Three Essays in Bayesian Microeconometrics: Embracing Causality and Heterogeneity

A thesis by

Duong Trinh-Thi-Thuy

Submitted in fulfilment of the requirements
for the Degree of Doctor of Philosophy

Adam Smith Business School

College of Social Science

University of Glasgow

June 2025

Abstract

This thesis leverages Bayesian methods to address econometric challenges in microeconomic settings, with a focus on causality and heterogeneity. The contributions are provided in three essays.

The first essay (Chapter 2) proposes a novel approach, Bayesian Analogue of Doubly Robust (BADR) estimation, to estimate unconditional Quantile Treatment Effects (QTEs) in observational studies. This estimand offers valuable insights into treatment effect heterogeneity across different outcome ranks. By incorporating Bayesian machine learning techniques, the framework can effectively handle high-dimensional covariates and nonlinear relationships to achieve better accuracy and appropriate uncertainty quantification. The simulation results show that BADR estimators yield a substantial improvement in bias reduction for QTE estimates compared with popular alternative estimators found in the literature. I revisit the role of microcredit expansion and loan access on Moroccan household outcomes, demonstrating how the new method adds value in characterising heterogeneous distributional impacts on outcomes and detecting changes in overall economic inequality, which is also appealing to other applied contexts.

The second essay (Chapter 3) introduces a new approach that harnesses network or spatial data to identify and estimate direct and indirect causal effects in the presence of selection-on-unobservables and spillovers. The proposed framework nests the Generalised Roy model to explicitly account for endogenous selection into treatment and goes beyond to capture spillovers through exposure mapping to neighbours' treatment. This allows for heterogeneous effects across individuals and enables exploration of various policy-relevant treatment effects. I develop Bayesian estimators based on data augmentation methods, offering efficient computation and proper uncertainty quantification, which is supported by simulation experiments. I apply the method to evaluate the Opportunity Zones (OZ) program, which aims to stimulate economic growth in distressed U.S. census tracts through tax incentives. The results show both direct and indirect positive impacts on housing unit growth in designated

Qualified Opportunity Zones (QOZs), but unselected tracts (non-QOZs) experience no beneficial spillovers, remaining at a disadvantage. Moreover, the model predicts that offering investment tax credits to non-QOZs would lead to negative outcomes, making the program's expansion to these areas ineffective.

The third essay (Chapter 4) is based on a joint work with Dr Santiago Montoya-Blandón. We develop a new econometric framework for modelling network interactions with heterogeneous effects, while addressing the issue of network endogeneity. The proposed Selection-corrected Heterogeneous Spatial Autoregressive (SCHSAR) model overcomes the limitations inherent in the standard spatial autoregressive (SAR) specification by achieving these dual objectives. We incorporate a finite mixture structure to capture rich heterogeneity in network interaction effects and explicitly model link formation, with latent variables playing a crucial role. For estimation and inference, our fully Bayesian approach effectively handles the computational challenges arising from the complex likelihood function and latent structure. We present a simulation study that validates the proposed approach. In the empirical application to an innovation network among American firms, we reveal significant positive yet heterogeneous interaction effects on corporate R&D investments, after accounting for endogenous network formation. The findings highlight different firm behaviours and reveal notable transmitters and absorbers in response to exogenous R&D policy shocks. This framework enables quantification of firm-level direct and spillover effects, thus providing valuable insights for evidence-based and targeted policy design.

By utilising recent developments in Bayesian econometrics, my research seeks to overcome the limitations of conventional methods, particularly in handling high-dimensional models, endogeneity, heterogeneity, and several forms of spillovers. Ultimately, the proposed methods enable more flexible and robust microdata analysis, contributing to a deeper understanding of individual and group differences in economic behaviour, as well as causal effects. This, in turn, can lead to more informed and effective policy decisions.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ACS	American Community Survey
AIC	Akaike Information Criterion
AICM	Akaike Information Criterion Monte Carlo
AIE	Average Indirect Effect
AIPW	Augmented Inverse Probability Weighted
APIE	Average Partial Indirect Effect
ASM	Adaptive Scaling Metropolis
ASP	After-School Program
ATE	Average Treatment Effect
ADTE	Average Direct Treatment Effect
ADTT	Average Direct Treatment Effect on the Treated
ADTUT	Average Direct Treatment Effect on the Untreated
BADR	Bayesian Analogue of Doubly Robust
BART	Bayesian Additive Regression Trees
BDR	Bayesian Doubly Robust
BDRS	Bayesian Doubly Robust with Shrinkage
BNP	Bayesian Nonparametric
BOM	Bayesian Outcome Modelling
BOMS	Bayesian Outcome Modelling with Shrinkage
BPSA	Bayesian Propensity Score Analysis
BvM	Bernstein–von Mises
CATE	Conditional Average Treatment Effect
CDF	Cumulative Distribution Function
CI	Credible Interval/Confidence Interval
DGP	Data Generating Process
DML	Double/Debiased Machine Learning
DPM	Dirichlet Process Mixture

FIPW	Firpo's Inverse Probability Weighted
GGRM-noSI	Gaussian Generalised Roy Model without Spillovers
GGRM-SI	Gaussian Generalised Roy Model with Spillovers
HSAR	Heterogeneous Spatial Autoregressive
IPW	Inverse Propensity Weighting
LATE	Local Average Treatment Effect
LDML	Localized Debiased Machine Learning
MAE	Mean Absolute Error
MCMC	Markov Chain Monte Carlo
MDA	Marginal Data Augmentation
MDTE	Marginal Direct Treatment Effect
MH	Metropolis-Hastings
MLE	Maximum Likelihood Estimator
NBER PDP	National Bureau of Economic Research Patent Data Project
OZ	Opportunity Zone
PAD	Patent Assignment Dataset
QOZ	Qualified Opportunity Zone
QTE	Quantile Treatment Effect
QTET	Quantile Treatment Effect on the Treated
R&D	Research and Development
RCT	Randomised Controlled Trial
RMSE	Root Mean Squared Error
SAR	Spatial Autoregressive
SCSAR	Selection-corrected Spatial Autoregressive
SCHSAR	Selection-corrected Heterogeneous Spatial Autoregressive
SNR	Signal-to-Noise Ratio
SUTVA	Stable Unit Treatment Value Assumption
TMLE	Targeted Maximum Likelihood Estimation
USPTO	United States Patent and Trademark Office

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisors for their unwavering guidance and support throughout my academic journey. I am deeply indebted to Professor Dimitris Korobilis, my principal supervisor, whose invaluable advice, wisdom, and motivation have been instrumental over the past five years. Your exceptional MRes econometric courses introduced me to the world of Bayesian inference and high-dimensional models, igniting my interest in this new terrain and inspiring me to embark on PhD research. I am profoundly grateful to Dr Kenichi Shimizu, my initial second supervisor, for his incredible patience and dedication in guiding me through abstract concepts and complex implementations in the early stages. Thank you for your compassionate understanding and encouragement that helped me embrace the process during moments of self-doubt and particularly challenging times. My heartfelt gratitude extends to Dr Santiago Montoya-Blándon, who became my second supervisor in the third year. Thank you for your confidence boosts and sharp insights at critical stages in realising my academic goals. In addition to research, your thorough help and exemplary approach to writing, presenting, teaching, and navigating the job market have had a significant impact on my development. The opportunity to learn from and collaborate with three Bayesian econometricians has been truly transformative, fundamentally shaping my own research pursuits presented in this thesis. I will carry forward the lessons you have imparted throughout my career and life. I would also like to thank my examining committee, Professor Dimitris Christelis, Professor Christoph Breunig, and Dr Maria Nareklishvili, for their excellent feedback on my thesis.

I gratefully acknowledge the Adam Smith Business School (ASBS) Scholarship offered by the University of Glasgow for my MRes and PhD studies, without which the completion of this thesis would not have been possible. During my time of study, I have been fortunate to interact with brilliant academic researchers and benefited enormously in lectures, seminars, and workshops. My special thanks go to Dr Alberto Ciancio (as well as Dr Cecilia Balocchi), Prof. Michele Battisti, Prof. Richard Dennis, Dr Constantine Sorokin, Dr John Levy, Dr Anthoulla Phella, Dr Tuuli Vanhapelto, Dr Zafer Kanik, and Dr Miguel Herculano. I am also grateful to Dr Wenya Cheng, Dr Marco Avarucci, and Dr Juhee Bae for their helpful advice during my initial years as Graduate Teaching Assistant. My sincere appreciation extends to Sophie Watson and other ASBS staff for their remarkable support of graduate students.

I am very grateful to Prof. Le Van Cuong for providing me with a recommendation letter; Dr Nguyen Ngoc Anh (& Depocen), and Dr Nguyen Vinh for mentoring me in undergraduate years. It was through their immeasurable support that my whole research journey began.

Graduate school would have been incredibly lonely without wonderful fellow friends and colleagues who have walked this path alongside me. I'll never forget the Southpark Metrics team. Thank you, Li Chen, for powerful computers and countless high-computing chats – your generosity saved my projects from futile attempts. Thank you, Keegan Chisha, for bringing a calm, humorous, and knowledgeable presence – your words stayed with me through tough moments. Special thanks to Xihao Song, Daniel Mittendorf, and Zhilang Xia, among many others at ASBS, for stimulating conversations. Great thanks to Ali Karabilgin, my awesome colleague who always went above and beyond. And to my beloved friend, Jo Jo – I'll always treasure the memories we made together. While I cannot name everyone, each companionship and gesture of support has enriched my experience immeasurably – thank you all!

While Glasgow is known for its unpredictable weather, my time in this beautiful city was also marked by many random acts of kindness. Mr Andy with Dear Garden, Lois, and Joseph brought meaningful life perspectives into my world. Through all the ups and downs, the West End became my home away from home, shaped by the camaraderie of e.Nguyet, c.Lana, c.Thu, c.Huyen, c.Ngoc, a.Tran, and e.Chi, with numerous joyful shared dinners and deep talks. Glasgow holds a special place in my heart.

