A New Empirical Bayes Estimation for Network Peer Effect Model

Siao Xu*†

September 8, 2026

Abstract

We propose a novel network-based peer effect model with unobserved unit-specific heterogeneity, in which the regression structure is augmented by a probabilistic graph that characterizes the latent dependence mechanism among all components of the model. This probabilistic graph, in turn, provides the foundation for the empirical Bayes estimator developed in the paper. The group structure embedded in the unit heterogeneity arises from a group prior, which we introduce to capture the latent heterogeneity across units. Although a least squares estimator exists for the proposed model, it performs poorly in the presence of sparsity in network. To address these challenges, we develop a machine learning–assisted empirical Bayes estimator. Building on this estimation framework and the proposed group prior, we further introduce a methodology for learning latent group structures in regression models. Simulation results demonstrate that the proposed empirical Bayes estimator substantially outperforms least squares based alternatives. Finally, we apply our framework to an empirical analysis of democratic spillovers across countries.

Keywords: random effect, empirical Bayes estimation, structural causal model, probabilistic graph, Bayesian belief network, spectral clustering, community detection.

JEL Codes: C18, C33, C38, C46, C51

*I am deeply indebted to my advisor Carsten Trenkler for his long-term support. I am also very grateful to Christopher Rothe, Mengshan Xu, Endong Wang for their very insightful suggestions and discussions. All errors are my own.

†Graduate School of Economic and Social Sciences, University of Mannheim.
E-mail: siao.xu@students.uni-mannheim.de

1 Introduction

Since the seminal contribution of Manski (1993), peer-effects models have become one of the central frameworks in empirical economics. The core idea is that units are influenced, on average, by other units within the same reference group, while interactions across groups are typically excluded. A subsequent strand of the literature extends this framework to social-network settings, in which units are influenced by others through observed network links. Calvó-Armengol et al. (2009) establishes a theoretical foundation for peer effects in networks, where peer influence is assumed to arise through links between nodes. Meanwhile, Bramoullé et al. (2009); Goldsmith-Pinkham and Imbens (2013); Boucher et al. (2014) formalize reference groups in network settings and propose network-based peer-effects models. These models are based on the assumption that peers' outcomes affect an individual's own outcome. However, in many settings, the observed correlation between the outcomes of connected units may instead arise from shared unobserved environments. For example, the strong performance of two students i and j may reflect their common teacher, curriculum, school quality, or neighborhood rather than causal peer influence. Similarly, the stock prices of connected firms may respond simultaneously to common shocks, such as oil-price shocks. In such cases, conventional peer-effects models may fail to capture the true underlying mechanism. To address this issue, we propose a novel heterogeneity-based network exposure model in which units are affected by network exposure to their peers' latent heterogeneity rather than by causal peer outcome influence. The model is motivated by recent developments in network fixed-effects models (Jochmans and Weidner, 2019; Cheng et al., 2025).

Broadly, this paper develops a linear regression framework that incorporates a latent generative mechanism through a probabilistic graphical model. The proposed model is a network-based peer-effects regression with random effects, in which each unit is exposed to the latent heterogeneity of other units through a block-structured social network. Network formation is assumed to be governed by latent group membership generated by unobserved group-level heterogeneity (Bonhomme and Manresa, 2015; Bester and Hansen, 2016). In contrast to traditional grouped fixed-effects models, which assign units in the same group a common fixed effect, our approach introduces a group prior that explicitly models latent group structure into distribution parameter. We embed these components in a structural causal model assisted with a directed acyclic graph, thereby obtaining a coherent representation of the latent data-generating process. Second, we

update structural causal model to Bayesian network by writing structural functions into conditional distributions. The resulting joint distribution specified by Bayesian network naturally yields a novel empirical Bayes estimator for the proposed regression model. We then use this estimator to develop a method for learning latent group structure in panel data.

The first contribution of this paper is to develop a network-based peer-effects model with unit-specific heterogeneity. Traditional peer-effects models (Manski, 1993; Duflo and Saez, 2002) assume that peer influence arises from interactions among individuals within a predefined reference group. Subsequent studies extend this framework to network settings, in which peer effects operate through observed links in a social network (Calvó-Armengol et al., 2009; Goldsmith-Pinkham and Imbens, 2013). Because peer effects are typically specified as linear combinations of peers' outcomes and characteristics, these models face severe identification challenges, most notably the reflection problem. We propose a new class of peer-effects models in which peer influence is captured by exposure to latent heterogeneity transmitted through the network. By avoiding the direct inclusion of peers' outcomes and characteristics in the regression equation, this formulation partially mitigates the identification problem. Our approach is motivated by recent advances in network regression models. In particular, Jochmans and Weidner (2019) propose a two-way fixed-effects model for bipartite network data, in which fixed effects from one group affect units in another group through network links, while links within groups are excluded. A canonical example is the effect of teachers on students' performance. In contrast, our model relaxes the bipartite-network restriction and allows peer influence to operate through general network links, thereby accommodating interactions among units within the same population.

The second contribution of this paper is to introduce a novel group prior for modeling group structure in regression models. Learning or exploiting group structure in econometric models has attracted substantial attention over the past decade (Bonhomme and Manresa, 2015; Su et al., 2016; Chetverikov and Manresa, 2022; Mugnier, 2025; Ando and Bai, 2016; Bonhomme et al., 2022, 2019). This literature typically assumes that group structure is reflected in slope coefficients, fixed effects, latent factors, or other parameters of interest. A common restriction is that units are heterogeneous across groups but homogeneous within groups. Although analytically convenient, this restriction can be implausible in empirical applications. For example, Acemoglu et al. (2008) study the relationship between income and democracy, and Bonhomme and Manresa (2015) revisit this question using a grouped fixed-effects panel model. In

their framework, countries classified into different democracy levels are assumed to share common group-specific fixed effects. This specification implies that all highly democratic countries have the same fixed effect, while all less democratic countries share another fixed effect. Such within-group homogeneity is restrictive. A more realistic specification would allow for meaningful heterogeneity even among countries assigned to the same group. For example, two highly democratic countries may have fixed effects $\alpha_1 = 1.1$ and $\alpha_2 = 0.9$, while two less democratic countries may have fixed effects $\alpha_3 = 0.10$ and $\alpha_4 = 0.02$, rather than imposing $\alpha_1 = \alpha_2 = 1.0$ and $\alpha_3 = \alpha_4 = 0$. To accommodate within-group heterogeneity while preserving group structure, we propose a novel group prior. In contrast to the standard grouped-regression framework, which imposes $\alpha_i = \sum_{g=1}^{G} \mathbb{I}\{g_i = g\}\alpha_g$ where $\mathbb{I}\{\cdot\}$ denotes the indicator function, we assume instead that α_i is drawn from a group-specific distribution: $\alpha_i \sim F\Big(\lambda_{g_i}\Big)$ or equivalently, $\alpha_i \sim F\Big(\sum_{g=1}^{G} \mathbb{I}\{g_i = g\}\lambda_g\Big)$, where λ_g denotes the distributional parameters associated with group g. For example, under a Gaussian group prior, $\alpha_i \sim \mathcal{N}\Big(\mu_{g_i}, \sigma^2_{g_i}\Big)$, or equivalently, $\alpha_i \sim \mathcal{N}\Big(\sum_{g=1}^{G} \mathbb{I}\{g_i = g\}\mu_g, \sum_{g=1}^{G} \mathbb{I}\{g_i = g\}\sigma^2_g\Big)$. This formulation preserves the central idea of grouped-regression models while relaxing the restrictive assumption of within-group homogeneity. It allows units in the same group to share a common distributional structure, rather than a common fixed-effect value.

The third contribution of this paper is to integrate structural causal models and Bayesian networks into a regression framework. Structural causal models provide a formal mathematical language for representing causal relationships among variables through graph-based structures (Pearl, 2009). By specifying conditional distributions associated with the structural relationships, such models can be naturally linked to Bayesian belief networks (Koller, 2009). In this paper, we use a directed acyclic graph to represent the latent mechanism underlying the proposed peer-effects model.

A Bayesian belief network is a directed acyclic graph (DAG) that encodes conditional dependence relationships among random variables. In our framework, the observed network structure is generated by latent group membership and auxiliary parameters, while the unit-specific effects are drawn from the group prior introduced above. We further allow for a latent dependence between the group structure governing network formation and the group structure underlying unit-specific heterogeneity. Together, these components define a latent data-generating process that is naturally represented by a Bayesian belief network. This representation provides a coherent probabilistic framework for characterizing the dependencies among the model's latent

and observed variables. Following the setup in Cheng et al. (2025), the covariates X and their associated coefficients β are treated as hyperparameters. The joint distribution implied by the Bayesian belief network then provides the basis for the empirical Bayes estimator developed in this paper, which constitutes our next contribution.

The fourth contribution of this paper is to develop a novel machine-learning-assisted empirical Bayes (EB) estimator. Although the proposed model admits a least-squares estimator, its finite-sample performance is often unsatisfactory. In particular, the effectiveness of least squares depends critically on the density and structural properties of the underlying network. When the network is sufficiently dense and exhibits no group structure, least squares can perform well, as discussed in Jochmans and Weidner (2019). However, such dense and unstructured networks are uncommon in empirical applications.

To address this limitation, Cheng et al. (2025) propose an empirical Bayes estimator for the network fixed-effects model of Jochmans and Weidner (2019), using shrinkage to trade a small amount of bias for variance reduction. In this paper, we relax two key assumptions underlying their framework by allowing the network to be sparse. Under these conditions, the least-squares estimator remains well defined but performs poorly, motivating an empirical Bayes approach. Our empirical Bayes estimator is derived from the proposed regression model together with a probabilistic graphical representation of the latent data-generating mechanism. The estimator uses machine learning methods to learn the relevant distributional components. Simulation results show that the proposed machine-learning-assisted empirical Bayes estimator substantially outperforms the corresponding least-squares estimator.

The final contribution of this paper is to develop a methodology for learning latent group structure in regression models using the proposed empirical Bayes estimator. The identification and estimation of group structure in econometric models have attracted increasing attention in recent years (Brownlees et al., 2022; Su et al., 2016; Yu et al., 2024). Our approach proceeds in two stages. In the first stage, we recover an initial group structure from the observed adjacency matrix A, treating the network as given.

Importantly, block formation in networks may be driven by physical constraints or interaction frictions rather than by the economic or behavioral heterogeneity of interest. Consequently, units belonging to the same latent type may be split across multiple smaller network-based groups. For example, in dormitory-assignment settings (Sacerdote, 2001; Carrell et al., 2009), students who smoke may be assigned to smoking apartments, while non-smoking students may be assigned to

non-smoking apartments. Because of room-capacity constraints, however, each population may be further divided into many small dormitory units, producing a network with multiple small groups. Although these groups are distinct in the observed network, they may correspond to a much smaller number of latent behavioral types. Using the proposed machine-learning-assisted empirical Bayes estimator, we estimate a parameter vector for each network-based group. We then stack these estimated group-level parameters and apply a clustering procedure. Under the assumption that groups associated with the same latent type have similar parameter values, we use k-means clustering to aggregate small network-based groups into a smaller number of economically meaningful groups, such as smoking and non-smoking students. This procedure addresses a limitation of traditional network-based group-detection methods, which may identify groups induced by physical constraints rather than by economically relevant heterogeneity. It therefore provides a new framework for learning latent group structure in regression models.

We first present the main methodological components of the paper, including the proposed model, the estimation procedure, and the theoretical analysis. In the simulation study, unit-specific effects are generated from the proposed group prior under a Gaussian specification, so that the latent group structure is reflected in their distribution. We then generate the network using a degree-corrected stochastic block model (Karrer and Newman, 2011) and simulate the outcome data conditional on the generated network and unit-specific effects. The results show that the proposed machine-learning-assisted empirical Bayes estimator consistently outperforms the least-squares estimator in finite samples.

In the empirical application, we apply our machine-learning-assisted empirical Bayes estimator to study the relationship between income and democracy. This relationship was first examined by Acemoglu et al. (2008) in a standard panel-data framework and later revisited by Bonhomme and Manresa (2015) using a grouped panel model. We revisit this question through the lens of our proposed framework, with the primary objective of identifying democratic spillovers across countries. We construct the network using international alliance groups. Our findings indicate that groups composed of countries with diverse cultural and historical backgrounds exhibit greater cross-sectional heterogeneity. This heterogeneity declines over time, particularly after the end of the Cold War and the acceleration of globalization. These patterns suggest that, as political barriers weaken and international interactions intensify, democratic norms and institutions increasingly diffuse across countries.

This paper builds on several strands of the literature. First, it contributes to the literature on

peer effects models, including Manski (1993), Calvó-Armengol et al. (2009), Bramoullé et al. (2009), Jochmans (2023), and Boucher et al. (2014). Second, it relates to the literature on econometric models with group structure, such as Su et al. (2016), Bonhomme and Manresa (2015), Guðmundsson and Brownlees (2021), and Zhang et al. (2019). Third, the paper builds on the empirical Bayes methodology, drawing on contributions including Robbins (1992), Efron (2012), Ignatiadis and Wager (2019), and Efron (2019). In addition, our framework incorporates probabilistic graphical models, specifically Bayesian belief networks, see e.g., Blei et al. (2003), Munro and Ng (2022), Koller (2009), and Borgelt et al. (2009). Finally, the paper is related to the growing literature on regression models for network data, including Zhu et al. (2017), Auerbach (2022), and Jochmans and Weidner (2019).

