and L_5 are summarized in Table D.1, for the upper percentiles $\overline{q}_1$, $\overline{q}_2$ and $\overline{q}_4$, and in Table D.2, for the lower percentiles, $\underline{q}_1$, $\underline{q}_2$ and $\underline{q}_4$, where two different percentile levels have been reported.

Figure 5 provides a graphical illustration of the largest components of the network of attractive and repulsive neighbors, constructed in accordance with the following procedure:

- For $\mathcal{E}^+$: $[\mathbf{1}(L_1 > \overline{q}_1) + \mathbf{1}(L_2 < \underline{q}_2) + \mathbf{1}(L_3 \geq 1) + \mathbf{1}(L_4 > \overline{q}_4) + \mathbf{1}(L_5 > \overline{q}_5)] \geq 3$;
- For $\mathcal{E}^-$: $[\mathbf{1}(L_1 < \underline{q}_1) + \mathbf{1}(L_2 > \overline{q}_2) + \mathbf{1}(L_3 = 0) + \mathbf{1}(L_4 < \underline{q}_4) + \mathbf{1}(L_5 < \underline{q}_5)] \geq 3$.

Two networks are generated setting two different percentile levels $\overline{q}_1$, $\overline{q}_2$ and $\overline{q}_4$: 80% and 90% percentiles, with $\underline{q}_1 = 100 - \overline{q}_1$, $\underline{q}_2 = 100 - \overline{q}_2$ and $\underline{q}_4 = 100 - \overline{q}_4$.

[20] https://databankfiles.worldbank.org/public/ddpext_download/GDP.pdf

[21] https://epi.yale.edu/

[22] https://www.equaldex.com/equality-index

[23] Afghanistan, United Arab Emirates, Burundi, Bangladesh, Brunei Darussalam, Cameroon, Comoros, Dominica, Algeria, Eritrea, Ethiopia, Guinea, Gambia, Iran, Kenya, Liberia, Libya, Sri Lanka, Morocco, Maldives, Mauritania, Malawi, Malaysia, Oman, Pakistan, Qatar, Saudi arabia, Sudan, Senegal, Solomon Islands, Somalia, Chad, Togo, Tunisia, United Republic of Tanzania, Uganda, Uzbekistan, Saint Vincent, Yemen, Zambia.

[24] Argentina, Australia, Austria, Belgium, Brazil, Canada, Switzerland, Chile, Colombia, Costa rica, Cuba, Germany, Denmark, Spain, Finland, France, United kingdom, Ireland, Iceland, Luxembourg, Malta, Netherlands, Norway, New zealand, Portugal, Slovenia, Sweden, Uruguay, United states, South africa.

[25] https://www.mipex.eu/

	Above 80-percentile				Above 90-percentile				
	covotes	contravotes	exports	imports	covotes	contravotes	exports	imports	alliances
singletons	46	5	0	0	71	8	0	0	55
edges	3847	7356	5806	5864	2109	4120	3090	3123	1934
degr. max	91 (PER)	181 (ISR)	183 (ARE)	183 (ARE)	75 (PER)	170 (ISR)	183 (GBR)	183 (GBR)	62 (BEL)
degr. min	0 (ARM)	0 (COM)	21 (LSO)	23 (TON)	0 (AFG)	0 (COM)	12 (GUY)	10 (PRK)	0 (ATG)

Table D.1: Networks generated by fixing a percentile levels in co-votes, contra-votes, export and import trade flow percentages, as well as by the existence of an alliance. If co-votes, contra-votes, and trade flow percentages are above the 80 or 90 percentiles, than an edge is included. The countries corresponding to the trigrams are: PER for Peru, ISR for Israel, ARE for united arab emirates, GBR for united kingdom, ARM for Armenia, COM for Comoros, LSO for Lesotho, TON for Tonga, AFG for Afghanistan, GUY for Guyana, PRK for North Korea, ATG for Antigua and barbuda.