Apart in distance but never in heart. Millions of thanks to my best friend, Linh Nau, for being my rock and confidant, who has put up with my truest self for over a decade; and to (Dr) Hieu Tran, for staying foolish like me (though always quicker and more positive!).

Words fail to convey my immense gratitude to my family. My parents and my little brother (Trung Kien) have supported and loved me so generously, unconditionally, and endlessly. Without that, I could never have reached this point. I love my family with all my heart –
Con cảm ơn cả gia đình đã luôn yêu thương, ủng hộ, sát cánh bên nhau và bên con!

Lastly, my boundless thanks go to the first person to encourage me to pursue this journey, Minh Tri. You believed in me even more than I believed in myself. Despite life's detours, you are my anchor, joy, and peace. Finding each other and growing together in this final chapter of my PhD has been the best thing to happen to me. You made this journey truly fulfilling.

To my family,

and

To the memory of my grandparents: your light lives on, a cherished flame in my heart.

You can't connect the dots looking forward; you can only connect them looking backwards. So you have to trust that the dots will somehow connect in your future. You have to trust in something – your gut, destiny, life, karma, whatever.

Steve Jobs, *Stanford Commencement Address*, June 12, 2005

Declaration

I declare that, except where explicit reference is made to the contribution of others, that this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

Duong Trinh-Thi-Thuy

Chapter 1

Introduction

1.1 General Background

Microeconometrics plays a vital role in developing sophisticated tools and methods that integrate microdata analysis with microeconomic modelling. At the heart of this field are two critical aspects: causality and heterogeneity.

Knowledge of causes and effects is profoundly important for decision-making, whether in government designing policies, firms allocating resources, or individuals making life choices. Causal inference has transformed our ability to move beyond mere correlation to establish genuine causal relationships. Within economic contexts, agents often operate in complex, interdependent systems where multiple factors simultaneously influence outcomes. Consequently, confounding factors can give rise to endogeneity, which poses a substantial challenge in identifying and estimating economic quantities of interest. Without proper attention to causality, applied microeconometric work risks producing misleading guidance that could lead to ineffective or even counterproductive interventions. Methodological advancements are thus crucial to equip researchers with rigorous tools for recovering causal effects. The field of causal inference in econometrics has experienced significant progress, culminating in the 2021 Nobel Prizes awarded to David Card, Josh Angrist, and Guido Imbens. As Imbens (2022) reflected in his Nobel Lecture, the field had rapidly evolved through fruitful dialogue between empirical practice and interdisciplinary research, enhancing both transparency and relevance. This recognition at the highest academic level underscores that embracing causality has become a defining feature of modern microeconometrics.

Heterogeneity is another crucial dimension well documented in microeconometrics, especially since Heckman (2001). He emphasised that "Accounting for heterogeneity and diversity and its implications for economics and econometrics is . . . a main theme" in the field. This marked a decisive departure from traditional economic models that

assumed homogeneous responses across individuals or units, a simplification that often obscured crucial variations in economic behaviour. The empirical reality reveals substantial heterogeneity: different individuals respond differently to identical policies, treatments, or economic shocks. This variation is not merely a statistical nuisance to be averaged out through aggregate measures. Rather, heterogeneity may contain essential information about the underlying economic processes and mechanisms that drive individual decision-making and outcomes. Understanding this heterogeneity is of great importance because it determines who benefits from policies, under what conditions interventions are effective, and how economic relationships vary across different populations or contexts. A growing focus on heterogeneous effects has been enabled by the availability of increasingly rich, granular datasets that allow researchers to examine variation across individuals, firms, regions, and time periods in unprecedented detail. These data developments necessitate methodological advancements to enhance the field's capacity to move beyond "one-size-fits-all" conclusions.

This dual focus on causality and heterogeneity has become a central theme in microeconometrics, however, creating substantial challenges that demand innovative methodological solutions.

1.2 Bayesian Inference in a Nutshell

Bayesian inference is a framework for statistical reasoning, where probability represents a degree of belief that is updated systematically as new data become available. Fundamentally, this is grounded in Bayes' theorem, which combines prior beliefs with empirical evidence to produce a posterior distribution. This posterior quantifies the uncertainty regarding the model parameters after observing the data in a principled, probabilistic manner. The formal setup is discussed below.

Likelihood function

The starting point of Bayesian analysis is a model of the data-generating process. Let $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subseteq \mathbb{R}^d$ denote the unknown parameter vector and $\mathbf{y} = [y_1, \ldots, y_n]$ be the observed data. The conditional probability density (or mass) function of the data given the parameters is

$$p(\mathbf{y}|\boldsymbol{\theta}) = \prod_{i=1}^{n} p(y_i|\boldsymbol{\theta}).$$

This is called the likelihood function when viewed as a function of $\boldsymbol{\theta}$ for fixed $\mathbf{y}$: $\mathcal{L}(\boldsymbol{\theta}) =$

$p(\mathbf{y}|\boldsymbol{\theta})$. According to the likelihood principle, all evidence about $\boldsymbol{\theta}$ provided by the data is contained in $\mathcal{L}(\boldsymbol{\theta})$. Bayesian inference proceeds by combining this likelihood with a prior distribution over $\boldsymbol{\theta}$.

Prior specification

The prior distribution $p(\boldsymbol{\theta})$ expresses beliefs about the values of $\boldsymbol{\theta}$ before observing any data, and must be explicitly stated by the researcher. When reliable prior knowledge is available, such as that from domain expertise, economic theory, or previous studies, informative priors can be used to incorporate that information. In the absence of such knowledge, diffuse (or non-informative) priors are often chosen to exert minimal influence on the inference process, allowing the data to "speak for themselves".

In complex econometric models, particularly those with heterogeneous agents or high-dimensional parameter spaces, simple prior choices may prove inadequate. In these settings, hierarchical priors introduce an additional layer of modelling, where the parameters themselves are drawn from distributions governed by higher-level (hyper-)parameters. For instance, rather than specifying a single prior for all parameters, we allow the prior distribution of each individual or group-specific parameter θ_i to depend on shared hyperparameters ψ, which are themselves estimated from the data

$$\theta_i \sim p(\theta_i|\psi), \quad \psi \sim p(\psi).$$

By borrowing strength across units, this structure enables partial pooling, a Bayesian compromise between no pooling (fully individual estimates) and complete pooling (homogeneous parameters). Another example is a class of shrinkage priors in high-dimensional settings (many parameters relative to observations). Here, hierarchical priors underpin modern Bayesian regularisation strategies to induce sparsity and shrinkage towards zero. These approaches are particularly powerful in microeconometrics, where data per unit (e.g., individual or firm) may be limited.

Bayes' theorem

Bayes' theorem provides the mechanism to update prior beliefs in light of the data, which translates into posterior beliefs

$$p(\boldsymbol{\theta}|\mathbf{y}) = \frac{p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\mathbf{y})} = \frac{p(\mathbf{y}|\boldsymbol{\theta}) \cdot p(\boldsymbol{\theta})}{\int_{\boldsymbol{\Theta}} p(\mathbf{y}|\boldsymbol{\theta}) \cdot p(\boldsymbol{\theta}) d\boldsymbol{\theta}}.$$

The resulting posterior distribution $p(\boldsymbol{\theta}|\mathbf{y})$ reflects updated uncertainty about $\boldsymbol{\theta}$ given observed data. The denominator $p(\mathbf{y}) = \int_{\boldsymbol{\Theta}} p(\mathbf{y}|\boldsymbol{\theta}) \cdot p(\boldsymbol{\theta}) d\boldsymbol{\theta}$ is the marginal likelihood, which serves as the normalising constant to ensure that the posterior integrates to one. In this step, we often use a shortcut and bypass the need to directly compute the integral. Specifically, we can work with the unnormalised posterior, noting that as a function of $\boldsymbol{\theta}$, the conditional density of $\boldsymbol{\theta}$ given $\mathbf{y}$ is proportional to

$$p(\boldsymbol{\theta}|\mathbf{y}) \propto p(\mathbf{y}|\boldsymbol{\theta}) \cdot p(\boldsymbol{\theta}) = \mathcal{L}(\boldsymbol{\theta}|\mathbf{y}) \cdot p(\boldsymbol{\theta}),$$

which only requires knowing the posterior density up to a normalising constant. This proportionality is sufficient for most inferential and computational tasks, particularly in simulation-based approaches.

Posterior summary

A key advantage of Bayesian procedures is their capacity to quantify uncertainty fully in the form of the entire posterior distribution of the parameters of interest. Researchers are left to decide which summary statistics to report, potentially on the basis of decision-theoretic criteria. Common point estimators include the posterior mean (minimising posterior expected squared loss), median (minimising absolute loss), or mode, known as the Maximum A Posteriori (MAP) estimator, which minimises the Dirac loss. Notably, with a flat prior density, the posterior mode coincides with the maximum likelihood estimator (MLE). Regarding interval estimation, Bayesian credible intervals (e.g., 95% highest posterior density (HPD) intervals), which directly express the posterior probability that parameters lie within a specific region, provide a natural interpretation of uncertainty.

Properties of Bayesian procedures in both large and small samples are generally at least as good as those of maximum likelihood-based procedures. Bayesian analysis, being conditional on the data, yields exact finite-sample inference, thereby obviating the need for finite-sample corrections. This is particularly attractive in econometric contexts involving limited data, weak identification, or complex hierarchical structures, where frequentist methods may be unreliable. Furthermore, as the sample size increases, the influence of the prior diminishes and the likelihood component of the posterior becomes dominant. Consequently, Bayesian estimators tend to align closely with their frequentist counterparts in large samples. This insight is formalised as the Bernstein-von Mises theorem in the literature

(see, e.g., Van der Vaart, 2000, chap. 10). Under suitable regularity conditions, the posterior distribution is asymptotically normal and centred at the MLE with variance equal to the inverse Fisher information matrix (the asymptotic variance of the MLE). From a frequentist viewpoint, this result implies that Bayesian methods can produce point estimators that are asymptotically efficient, as well as confidence intervals that have asymptotically correct coverage probability. Asymptotically, frequentist and Bayesian inferences rely on the same limiting multivariate normal distribution. Thus, in regular cases and large samples, there is no significant discrepancy between the two approaches.