The rest of the paper is organized as follows. Section 2 presents the main model. Section 3 is devoted to estimation method. Section 4 is about theory. Section 5 and 6 are left to simulated and empirical studies, respectively. Finally, Section 7 concludes. Additional materials and proofs are provided in the Appendix and supplementary materials.

1.1 Notation and the definition for Graph

Let N denote the number of time series and T represent the number of observations. We define N_0 as the number of time series with pure memberships and $N - N_0$ as the number of time series with mixed membership. Lowercase letters, such as y, denote scalars, while capital letters, like Y, represent matrices. For any matrix Y, we use $y_{i\cdot}$, $y_{\cdot j}$ and $y_{i,j}$ to denote its i-th row, j-th column and ij-th entry, respectively. For an arbitrary square matrix $Y \in \mathbb{R}^{n\times n}$, $\lambda_i(Y)$ refers to the i-th eigenvalue of Y with $|\lambda_1(Y)| \geq |\lambda_2(Y)| \geq \cdots \geq |\lambda_n(Y)|$. The spectral radius of a square matrix Y is defined as $\rho(Y) = |\lambda_1(Y)|$. For a symmetric matrix Y, $\lambda_{max}(Y)$ and $\lambda_{min}(Y)$ denote the largest and smallest eigenvalues respectively. The following matrix norms are used: ℓ_2-norm: $\|Y\|$ and Frobenius norm: $\|Y\|_F = Tr(Y'Y)$. $\mathbb{I}$ stands for indicator function.

A graph is defined as a triple: $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{W})$ where $\mathcal{V} = 1, \cdots, N$ stands for the vertices of the graph, $\mathcal{E} \in \mathcal{V} \times \mathcal{V}$ the set of edges, and, accordingly, $\mathcal{W} \in \mathbb{R}$ the weights of all edges if necessary. A is the adjacency matrix which is the matrix representation of the graph. In this work, I assume that the adjacency matrix A is a $N \times N$ matrix where the elements $[A]_{ij} = w_{ij} \neq 0$ if there exists an edge connecting node j and node i and $[A]_{ij} = w_{ij} = 0$ otherwise.

2 Network Fixed Effect Regression Model

2.1 Main Model

In this paper, we propose a novel network-based peer-effects model with unit-specific heterogeneity. The model assumes that each unit is affected by network exposure to the latent heterogeneity of other units, rather than by peers' realized outcomes. The main specification is given by:

$$Y = \boldsymbol{\alpha} I_N + A\boldsymbol{\alpha} + X^{'}\beta + \epsilon, \tag{1}$$

where $Y = (y_1, y_2, \cdots, y_N)^{'}$, $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_N)^{'}$, I_N denotes the $N \times N$ identity matrix, and A is a row-normalized adjacency matrix. X denotes a vector of control variables, and ϵ is an $i.i.d.$ error term following a multivariate normal distribution. Combining the terms $\boldsymbol{\alpha} I_N$ and $A\boldsymbol{\alpha}$, equation (1) can be rewritten as follows:

$$Y = A^{*}\boldsymbol{\alpha} + X^{'}\beta + \epsilon, \ where \ A^{*} = A + I_N. \tag{2}$$

Moreover, in univariate form, equation (1) can be rewritten as

$$y_i = \alpha_i + N_i^{-1} \sum_{j \neq i}^{N} a_{i,j}\alpha_j + X_i^{'}\beta + \epsilon_i, \tag{3}$$

where $N_i = \sum_{j \neq i}^{N} a_{i,j}$ denotes the in-degree of unit i, and α_i represents unit-specific heterogeneity. The term $N_i^{-1} \sum_{j \neq i}^{N} a_{i,j}\alpha_j$ captures the average exposure of unit i to the latent heterogeneity of its neighbors. X_i denotes a vector of covariates, and β represents the associated slope coefficients. The error term ϵ_i is assumed to be independent of X_i and to follow a standard normal distribution. Although allowing for correlated error terms would be more flexible, doing so would introduce a large number of additional parameters. To reduce dimensionality, we therefore impose a diagonal covariance structure.

In addition to the main specification, a key novelty of our framework is its augmentation with a directed acyclic graph (DAG) that represents the latent data-generating mechanism. The DAG allows the proposed model and its latent generative process to be expressed as a system of structural functions. This construction follows the logic of structural causal models, which encode causal relationships among variables through directed acyclic graphs (Pearl, 2009). Figure 1 illustrates the proposed generative process and the associated latent relationships among the

model components. We assume that the unit-specific heterogeneity α_i is drawn from a prior distribution governed by a group-level hyperparameter λ_g. To formalize this mechanism, we introduce a novel group prior. Details of the group prior are provided in the next subsection. At this stage, it suffices to note that the group prior induces latent group structure in $\boldsymbol{\alpha}$. For example, consider a Gaussian prior with two latent groups, one with mean 10 and the other with mean -10. Units assigned to the first group have heterogeneity parameters concentrated around 10, whereas those assigned to the second group have heterogeneity parameters concentrated around -10.

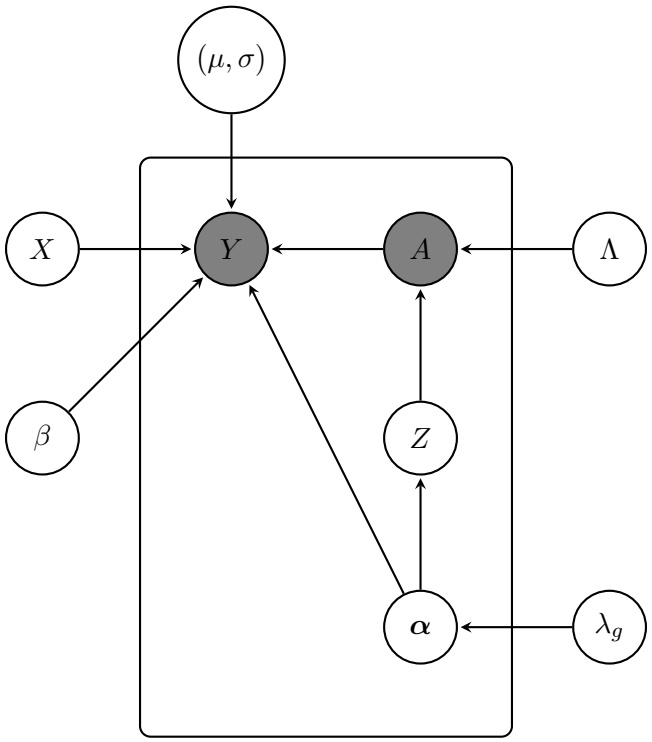


Figure 1: Graphical Representation of the Model in equation (1). Circles outside the plates represent hyperparameters, while circles inside the plates denote latent or observed variables. Shaded (Grey) circles indicate observed quantities. The adjacency matrix A and the dependent variable Y are observed. The set of control variables X and their associated slope coefficients β are treated as hyperparameters.

Next, conditional on $\boldsymbol{\alpha}$, we infer a latent group structure Z. For each unit $i = 1, \cdots, N$, Z_i is represented as a one-hot membership vector, with one entry equal to one and all other entries equal to zero. The vector Z_i therefore encodes the latent group membership of unit i. For example, in a three-group setting, $Z_i = [0, 1, 0]$ indicates that unit i is assigned to group 2.

The group memberships Z and hyperparameters Λ jointly give rise to a network A with latent group structure. A canonical example is the degree-corrected stochastic block model, in which $\Lambda = [B, \theta]$, where B denotes the group-level connection probability matrix and θ captures

unit-specific degree heterogeneity. Importantly, we do not impose a specific parametric model on A. Instead, we assume that the network is generated jointly with the group assignments Z, under the assumption that units belonging to the same group are more likely to interact than units belonging to different groups. Consequently, the structure of A depends on the latent membership structure Z as well as other governing parameters.

Finally, conditional on the network A and the heterogeneity vector $\boldsymbol{\alpha}$, we treat the covariates X and their associated slope coefficients β as given hyperparameters. We further assume that the error term ϵ follows a multivariate normal distribution with mean $\mu\mathbf{1}$ and variance $\sigma^2 I_N$, where μ and σ are also treated as hyperparameters. It follows that $Y \sim \mathcal{N}\Big(\mu + X^{'}\beta + AI + A\boldsymbol{\alpha}, \sigma^2\Big)$ since $\epsilon \sim \mathcal{N}\Big(\mu, \sigma^2\Big)$.

The latent mechanism underlying the proposed model can be represented by a directed acyclic graph (DAG) with a set of structural functions. The set of functions can be rewritten into a sequence of conditional distributions. Then, DAG accordingly is updated to probabilistic graphical model or, more specifically, a Bayesian belief network. This representation offers several advantages. First, it provides a clear and concise description of the underlying data-generating process, rendering the structural assumptions of the regression model transparent and easy to interpret. Second, the framework is highly flexible and can be readily adapted to alternative model specifications and empirical settings. Third, the probabilistic graphical formulation naturally facilitates Bayesian estimation, which is particularly advantageous in high-dimensional environments. Details of the Bayesian estimation procedure are discussed in the estimation section.

2.2 Distribution Specification

The joint distribution and the associated probabilistic graphical representation are equivalent characterizations of the same model. Accordingly, the probabilistic graph shown in Figure 1 can be expressed analytically using equation (4). Following the principles of Bayesian belief networks (Koller, 2009), equation (4) admits the following factorization of the joint distribution:

$$\begin{aligned}&\mathbb{P}\Big(Y, A, Z, \boldsymbol{\alpha} | X, \beta, \lambda, \Lambda, \mu, \sigma\Big) \\ &= \mathbb{P}\Big(Y | X, \beta, \boldsymbol{\alpha}, A, \mu, \sigma\Big)\mathbb{P}\Big(A | Z, \Lambda\Big)\mathbb{P}\Big(Z | \boldsymbol{\alpha}\Big)\mathbb{P}\Big(\boldsymbol{\alpha} | \lambda\Big)\end{aligned} \tag{4}$$

This joint distribution forms the basis for empirical Bayes estimation, as the posterior distri-

bution is proportional to the joint distribution and can be factorized into a collection of lower-dimensional conditional distributions. The empirical Bayes estimator is defined as the posterior conditional expectation of the unit-level heterogeneity after integrating out the latent group membership Z. The estimator can be computed from the joint distribution in equation (4), which in turn is implied by the Bayesian belief network depicted in Figure 1. Details of the estimation procedure are provided in the estimation section.

Remark 2.1. *The decomposition of* $\mathbb{P}\Big(Y, A, Z, \boldsymbol{\alpha}|X, \beta, \lambda, \Lambda, \mu, \sigma\Big)$*, as shown in equation (4), proceeds in two steps. First, by applying the chain rule of probability (Bayes' law), the joint distribution can be expressed as a product of conditional distributions. Second, several of these conditional distributions can be further simplified by leveraging conditional independence relationships among variables. These independence assumptions are encoded in the Bayesian network illustrated in Figure 1. By systematically applying these two steps—factorization and simplification via independence—we arrive at the representation given in Equation (4).*

The distributional assumptions of the proposed model are formally specified as follows:

Definition 1. *Given the the model as defined in equation (1) assisted with the probabilistic graph as defined in figure 1, the component distributions are defined as follows:*

$$\boldsymbol{\alpha}_i \sim \mathcal{N}(\lambda_{i,g}) \ for \ i = 1, \ldots, N \ and \ g = 1, \cdots, K$$

$$Z_i = f_Z(\alpha_i) \ for \ i = 1, \ldots, N$$

$$A_{i,j} = f_A(Z_i, Z_j, \Lambda) \ for \ i, j = 1, \ldots, N$$

$$Y|\mu, \sigma, X, \beta, \boldsymbol{\alpha}, A \sim \mathcal{N}\Big(\mu + X\beta' + \boldsymbol{\alpha}_i + \frac{1}{N_i}\sum_{j=1, j\neq i}^{N} A_{i,j}\boldsymbol{\alpha}_j, \sigma^2\Big)$$

Throughout, we assume normal priors for the model parameters. Conditional on the covariates and latent variables, the outcome vector Y is also assumed to follow a normal distribution, consistent with the assumption of normally distributed innovations. While the conditional distribution of Y need not be normal in general, we maintain the normality assumption in this paper for tractability. One advantage of this choice is that the normal likelihood is conjugate to the normal prior, which substantially simplifies estimation. Alternative distributional assumptions can also be accommodated using simulation-based methods, including non-conjugate specifications. We leave the conditional distributions of the latent group memberships Z and the network A unspecified, as the prior and likelihood specifications are sufficient for implementing Bayesian

estimation. This choice keeps the model parsimonious and limits the dimensionality of the parameter space. Although we do not explicitly specify the latent distributions of Z and A, we assume that group membership is related to unit-level heterogeneity $\boldsymbol{\alpha}$, reflecting the goal of inferring Z given $\boldsymbol{\alpha}$. Moreover, the network A is assumed to be generated as a function of latent group memberships, ensuring a block structure consistent with standard peer effects models. Potential specifications include, but are not limited to, the degree-corrected stochastic block model or graph-based generative models.