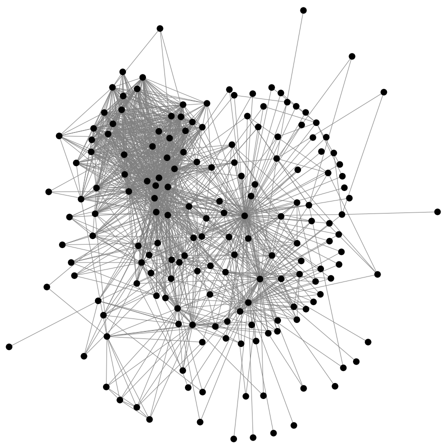
	Below 20-percentile				Below 10-percentile				
	covotes	contravotes	exports	imports	covotes	contravotes	exports	imports	allinaces
singletons	0	0	0	0	0	0	0	0	0
edges	4053	13125	6207	6505	2030	13125	3350	3485	16788
degr. max	194 (COM)	194 (AFG)	183 (TUV)	182 (TUV)	194 (COM)	194 (AFG)	179 (TUV)	177 (TUV)	194 (ATG)
degr. min	12 (CRI)	70 (FSM)	16 (NGA)	14 (NGA)	8 (ARG)	70 (FSM)	6 (ETH)	7 (CIV)	132 (BEL)

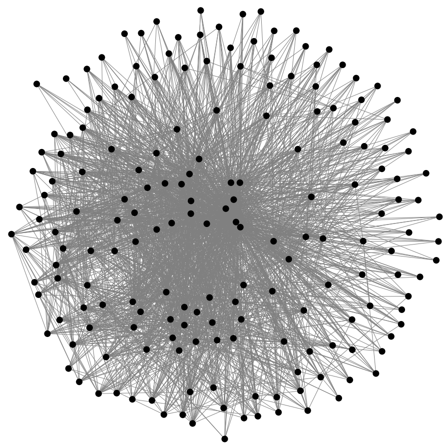
Table D.2: Networks generated by fixing a percentile levels in co-votes, contra-votes, trade flow percentages, as well as by the existence of an alliance. If co-votes, contra-votes, and trade flow percentages are above the 75 or 90 percentiles, than an edge is included. The countries corresponding to the trigrams are: COM for Comoros, AFG for Afghanistan, TUV for Tuvalu, ATG for Antigua and barbuda, CRI for Costa Rica, FSM for Federated States of Micronesia, NGA for Nigeria, ARG for Argentina, ETH for Ethiopia, CIV for Cote d'Ivoire and BEL for Belgium.

Figure 6 provides a graphical illustration of the largest components of the weighted network of attractive and repulsive neighbors, as defined in (12).

Table D.3 provides the objective function, CPU time and linear relaxation value for 48 instances (namely 24 instances for the social welfare problem (7) plus 24 instances for the Nash equilibrium problem (8)) of the GVG-CNC model.

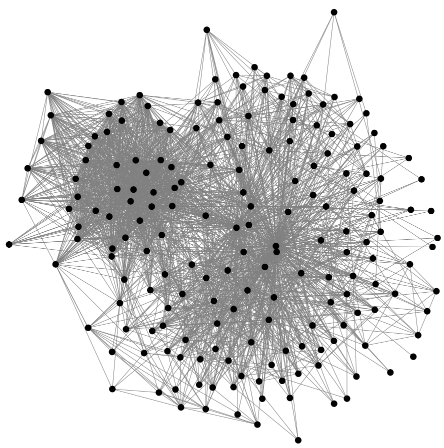
The first insight from this results is that in most analyzed instances the Nash equilmibrium solution differs from the social welfare. This is particularly evident for parametrizations in which the range of feasible threshold levels is shrank as $\Theta_i = \{(1 + \lfloor 0.3(d_i - 1)/2 \rfloor), \ldots, (d_i - \lceil 0.3(d_i - 1)/2 \rceil)\}$. Secondly, the LP relaxations of problem (7) coincides with the one of problem (8), entailing that the system of constraints (9a)-(9k) and (10a)-(10c) is satisfied by the optimal solution of the LP relaxation of problem (7). Thirdly, the computation time required to characterize the social welfare solution ranges between 6.02 seconds and 431.36 seconds, whereas the one required to characterize the Nash equilibrium solution rages from 9.6 seconds to 194.49. These computational times are mostly aligned with no clear pattern in relation with the specific GVG-CNC parametrization.