Bayesian computation

The appealing theoretical properties of Bayesian methods have been acknowledged for many years, but traditionally computational difficulties held back their practical applications. Bayesian inference fundamentally involves evaluating integrals of the form

$$\mathbb{E}_{\boldsymbol{\theta}|\mathbf{y}}[h(\boldsymbol{\theta})] = \int_{\boldsymbol{\Theta}} h(\boldsymbol{\theta}) p(\boldsymbol{\theta}|\mathbf{y}) d\boldsymbol{\theta},$$

which typically have no closed-form solution, especially in high-dimensional parameter spaces common in econometric models. The high dimensionality and intractability of posterior distributions pose substantial barriers, necessitating numerical approximation techniques.

Computational advances over the past decades – most notably, the advent and refinement of Markov chain Monte Carlo (MCMC) algorithms – have significantly mitigated earlier computational challenges. MCMC methods simulate a dependent random sequence, known as a "Markov chain", of parameter draws $\{\boldsymbol{\theta}^{(1)}, \ldots, \boldsymbol{\theta}^{(B)}\}$ whose stationary distribution matches the posterior distribution $p(\boldsymbol{\theta}|\mathbf{y})$. The simulated chain facilitates empirical approximation of posterior expectations as

$$\mathbb{E}_{\boldsymbol{\theta}|\mathbf{y}}[h(\boldsymbol{\theta})] \approx \frac{1}{B} \sum_{b=1}^{B} h(\boldsymbol{\theta}^{(b)}).$$

Metropolis-Hastings (MH) is a versatile and widely-used class of MCMC algorithms that constructs a Markov chain by iteratively proposing parameter values from a chosen distribution and correcting potential inaccuracies through acceptance-rejection steps. A common implementation, the random-walk Metropolis algorithm generates a candidate parameter draw by sampling from a Gaussian proposal distribution centred at the current state. Formally, given a current state is $\boldsymbol{\theta}^{\text{curr}}$, a candidate draw $\boldsymbol{\theta}^{\text{cand}}$ is generated as

$\boldsymbol{\theta}^{\text{cand}} \sim \mathcal{N}(\boldsymbol{\theta}^{\text{curr}}, \sigma^2\mathbf{I})$. The move from $\boldsymbol{\theta}^{\text{curr}}$ to $\boldsymbol{\theta}^{\text{cand}}$ is accepted with probability

$$\alpha = \min\left\{\frac{p(\boldsymbol{\theta}^{\text{cand}}|\mathbf{y})}{p(\boldsymbol{\theta}^{\text{curr}}|\mathbf{y})}, 1\right\},$$

and rejected (i.e., remaining at $\boldsymbol{\theta}^{\text{curr}}$) with probability $1 - \alpha$. Despite being robust and broadly applicable, the efficiency of MH depends critically on properly tuning the proposal distribution. Gibbs sampling, a special case of the Metropolis–Hastings algorithm with an acceptance probability of one, breaks the curse of dimensionality by exploiting low-dimensional conditional distributions. The Gibbs sampler cycles sequentially through parameter blocks, drawing each component from its full conditional posterior distribution

$$\theta_j^{(b)} \sim p(\theta_j|\theta_{-j}^{(b-1)}, \mathbf{y}).$$

This significantly simplifies the computation when these conditional posteriors have a convenient analytical form. When facing complex conditional targets, Gibbs sampling is often combined with the Metropolis-Hastings steps, resulting in a hybrid MH-within-Gibbs approach utilised efficiently in numerous applications. Notably, efficient Bayesian computation remains an active interdisciplinary research area, although detailed discussion exceeds the scope of this overview.

In summary, the computational revolution has made Bayesian inference widely feasible across econometric settings. There are now few restrictions regarding the choice of priors, complexity of likelihood functions, or dimensionality of parameter spaces. Consequently, Bayesian methods have evolved to serve as superior alternatives to classical methods, delivering innovative computational and modelling solutions to complex econometric problems. Early examples of significant microeconometric issues, especially related to causality and heterogeneity, solved using Bayesian methods include: (1) Discrete choice models and panel data models with individual heterogeneity (e.g., Athey and Imbens, 2007; Chamberlain and Hirano, 1999; McCulloch and Rossi, 1994), and (2) Causal inference, treatment effects, and selection models (e.g., Chamberlain and Imbens, 2003; Chib and Jacobi, 2007; Jacobi et al., 2016). This list, of course, is not exhaustive. Bayesian econometric applications have kept expanding rapidly. In microeconometrics, the Bayesian approach enables flexible, coherent, and computationally feasible inference in settings where heterogeneity, dynamics, and complex data structures matter.

This brief overview can only provide a preliminary map to the major ideas and developments of Bayesian methods in microeconometrics. For more comprehensive discussions, the reader is referred to Li and Tobias (2011), Cameron and Trivedi (2005; 2022), Chan et al. (2019), and Rossi et al. (2024). The merits of the Bayesian approach combined with efficient computational techniques are manifold. Importantly, there is an ongoing synthesis of Bayesian and frequentist perspectives, shifting emphasis from philosophical debate toward methodological practicality regarding what works best for different types of problems. As both approaches have appealing features, recognising this synergy is beneficial for empirical research. In the golden age of algorithmic development and the rising popularity of probabilistic programming, Bayesian methods show great potential to meet the demands of modern microeconometrics and deserve further exploration.

1.3 Thesis Contribution

In light of the motivation outlined above, this thesis develops methods for three distinct – yet methodologically interconnected – settings in microeconometrics, with a consistent focus on both causality and heterogeneity. The common vantage point is a Bayesian inferential framework for handling the specific technical challenges that arise in each setting.

The first essay addresses the challenges associated with estimating unconditional Quantile Treatment Effects (QTEs) in observational studies. This causal estimand offers valuable insights into treatment effect heterogeneity by examining how effects vary across different ranks of the outcome distribution, moving beyond average treatment effects (ATEs), which can mask important distributional impacts. However, applied researchers often encounter a vast set of possible covariates in observed datasets yet remain uncertain about which specific ones are necessary to control for when recovering treatment effects. Additionally, the QTE estimation problem involves nuisance parameters, including the entire conditional cumulative distribution function (CDF) of each potential outcome conditional on potentially high-dimensional covariates, which increases computational complexity. To circumvent such obstacles, I propose the Bayesian Analogue of Doubly Robust (BADR) approach with two key ingredients. First, to effectively accommodate high-dimensional covariates and nonlinear relationships while achieving proper uncertainty quantification, I incorporate Bayesian regularisation methods with attractive prediction performance to generate auxiliary estimators for both the propensity score and conditional outcome distribution. I leverage

multiple Bayesian quantile regressions augmented with shrinkage priors to address the unique challenge of quantile estimation. Second, I derive the estimator for target QTEs by solving an estimating equation built upon an efficient influence function specifically tailored to quantile functionals. This results in double robustness, ensuring the final estimator remains consistent if either the treatment assignment model or the outcome regression model is consistently estimated but not necessarily both. The framework features a highly flexible Bayesian modelling scheme that showcases favourable frequentist properties in finite samples for QTEs, which has not been explored before. The simulation results show that BADR estimators yield a substantial improvement in bias reduction for QTE estimates compared with popular alternative estimators found in the literature.

While the first essay maintains selection-on-observables – stipulating that treatment is as good as randomly assigned once we condition on observables – this assumption may not hold in economic scenarios when unobservable factors simultaneously influence both individual choices and their outcomes. The second and third essays delve into these more complex sources of endogeneity in contexts such as noncompliance and network interference, while continuing to model heterogeneous effects.

The second essay develops a new approach that utilises network or spatial data to identify and estimate direct and indirect causal effects when both selection-on-unobservables and spillovers (also known as interference) are present. The endogeneity challenge here is twofold: unobservable characteristics affect treatment selection, and spillovers create interdependence among individuals, which violates standard causal inference assumptions. The proposed framework nests the Generalised Roy model to explicitly account for endogenous selection into treatment and captures spillovers through exposure mapping to neighbours' treatment. Crucially, this allows for heterogeneous effects both across individuals and in terms of how they respond to neighbours' treatments, enabling the exploration of various policy-relevant treatment effects. Given the inherent nature of the setting as a missing data problem, I develop Bayesian estimators based on data augmentation methods, offering efficient computation and proper uncertainty quantification.

A natural extension arises when the network structure that shapes the interference pattern itself is endogenous. This is relevant to models of network interactions, where individual outcomes depend on their peers in a network, and the network structure itself is endogenously formed. Specifically, the decision to form links and the outcome of interest may be jointly

determined by unobservable factors, creating a complex source of endogeneity. This concern motivates the third essay, which shifts the focus to modelling network interactions with heterogeneous effects while addressing the issue of network endogeneity. The third essay introduces Selection-corrected Heterogeneous Spatial Autoregressive (SCHSAR) model, a new econometric framework that overcomes the limitations inherent in the standard spatial autoregressive (SAR) specification. We incorporate a finite mixture structure to capture rich heterogeneity in network interaction effects and explicitly model endogenous link formation, with latent variables playing a crucial role. The fully Bayesian approach effectively handles the computational challenges arising from the complex likelihood function and latent structure.

Taken together, each essay demonstrates how flexible Bayesian modelling can overcome limitations of conventional approaches. Beyond methodological contributions, these essays provide empirical analyses using proposed methods, yielding policy-relevant insights across diverse economic contexts. From understanding distributional impacts of financial expansion (the first essay) to evaluating place-based economic development policies with spillovers (the second essay), to informing targeted innovation policies based on firm-level network formation and interaction (the third essay), the findings illustrate how accounting for both causality and heterogeneity can enhance policy relevance.