2.3 Group Prior

In this subsection, we introduce the novel group prior proposed in this paper. Learning latent group structure in regression models has recently attracted considerable attention in the literature. Existing approaches typically impose group structure on parameters of interest—such as unit-level heterogeneity, slope coefficients, or both—by assuming homogeneity within groups and heterogeneity across groups. While convenient, this assumption can be overly restrictive and may not hold in practice. For example, social groups in educational settings are not formed solely on the basis of individual ability, intelligence, or academic performance. When the outcome of interest is grade point average (GPA), imposing a common group-level heterogeneity on all students within a small peer group may therefore be inappropriate. Instead, it is often more realistic to allow for within-group heterogeneity, even among units assigned to the same latent group.

In contrast to traditional specifications that impose within-group homogeneity and across-group heterogeneity—such as $\alpha_i = \sum_{g=1}^{G} \mathbb{I}\{g_i = g\}\alpha_g$, we propose a group prior that assumes homogeneity in distributional parameters within groups but allows for heterogeneity at the unit level. Specifically, conditional on group membership, unit-level heterogeneity is drawn from a group-specific distribution, $\alpha_i \sim \mathcal{D}\Big(\sum_{g=1}^{G} \mathbb{I}\{g_i = g\}\boldsymbol{\theta}_g\Big)$, where $\boldsymbol{\theta}_g$ denotes the distributional parameters associated with group g. A canonical example is the Gaussian group prior, $\alpha_i \sim \mathcal{N}\Big(\sum_{g=1}^{G} \mathbb{I}\{g_i = g\}\mu_g, \sum_{g=1}^{G} \mathbb{I}\{g_i = g\}\sigma_g^2\Big)$. This specification permits heterogeneity among units within the same group while inducing clustering around group-specific moments. Consequently, fixed effects for units belonging to the same group tend to be similar—though not identical—while remaining distinct from those of units in other groups. The proposed prior therefore preserves latent group structure while substantially increasing modeling flexibility.

3 Least Square and Empirical Bayes Estimators

In this subsection, we describe the estimation procedure. Standard OLS-based estimators perform poorly in settings where the adjacency matrix A is sparse and exhibits latent group structure. To address this limitation, we propose a novel empirical Bayes estimator derived from the probabilistic graphical representation of the model and its associated joint distribution. The model setup has been detailed in the preceding sections; here, we focus on implementation. The core component of the estimation procedure is the empirical Bayes estimator, defined as the posterior conditional expectation of unit-level heterogeneity. The joint distribution corresponding to the graphical model in Figure 1—omitting hyperparameters for brevity—is given by the following expression.

$$\mathbb{P}\Big(Y, A, Z, \boldsymbol{\alpha}\Big) = \mathbb{P}\Big(Y|\boldsymbol{\alpha}, A\Big)\mathbb{P}\Big(A|Z\Big)\mathbb{P}\Big(Z|\boldsymbol{\alpha}\Big)\mathbb{P}\Big(\boldsymbol{\alpha}\Big) \tag{5}$$

The posterior distribution of interest is proportional to the joint distribution in equation (5), which corresponds to the graphical model depicted in figure 1. Exploiting the conditional independence structure encoded by the probabilistic graph, the joint distribution can be factorized into prior and likelihood by using machine learning as follows:

$$\begin{aligned}
\mathbb{P}\Big(\boldsymbol{\alpha}|Y, A\Big) &\propto \sum_{Z} \mathbb{P}\Big(Y, A, Z, \boldsymbol{\alpha}\Big) \\
&= \sum_{Z} \mathbb{P}\Big(Y|\boldsymbol{\alpha}, A\Big)\mathbb{P}\Big(A|Z\Big)\mathbb{P}\Big(Z|\boldsymbol{\alpha}\Big)\mathbb{P}\Big(\boldsymbol{\alpha}\Big) \\
&\propto \sum_{Z} \mathbb{P}\Big(Y|\boldsymbol{\alpha}, A\Big)\mathbb{P}\Big(Z|A\Big)\mathbb{P}\Big(\boldsymbol{\alpha}|Z\Big) \\
&= \mathbb{P}\Big(Y|\boldsymbol{\alpha}, A\Big)\mathbb{P}\Big(\boldsymbol{\alpha}|Z^*\Big), \quad where \;\; \mathbb{P}\Big(Z^*|A\Big) = 1,
\end{aligned} \tag{6}$$

where $\propto$ denotes proportionality. Equation (6) illustrates how the posterior distribution $\mathbb{P}\Big(\boldsymbol{\alpha}|Y, A\Big)$ is decomposed into the product of the likelihood $\mathbb{P}\Big(Y|\boldsymbol{\alpha}, A\Big)$ and the group prior $\mathbb{P}\Big(\boldsymbol{\alpha}|Z^*\Big)$. By Bayes' rule, the posterior $\mathbb{P}\Big(\boldsymbol{\alpha}|Y, A\Big)$ is proportional to the joint distribution $\mathbb{P}\Big(\boldsymbol{\alpha}, Y, A\Big)$, which is obtained by integrating out the latent group memberships Z from the full joint distribution $\mathbb{P}\Big(Y, A, Z, \boldsymbol{\alpha}\Big)$. In turn, the joint distribution $\mathbb{P}\Big(Y, A, Z, \boldsymbol{\alpha}\Big)$, corresponding to the graphical model in Figure 1, factorizes into a sequence of conditional distributions as shown in equation (5). Straightforward algebraic manipulation yields the proportionality between the second and third lines of equation (6). If a group structure Z^* can be consistently recovered from the net-

work A such that $\mathbb{P}\Big(Z^*|A\Big) \to 1$, then the empirical Bayes estimator $\mathbb{E}\Big(\boldsymbol{\alpha}\Big|A^*, Y, Z^*\Big)$ can be constructed using the likelihood–prior product in the final line of equation (6). The following figure illustrates the workflow underlying this key step and the procedure used to recover Z^*.

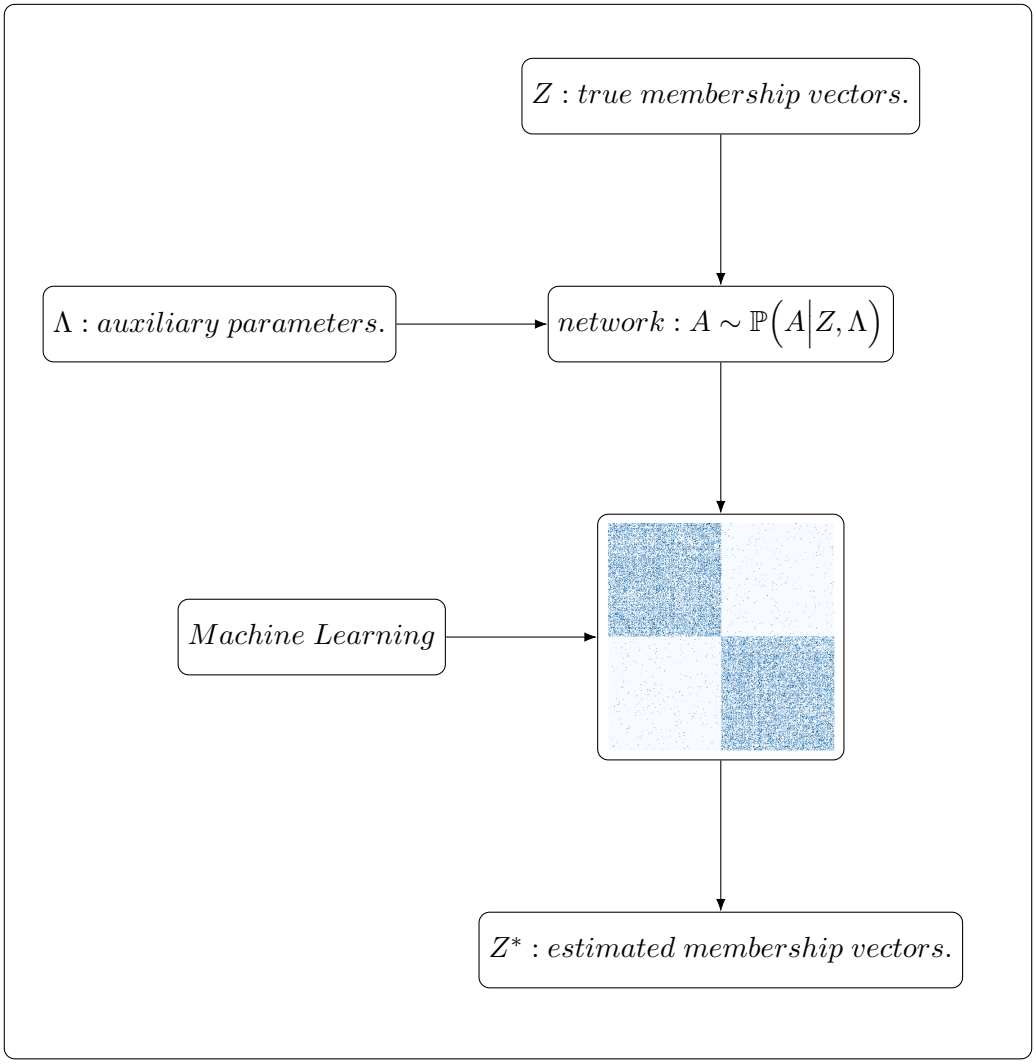


Figure 2: Given the group structure Z and other auxiliary parameters, a network with latent group structure is generated. Conversely, given a network exhibiting group structure, the underlying group assignments can be consistently recovered using suitable machine learning or network clustering methods. We denote the estimated group structure by Z^*.

The transition from the second to the third line in equation (6) shifts the focus from the generative distribution $\mathbb{P}\Big(A|Z\Big)$ to the inverse problem $\mathbb{P}\Big(Z|A\Big)$. Rather than modeling the network A as generated by latent group memberships Z, our objective is to infer the group structure given the observed network. A large body of literature has proposed algorithms for this task, many of which are consistent in the sense that the estimated group assignments Z^* satisfy $\mathbb{P}\Big(Z^*|A\Big) \to 1$ as the network size grows.

Given the membership vectors Z^*, the expression $\sum_{g=1}^{G} \mathbb{I}\{g_i = g\}\lambda_g$ can be written com-

pactly as $Z_i^* \mu_g$. An analogous representation applies to the variance parameters σ_g. Under the assumption that the group recovery procedure is consistent, $\mathbb{P}\Big(Z^* | A\Big) \to 1$, and the integration over Z is no longer required. Consequently, the final line of equation (6) follows directly.

3.1 Detecting Z Using Machine Learning

In this subsection, we describe the procedure used to recover the latent group structure Z^* from the observed network A. Learning Z^* is a key step in deriving equation (6), as it enables the reduction of the joint distribution to the product of the likelihood and the group prior. Our objective is to apply a network clustering method that consistently recovers group assignments from A, in the sense that $\mathbb{P}\Big(Z^* | A\Big) \to 1$. Under this condition, the final line of equation (6) follows directly. We emphasize that any consistent group detection algorithm is admissible within our framework, provided it satisfies the above consistency requirement. In this paper, we adopt spectral clustering, which is among the most widely used and computationally efficient methods for detecting latent group structure in networks (see, e.g., Ng et al. (2001); Von Luxburg (2007); Rohe et al. (2011),). Figure 3 illustrates examples of degree-corrected stochastic block models under different parameter configurations.

We then apply spectral clustering to networks generated under the same settings as those shown in Figure 3. The results, reported in the following table, demonstrate the effectiveness of spectral clustering in recovering the latent group structure. Reported precision values are averages computed over 500 simulation replications.

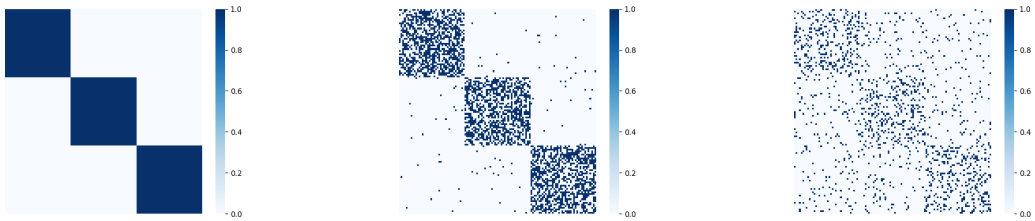


Figure 3: Three simulated networks with a common membership structure Z. The number of groups is fixed at three. Left: A network in which units within the same group are fully connected and there are no connections across groups. Middle: A network in which within-group connections are more likely than between-group connections. Right: A network generated under the same setup as the middle panel, but with weaker block structure, controlled by a lower contrast between within- and between-group connection probabilities.