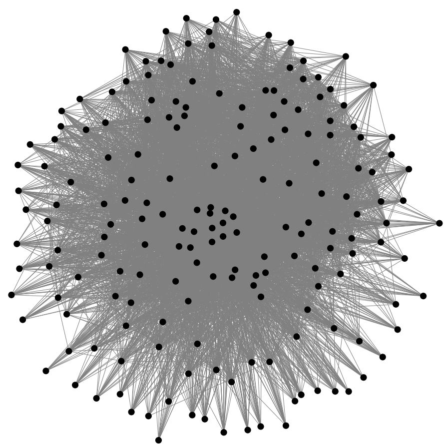




(a) The largest component of the friends 90 network, containing 174 countries and 1280 links.

(b) The largest component of the enemies 90 network, containing 184 countries and 2367 links.

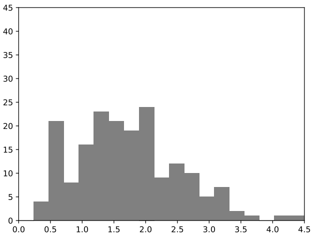




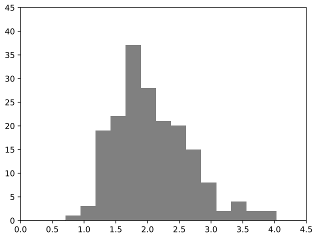
(c) The largest component of the friends 80 network, containing 184 countries and 2437 links.

(d) The largest component of the enemies 80 network, containing 184 countries and 5653 links.

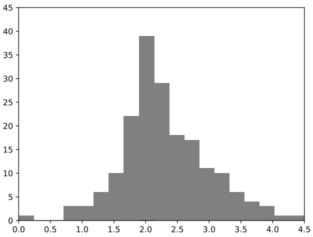
Figure 5: The largest component of the friends network (left panel) and the enemies network (right panel).



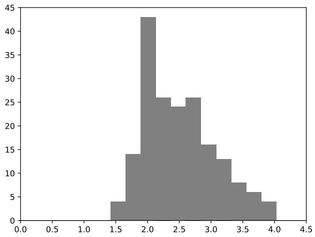
(a) in-degree for the 90-percentile network.



(b) out-degree for the 90-percentile network.



(c) in-degree for the 80-percentile network.



(d) out-degree for the 80-percentile network..

Figure 6: The distribution of weighted in-degree and weighted out-degrees for each nodes in the 80-percentile and 90-percentile network respectively. The x-axis is reported with base-10 logarithmic scale. We can notice a general trend of having fewer weighted degree in 90-percentile network, matching the fact that the criteria for being an edge in 90 percentile is more demanding. In addition, we notice all logarithmic distribution of the weighted degrees are bell-shaped, consistently with the power law distribution of multiple historically observed social network data.