The remainder of this thesis is organised as follows. Chapter 2 introduces the first essay, *"Causal Inference on Quantiles in High Dimensions: A Bayesian Approach"*. Chapter 3 presents the second essay, *"Bayesian Causal Inference in the Presence of Endogenous Selection into Treatment and Spillovers"*. Chapter 4 covers the third essay, *"Modelling Interactions with Heterogeneous Effects and Endogenous Network Formation"*. Each essay is self-contained, providing necessary information for the reader to understand the setting and rationale underlying the proposed method. Finally, the conclusions summarise the findings and offer broader implications. The technical details and supplementary results are relegated to the appendices. Notation is introduced when appropriate.

Chapter 2

Causal Inference on Quantiles in High Dimensions: A Bayesian Approach

2.1 Introduction

When evaluating the causal effect of policy interventions, the distributional impact appeals to researchers and policymakers rather than the average impact alone. It helps to gain more comprehensive and nuanced understanding of the complex effects, ultimately leading to more effective decision-making. In many instances, uniform policies may benefit certain individuals while adversely affecting others. If the effects are considerably heterogeneous, the average treatment effect may not be a sufficient measure, as it likely masks substantial positive and negative effects. Consequently, it is crucial to determine whether certain individuals are worse off as a result of the policy. Even if multiple programs generate positive effects for all individuals, the one that offers the greatest benefits to those at the lower tail of the distribution of the outcome variable might be the most favourable. To illustrate, consider two job training programs with identical mean net impact that is positive. The first program, which increases wages at the bottom of the wage distribution, would be more appreciated than the second program, which only raises the top of the wage distribution. This necessitates the advancement of econometric techniques to enable studies on distributional treatment effects in the presence of heterogeneity. This goal has received special interest and has become increasingly relevant in economic applications[1].

Sets of quantile treatment effects (QTEs) can characterise the heterogeneous impacts of the treatment on different points of the outcome distribution. With a binary treatment, as originally defined by Doksum (1974) and Lehmann et al. (1974), QTEs measure the difference

[1] A variety of relevant applications include, but are not limited to, financial interventions (Callaway and Li, 2019; Meager, 2022), educational subsidies (Duflo et al., 2021), public health policies (Schiele and Schmitz, 2016), and local migration incentives (Bryan et al., 2014; Chetty et al., 2016); all entail social welfare implications and garner substantial attention in public discourse.

between the unconditional quantiles of the potential outcome distribution under treatment and the potential outcome distribution under nontreatment. Put differently, this captures any difference between the two cumulative distribution functions of treated and untreated potential outcomes. Moreover, the quantile method may be employed, even by those not primarily interested in distributional consequences, to enhance the robustness of their analysis. This is particularly relevant in light of the well-established fact that median regression is more resistant to outliers than mean regression, while many economic data sets involve heavy tails. One notable example in development economics is household welfare measures, including consumption and business outcomes.

The difference between these two unconditional distributions of the potential outcomes itself might be appealing to policymakers. This reflects the change in the distribution function as a whole when the treatment could be exogenously shifted between two distinct counterfactual scenarios: universal treatment and no treatment. As the entire distribution function often yields insights into inequality or social welfare analysis, computing QTEs serves as a convenient way to summarise noteworthy aspects. For instance, this enables the detection of changes in overall inequality in the distribution of outcomes, which is a critical concern given the potentially negative consequences of social and economic inequality in the contemporary world[2].

By definition, QTEs can reveal heterogeneity in the causal effects on different quantiles. However, individual results are only interpretable under a rank preservation assumption on the underlying treatment effect distribution. This assumption asserts that an observed individual would maintain their position (rank) in the distribution regardless of their treatment status. As a result, the set of quantile treatment effects is equivalent to the quantiles of the distribution of individual treatment effects. Nonetheless, rank preservation is a strong assumption because it requires the relative value of the potential outcome for a given individual to be unchanged, whether that individual is treated or untreated. Even when rank preservation is violated, heterogeneity in the effects across various quantiles shows evidence of heterogeneity in these individual effects, making QTEs remain a meaningful parameter for policy purposes (see, e.g., Angrist and Pischke, 2009; Meager, 2022).

In this paper, the primary focus is on unconditional QTEs, which are separate from conditional QTEs. Although both are standard parameters of interest in the program evaluation literature, it is important to highlight the distinction between them. An unconditional

[2]This interest is at the core of the econometric literature strand on distributional counterfactual analysis (see, e.g., Chernozhukov et al., 2013; Firpo and Pinto, 2016; Rothe, 2010).

(marginal) quantile function is a one-dimensional function of the quantile level τ only. Defined as the difference between the unconditional quantiles of the treated and untreated potential outcome distributions, unconditional QTEs describe the effects of treatment status on the overall outcome distribution without conditioning on the covariates. In contrast, conditional quantile functions are multi-dimensional, depending on not only a chosen quantile level but also values of the covariates. Conditional QTEs thereby express the effects on the outcome distribution within sub-populations characterised by covariate values. More specifically, an individual may rank high in the unconditional distribution of the outcome, meanwhile possessing a low rank in the conditional distribution of the outcome. This is possible if that person has values of observed characteristics that are associated with a large value of outcome overall, yet within the group of people sharing identical values of the observed characteristics, he or she has a comparatively low outcome[3]. Conditional QTEs enable examination of the heterogeneity of the effects with respect to the observables; however, they might be sensitive to the choice of covariates to be included. Unconditional QTEs, on the other hand, aggregate the conditional effects across the entire population, thereby being more easily conveyed to the policymakers and the public, at the cost of not providing any information about the relationship between the covariates and the outcome. Further discussion can be found in surveys by Glewwe and Todd (2022) and Frölich and Melly (2013). The unconditional quantile treatment effects are appropriate estimands to focus on when the ultimate objective is related to the marginal distribution, for example, the welfare of the (unconditionally) poor. This unconditional effect has been a central estimand of interest in the literature on micro-credit expansion and housing outcomes, reinforcing the need for robust methods suited to recovering and exploring heterogeneity in these effects.

When the target causal estimand is an unconditional quantile treatment effect (QTE), several identification strategies have been developed in the literature. One common approach relies on the assumption of exogenous treatment, typically in the form of a randomised controlled experiment where all participants comply with their treatment assignment. In this ideal scenario, implementing an unconditional QTE estimator is a straightforward process, similar to estimating the average treatment effect (ATE) directly from the realised outcomes of control and treatment groups. However, when such experimental data is unavailable, inferring causal relationships from observational data poses challenges because the observed

[3]Consider a simple example involving wages and years of education, the median income of all individuals with doctoral degrees may be greater than the top quantile for high school dropouts, presuming a strong positive association between education levels and earnings.

treatment status is not assigned randomly. This gives rise to the second approach based on the *selection-on-observables* assumption, which implies that the treatment is as good as randomly assigned once we condition on observables. It is worth noting that, although our ultimate goal is to obtain an unconditional QTE, covariate information serves to correctly identify the unconditional quantiles and remove selection bias.

In this paper, we maintain the identifying assumption of *selection-on-observables*, which is widely applicable to empirical studies in economics. This is due to the fact that randomised controlled trials (RCTs) are often intricate and expensive, rendering them infeasible in many cases. In addition, this assumption is justifiable in various contexts, such as when the treatment is randomly allocated within demographic groups. As elaborated in the subsequent sections, employing covariates for the sake of identification involves identifying the entire conditional cumulative distribution function (CDF) of each potential outcome conditional on potentially high-dimensional covariates. This CDF is then a nuisance function for the identification of QTEs. More often than not, applied researchers encounter a vast set of possible covariates, but they are uncertain about which specific ones are necessary to control for when recovering treatment effects. In addition, the conditional CDF can itself be a complex function. This necessitates the consideration of high-dimensional models to estimate quantile treatment effects.

This paper aims to circumvent such obstacles and contribute to the emerging econometric literature on identification and estimation of QTEs. We propose a novel framework, the Bayesian Analogue of Doubly Robust (BADR) approach, for estimating QTEs in an observational study while accounting for the presence of potentially high-dimensional covariates. Briefly, we employ Bayesian techniques to specify and estimate both the treatment assignment and the outcome models, obtaining posterior draws that are then plugged into the doubly robust estimator for QTEs. This estimator is derived as the solution to efficient influence functions, leading to its double-robustness property. The resulting BADR framework comprises two ingredients. First, to effectively accommodate high-dimensional covariates and nonlinear relationships while achieving proper uncertainty quantification, we incorporate various Bayesian regularisation methods including sparsity-inducing priors and Bayesian nonparametric methods to generate auxiliary estimators for both the propensity score and the conditional distribution of the outcome variable. We leverage multiple Bayesian quantile regressions to address the unique challenge in quantile estimation, which differs

from previous studies on ATE. Second, our method provides double protection against model specification by employing posterior predictive distributions of parameters from both the treatment assignment and outcome models. In the absence of high-dimensional covariates, this approach collapses to a doubly-robust Bayesian estimator for the QTEs without shrinkage priors, which itself has not been explored in previous literature. Overall, the proposed strategy enables us to obtain QTE estimators which showcase highly flexible Bayesian modelling manner coupled with favourable frequentist properties in finite samples.