	$p_{i,i}/p_{i,j}$: 1.0/0.0	$p_{i,i}/p_{i,j}$: 0.5/0.01	$p_{i,i}/p_{i,j}$: 0.2/0.05
precision	100%	100%	100%

Table 1: The table reports the classification precision of spectral clustering applied to the simulated networks. Across all parameter configurations considered, spectral clustering achieves perfect classification accuracy when a block structure is present, with precision equal to 100%.

Table 1 shows that spectral clustering achieves 100% classification precision across all simulated networks considered, provided that a latent group structure is present. Accordingly, the recovered group assignments satisfy $\mathbb{P}\Big(Z^*\big|A\Big) = 1$ in these settings, and equation 6 holds. These results confirm that spectral clustering is an appropriate and effective choice for group recovery in the empirical setup considered in this paper.

3.2 Machine Learning Assisted Empirical Bayes Estimator

In this subsection, we introduce the proposed machine-learning–assisted empirical Bayes estimator.

Without loss of generality, we assume that the disturbance term satisfies $\epsilon \sim \mathcal{N}\Big(0, I_N\Big)$. Define the residualized outcome $\tilde{Y} = Y - X^{'}\hat{\beta}$ where $\hat{\beta}$ denotes the OLS estimator of the slope coefficients. Conditional on the recovered network A^* and unit-level heterogeneity $\boldsymbol{\alpha}$, we then have

$$\tilde{Y}\Big|A^*, \boldsymbol{\alpha} \sim \mathcal{N}\Big(A^*\boldsymbol{\alpha}, I\Big).$$

We further assume the prior

$$\boldsymbol{\alpha}\Big|Z^* \sim \mathcal{N}\Big(\mu_{\boldsymbol{\alpha}}, diag(\sigma^2_{\boldsymbol{\alpha}})\Big).$$

Given the estimated group memberships Z^*, the prior mean and variance can be written compactly as $\mu_{\boldsymbol{\alpha}} = Z^*\mu_g$ and $\sigma^2_{\boldsymbol{\alpha}} = Z^*\sigma^2_g$. Consequently, the likelihood and prior jointly imply the following multivariate normal structure.

$$\begin{bmatrix}\tilde{Y}\\ \boldsymbol{\alpha}\end{bmatrix}\Bigg|A, Z^* \sim \mathcal{N}\Bigg(\begin{bmatrix}A^*Z^*\mu_g\\ Z^*\mu_g\end{bmatrix}, \begin{bmatrix}A^*\Big(diag(Z^*\sigma_g)\Big)A^{*'} + I & A^*\Big(diag(Z^*\sigma_g)\Big)\\ \Big(diag(Z^*\sigma_g)\Big)A^{*'} & diag(Z^*\sigma_g)\end{bmatrix}\Bigg). \tag{9}$$

Using standard results for conditional distributions of jointly normal random variables, the empirical Bayes estimator $\hat{\boldsymbol{\alpha}} = \mathbb{E}\Big(\boldsymbol{\alpha}\Big|A^*, \tilde{Y}, Z^*\Big)$ admits the following closed-form expression:

$$Z^*\mu_g + A^*\Big(diag(Z^*\sigma_g)\Big)\Big(A^*\Big(diag(Z^*\sigma_g)\Big)A^{*'} + I\Big)^{-1}\Big(Y - A^*Z^*\mu_g\Big). \tag{10}$$

Given Z^*, the group-level parameters μ_g and σ_g can be estimated via maximum likelihood. Following standard practice in the empirical Bayes literature, we obtain these estimates by maximizing the marginal likelihood function in equation (11).

$$\mathbb{P}\Big(\tilde{Y}\Big|A^*, Z^*, \lambda_g\Big) = \int_{\boldsymbol{\alpha}} \mathbb{P}\Big(\tilde{Y}|\boldsymbol{\alpha}, A^*\Big)\mathbb{P}\Big(\boldsymbol{\alpha}|Z^*, \lambda_g\Big)\,\mathrm{d}\boldsymbol{\alpha}, \tag{11}$$

which give us the estimated λ_g

$$\lambda_g^{MLE} = \underset{\lambda\in\mathcal{J}}{\arg\min}\,\mathbb{P}\Big(\tilde{Y}\Big|A^*, Z^*, \lambda_g\Big). \tag{12}$$

Substituting the maximum likelihood estimates λ_g^{MLE} into equation 10) yields the empirical Bayes estimator yields the empirical Bayes estimator. We summarize the empirical Bayes estimation procedure in the following algorithm.

Algorithm 1 ML-Assisted Empirical Bayes Estimation

Input: Y, X **and adjacency matrix** A.

1: Using existing algorithm, determine the number of groups K in adjacency A.

2: Using OLS, estimate $\hat{\beta}$ and calculate $\tilde{Y} = Y - X\hat{\beta}'$.

3: Detecting group membership structure Z^* by applying spectral clustering to A.

4: Construct EB estimator $\mathbb{E}\Big(\boldsymbol{\alpha}\Big|A, \tilde{Y}, Z^*, \lambda_g\Big)$.

5: Estimate $\hat{\lambda}_g$ in $\int_{\boldsymbol{\alpha}} \mathbb{P}\Big(\tilde{Y}|\boldsymbol{\alpha}, A, \lambda\Big)\mathbb{P}\Big(\boldsymbol{\alpha}|Z^*, \lambda\Big)\,\mathrm{d}\boldsymbol{\alpha}$.

6: Putting $\hat{\lambda}$ back to EB estimator.

Output: ML-Assisted Empirical Bayes estimator $\hat{\boldsymbol{\alpha}}^{EB} = \mathbb{E}\Big(\boldsymbol{\alpha}\Big|A, \tilde{Y}, Z^*, \hat{\lambda}\Big)$.

Algorithm 1 summarizes the procedure for implementing the proposed machine-learning–assisted empirical Bayes estimator. The observed data consist of the outcome variable Y, a set of covariates X, and the network A. The first step is to determine the number of latent groups K in the network A which corresponds to the number of parameter sets in the group prior. A variety of methods have been proposed for estimating the number of communities in networks (see, e.g., Le and Levina (2015); Li et al. (2022); Karrer and Newman (2011); Gao et al. (2018)).

Since community number selection is not the focus of this paper, we allow any computationally efficient and theoretically justified method to be used in this step. Next, we estimate the slope coefficients β associated with the covariates X using ordinary least squares and compute the residualized outcome $\tilde{Y} = Y - X\hat{\beta}'$. This step effectively treats X and β as hyperparameters in the subsequent analysis. We then recover the latent group structure Z^* from the network A using spectral clustering. Alternative clustering methods are also admissible, provided they yield consistent group recovery in the sense that $\mathbb{P}\Big(Z^*|A\Big) \to 1$. As with community number selection, the technical details of network clustering are outside the scope of this paper. Finally, given $\tilde{Y}$, A and Z^*, we construct the empirical Bayes estimator. Under the normality assumption, we first estimate the group-level hyper-parameters $\hat{\lambda}_g = (\hat{\mu}_g, \hat{\sigma}_g)$ by maximizing the marginal likelihood in equation (11). Substituting $\hat{\lambda}_{MLE}$ into equation (10) yields the empirical Bayes estimator, which constitutes the final output of Algorithm 1.

3.3 Learning Group Structure in Panel

In this subsection, we propose a methodology for learning latent group structure in regression models using the group prior defined above. The key idea is to infer group memberships by combining information from the group-structured prior on model parameters with the outputs of the empirical Bayes estimation procedure described in Algorithm 1.

Algorithm 2 Empirical Bayes Assisted Classification

Input: Y, X, **adjacency matrix** A **and the numbers of groups** K_A **for** A **and** K_Y **for** Y**.**

1: Using spectral clustering to find K_A-group memberships Z^* in A.

2: Using ML-Assisted Empirical Bayes Estimation to find K_A sets of group parameters for Y.

3: Vectorize the groups parameters for each unit.

4: Applying k-means to vectors of group parameters for each unit.

Output: K_Y-group memberships for all units.

Algorithm 2 formalizes the proposed methodology for learning latent group structure in the regression model. In addition to the observed outcome Y, covariates X, and the network A, we assume that the number of network groups K_A and the number of regression groups K_Y are taken as given. A large literature studies group determination in networks, cross-sectional,

time-series, and panel data; addressing this issue is beyond the scope of the present paper. Accordingly, we treat them as given. The procedure proceeds as follows. First, we recover the network group structure in A using spectral clustering with K_A groups and then implement Algorithm 1. This step yields group-specific prior parameter estimates. Under the normality assumption, these parameters take the form (μ_g, σ_g) for $g = 1, \cdots, K_A$. Next, we vectorize the group-level prior parameters, forming two-dimensional vectors (μ_g, σ_g) for each network group. In the final step, we apply k-means clustering to these vectors to further aggregate them into K_Y groups. The resulting clusters define a coarser latent group structure for the regression model, which constitutes the final output of Algorithm 2.

4 Theory

This section presents the theoretical analysis. We first establish the consistency of the ordinary least-squares (OLS) estimator under the assumption that the true adjacency matrix A^* is dense. We then examine why the OLS estimator deteriorates in sparse networks with latent group structure. Finally, we compare OLS with the empirical Bayes (EB) estimator in sparse-network settings and derive an upper bound for $\mathbb{P}\Big(MSE(EB) \leq MSE(OLS)\Big)$.

4.1 Consistency of OLS

We establish the consistency of the least-squares-based estimator for the proposed model. We then show that this estimator performs well in dense networks with latent group structure, but deteriorates in sparse networks without latent group structure.

We begin by introducing the assumptions used in the theoretical analysis. These assumptions are standard in the literature and are useful for establishing the main theorem.

Assumption 1. *Given a adjacency A^* as defined above and a model as defined in equation (1), then, we can always find a $c \in [0, 1]$ such that*

$d_{min} \geq cH_d$, where d_{min} is the minimal degree of A^* and H_d is the harmonic mean of A^*.

Assumption 2. *Given a adjacency A as defined above, A is without bipartite structure which cause λ_2 away from -1.*

Assumption 3. *Given a adjacency A, there exists a constant $c \in [0, 1]$ such that $d_{min} \geq cH_d$ where $d_{min} = \arg\min_{d_i} \Big\{d_1, \cdots, d_N\Big\}$ and H_d is the harmonic mean of A.*

The following theorem establishes the existence of the least-squares–based estimator defined in equation (1).

Theorem 1. *Given a model defined as in equation (1) satisfying assumption (1), then*

$$\hat{\boldsymbol{\alpha}} = \Big(A^{*'} M_X A^*\Big)^{-1}\Big(A^{*'} M_X Y\Big),$$

$$where\ M_X = I - X\Big(X^{'}X\Big)^{-1}X^{'}$$

exists and is unique.

Theorem 1 establishes the existence of the least-squares estimator for the proposed model in equation (1). The following theorem provides an upper bound on the variance of the least-squares–based estimator defined in equation (1).

The next theorem provides a bound between OLS estimator $\hat{\alpha}$ and α.

Theorem 2. *Given $Y = I_N\boldsymbol{\alpha} + A\boldsymbol{\alpha} + X\beta^{'} + \epsilon$ and $A*$ satisfies Assumption 1 and 2, where $A* = I_N + A$. Denote $\tilde{Y} = Y - X\beta^{'}$, $L = D^{-1/2}A^*D^{-1/2}$ and $\hat{\boldsymbol{\alpha}}_{OLS} = \Big(A^{*'}A^*\Big)^{-1}\Big(A^{*'}\tilde{Y}\Big)$, we have*

$\Big\|\hat{\boldsymbol{\alpha}}_{OLS} - \boldsymbol{\alpha}\Big\| \leq O_p\Big(\frac{\|\sqrt{N}\|}{cH_d|\lambda_2|(L)}\Big)$,

where $|\lambda_2|(L)$ is second largest eigenvalue re-ordered in absolute value, H_d harmonic mean of A^* and $c \in$[0, 1].

Theorem 2 establishes the consistency of the least-squares estimator $\hat{\alpha}_i$ for the true unit-specific parameter α_i. Moreover, the bound for $\hat{\alpha}_i - \alpha_i$ becomes small when the network is dense, as reflected by large node degrees d_i, and when the network is well connected, as indicated by a large second-smallest eigenvalue of the graph Laplacian, λ_2.

4.2 $MLE(OLS)V.SMLE(EB)$

Without loss of generality, given $\tilde{Y} = Y - X^{'}\beta$, then we have $\tilde{Y} = A^*\boldsymbol{\alpha} + \epsilon$. Define Laplacian $L = D^{-1/2}A^*D^{-1/2}$, where $D = diag(d_1, \cdots, d_N)$ and $d_i = \sum_{j=1}^{N} A_{i,j}$, and the harmonic mean degree $H_d = \Big(\frac{1}{N}\sum_{i=1}^{N}\frac{1}{d_i}\Big)^{-1}$.