			social welfare (problem (7))			Nash equilibrium (problem (8))		
κ	Percentile	ρ	Obj. fun.	CPU time	LR	Obj. fun.	CPU time	LR
1.0	80	0.1	1	2.78	0.26	1	8.32	0.26
1.0	80	0.5	1	3.85	0.06	1	4.96	0.06
1.0	80	1.0	1	60.64	0.0	1	16.27	0.0
1.0	90	0.1	2	3.3	1.12	2	1.06	1.12
1.0	90	0.5	2	3.1	0.85	2	5.18	0.85
1.0	90	1.0	2	3.62	0.59	2	0.94	0.59
0.7	80	0.1	1	15.0	0.26	1	15.94	0.26
0.7	80	0.5	1	17.25	0.06	1	4.76	0.06
0.7	80	1.0	1	85.68	0.0	1	16.77	0.0
0.7	90	0.1	5	4.29	1.16	5	1.63	1.16
0.7	90	0.5	3	65.84	0.91	3	3.94	0.91
0.7	90	1.0	3	16.2	0.66	3	5.74	0.66
0.8	80	0.1	1	8.53	0.26	1	9.01	0.26
0.8	80	0.5	1	8.96	0.06	1	5.81	0.06
0.8	80	1.0	1	59.92	0.0	1	22.07	0.0
0.8	90	0.1	2	9.8	1.12	2	4.14	1.12
0.8	90	0.5	2	6.27	0.87	2	3.81	0.87
0.8	90	1.0	2	11.74	0.61	2	8.07	0.61
0.9	80	0.1	1	9.48	0.26	1	8.78	0.26
0.9	80	0.5	1	7.6	0.06	1	4.11	0.06
0.9	80	1.0	1	56.02	0.0	1	18.64	0.0
0.9	90	0.1	2	3.77	1.12	2	1.9	1.12
0.9	90	0.5	2	6.21	0.85	2	3.73	0.85
0.9	90	1.0	2	9.06	0.59	2	3.59	0.59

Table D.3: social welfare and Nash equilibrium solutions under different parametrization of the DVG-CNC, for the country-by-country geopolitical alignments application. From left to right, columns represent: the shrinkage threshold interval κ, the network topology (generated either with 80% or 90% percentile), the weighted function parameter ρ, the optimal objective function of problem (7), the corresponding CPU time (in second) and value of the LP relaxation, the optimal objective function of problem (8), the corresponding CPU time (in second) and value of the LP relaxation.

			social welfare (problem (7))			Nash equilibrium (problem (8))			
κ	Per	ρ	$t = 1(23)$	$t = 2(52)$	$t = 3(56)$	$t = 1(23)$	$t = 2(52)$	$t = 3(56)$	SW-NE
1.0	80	0.1	1.00	0.83	0.88	0.95	0.69	0.79	
1.0	80	0.5	0.70	0.71	0.67	0.70	0.65	0.63	
1.0	80	1.0	0.70	0.81	0.78	0.65	0.77	0.72	
1.0	90	0.1	0.70	0.79	0.76	0.75	0.69	0.63	*
1.0	90	0.5	0.90	0.83	0.83	0.45	0.58	0.48	
1.0	90	1.0	0.70	0.81	0.76	0.45	0.65	0.59	
0.7	80	0.1	0.65	0.67	0.75	0.70	0.75	0.71	**
0.7	80	0.5	0.75	0.69	0.76	0.85	0.75	0.80	***
0.7	80	1.0	0.65	0.69	0.67	0.65	0.71	0.75	**
0.7	90	0.1	0.65	0.69	0.64	0.65	0.60	0.56	
0.7	90	0.5	0.65	0.73	0.65	0.50	0.56	0.48	
0.7	90	1.0	0.70	0.71	0.73	0.65	0.65	0.61	
0.8	80	0.1	0.70	0.81	0.75	0.95	0.71	0.82	**
0.8	80	0.5	0.90	0.77	0.73	0.95	0.83	0.83	***
0.8	80	1.0	0.80	0.67	0.77	0.75	0.77	0.74	*
0.8	90	0.1	0.75	0.73	0.77	0.80	0.65	0.64	*
0.8	90	0.5	0.90	0.75	0.73	0.60	0.69	0.61	
0.8	90	1.0	0.75	0.73	0.71	0.90	0.69	0.63	*
0.9	80	0.10	0.9	0.75	0.77	0.85	0.75	0.79	*
0.9	80	0.50	0.85	0.77	0.78	0.70	0.67	0.70	
0.9	80	1.0	0.75	0.77	0.74	0.75	0.75	0.76	*
0.9	90	0.10	0.9	0.81	0.8	0.80	0.67	0.64	
0.9	90	0.50	0.75	0.69	0.7	0.50	0.42	0.40	
0.9	90	1.0	0.70	0.77	0.75	0.70	0.71	0.74	

Table D.4: Percentage of matches between observed choice and choice predicted by the DVG-CNC, with intermediate level countries filtered out at %70 percentile.