The advantages of the proposed estimators are demonstrated in Monte Carlo simulations, which consider difficult settings such as high-dimensional covariate spaces or complex nonlinear effects of covariates. Through numerical evidence, we observe substantial gains in bias reduction for QTE estimates across all scenarios, highlighting the strong estimation and inferential features of our methodology in comparison with the naive estimator and popular approaches considered in the literature. The proposed methodology introduces a fresh perspective to empirical research by offering a novel approach for estimating unconditional QTEs in microeconomic applications. The new estimator, Bayesian Doubly Robust with shrinkage prior, allows us to revisit the microcredit experiment originating from the work of Crépon et al. (2015) and explore the impact of household financial access on household welfare. Unlike previous studies that strictly rely on the randomisation of microcredit availability at the village level, we employ a new causal estimand and identification strategy that utilises observed, non-random borrowing patterns at the household level as well as observable household characteristics. Our findings indicate an overall positive effect, with heterogeneous impacts across the different points of each outcome distribution of interest. It is anticipated that universal financial access will result in an ex-post increase in economic inequality among households, mostly attributed to the significant improvements in consumption and business outcomes at the upper quantiles. Notably, there is evidence of systematic harm in terms of total profit, as a segment of households may experience adverse effects that extend the left tail of the distribution to the left.

The remainder of this paper is organised as follows. Section 2.2 presents a brief review of existing studies relevant to our paper and situates the paper within existing literature. In Section 2.3, we formally define quantile treatment effects in a causal framework along with key identification assumptions. In Section 2.4, we present the proposed approach for estimating quantile treatment effects. Next, we evaluate the performance of our methods using

simulations in Section 2.5 and use the proposed method to examine the causal impact of loan access on the distribution of household outcomes in Section 2.6. Finally, we conclude the paper in Section 2.7 with brief final remarks on the method and policy recommendations based on our results.

2.2 Related Literature

2.2.1 Causal Inference on Quantiles

Firpo (2007) first considered efficient estimation of unconditional quantile treatment effects (QTEs) and proposed an inverse propensity weighting (IPW) estimator based on propensity scores estimated using a sieve approach, specifically a logistic power series approximation. Under strong smoothness conditions, this IPW estimator is $\sqrt{N}$-consistent[4] and achieves the semiparametric efficiency bound, which is reminiscent of analogous results for the IPW estimator for the average treatment effect (ATE) with nonparametrically estimated propensities (Hirano et al., 2003). Although these purely weighted methods circumvent the estimation of nuisances that depend on the estimand, their desired behaviour is restricted to certain nonparametric weight estimators and requires strong smoothness assumptions. Extending the IPW estimator to high-dimensional settings runs into issues due to the fact that its convergence rate can be slowed down by that of the propensity score and its error may depend heavily on the particular method used to learn the propensity score. The properties prohibit the use of general machine learning methods and potentially leading to unstable estimates. In this sense, Firpo's (2007) IPW estimator lacks the double robustness and flexibility of our proposal.

Zhang et al. (2012) developed several nonparametric methods that resemble those used for ATE and proved that the augmented inverse probability weighted (AIPW) estimator, by augmenting a term that involves the residuals from the outcome regression model, enhances the efficiency of the IPW estimator. The AIPW estimator is expected to be locally efficient and doubly robust under regularity conditions. Díaz (2017) proposed a semiparametric approach using targeted maximum likelihood estimation (TMLE) for marginal quantiles. While sharing the same asymptotic properties as the standard AIPW, the TMLE estimator demonstrates better finite-sample performance when analysing causal effects on the quantiles, similar to improvements in the mean effect (e.g., Van der Laan et al., 2011). Our proposed approach relates to both AIPW and TMLE estimators when solving the estimating equation

[4] N is the sample size.

derived from the efficient influence function, a core concept for achieving double robustness. However, Zhang et al.'s (2012) AIPW method assumes strong distributional assumption (e.g., a normal linear model after a Box-Cox transformation of the outcome for each treatment), limiting its application to cases of positive outcomes. In contrast, our Bayesian Analogue of Doubly Robust estimation framework employs Bayesian data-adaptive estimation to flexibly fit nuisance functions. Diaz's (2017) TMLE approach can be considered quite general and closest to our approach among frequentist methods. While both opt for estimating the conditional distribution as an important middle step, the distinct feature of our modelling strategy lies in utilising multiple Bayesian quantile regressions and thus can incorporate regularisation seamlessly.

Unlike previous studies that did not explicitly consider the case of potentially high-dimensional covariates, Kallus et al. (2024) proposed Localized Debiased Machine Learning (LDML) to enable efficient inference on QTEs in this scenario. For ATE estimation, nuisance functions do not depend on the estimand and can therefore be estimated independently using flexible, data-driven, machine learning methods and plugged into the estimating equation. This Debiased Machine Learning (DML) approach is, however, far more challenging for QTEs estimation, as the efficient influence function involves nuisances that depend on the estimand of interest. Specifically, DML requires we learn the whole conditional cumulative distribution function of a real-valued outcome, potentially conditioned on high-dimensional covariates, evaluated at the quantile of interest. To obviate this cumbersomeness, Kallus et al. (2024) localise the nuisance estimation step to a single initial rough guess of the estimand, such as the IPW estimate, thereby enabling the standard use of machine learning methods in this DML-extended framework. Despite also aiming for a flexible modelling manner, our paper takes a different approach, estimating the whole continuum of the estimand-dependent nuisances by discretising a hypothetical continuum of quantile regression estimators. The rationale for our choice is based on the advantages of Bayesian quantile regression over the frequentist alternative.

While most studies on unconditional QTEs are based on frequentist methods, Xu et al. (2018) proposed a Bayesian nonparametric approach (BNP) that utilises a Bayesian additive regression tree (BART) model to estimate the propensity score, followed by a Dirichlet process mixture (DPM) of normals model to construct the distribution of potential outcomes conditional on the estimated propensity score. A key advantage of this approach

over frequentist methods is the simultaneous estimation of multiple quantiles of interest. However, it can be regarded as a propensity score analysis which avoids directly modelling the conditional distribution of potential outcomes given the covariates. In contrast, we propose Bayesian Analogue of Doubly Robust estimators that can handle a large number of covariates and are more robust to misspecifications.

2.2.2 Causal Inference in High Dimensions

This paper fits into a broader literature on high-dimensional causal inference with observational data. High-dimensional settings are becoming increasingly prevalent, presenting challenges for causal inference. This problem involves either a large number of available covariates or an outcome model with an infinite or large number of parameters, such as nonparametric and semiparametric models. regularisation, a popular technique originally designed to perform prediction in high-dimensional data analysis, has garnered substantial attention in causal inference. It gives rise to numerous causal machine-learning techniques which provide high-quality inference on treatment effects (Athey et al., 2018; Belloni et al., 2014; Chernozhukov et al., 2018; Farrell, 2015). While the majority of studies are frequentist regularisation-based approaches, there has been growing interest in adopting Bayesian regularisation-based techniques into causal inference, as Bayesian inference is a natural probabilistic framework for quantifying uncertainty and learning about model parameters. It is known that many frequentist penalised likelihood estimators can be considered equivalent to the posterior modes of Bayesian estimators under certain choices of shrinkage priors such as spike-and-slab prior (Ishwaran and Rao, 2005; Mitchell and Beauchamp, 1988), Bayesian Lasso prior (Park and Casella, 2008), and Horseshoe prior (Carvalho et al., 2010, 2009).[5] Recent studies have successfully deployed these techniques for confounding adjustment to estimate average treatment effects in the presence of high-dimensional controls (Antonelli et al., 2022, 2019; Hahn et al., 2018). Bayesian nonparametric methods are also powerful tools utilised for regularisation within the Bayesian paradigm. Among them, the Bayesian Additive Regression Tree (BART) has emerged as a workhorse widely used for causal inference. Introduced by Chipman, George, and McCulloch (2006; 2012), BART models offer several advantages over linear models, such as automatic adaptation to nonlinearity. Regarding implementation, BART is also preferred due to its fast computation, good performance of

[5]Thorough reviews and well-designed simulations could be found in Korobilis and Shimizu (2022), Van Erp et al. (2019) and Polson and Sokolov (2019), who advocate the merits of Bayesian sparsity-inducing priors in comparison to frequentist counterparts.

default choices of hyperparameters and available software (Linero and Antonelli, 2023). When there is sufficient covariate overlap, BART has been shown to outperform numerous Frequentist machine learning methods in prediction problems, including random forests. Hill (2011) further proposed the use of BART in causal inference and demonstrated its efficacy in flexibly modelling the response surface. To mitigate the regularisation-induced confounding issue (Hahn et al., 2018) when using a BART outcome model, Hahn et al. (2020) developed the Bayesian causal forest model, a BART-based approach that includes a fixed estimate of the propensity score for additional adjustment in the outcome model. This model yields excellent performance in estimating heterogeneous treatment effects, making BART a strong default choice for integrating Bayesian nonparametric methods into causal effect modelling. Subsequent studies by Spertus and Normand (2018) and Xu et al. (2018) employed BART models to fit the propensity score in the first stage, enabling the use of Bayesian propensity score analysis to estimate ATE (with high-dimensional data) and QTEs, respectively.

2.2.3 Double Robustness

This paper also adds to the development of doubly robust estimators, which have gained extensive use in the causal inference literature (Bang and Robins, 2005; Scharfstein et al., 1999) owing to their desirable property of providing consistent inference even under misspecification of either the treatment assignment or outcome regression models (but not both). For a comprehensive survey of doubly robust estimators and their properties, we refer interested readers to Daniel (2014).

Doubly robust estimators have been extended to accommodate nonparametric or high-dimensional settings by enabling data-adaptive estimation of treatment and outcome models. This includes doubly robust estimators with the group Lasso (Farrell, 2015), double machine learning estimators (Chernozhukov et al., 2018), doubly robust matching estimators (Antonelli et al., 2018) and targeted maximum likelihood estimators (Van der Laan et al., 2011), among others. In these complex settings, doubly robust estimators offer an extra benefit that parametric convergence rates ($\sqrt{N}$) can be achieved even when each of the propensity score or outcome regression models converges at slower rates ($\sqrt[4]{N}$ or faster). Roughly speaking, this echoes the insights from recent advances in the double machine learning literature (Chernozhukov et al., 2022). Specifically, penalising either the propensity score model or the outcome model alone would be insufficient for valid causal inference, but combining the two as nuisance functions achieves a desirable convergence rate and finite-sample performance in

high-dimensional causal analyses.