We have respectively

$\hat{\boldsymbol{\alpha}}_{OLS} - \boldsymbol{\alpha} = \Big(A^{*'}A^*\Big)^{-1}A^*\epsilon$

and

$\hat{\boldsymbol{\alpha}}_{EB} - \boldsymbol{\alpha} = -\boldsymbol{\eta}_g\Big(A^{*'}A^* + \boldsymbol{\eta}_g I_N\Big)^{-1}\Big(\boldsymbol{\alpha} - \boldsymbol{\mu}_g\Big) + \Big(A^{*'}A^* + \boldsymbol{\eta}_g I_N\Big)^{-1}A^{*'}\epsilon.$

Let's decompose $A^{*'}A^*$ into $V\Lambda V'$, where V is the eigenvectors of $A^{*'}A^*$ and Λ is the diagonal eigenvalue matrix of $A^{*'}A^*$. Next, we calculate mean sqaure error of OLS and EB as follows:

$MLE_{OLS} = \sigma^2 \sum_{k=2}^{N} \frac{1}{q_k}$

and

$MLE_{EB} = \sum_{k=2}^{N} \frac{\eta_k^2 b_k^2}{\left(q_k+\eta_k\right)^2} + \sigma^2 \sum_{k=2}^{N} \frac{q_k}{\left(q_k+\eta_k\right)^2}$,

where $q_2, q_3, \cdots, q_N$ are the eigenvalues of $A^{*'}A^*$ except for first one. $\boldsymbol{b}_k = V_k\Big(\boldsymbol{\alpha} - \boldsymbol{\mu}_g\Big)$ where V is the eigenvector matrix of $A^{*'}A^*$. And $\boldsymbol{\eta}_g = \Big(\sigma/\sigma_{g_1}, \cdots, \sigma/\sigma_{g_N}\Big)$.

The following Lemma build that all eigenvalues of $A^{*'}A^*$ are smaller than $1-\lambda_2$ where λ_2 is the second largest eigenvalue of L in absolute value, where $L = D^{-1/2}A^*D^{-1/2}$.

Lemma 1. *Given a adjacency A and its Laplacian $L = D^{-1/2}AD^{-1/2}$, We have $q_k = \lambda_k(A'A) \geq c^2 H_d^2(1-\lambda_2^2)(L)$, where c is a constant between 0 and 1, H_d is the harmonic mean of the A, s_2 is k-th singular value and $\lambda_k(\,\cdot\,)$ is k-th eigenvalue in Λ.*

Lemma 1 builds the lower bound for all eigenvalues of $A^{*'}A^*$. It is used to proving Theorem 3 in the following.

Theorem 3. *Given $Y = I_N\boldsymbol{\alpha} + A\boldsymbol{\alpha} + X\beta' + \epsilon$ and $A*$ satisfying Assumption 1, where $A* = I_N + A$. Denote $\tilde{Y} = Y - X\beta'$, $\hat{\boldsymbol{\alpha}}$, then, $\mathbb{P}\Big(MSE(EB) \leq MSE(OLS)\Big)$ is bounded above by $2\Phi\left(\sqrt{\frac{\sigma^2}{\eta_k\omega_k^2}\left(2+\frac{\eta_k}{cH_d\lambda_2}\right)}\right) - 1$.*

Theorem 3 shows that as the true network A^* becomes increasingly sparse, the upper bound on $\mathbb{P}\Big(MSE(EB) \leq MSE(OLS)\Big)$ converges to 1. Conversely, as A^* becomes increasingly dense, this upper bound converges to 0. Theorem 3 therefore suggests that, in sparse-network settings, the empirical Bayes estimator is preferable to the OLS estimator.

5 Simulated Study

5.1 Simulation Design

In this section, we conduct a simulation study to examine the finite-sample performance of the proposed machine-learning–assisted empirical Bayes (ML-EB) estimator. The simulation design is as follows. The number of latent groups in the network is set to $K_A = 3$, and the total number of units is $N = 3000$, evenly divided across groups with $N_1 = N_2 = N_3 = 1000$. For the group prior, the group means are set to $\mu_{g_1} = 1.0$, $\mu_{g_2} = 0.5$ and $\mu_{g_3} = 0.0$, while the group

variances are fixed at $\sigma_{g_1} = \sigma_{g_2} = \sigma_{g_3} = 0.01$. The group membership vectors are defined as $Z_1 = [1, 0, 0]$, $Z_2 = [0, 1, 0]$, and $Z_3 = [0, 0, 1]$. Unit-level heterogeneity $\boldsymbol{\alpha}$ is generated from group-specific normal distributions with corresponding means μ_g and variances σ_g^2. To induce additional separation across groups, units with $\alpha_i > 0.75$ are assigned to group 1, units with $\alpha_i < 0.25$ are assigned to group 3, and the remaining units in between are assigned to group 2. The network A is generated using a degree-corrected stochastic block model. Without loss of generality, node-specific degree heterogeneity is set to $\Theta = I_N$. The within-group connection probabilities are set to $b_{i,i} = 0.05$, while the between-group probabilities are set to $b_{i,j} = 0.001$ for $i \neq j$. The adjacency matrix A is sampled from a Bernoulli distribution with parameter matrix $\Theta Z B Z' \Theta$ and is subsequently normalized by row sums. The disturbance term ϵ is independently drawn from a standard normal distribution. We include one control variable $X \sim \mathcal{N}\Big(0, 1\Big)$ with slope coefficient fixed at 0.3. Finally, the outcome variable is generated according to $Y_i = 0.3 * X_i + \boldsymbol{\alpha}_i + \frac{1}{N_i} \sum_{j=1, j \neq i}^{N} A_{i,j} \boldsymbol{\alpha}_j + \epsilon_i$ for $i = 1, \cdots, N$. Figure 4 illustrates the network A generated under this simulation design.

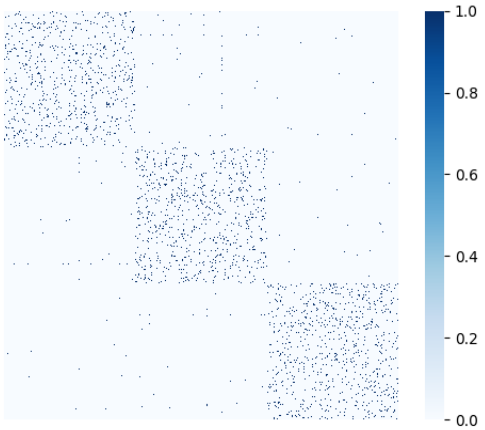


Figure 4: Simulated Network A

5.2 Main Result

We repeat the simulation 500 times. In each replication, given the simulated network A^*, we recover the group membership vectors Z_i^* for $i = 1, \cdots, N$, using spectral clustering as described above. Conditional on Z^* and A^*, we then estimate the group-level parameters μ_g and σ_g in equation (11) via maximum likelihood. Equation (6) implies that maximizing the right-hand side of the joint distribution is equivalent to maximizing the left-hand side, so the

resulting estimates $(\hat{\mu_g}, \hat{\sigma_g^2})$ coincide with those obtained from the marginal likelihood. Finally, we compute the empirical Bayes estimator by substituting Z^*, A^*, μ_g and σ_g into Equation (10).

Table 2: **First part reports the setup and estimated means in group prior using MLE. Second part reports the root mean square error of mean and median for α over 500 simulation replications, using $\frac{1}{N}\sum_{i=1}^{N}\sqrt{(\alpha_i - \hat{\alpha}_i)^2}$.**

			$ML-EB_{MLE}$	
			Mean	Std.
μ_1	1.0	$\hat{\mu}_1$	0.997	0.015
μ_2	0.5	$\hat{\mu}_2$	0.505	0.016
μ_3	0.0	$\hat{\mu}_3$	0.001	0.017

	$ML-EB_{MLE}$		OLS	
	Mean	Median	Mean	Median
$RMSE$	0.097	0.095	1.0104	1.0102

The first panel of Table 2 reports the results for estimating the group-level parameters using maximum likelihood. As shown in the first two columns, the true group means are set to $\mu_1 = 1.0$, $\mu_2 = 0.5$ and $\mu_3 = 0$. The fourth column reports the average estimated group means $\hat{\mu_g}$ across 500 simulation replications. The estimated values closely match the true parameter values used in the data-generating process. The fifth column reports the standard deviation of $\hat{\mu_g}$ across the 500 simulations, indicating that the estimators are both accurate and stable.

The second panel of Table 2 presents the performance of the proposed machine-learning–assisted empirical Bayes (ML-EB) estimator. For comparison, we also report results from the ordinary least squares (OLS) estimator. Estimation accuracy is evaluated using the root mean squared error (RMSE), $\frac{1}{N}\sum_{i=1}^{N}\sqrt{(\alpha_i - \hat{\alpha}_i)^2}$. RMSEs are computed for both the ML-EB estimator and the OLS estimator. As shown in the final row of Table 2, the ML-EB estimator substantially outperforms OLS in terms of RMSE, demonstrating the effectiveness and finite-sample advantages of the proposed approach.

6 Empirical Study

6.1 Income and Democracy

The statistical relationship between income and democracy has long been a central topic in the empirical political economy literature. Acemoglu et al. (2008) show that the statistically significant positive relationship between income and democracy disappears once country-specific fixed effects are taken into account. More recently, Bonhomme and Manresa (2015); Okui and Wang (2021); Mugnier (2025) revisit this relationship under the assumption that country-specific heterogeneity exhibits latent group structure. These studies propose alternative estimators that jointly recover the income–democracy relationship and the underlying group structure in panel data. Specifically, they consider a regression of democracy—measured by the Freedom House index—on lagged democracy and lagged log GDP per capita, allowing for unrestricted group-specific time-varying heterogeneity $\alpha_{g_i,t}$. In this study, we follow the empirical setup of the aforementioned studies while explicitly incorporating peer effects into the model. The resulting specification is given by:

$$\textbf{democracy}_{i,t} = \theta_1 \textbf{democracy}_{i,t-1} + \theta_2 \textbf{income}_{i,t-1} + \boldsymbol{\alpha}_{i,t} + \frac{1}{N_i} \sum_{j=1, j\neq i}^{N} A_{i,j} \boldsymbol{\alpha}_j + \epsilon_{i,t},$$

In its matrix form, we have

$$\textbf{democracy}_{t} = \theta_1 \textbf{democracy}_{t-1} + \theta_2 \textbf{income}_{t-1} + \boldsymbol{\alpha} I_N + A\boldsymbol{\alpha} + \epsilon_{t},$$

In a more compact form, we have

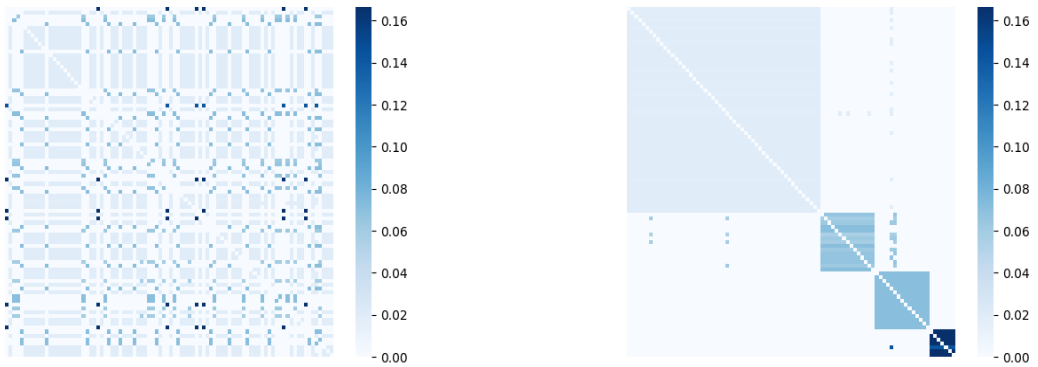
$$\textbf{democracy}_{t} = \theta_1 \textbf{democracy}_{t-1} + \theta_2 \textbf{income}_{t-1} + A^* \boldsymbol{\alpha} + \epsilon_{t},$$

where $A^* = I_N + A$.

The dataset used is a balanced panel consisting of $N = 90$ countries observed over $T = 7$ time periods at five-year intervals from 1970 to 2000. The number of latent groups is set to $G = 4$. This sample corresponds to a balanced subsample of the data used in Acemoglu et al. (2008).

The network A constructed in the empirical analysis is defined as follows. The 90 countries in the panel are classified into four groups: (1) members of the North Atlantic Treaty Organi-

zation (NATO), (2) members of the Arab League, (3) global allies of the United States outside NATO, and (4) a residual group consisting primarily of developing economies. NATO countries are characterized by broadly similar cultural, political, and institutional backgrounds and are primarily located in Europe and North America, with Turkey constituting a notable exception. The second group comprises countries that are members of the Arab League, all of which are located in West Asia and North Africa. The third group includes U.S. military allies outside NATO, such as Japan, South Korea, Thailand, Singapore, and the Philippines in Asia; Australia and New Zealand in Oceania; and Puerto Rico in the Americas. The remaining countries form the fourth group, consisting largely of developing economies, including India, China, Indonesia, Malaysia, Brazil, and Kenya. In the constructed network, all countries belonging to the same group are mutually connected. In addition, the United States is unilaterally connected to its allied countries, and the United Kingdom is unilaterally connected to countries within the Commonwealth of Nations. This procedure yields a 90×90 adjacency matrix representing the network structure. The left panel of figure 5 displays a heatmap of the constructed adjacency matrix. We then apply spectral clustering to A to recover the group assignments Z^*, which are subsequently used in the machine-learning–assisted empirical Bayes estimation. The right panel of figure 5 shows the reordered adjacency matrix based on the estimated classification, revealing a clear block structure consistent with the imposed network design. The country-level classification is reported in the Appendix.



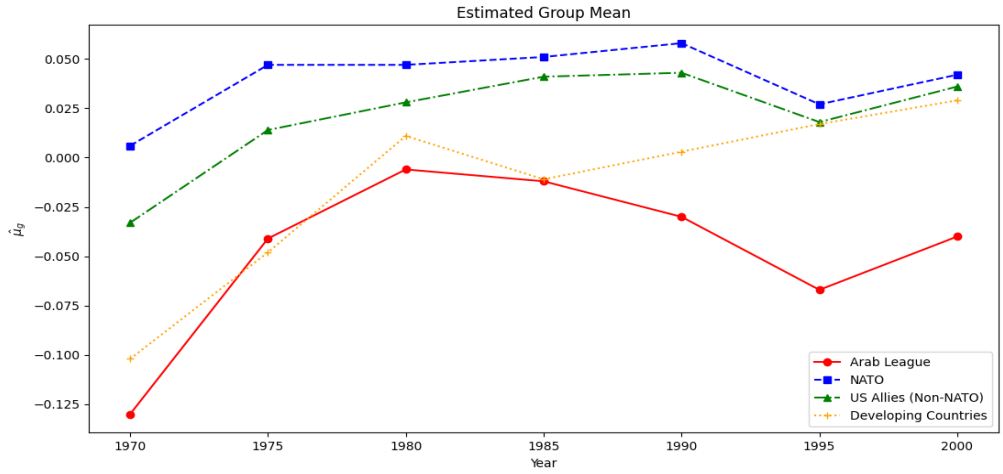
(a) (b)

Figure 5: Panel (a) displays the heatmap of the adjacency matrix constructed for the 90-country network with four groups. Panel (b) shows the heatmap of the same network after reordering according to the group partition obtained via spectral clustering.

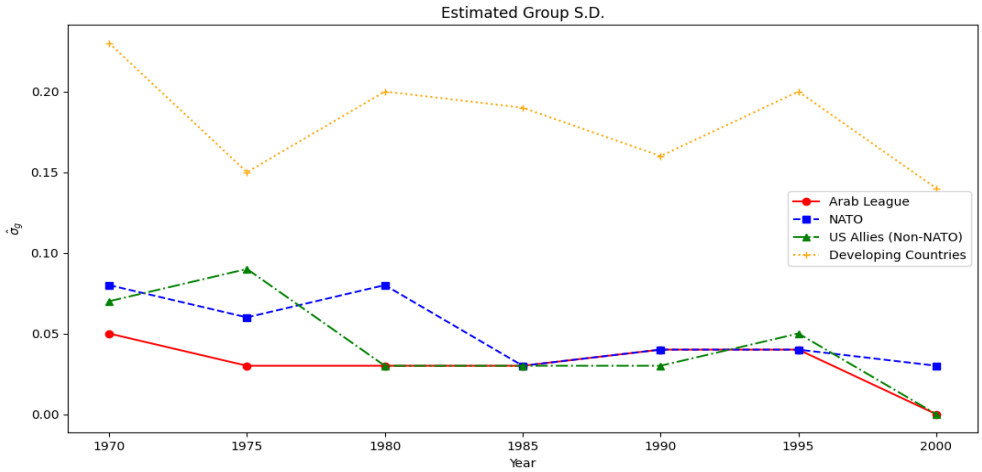
In implementing the machine-learning–assisted empirical Bayes (ML-EB) estimator, we first estimate the slope coefficients using ordinary least squares. The estimated coefficient on lagged democracy, $\hat{\theta}_1$, is 0.801, while the estimated coefficient on lagged log GDP per capita, $\hat{\theta}_2$, is 0.056. Accordingly, we obtain

$$\overline{\textbf{democracy}}_t = \textbf{democracy}_t - \hat{\theta}_1 \textbf{democracy}_{t-1} - \hat{\theta}_2 \textbf{income}_{t-1}.$$

To obtain a balanced panel, Bonhomme and Manresa (2015); Okui and Wang (2021); Mugnier (2025) exclude countries that experienced major political unification or dissolution events leading to missing observations, such as East and West Germany, the Soviet Union, Vietnam, and Czechoslovakia. After this careful sample selection, the resulting adjacency matrix is politically stable over the period 1970–2000, and country memberships in the four groups remain unchanged. We therefore treat the network as static and use it for regression analysis over the full sample period. Given the averaged democracy measure $\overline{\textbf{democracy}}_t$, the recovered network A^*, and the estimated group assignments Z^*, we proceed with the remaining steps of the ML-EB estimation. This procedure yields estimates of country-specific heterogeneity $\hat{\alpha}_i$ for $i = 1, \cdots, N$.



(a)



(b)

Figure 6: Panel (a) shows the estimated group-level means over time for the four groups. Panel (b) shows the corresponding estimated group-level variances over time.

The estimation results are summarized in figure reffigure 6. The left panel of figure reffigure 6 reports the estimated group-level means from 1970 to 2000, while the right panel displays the corresponding group-level variances over the same period. Overall, we observe a stable but gradual increase in group means across all groups. For the U.S. global allies group, the estimated group mean exhibits mild oscillations over time. This pattern is unsurprising, given the substantial heterogeneity among countries in this group in terms of historical background, political institutions, religion, language, and geographic location. Several of these countries also experienced major political transitions during this period. Notably, however, the group vari-

ance declines steadily over time, indicating convergence in country-specific heterogeneity. While member countries displayed markedly different levels of democracy in the 1970s, by the end of the twentieth century they had converged toward similarly high levels of democratic governance. This pattern provides evidence of spillover effects operating through international peer interactions. For the NATO and Arab League groups, whose member countries share relatively homogeneous cultural, linguistic, and institutional characteristics, we observe both steadily increasing group means and declining group variances. These patterns indicate gradual convergence within each group. Finally, the Third World group—comprising countries with highly diverse cultural, religious, and historical backgrounds across multiple continents—exhibits pronounced oscillations in both group means and group variances. This behavior is consistent with the substantial political upheavals and reform episodes experienced by many of these countries between 1970 and 2000.

Overall, we observe an increase in country-specific heterogeneity in democracy over the period 1970–2000, alongside a growing tendency toward cross-country convergence. This pattern is consistent with broader historical developments following the end of the Cold War and the onset of globalization. Advances in technology, transportation, and communication have substantially increased interconnectedness among countries, facilitating more frequent interactions and information diffusion. As a result, spillover effects and convergence in political institutions naturally emerge in an increasingly globalized world.

7 Conclusion

We propose a novel network peer-effects model driven by unit-specific heterogeneity, in which units affect one another through network exposure to latent heterogeneity. The model is formalized using a Bayesian belief network that explicitly characterizes the latent data-generating mechanism. Building on this probabilistic graphical representation, we develop a machine-learning-assisted empirical Bayes estimator designed to address the poor performance of ordinary least squares in sparse networks. We also establish theoretical results characterizing the properties of the proposed estimator.

In addition, we introduce a novel group prior that allows for within-group heterogeneity while preserving latent group structure. Applying the proposed methodology to the relationship between income and democracy, we find evidence of global spillovers and evolving dependence

patterns after the end of the Cold War and during the subsequent era of globalization.

References

Acemoglu, D., Johnson, S., Robinson, J. A., and Yared, P. (2008). Income and democracy. American economic review, 98(3):808–842.

Ando, T. and Bai, J. (2016). Panel data models with grouped factor structure under unknown group membership. Journal of Applied Econometrics, 31(1):163–191.

Auerbach, E. (2022). Identification and estimation of a partially linear regression model using network data. Econometrica, 90(1):347–365.

Bester, C. A. and Hansen, C. B. (2016). Grouped effects estimators in fixed effects models. Journal of Econometrics, 190(1):197–208.

Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent dirichlet allocation. Journal of machine Learning research, 3(Jan):993–1022.

Bonhomme, S., Lamadon, T., and Manresa, E. (2019). A distributional framework for matched employer employee data. Econometrica, 87(3):699–739.

Bonhomme, S., Lamadon, T., and Manresa, E. (2022). Discretizing unobserved heterogeneity. Econometrica, 90(2):625–643.

Bonhomme, S. and Manresa, E. (2015). Grouped patterns of heterogeneity in panel data. Econometrica, 83(3):1147–1184.

Borgelt, C., Steinbrecher, M., and Kruse, R. R. (2009). Graphical models: representations for learning, reasoning and data mining. John Wiley & Sons.

Boucher, V., Bramoullé, Y., Djebbari, H., and Fortin, B. (2014). Do peers affect student achievement? evidence from canada using group size variation. Journal of applied econometrics, 29(1):91–109.

Bramoullé, Y., Djebbari, H., and Fortin, B. (2009). Identification of peer effects through social networks. Journal of econometrics, 150(1):41–55.

Brownlees, C., Guðmundsson, G. S., and Lugosi, G. (2022). Community detection in partial correlation network models. Journal of Business & Economic Statistics, 40(1):216–226.

Calvó-Armengol, A., Patacchini, E., and Zenou, Y. (2009). Peer effects and social networks in education. The review of economic studies, 76(4):1239–1267.

Carrell, S. E., Fullerton, R. L., and West, J. E. (2009). Does your cohort matter? measuring peer effects in college achievement. Journal of Labor Economics, 27(3):439–464.

Cheng, X., Ho, S. C., and Schorfheide, F. (2025). Optimal estimation of two-way effects under limited mobility. arXiv preprint arXiv:2506.21987.

Chetverikov, D. and Manresa, E. (2022). Spectral and post-spectral estimators for grouped panel data models. arXiv preprint arXiv:2212.13324.

Duflo, E. and Saez, E. (2002). Participation and investment decisions in a retirement plan: The influence of colleagues' choices. Journal of public Economics, 85(1):121–148.

Efron, B. (2012). Large-scale inference: empirical Bayes methods for estimation, testing, and prediction, volume 1. Cambridge University Press.

Efron, B. (2019). Bayes, oracle bayes and empirical bayes. Statistical science, 34(2):177–201.

Gao, C., Ma, Z., Zhang, A. Y., and Zhou, H. H. (2018). Community detection in degree-corrected block models.

Goldsmith-Pinkham, P. and Imbens, G. W. (2013). Social networks and the identification of peer effects. Journal of Business & Economic Statistics, 31(3):253–264.

Guðmundsson, G. S. and Brownlees, C. (2021). Detecting groups in large vector autoregressions. Journal of Econometrics, 225(1):2–26.

Ignatiadis, N. and Wager, S. (2019). Covariate-powered empirical bayes estimation. Advances in Neural Information Processing Systems, 32.

Jochmans, K. (2023). Peer effects and endogenous social interactions. Journal of Econometrics, 235(2):1203–1214.

Jochmans, K. and Weidner, M. (2019). Fixed-effect regressions on network data. Econometrica, 87(5):1543–1560.

Karrer, B. and Newman, M. E. (2011). Stochastic blockmodels and community structure in networks. Physical Review E—Statistical, Nonlinear, and Soft Matter Physics, 83(1):016107.

Koller, D. (2009). Probabilistic graphical models: Principles and techniques.

Le, C. M. and Levina, E. (2015). Estimating the number of communities in networks by spectral methods. arXiv preprint arXiv:1507.00827.

Li, T., Lei, L., Bhattacharyya, S., Van den Berge, K., Sarkar, P., Bickel, P. J., and Levina, E. (2022). Hierarchical community detection by recursive partitioning. Journal of the American Statistical Association, 117(538):951–968.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. The review of economic studies, 60(3):531–542.

Mugnier, M. (2025). A simple and computationally trivial estimator for grouped fixed effects models. Journal of Econometrics, 250:106011.

Munro, E. and Ng, S. (2022). Latent dirichlet analysis of categorical survey responses. Journal of Business & Economic Statistics, 40(1):256–271.

Ng, A., Jordan, M., and Weiss, Y. (2001). On spectral clustering: Analysis and an algorithm. Advances in neural information processing systems, 14.

Okui, R. and Wang, W. (2021). Heterogeneous structural breaks in panel data models. Journal of Econometrics, 220(2):447–473.

Pearl, J. (2009). Causal inference in statistics: An overview.

Robbins, H. E. (1992). An empirical bayes approach to statistics. In Breakthroughs in Statistics: Foundations and basic theory, pages 388–394. Springer.

Rohe, K., Chatterjee, S., and Yu, B. (2011). Spectral clustering and the high-dimensional stochastic blockmodel. Annals of Statistics.

Sacerdote, B. (2001). Peer effects with random assignment: Results for dartmouth roommates. The Quarterly journal of economics, 116(2):681–704.

Su, L., Shi, Z., and Phillips, P. C. (2016). Identifying latent structures in panel data. Econometrica, 84(6):2215–2264.

Von Luxburg, U. (2007). A tutorial on spectral clustering. Statistics and computing, 17:395–416.

Yu, L., Gu, J., and Volgushev, S. (2024). Spectral clustering with variance information for group structure estimation in panel data. Journal of Econometrics, 241(1):105709.