There have been attempts to propose doubly robust Bayesian recipes, however, this area is still less established than its Frequentist counterpart. This is mainly due to a lack of consensus on propensity score adjustment in the Bayesian causal modelling framework (Li et al., 2023; Robins et al., 2015; Robins and Ritov, 1997), despite its central role being well recognised in the literature (Rosenbaum and Rubin, 1983; Zigler, 2016). Pioneering work on the Bayesian approach for doubly robust causal inference was done by Saarela et al. (2016), in which the authors formalised the problem and addressed it by combining the posterior predictive distribution of parameters with the Bayesian bootstrap. The idea of utilising posterior predictive distribution was later advanced in the line of work by Antonelli et al. (2022) and Shin and Antonelli (2023), who aimed to improve inference for doubly robust estimators for the average treatment effect (ATE) and the conditional average treatment effect (CATE), respectively. The general strategy involves estimating both the propensity score and the conditional outcome mean using Bayesian methods. Posterior draws from their respective posterior predictive distributions are then obtained and plugged into a doubly robust estimator. While this approach is not fully Bayesian because there is no joint likelihood for all parameters stated, it effectively integrates Bayesian modelling techniques and Frequentist inferential procedures for causal analysis with proper uncertainty quantification. These features are particularly important in high-dimensional scenarios, where handling large numbers of covariates and quantifying uncertainty can be challenging. Recent theoretical developments have established Bernstein-von Mises (BvM) results when examining the asymptotic behaviours of Bayesian inference procedures for causal effects. Ray and Vaart (2020) attained BvM for ATE using propensity score-adjusted priors, however, that prior adjustment alone leaves a bias when smoothness is traded off between the propensity score and the conditional outcome mean functions. Breunig et al. (2025) made this issue explicit and introduced a posterior correction to eliminate the bias term, yielding a semiparametric BvM under double-robust smoothness conditions. Additionally, Yiu et al. (2025) demonstrated that using double robust estimands still requires posterior corrections for valid inference, even when their conditions for the BvM theorem impose Donsker properties on the propensity score and the conditional mean functions – a more restrictive assumption than the prior correction of Ray and Vaart (2020) and particularly, the double robust method of Breunig et al. (2025). A comparison between the double robust method and semiparametric Bayesian methods was also provided in Breunig et al. (2024) for average treatment effects on the treated

(ATT) in the Difference-in-Differences (DiD) context. Both satisfy the BvM, but only the regularized version achieves it under double robust smoothness conditions. In more regular, low-dimensional settings, the semiparametric Bayesian approach is very competitive, but its performance declines with increasing model complexity. These findings serve as an interesting motivation for doubly robust approaches combined with posterior corrections under more complex designs.

In spite of the progress made in Bayesian literature, most studies have focused on the doubly robust estimation of either unconditional or conditional average treatment effect. Our paper presents a distinctive contribution by concentrating on unconditional quantile treatment effects (QTE). This causal estimand is of independent interest as it offers a different and complementary approach to uncover treatment effect heterogeneity. Although the conditional average treatment effect (CATE) also characterises treatment effect heterogeneity, the effects vary across individuals or subgroups defined by observed characteristics. In contrast, unconditional QTE focuses on the effect heterogeneity of the treatment across different outcome ranks without conditioning on individual characteristics or covariates. Evaluating the impact on the entire outcome distribution of interest makes this approach particularly relevant to distributional concerns and inequality, offering valuable insights when used alongside other estimands, such as CATE, to understand the potential consequences of treatments and policies. With regard to methodology, we build on the work of Antonelli et al. (2022), who combined the posterior predictive distribution of nuisance parameters with the Frequentist doubly robust estimator initially proposed for the ATE. We develop a Bayesian Analogue of Doubly Robust estimators for the QTE, tackling unique challenges that arise in the quantile setting. First, we address the need for different doubly robust estimators for the QTE compared to the ATE setting by solving an estimating equation built upon an efficient influence function specifically tailored to quantile functionals. Second, the QTE estimation problem involves new nuisance parameters, including the entire conditional cumulative distribution function (CDF) of each potential outcome conditional on potentially high-dimensional covariates, which increases the computational complexity. We overcome this hurdle by employing multiple Bayesian quantile regressions that incorporate shrinkage priors. This helps us explicitly estimate the conditional distribution while accounting for high dimensionality. This approach has not been pursued in previous studies, making our contribution unique in the literature.

2.3 Notation and Causal Estimand

2.3.1 Notation

Let T and Y be the treatment and outcome of interest, respectively, while $\mathbf{X}$ is a p-dimensional vector of potential controls. Denote $\mathcal{P}_o$ as the joint distribution of the observed data. Assume that we observe an independent and identically distributed (i.i.d.) sample $\mathbf{Z}_i = \{Y_i, T_i, \mathbf{X}_i\}$ for $i = 1, \ldots, N$ with empirical distribution $\mathcal{P}_N$, where we collect all observations into $\mathbf{Z} = (\mathbf{Z}_1, \ldots, \mathbf{Z}_N)$. For $t \in \{0, 1\}$, let

- $Y^{(t)}$ be the potential outcome for a generic subject under treatment t.
- $F_t(y) \coloneqq \mathbb{P}[Y^{(t)} \leq y]$ be the cumulative distribution function (CDF) of $Y^{(t)}$, and $q_t(\tau) \coloneqq F_t^{-1}(\tau) = \inf\{y \mid F_t(y) \geq \tau\}$ be its τ^{th} quantile, where $\tau \in (0, 1)$.
- $F_{t|1}(y) \coloneqq \mathbb{P}[Y^{(t)} \leq y \mid T = 1]$ be the cumulative distribution function (CDF) of $Y^{(t)}$ given $T = 1$, and $q_{t|1}(\tau) \coloneqq F_{t|1}^{-1}(\tau) = \inf\{y \mid F_{t|1}(y) \geq \tau\}$ be its τ^{th} quantile, where $\tau \in (0, 1)$.

2.3.2 Causal Estimand

Quantile Treatment Effects (QTEs) are defined as the difference between the τ^{th} quantiles (for a particular value of τ) of the treated potential outcome distribution and the untreated potential outcome distribution. For $\tau \in (0, 1)$,

$$QTE(\tau) \coloneqq F_1^{-1}(\tau) - F_0^{-1}(\tau) = q_1(\tau) - q_0(\tau). \tag{2.1}$$

For identification, a fundamental issue is whether each of the potential outcome distributions, $F_1(y)$ and $F_0(y)$, is identified. We therefore make the following assumptions in our setup:

1. *Unconfoundedness (Selection-on-Observables)*

$$Y^{(1)}, Y^{(0)} \perp\!\!\!\perp T|\mathbf{X}, \tag{2.2}$$

where $\perp\!\!\!\perp$ denotes statistical independence.

2. *Strong Overlap*

$$\exists \delta \in \mathbb{R}: \quad 0 < \delta < \mathbb{P}[T = 1 \mid \mathbf{X}] < 1 - \delta < 1. \tag{2.3}$$

3. *Stable Unit Treatment Value Assumption (SUTVA)*

$$T_i = t \text{ implies } Y_i^{obs} = Y_i^{(t)}, \quad \text{for } t \in \{0, 1\}. \tag{2.4}$$

The conditional distributions of potential outcomes are therein identified by determining the conditional distribution of observed outcomes for individuals within each group, as expressed by: $\mathbb{P}[Y(t) \leq y \mid \mathbf{X}] = \mathbb{P}[Y \leq y \mid T = t, \mathbf{X}]$. Consequently, the marginal distribution of potential outcomes can be identified and calculated as

$$F_t(y) = \int \mathbb{P}[Y \leq y \mid T = t, \mathbf{X} = \mathbf{x}] dF_{\mathbf{X}}(\mathbf{x}), \quad \text{for } t \in \{0, 1\}, \tag{2.5}$$

where $F_{\mathbf{X}}(\mathbf{x})$ is the marginal distribution of covariates $\mathbf{X}$ in the population of interest.

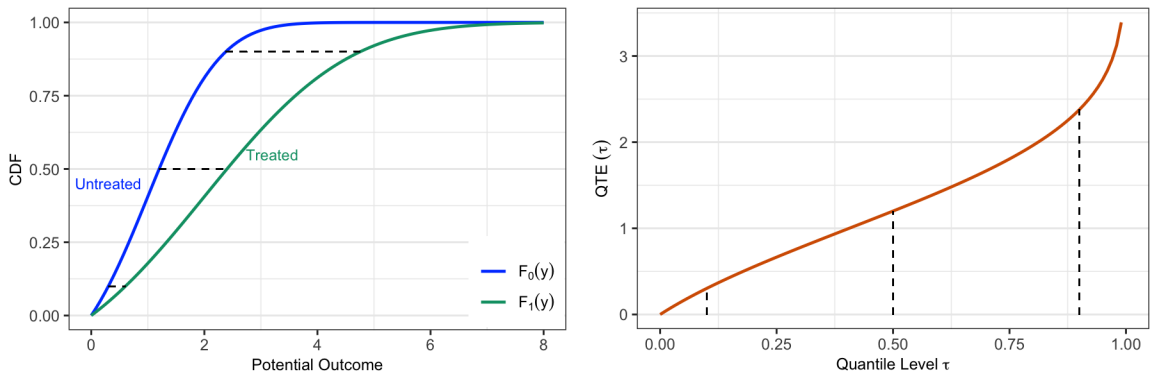

Figure 2.1: Illustration of Quantile Treatment Effects (QTEs). The left-hand figure demonstrates unconditional distributions of Treated and Untreated potential outcomes, which are colored in green and blue, respectively. The horizontal distance between these two distributions yields QTEs. For instance, QTE(0.95), QTE(0.5), and QTE(0.05) are represented by three dashed lines in the figure. QTEs across all values of quantile levels are plotted in the right-hand figure.