Zhang, Y., Wang, H. J., and Zhu, Z. (2019). Quantile-regression-based clustering for panel data. Journal of Econometrics, 213(1):54–67.

Zhu, X., Pan, R., Li, G., Liu, Y., and Wang, H. (2017). Network vector autoregression.

Appendix A. Proof

Proof of Theorem 1. The standard estimator of $\boldsymbol{\alpha}$ is the constrained least-square estimator as follows:

$$\hat{\boldsymbol{\alpha}} = (\alpha_1, \cdots, \alpha_N)' = \underset{\{\boldsymbol{\alpha}\in\mathbb{R}^N:\ d'\boldsymbol{\alpha}=0\}}{\arg\min} \|M_X y - M_X A^* \boldsymbol{\alpha}\|^2 \tag{A1}$$

Under the assumption 1, we have $(d'\boldsymbol{\alpha})^2 = 0$. By introducing Lagrange multiplier $\lambda > 0$, we have

$$\hat{\boldsymbol{\alpha}} = \underset{\{\boldsymbol{\alpha}\in\mathbb{R}^N\}}{\arg\min} (y - A^*\boldsymbol{\alpha})' M_X (y - A^*\boldsymbol{\alpha}) - \lambda (d'\boldsymbol{\alpha})^2. \tag{A2}$$

After solving A2 using first-order condition, we have

$$\begin{aligned} \hat{\boldsymbol{\alpha}} &= (A^{*'} M_X A^* + \lambda d' d)^{-1} A^{*'} M_X y \\ &= D^{-1/2}\Big(D^{-1/2} A^{*'} M_X A^* D^{-1/2} + D^{-1/2} d\lambda D^{-1/2} d\Big)^{-1} D^{-1/2} A^{*'} M_X y, \end{aligned} \tag{A3}$$

where $D = diag(\{d_i\}_{i=1}^N)$. Under assumption 1, we have $D^{-1/2}d = D^{1/2}\kappa\mathbf{1}$. Also in assumption 1, the graph is connected which implies that $d_i > 0$ for i = $i = 1, \cdots, N$. Then, D is invertible. By construction, we have $A^*\kappa = 0$ where κ is a N by N matrix in which all diagonals are 1 and all non-diagonals are -1. Under assumption 1, rank(X) = p and rank($A^{*'} M_X A^*$) = $N - 1$, it's obvious that zero-eigenvector of $A^{*'} M_X A^*$ is constructed using κ. Then, $D^{1/2}\kappa\mathbf{1}$ is the eigenvector corresponding to zero eigenvalue of $D^{-1/2} A^{*'} M_X A^* D^{-1/2}$. It implies that $\Big(D^{-1/2} A^{*'} M_X A^* D^{-1/2} + D^{-1/2} d\lambda D^{-1/2} d\Big)$ is invertible. Furthermore, because $D^{-1/2} A^{*'} M_X A^* D^{-1/2}$ times $D^{-1/2} d\lambda D^{-1/2} d$ equals 0, then, using pseudo-inverse, we have the following identity

$$\begin{aligned} &\Big(D^{-1/2} A^{*'} M_X A^* D^{-1/2} + D^{-1/2} d\lambda D^{-1/2} d\Big)^{-1} \\ &= \Big(D^{-1/2} A^{*'} M_X A^* D^{-1/2}\Big)^{+} + \lambda^{-1}\Big(D^{-1/2} d D^{-1/2} d\Big)^{+}. \end{aligned} \tag{A4}$$

Because $A^*\kappa = 0$ and $D^{-1/2}d = D^{1/2}\kappa$, it's obvious that $\lambda^{-1}\Big(D^{-1/2} d D^{-1/2} d\Big)^{+} D^{-1/2} A^{*'} M_X Y = 0$ in (A3). It implies that

$$\hat{\boldsymbol{\alpha}} = \Big(A^{*'} M_X A^*\Big)^{+} A^{*'} M_X Y.$$

It means $\hat{\boldsymbol{\alpha}}$ does not depend on the choice of λ and constraint $d'\boldsymbol{\alpha}$ is also satisfied. □

Proof of Theorem 2. For brevity, let's define $D^{-1/2} A^{*'} A^* D^{-1/2} = F_1$ and $D^{-1/2} d = F_2$. Then,

(A4) can be rewritten as $\Big(F_1\Big)^+ + \lambda^{-1}\Big(F_2F_2'\Big)^+$. Then, we have

$$I_N = \Big(F_1^+ + \lambda^{-1}(2N)^{-2}F_2F_2'\Big)\Big(F_1 + \lambda F_2F_2'\Big)$$

$$= F_1^+F_1 + (2N)^{-1}F_2F_2',$$

since $F_1F_2 = 0$ and $F_2'F_2 = 2N$. Then, it follows that $F_1^+F_1 = I_N - (2N)^{-1}F_2F_2'$ from above. It is an idempotent matrix that orthogonally projects to F_2. Then, $L^+L = D^{-1/2}F_1^+F_1D^{-1/2} = I_N - (2N)^{-1}F_2F_2' = I_N - (2N)^{-1}(D^{1/2}\mathbf{1})'D^{1/2}\mathbf{1}$. Plugging in $L = I_N - D^{-1/2}A^{*'}A^*D^{-1/2}$ and solving for L^+, we have

$$L^+ = L^+D^{-1/2}A^{*'-1}A^{*-1}D^{-1/2} + I_N - (2N)^{-1}(D^{1/2}\mathbf{1})'D^{1/2}\mathbf{1}.$$

First, because $0 \leq L^+ \leq \lambda_2^{-1}D^{-1}$, we have

$$0 \leq L^+ - I_N + (2N)^{-1}(D^{1/2}\mathbf{1})'D^{1/2}\mathbf{1} \leq \lambda_2^{-1}D^{-1}D^{-1/2}A^{*'-1}A^{*-1}D^{-1/2}$$

Next, multiply both sides by σ^2, we have

$$0 \leq var(\hat{\boldsymbol{\alpha}}) - \sigma^2I_N + \sigma^2(2N)^{-1}(D^{1/2}\mathbf{1})'D^{1/2}\mathbf{1} \leq \sigma^2\lambda_2^{-1}D^{-1}D^{-1/2}A^{*'-1}A^{*-1}D^{-1/2}$$

Define $c = (2N)^{-1}(D^{1/2}\mathbf{1})'D^{1/2}\mathbf{1} = \frac{1}{2N}\mathbf{1}'D\mathbf{1} = \frac{1}{2N}\sum_{i=1}^N d_i = \frac{\overline{d}}{2}$. Hence, c is mean degree. Define $M = D^{-1/2}A^{*'-1}A^{*-1}D^{-1/2} \geq 0$. Then, the inequality turns out

$$0 \leq var(\hat{\boldsymbol{\alpha}}) - \sigma^2I_N + \sigma^2cI_N \leq \sigma^2\lambda_2^{-1}D^{-1}M.$$

Furthermore, since $M \geq 0$, we have the operator bound $M \leq \lambda_{max}(M)I_N$. Then, $D^{-1}M \leq \lambda_{max}(M)D^{-1}$. So the original inequality implies the simpler one

$$0 \leq var(\hat{\boldsymbol{\alpha}}) - \sigma^2I_N + \sigma^2cI_N \leq \sigma^2\lambda_2^{-1}\lambda_{max}(M)D^{-1}$$

Take $i-th$ diagonal (use $e_i(\cdot)e_j$):

$$var(\hat{\alpha}_i) \leq \sigma^2(1-c) + \sigma^2\frac{\lambda_{max}(M)}{\lambda_2}\frac{1}{d_i}.$$

Or, equivalently, we have

$$tr\Big((var(\hat{\boldsymbol{\alpha}})\Big) \leq \sigma^2(1-c) + \sigma^2 \frac{\lambda_{max}(M)}{\lambda_2}\frac{1}{d_H},$$

where $d_H = \left(\frac{1}{N}\sum_{i=1}^{N}\frac{1}{d_i}\right)^{-1}$ is harmonic mean degree. □

Proof of Theorem 3. The constrained least-squares estimator in (A1) can be expressed as

$$\hat{\boldsymbol{\alpha}} = (\alpha_1, \cdots, \alpha_N)^{'} = \underset{\{\boldsymbol{\alpha}\in\mathbb{R}^N:\ d'\boldsymbol{\alpha}=0\}}{\arg\min} \|Y - X\hat{\beta} - A^*\boldsymbol{\alpha}\|^2 \tag{A5}$$

Let $\hat{\beta} = \Big(X^{'}M_{A^*}X\Big)^{-1}X^{'}M_{A^*}Y$, then, $\hat{\beta} - \beta = \Big(X^{'}M_{A^*}X\Big)^{-1}X^{'}M_{A^*}\epsilon$. And $\mathbb{E}\Bigg(\Big(\hat{\beta}-\beta\Big)\Big(\hat{\beta}-\beta\Big)^{'}\Bigg) = \sigma^2\Big(X^{'}M_{A^*}X\Big)^{-1}$.

Furthermore, from (A5), we have

$$A^{*'}A^*\boldsymbol{\alpha} = A^{*'}\Big(Y - X\hat{\beta}\Big) \tag{A6}$$

Assume for simplicity an undirected network so $A^{*T} = A^*$ and then $A^{*'}A^* = A^{*2}$. Let $D = diag(d_i)$ and define $W = D^{-1/2}A^*D^{-1/2}$. Then, $A^* = D^{1/2}WD^{1/2}$ and multiply (A6) by $D^{-1/2}$, we have

$$WDWD^{1/2}\hat{\boldsymbol{\alpha}} = WD^{1/2}\Big(Y - X\hat{\beta}\Big).$$

Plugging in $Y = X\beta + A^*\boldsymbol{\alpha} + \epsilon$, we have

$$WDWD^{1/2}\hat{\boldsymbol{\alpha}} - WD^{1/2}A^*\boldsymbol{\alpha} = WD^{1/2}X\Big(\beta - \hat{\beta}\Big) + WD^{1/2}\epsilon.$$

Because $WD^{1/2}A^*\boldsymbol{\alpha} = WD^{1/2}\Big(D^{1/2}WD^{1/2}\Big)\boldsymbol{\alpha} = WDWD^{1/2}\boldsymbol{\alpha}$, then, we have

$$\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha} = D^{-1/2}\Big(WDW\Big)^{+}WD^{1/2}X\Big(\beta - \hat{\beta}\Big) + D^{-1/2}\Big(WDW\Big)^{+}WD^{1/2}\epsilon. \tag{A7}$$

Define the operator $H = D^{-1/2}\Big(WDW\Big)^{+}WD$, then (A7) becomes

$$\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha} = HX\Big(\beta - \hat{\beta}\Big) + H\epsilon.$$

Since $\epsilon \sim \mathcal{N}\Big(0, I_N\sigma^2\Big)$ and define $e = H\epsilon$, $Var\Big(e|A^*, X\Big) = \sigma^2HH^T$. Now, to calculate

HH^T:

$$HH^T = D^{-1/2}\Big(WDW\Big)^+ WDW\Big(WDW\Big)^+ D^{-1/2}.$$

Now use the Moore–Penrose property for PSD matrices: if $M = WDW \geq 0$, then $M^+MM^+ = M^+$. Hence, $\Big(WDW\Big)^+ WDW\Big(WDW\Big)^+ = \Big(WDW\Big)^+$ which implies

$$Var\Big(e|A^*, X\Big) \leq \sigma^2 D^{-1/2}\Big(WDW\Big)^+ D^{-1/2}. \tag{A8}$$

On the identifiable subspace (orthogonal to the null space; usually constants), a spectral gap inequality of the form $WDW \geq \lambda_2 D$. By order-reversal under inversion/pseudoinversion on that subspace, $\Big(WDW\Big)^+ \leq \frac{1}{\lambda_2} D^{-1}$. Plug it into (A8), we have

$$\begin{aligned} Var\Big(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}|A^*, X\Big) &\leq \sigma^2 D^{-1/2}\Big(\frac{1}{\lambda_2} D^{-1}\Big) D^{-1/2} \\ Var\Big(\hat{\boldsymbol{\alpha}}_i - \boldsymbol{\alpha}_i|A^*, X\Big) &= O_p\left(\frac{\sigma^2}{\lambda_2}\frac{1}{d_i^2}\right). \end{aligned} \tag{A9}$$

□

Proof of Theorem 4. Given $\tilde{Y} = Y - X'\beta$, then we have $\tilde{Y} = A^*\boldsymbol{\alpha} + \epsilon$, $\epsilon \sim \mathcal{N}(0, \sigma^2 I_N)$ let's define

$$\hat{\boldsymbol{\alpha}} = \Big(A^{*'}A^*\Big)^{-1}\Big(A^{*'}\tilde{Y}\Big).$$

Define Laplacian $L = D^{-1/2}A^*D^{-1/2}$ where $D = diag(d_1, \cdots, d_N)$ and $d_i = \sum_{j=1}^N A_{i,j}$, and the harmonic mean degree $H_d = \Big(\frac{1}{N}\sum_{i=1}^N \frac{1}{d_i}\Big)^{-1}$.