2.4 Proposed Estimation Approach

2.4.1 Justification

With the primary parameter of interest being Quantile Treatment Effects (QTE), we develop the Bayesian Analogue of Doubly Robust (BADR) estimation framework for this target causal estimand. Our approach draws inspiration from the work by Antonelli et al. (2022), originally proposed for the Average Treatment Effect (ATE), to combine Bayesian modelling methods for treatment assignment and outcome models with Frequentist doubly robust estimators using posterior predictive distributions. By tackling extra unique challenges that arise in the quantile setting, our framework aims to offer enhanced finite-sample performance without strict reliance on correct model specifications.

The implementation procedure is straightforward as follows:

1. Specify separate Bayesian treatment assignment and outcome models;
2. Draw the propensity score and the conditional distribution of each potential outcome from their corresponding posterior predictive distributions; and
3. Plug these values into a doubly robust estimator associated with the parameter of interest.

While our estimation approach is applicable in general, it is particularly useful in handling high dimensionality, addressing challenges posed by potentially large numbers of controls and the involvement of the entire conditional potential outcome distribution as a nuisance parameter. In subsection 2.4.2, we propose a modelling framework in high dimensions that flexibly incorporates Bayesian regularisation techniques. Before doing so, we first establish a foundation by defining the doubly robust estimator of QTE. Then, we demonstrate a promising avenue for estimation and inference within a Bayesian framework.

2.4.1.1 The Doubly Robust Estimator for Quantile Treatment Effects

Let $\pi(\mathbf{X}) \coloneqq \mathbb{P}(T = 1 \mid \mathbf{X}; \Theta_\pi)$ be the propensity score (i.e., the probability of receiving active treatment given covariates $\mathbf{X}$), which is associated with the treatment assignment model; and $G(y \mid t, \mathbf{X}) \coloneqq \mathbb{P}[Y \leq y \mid T = t, \mathbf{X}; \Theta_G]$ (for $t \in \{0, 1\}$) be the conditional distributions of Y given the treatment status and covariates, which is associated with the outcome model. Let $\Theta = \Theta_\pi \cup \Theta_G$ represent the parameters of the treatment assignment and outcome models. A general QTE estimation problem involves $\pi(\mathbf{X})$ and $G(y \mid t, \mathbf{X})$ as

nuisance functions. We denote $\hat{\pi}(\mathbf{X})$ and $\hat{G}(y \mid t, \mathbf{X})$ as estimators of $\pi(\mathbf{X})$ and $G(y \mid t, \mathbf{X})$ (for $t \in \{0, 1\}$), respectively.

For a chosen quantile level $\tau \in (0, 1)$, a doubly robust (DR) estimator of the QTE for binary treatments is given by

$$\widehat{QTE}^{dr}(\tau) = \hat{q}_1^{dr}(\tau) - \hat{q}_0^{dr}(\tau), \tag{2.6}$$

where for a sample of size N

- $\hat{q}_1^{dr}(\tau)$ is a DR estimator of the τ-quantile of *treated* potential outcome and can be derived as the solution to

$$N^{-1} \sum_i \left\{ \frac{T_i}{\hat{\pi}(\mathbf{X}_i)} \left[\mathbb{1}\{Y_i \leq q_1\} - \hat{G}(q_1 \mid 1, \mathbf{X}_i) \right] + \hat{G}(q_1 \mid 1, \mathbf{X}_i) \right\} = \tau. \tag{2.7}$$

- $\hat{q}_0^{dr}(\tau)$ is a DR estimator of the τ-quantile of *untreated* potential outcome and can be derived as the solution to

$$N^{-1} \sum_i \left\{ \frac{1 - T_i}{1 - \hat{\pi}(\mathbf{X}_i)} \left[\mathbb{1}\{Y_i \leq q_0\} - \hat{G}(q_0 \mid 0, \mathbf{X}_i) \right] + \hat{G}(q_0 \mid 0, \mathbf{X}_i) \right\} = \tau. \tag{2.8}$$

Formal derivation and discussion regarding $\hat{q}_1^{dr}(\tau)$ are presented in Appendix A.1, where (A.14) is equivalent to equation (2.7). Estimating equations (2.7) and (2.8) are built upon the efficient influence function for quantiles of each potential outcome distribution. The efficient influence function captures the first-order sensitivity of the target parameter to small perturbations in the underlying distributions. In our estimation problem, which involves two nuisance models – treatment assignment and outcome – the efficient influence function exhibits a double robustness property, resulting in a doubly robust estimator. This estimator is consistent provided that either the propensity score $\hat{\pi}(\mathbf{X})$ or the conditional outcome distribution $\hat{G}(y \mid t, \mathbf{X})$ is consistent, but not necessarily both.

Estimating equations (2.7) and (2.8) are also closely connected to the Neyman orthogonal moment conditions, which is extensively leveraged in the debiased machine learning literature (see, e.g., Belloni et al., 2017; Chernozhukov et al., 2018; Kallus et al., 2024). Neyman orthogonality is a desirable property that ensures the final estimate of the target parameter remains robust even when there are small errors in the estimation of nuisance parameters. This property is particularly relevant when regularisation methods are needed to handle

high-dimensional covariates in estimating the nuisances. In such cases, employing Neyman orthogonal moment conditions helps correct for the first-order biases that may arise from plugging in estimates of the nuisance parameters.

2.4.1.2 The Bayesian Analogue of Doubly Robust Estimator

The population parameters Θ are typically unknown and need to be estimated. In this setting, we consider a Bayesian framework to estimate the parameters associated to both the treatment assignment and outcome models. This enables uncertainty in parameter estimation to be directly captured from the posterior distribution.

Let $\mathbb{P}_{\Theta|\mathbf{Z}}$ denote the posterior distribution, and $\{\Theta^{(b)}\}_{b=1}^{B}$ be a sequence of B draws obtained from this posterior distribution. The point estimator $\hat{\Delta}$ for the estimand of interest Δ takes a form of the posterior mean

$$\hat{\Delta} := \mathbb{E}_{\Theta|\mathbf{Z}}[\Delta(\mathbf{Z}, \Theta)] \approx B^{-1} \sum_{b} \Delta(\mathbf{Z}, \Theta^{(b)}), \tag{2.9}$$

where $\Delta(\mathbf{Z}, \Theta^{(b)})$ is evaluated using the observed data $\mathbf{Z}$ and parameters $\Theta^{(b)}$. Therefore, our point estimator for the QTE at a chosen quantile level $\tau \in (0, 1)$ is the average value of the quantity in (2.6) with respect to the posterior distribution of model parameters.

Regarding inference, our variance of interest corresponds to the variance of the estimator's sampling distribution and can be defined as follows

$$\mathbb{V}_{\mathbf{Z}}(\hat{\Delta}) := \mathbb{V}_{\mathbf{Z}}(\mathbb{E}_{\Theta|\mathbf{Z}}[\Delta(\mathbf{Z}, \Theta)]). \tag{2.10}$$

Variance estimation can be implemented using the nonparametric bootstrap (Tibshirani and Efron, 1993) to account for uncertainty in all stages of the estimator. Specifically, L multiple datasets $\{\mathbf{Z}^{(l)}\}_{l=1}^{L}$ are created by sampling with replacement from the empirical distribution of the data. For each resampled dataset, we re-estimate the posterior distribution of Θ and then recalculate $\mathbb{E}_{\Theta|\mathbf{Z}}[\Delta(\mathbf{Z}, \Theta)]$ accordingly. Finally, we compute variance of this quantity across all bootstrap samples. It is worth noting that there are two main sources of uncertainty arising throughout our QTE estimation procedure: first, the sampling variability stemming from the data even if we know the true outcome and treatment assignment models; second, the variability in parameter estimation for the propensity score and conditional outcome

distributions. An alternative inference scheme from Antonelli et al. (2022) can be adopted by targeting these two parts separately

$$\mathbb{V}_{\mathbf{Z}}(\hat{\Delta}) = \underbrace{\mathbb{V}_{\mathbf{Z}^{(l)}}\{\mathbb{E}_{\Theta|\mathbf{Z}}[\Delta(\mathbf{Z}^{(l)}, \Theta)]\}}_{\substack{\text{uncertainty stemming} \\ \text{from the data}}} + \underbrace{\mathbb{V}_{\Theta|\mathbf{Z}}[\Delta(\mathbf{Z}, \Theta)]}_{\substack{\text{uncertainty in estimation} \\ \text{of nuisance parameters}}} . \tag{2.11}$$

The first term resembles the true variance, except for the outer moment, which is associated with a resampled version $\mathbf{Z}^{(l)}$ of the original data $\mathbf{Z}$. The posterior samples of parameters Θ estimated using the original data are maintained, but the point estimator $\mathbb{E}_{\Theta|\mathbf{Z}}[\Delta(\mathbf{Z}^{(l)}, \Theta)]$ is recalculated for each resampled dataset. In this way, it captures only the uncertainty of the data, not that resulting from parameter estimation. In contrast, the second term accounts for the latter type of uncertainty based on the variability of the full posterior samples of $\Delta(\mathbf{Z}, \Theta)$ given the observed data $\mathbf{Z}$.

2.4.2 Modelling Framework in High Dimensions

At this juncture, we present the high-dimensional modelling framework we use to provide scalable estimation algorithms. We employ Bayesian techniques to specify and estimate the treatment assignment and outcome models separately. To address the challenges posed by high-dimensional feature spaces, we integrate various Bayesian regularisation methods into our proposed framework to yield estimators of the nuisance functions corresponding to the propensity score and the conditional distribution of the outcome, denoted as $\hat{\pi}(\mathbf{X})$ and $\hat{G}(y \mid t, \mathbf{X})$, respectively. For the treatment assignment model, we adopt Bayesian Additive Regression Trees (BART) priors, whose merits have been increasingly recognised (see, e.g., Chipman et al., 2012; Hahn et al., 2020; Hill, 2011; Linero and Antonelli, 2023). For the outcome model, we leverage multiple Bayesian quantile regressions combined with shrinkage priors to explicitly estimate the conditional distribution while accommodating potentially high-dimensional covariates. This combined strategy allows us to develop a flexible and doubly-robust Bayesian estimator with desirable finite-sample frequentist properties. To the best of our knowledge, this modelling framework has not been previously pursued in the literature, resulting in a novel approach to estimation of heterogeneous treatment effects.