Then

$$\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha} = \Big(A^{*'}A^*\Big)^{-1} A^{*'}\epsilon$$

Next, we have

$$\Big\|\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}\Big\| = \Big\|\Big(A^{*'}A^*\Big)^{-1} A^{*'}\epsilon\Big\| \leq \Big\|\Big(A^{*'}A^*\Big)^{-1} A^{*'}\Big\|_{op}\Big\|\epsilon\Big\|,$$

by using the operator norm inequality.

The matrix

$$\Big(A^{*'}A^*\Big)^{-1} A^{*'}$$

is the Moore-Penrose inverse of A^*, usually written as

$$A^+$$

So

$$\left(A^{*'}A^*\right)^{-1}A^{*'} = A^+$$

Thus,

$$\left\|\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}\right\| \leq \left\|A^+\right\|_{op}\left\|\epsilon\right\|$$

Now use singular values. Suppose the singular value decomposition of A is $A = V\Sigma U^{'}$ where $\Sigma = diag(s_1, \cdots, s_N)$ and s_i's are singular values such that $s_1 \geq s_2 \geq \cdots, \geq s_N \geq 0$. Then, $A^+ = V\Sigma^{-1/2}U^{'}$. We have instead $1/s_N \geq \cdots, \geq 1/s_1 \geq 0$ and then,

$$\left\|\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}\right\| \leq \left\|A^+\epsilon\right\|_{op} \leq \left\|A^+\right\|\left\|\epsilon\right\| = \frac{\|\epsilon\|}{s_{min}\left(A^*\right)},$$

following a basic fact that $\|A\|_{\text{op}} = \sup_{\|x\|=1}\|Ax\| = s_{\max}(A)$.

Now, since $L = D^{-1/2}A^*D^{-1/2}$, we have $A^* = D^{1/2}LD^{1/2}$.

For any vector x in the identified subspace,

$$\left\|A^*\boldsymbol{\alpha}\right\| = \left\|D^{1/2}LD^{1/2}\boldsymbol{\alpha}\right\|.$$

Let

$$Z = D^{1/2}\boldsymbol{\alpha}$$

$$\left\|D^{1/2}LD^{1/2}\boldsymbol{\alpha}\right\| = \left\|D^{1/2}LZ\right\|$$

Because $D^{1/2}$ is diagonal with entries $\sqrt{d_i}$,

$$\left\|D^{1/2}y\right\| \geq \sqrt{d_{min}}\left\|y\right\|$$

for any vector y where $d_{min} = \min_i d_i$. Therefore,

$$\left\|D^{1/2}LZ\right\| \geq \sqrt{d_{min}}\left\|LZ\right\|$$

Now use the singular-value lower bound for B. On the identified subspace,

$$\Big\|LZ\Big\| \geq s_2\Big(L\Big)\Big\|Z\Big\|$$

Therefore,

$$\Big\|D^{1/2}LZ\Big\| \geq \sqrt{d_{min}}s_2\Big(L\Big)\Big\|Z\Big\|.$$

Since $Z = D^{1/2}\boldsymbol{\alpha}$, we have

$$\Big\|Z\Big\| = \Big\|D^{1/2}\boldsymbol{\alpha}\Big\| \geq \sqrt{d_{min}}\Big\|\boldsymbol{\alpha}\Big\|.$$

Combining these inequalities givens

$$\Big\|D^{1/2}LD^{1/2}\boldsymbol{\alpha}\Big\| \geq d_{min}s_2\Big(L\Big)\Big\|\boldsymbol{\alpha}\Big\|.$$

Equally,

$$\Big\|A^*\boldsymbol{\alpha}\Big\| \geq d_{min}s_2\Big(L\Big)\Big\|\boldsymbol{\alpha}\Big\|.$$

Since $s_2\Big(A^*\Big)$ is the smallest nonzero singular value of A^* on the same subspace,

$$s_2\Big(A^*\Big)\Big\|x\Big\| = \inf_{\Big\|x\Big\|=1, x\in\mathcal{S}} \|A^*x\|.$$

Then,

$$s_2\Big(A^*\Big) \geq d_{min}s_2\Big(L\Big).$$

Under mild degree regularity condition that $d_{min} \geq cH_d$ where $1 \geq c \geq 0$. Then,

$$s_2\Big(A^*\Big) \geq cH_ds_2\Big(L\Big).$$

Finally,

$$\Big\|\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}\Big\| \leq \frac{\|\epsilon\|}{cH_d\lambda_2},$$

since $\lambda_2 = s_2\Big(L\Big)$ because L is Laplacian. $\square$

Proof of Theorem 5. Given $\tilde{Y} = Y - X^{'}\beta$, then we have $\tilde{Y} = A^*\boldsymbol{\alpha} + \epsilon$, $\epsilon \sim \mathcal{N}(0, \sigma^2 I_N)$.

For OLS,

$$\hat{\boldsymbol{\alpha}}_{OLS} - \boldsymbol{\alpha} = \Big(A^{*'}A^*\Big)^{-1}A^*\epsilon.$$

For EB, let $\eta_{g_i} = \sigma/\sigma_{g_i}$ and $\boldsymbol{\eta}_g = \Big(\eta_{1(g)}, \cdots, \eta_{N(g)}\Big)$,

$$\hat{\boldsymbol{\alpha}}_{EB} - \boldsymbol{\alpha} = -\boldsymbol{\eta}_g\Big(A^{*'}A^* + \boldsymbol{\eta}_g I_N\Big)^{-1}\Big(\boldsymbol{\alpha} - \boldsymbol{\mu}_g\Big) + \Big(A^{*'}A^* + \boldsymbol{\eta}_g I_N\Big)^{-1}A^{*'}\epsilon.$$

Let $S = A^{*'}A^*$. Suppose the eigenvalues of S on the identified subspace are $q_2, q_3, \cdots, q_N$ and let $\boldsymbol{b} = V\Big(\boldsymbol{\alpha} - \boldsymbol{\mu}_g\Big)$.

Then, the exact OLS MLE is

$$MLE_{OLS} = \sigma^2 \sum_{k=2}^{N} \frac{1}{q_k}.$$

$$MLE_{EB} = \sum_{k=2}^{N} \frac{\eta_k^2 b_k^2}{(q_k + \eta_k)^2} + \sigma^2 \sum_{k=2}^{N} \frac{q_k}{(q_k + \eta_k)^2}.$$

Therefore,

$$MLE_{EB} \leq MLE_{OLS}$$

if and only if

$$\sum_{k=2}^{N} \frac{\eta_k^2 b_k^2}{(q_k + \eta_k)^2} \leq \sigma^2 \sum_{k=2}^{N} \left(\frac{1}{q_k} - \frac{q_k}{(q_k + \eta_k)^2}\right).$$

The left-hand side is the squared bias introduced by EB and right-hand side is the variance reduction relative to OLS. In other words, EB outperforms OLS exactly when squared EB bias is smaller than variance reduction from shrinkage.

After simplifying the variance reduction term by some simple algebra, we have sufficient condition

$$\sum_{k=2}^{N} \frac{\eta_k^2 b_k^2}{(q_k + \eta_k)^2} \leq \sigma^2 \sum_{k=2}^{N} \frac{\eta_k(2q_k + \eta_k)}{q_k(q_k + \eta_k)^2}.$$

For each direction k, after some simple algebra, a sufficient conditions becomes

$$b_k^2 \leq \frac{2\sigma^2}{\eta_k} + \frac{\sigma^2}{q_k}.$$

Recall $b_k = v_k(\boldsymbol{\alpha} - \mu_g)$, and $\boldsymbol{\alpha} \sim \mathcal{N}\Big(\mu_g, \sigma_g^2 I_N\Big)$, then $b_k \sim \mathcal{N}\Big(0, \omega_k^2\Big)$ where $\omega_k = \sum_i^N \sigma_{g_i}^2 v_{k,i}^2$. Since $\frac{b_k}{\omega_k} \sim \mathcal{N}\Big(0, 1\Big)$, Then,

$$\frac{b_k^2}{\omega_k^2} \sim \mathcal{X}_1^2.$$

Therefore,

$$\mathbb{P}\left(\mathcal{X}_1^2 < \frac{\sigma^2}{\eta_k \omega_k^2}\left(2 + \frac{\eta_k}{q_k}\right)\right) = 2\Phi\left(\sqrt{\frac{\sigma^2}{\eta_k \omega_k^2}\left(2 + \frac{\eta_k}{q_k}\right)}\right) - 1.$$

Since, for each k, $q_k \geq c^2 H_d^2 \lambda_2^2$, the conclusion follows up. $\square$

Proof of Lemma 1. Given a adjacency A and its Laplacian $L = D^{-1/2} A D^{-1/2}$, We want to prove

$$\lambda_k(A'A) \geq c^2 H_d^2 \lambda_2^2(L),$$

where c is a constant between 0 and 1, H_d is the harmonic mean of the A, s_2 is k-th singular value and $\lambda_k(\,\cdot\,)$ is k-th eigenvalue.

For any x in the identified subspace,

$$\Big\|Ax\Big\| = \Big\|D^{1/2} L D^{1/2} x\Big\|.$$

Let $Z = D^{1/2}x$, then,

$$\Big\|Ax\Big\| = \Big\|D^{1/2} L z\Big\|.$$

Since $D^{1/2}$ is diagonal with smallest entry $\sqrt{d_{min}}$, we have $\Big\|D^{1/2} y\Big\| \geq \sqrt{d_{min}}\Big\|y\Big\|$, and

$$\Big\|D^{1/2} L z\Big\| \geq \sqrt{d_{min}}\Big\|Lz\Big\|.$$

Because $\Big\|Lz\Big\| \geq s_2(L)\Big\|z\Big\|$, combining all together, we have

$$\Big\|Ax\Big\| \geq \sqrt{d_{min}} s_2(L)\Big\|z\Big\|.$$

On the other hand, $\Big\|z\Big\| = \Big\|D^{-1/2} x\Big\| \geq \sqrt{d_{min}}\Big\|x\Big\|$, then, we have

$$\Big\|Ax\Big\| \geq d_{min} s_2(L)\Big\|x\Big\|.$$

And, then

$$s_k(A) \geq d_{min} s_2(L).$$

By Assumption 2, we have $d_{min} \geq c H_d$ and square both sides,

$$\lambda_k(A'A) \geq c^2 H_d^2 \Big(|\lambda_2|(L)\Big)^2,$$

since $s_k^2(A) = \lambda_k(A'A)$ and $s_2(L) = |\lambda_2(L)|$. $\square$

Appendix B. Supplementary Material

SM.1.1 Classification in Empirical Study

Table 3: Partitions for Nations Network

Group Number		Countries		
NATO	Belgium	Canada	Denmark	France
	Greece	Iceland	Italy	Norway
	Spain	Luxembourg	Netherlands	Portugal
	Turkey	United Kingdom	United States	
League of Arab States	Algeria	Egypt, Arab Rep	Jordan	Syrian Arab Rep.
	Tunisia	Morocco	Mauritania	
US Allies (Non-NATO)	Australia	Austria	Thailand	Taiwan
	Cyprus	Finland	Ireland	Israel
	Japan	Korea, Rep.	New Zealand	Philippines
	Singapore	Sweden	Switzerland	
Third World	Argentina	Benin	Bolivia	Brazil
	Burkina Faso	Burundi	Cameroon	Central African Rep.
	Chad	Chile	China	Colombia
	Congo, Dem. Rep	Congo, Rep.	Costa Rica	Cote d'Ivoire
	Dominican Rep.	Ecuador	El Salvador	Gabon
	Ghana	Guatemala	Guinea	Honduras
	India	Indonesia	Iran	Jamaica
	Kenya	Madagascar	Malawi	Malaysia
	Mali	Mexico	Nepal	Nicaragua
	Niger	Nigeria	Panama	Paraguay
	Peru	Romania	Rwanda	Sierra Leone
	Zambia	South Africa	Sri Lanka	Tanzania
	Togo	Trinidad Tobago	Uganda	Uruguay
	Venezuela, RB			

SM.2.2 Empirical Study

Table 4: **Using OLS, we have that** $\hat{\theta}_1 = 0.801$**,** $\hat{\theta}_2 = 0.056$

$\hat{\mu}$	1970	1975	1980	1985	1990	1995	2000
Arab League	-0.13	-0.041	-0.006	-0.012	-0.03	-0.067	-0.04
NATO	0.006	0.047	0.047	0.051	0.058	0.027	0.042
US Global Alley (Non-NATO)	-0.033	0.014	0.028	0.041	0.043	0.018	0.036
Third World (Developing Country)	-0.102	-0.048	0.011	-0.011	0.003	0.017	0.029

$\hat{\sigma}^2$	1970	1975	1980	1985	1990	1995	2000
Arab League	0.003	0.001	0.001	0.001	0.002	0.002	0.000
NATO	0.006	0.004	0.007	0.001	0.002	0.002	0.001
US Global Alley (Non-NATO)	0.005	0.008	0.001	0.001	0.001	0.003	0.000
Third World (Developing Country)	0.054	0.024	0.04	0.038	0.016	0.04	0.021