2.4.2.1 Treatment Assignment Model

We fit a binary Bayesian Additive Regression Trees (BART) model on the observations $\{T_i, \mathbf{X}_i\}_{i=1}^n$ to model the regression of the treatment assignment on control variables, that is,

$$\pi(\mathbf{X}_i) = \mathbb{P}\left(T_i = 1 \mid \mathbf{X}_i\right) = H\left[f_{\text{BART}}\left(\mathbf{X}_i\right)\right], \tag{2.12}$$

where the link function H is either the CDF of the standard normal distribution for probit BART or the CDF of the logistic distribution for the logit BART, and

$$f_{\text{BART}}(\mathbf{X}_i) = \sum_{m=1}^{M} f_{\text{tree}}\left(\mathbf{X}_i; \Gamma_m, \mu_m\right) \text{ are sum of } M \text{ Bayesian regression trees.}$$

For $m \in \{1, \ldots, M\}$, Γ_m is a tree structure that consists of a set of splitting rules and a set of terminal nodes; and $\mu = (\mu_{m,1}, \ldots, \mu_{m,b_m})$ is a vector of parameters associated with b_m terminal nodes of Γ_m, such that $f_{\text{tree}}(\mathbf{X}_i; \Gamma_m, \mu_m) = \mu_{m,l}$ if $\mathbf{X}_i$ is corresponding to the l^{th} terminal node of Γ_m.

The modelling choices used to implement the BART specification are presented in Appendix A.2. Once a sequence of B posterior draws for the underlying BART parameters has been obtained, B posterior samples of the fitted propensity score $\{\pi^{(b)}(\mathbf{X})\}_{b=1}^B$ can be calculated by

$$\pi^{(b)}(\mathbf{X}_i) = H\left[\sum_{m=1}^{M} f_{\text{tree}}\left(\mathbf{X}_i; \Gamma_m^{(b)}, \mu_m^{(b)}\right)\right] \quad \text{for } i = 1, \ldots, N \text{ and } b = 1, \ldots, B. \tag{2.13}$$

2.4.2.2 Outcome Model

In contrast to the literature's extensive coverage of conditional expectation estimation, data-adaptive estimation of conditional distributions has received considerably less attention. For each $t \in \{0, 1\}$, we estimate the conditional distribution based on fitted conditional quantiles, employing the sample analogue of the following alternative representation of the conditional distribution

$$F_{Y|\mathbf{X}}(y) = \int_0^1 \mathbb{1}\left\{F_{Y|\mathbf{X}}^{-1}(\tau) \leq y\right\} d\tau = \int_0^1 \mathbb{1}\left\{\mathcal{Q}_{Y|\mathbf{X}}(\tau) \leq y\right\} d\tau, \tag{2.14}$$

where $F_{Y|\mathbf{X}}(\cdot)$ and $\mathcal{Q}_{Y|\mathbf{X}}(\cdot)$ are conditional distribution and conditional quantiles, respectively.

The corresponding estimator is

$$\begin{aligned}\hat{F}_{Y|\mathbf{X}}(y) &= \int_0^1 \mathbb{1}\left\{\hat{\mathcal{Q}}_{Y|\mathbf{X}}(\tau) \leq y\right\} d\tau \\ &\approx \epsilon + \int_\epsilon^{1-\epsilon} \mathbb{1}\left\{\hat{\mathcal{Q}}_{Y|\mathbf{X}}(\tau) \leq y\right\} d\tau \\ &\approx \epsilon + \sum_{s=1}^{S} \delta_s \mathbb{1}\left\{\hat{\mathcal{Q}}_{Y|\mathbf{X}}(\tau_s) \leq y\right\},\end{aligned} \tag{2.15}$$

where $\hat{\mathcal{Q}}_{Y|\mathbf{X}}(\tau_s) = \mathbf{X}^\top \hat{\beta}_{(\tau_s)}$ can be obtained by estimating S Bayesian quantile regression model for each $\{\tau_s\}_{s=1}^S$, where $\epsilon \leq \tau_0 < \ldots < \tau_s \leq 1 - \epsilon$ and the width $\delta_s = \tau_s - \tau_{s-1} \to 0$ as $S \to \infty$. The second equation is adapted for tail trimming. The third equation aims to avoid estimating the whole quantile regression process. Our discretisation technique is similar to some previous studies (Belloni et al., 2017; Chernozhukov et al., 2013; Frölich and Melly, 2013), however, we use the Bayesian quantile regression model rather than Koenker and Bassett's (1978) quantile regression from a frequentist viewpoint.

There are several advantages of this computational approach to the problem of estimating conditional distributions. First, it enables us to leverage the Bayesian quantile regression model, which not only suits our overall framework but also offers more flexibility than its frequentist counterpart. Especially, Bayesian shrinkage priors can be readily applied to this parametric quantile model with minor modifications, thereby handling better high-dimensional covariates. This feature is thoroughly reviewed by Korobilis and Shimizu (2022). In addition, while a very fine grid for values of τ_s (i.e., large S) is often required to gain accuracy, we can make use of parallel computation because the conditional posteriors are applied in each quantile level independently. While crossing or non-monotonic estimated quantiles are a valid concern when the regression for each quantile is estimated separately[6], the algorithm presented above is originally designed for the rearrangement of crossing quantiles. This ensures that our primary objective of interest, the conditional distribution, remains unaffected by these potential estimation issues.

Further discussion on the Bayesian shrinkage priors and Bayesian quantile regression can be found in Appendices A.3 and A.4, respectively. By drawing a sequence of B posterior draws for the quantile regression parameters, we can obtain B posterior samples of the fitted conditional outcome distributions $\{G^{(b)}(y \mid 0, \mathbf{X})\}_{b=1}^B$ and $\{G^{(b)}(y \mid 1, \mathbf{X})\}_{b=1}^B$.

[6]The estimated conditional quantile functions may be non-monotonic in the sense that $\bar{\tau} > \tilde{\tau}$ does not necessarily imply $\hat{\mathcal{Q}}_{Y|\mathbf{X}}(\bar{\tau}) > \hat{\mathcal{Q}}_{Y|\mathbf{X}}(\tilde{\tau})$.

2.4.2.3 Algorithms for the Bayesian Analogue of Doubly Robust (BADR) Estimation

Upon acquiring sequences of B posterior draws of the fitted propensity score $\{\pi^{(b)}(\mathbf{X})\}_{b=1}^{B}$ and the fitted conditional outcome distributions $\{G^{(b)}(y \mid t, \mathbf{X})\}_{b=1}^{B}$ for $t \in \{0, 1\}$, we can compute B values of the corresponding Quantile Treatment Effect (QTE) based on the full posterior distribution of these nuisance parameters. The BADR point estimate of the QTE used in this paper is derived as the average of these B values. The details of the implementation are presented in Algorithm 2.1. Utilising B posterior samples of QTE, variance estimation can be proceeded according to (2.11).

It is noteworthy that following Algorithm 2.1 requires solving two estimation equations, (2.7) and (2.8), B times. This step may lead to intensive computations, particularly when bootstrapping is involved. An alternative approach to combining the posterior distribution of model parameters and the doubly robust estimator in Algorithm 2.1 would replace the nuisance parameters $\pi(\mathbf{X})$ and $G(y \mid t, \mathbf{X})$ with plug-in estimates, such as their posterior means, as outlined in Algorithm 2.2. This aligns with a frequentist modelling approach where doubly robust estimators are evaluated using plug-in estimates of the parameters Θ. While Algorithm 2.2 uses more compact information, there is a clear computational gain due to the fact that estimation equations only need to be solved once. The inference procedure is conducted using the original bootstrap variance estimation in (2.10) for ease of implementation. Furthermore, our pilot Monte Carlo findings suggest that the alternative estimator yields similar results.

Algorithm 2.1:

Bayesian Analogue of Doubly Robust (BADR) estimation for QTEs

(*Full posterior samples*)

Data: $\{Y_i, T_i, \mathbf{X}_i\}_{i=1}^n, \tau \in (0,1)$

Result: $\widehat{QTE}^{dr}(\tau)$

Fit treatment assignment model on $\{T_i, \mathbf{X}_i\}_{i=1}^n$ and obtain B posterior samples $\{\pi^{(b)}(\mathbf{X})\}_{b=1}^B$.

for $t = 0, 1$ **do**

  Fit outcome model on $\{Y_i, \mathbf{X}_i\}_{i:T_i=t}$ and obtain B posterior samples $\{G^{(b)}(y \mid t, \mathbf{X})\}_{b=1}^B$.

end

for $b = 1, \ldots, B$ **do**

  Solve $q_1^{(b)}(\tau), q_0^{(b)}(\tau)$ based on $\pi^{(b)}(\mathbf{X})$ and $G^{(b)}(y \mid t, \mathbf{X})$. ▷ (2.7) and (2.8)

  Calculate $QTE^{(b)}(\tau) = q_1^{(b)}(\tau) - q_0^{(b)}(\tau)$. ▷ (2.6)

end

Calculate $\hat{\Delta}_\tau \coloneqq \widehat{QTE}^{dr}(\tau) = \frac{1}{B}\sum_{b=1}^B QTE^{(b)}(\tau)$. ▷ (2.9)

Algorithm 2.2:

Bayesian Analogue of Doubly Robust (BADR) estimation for QTEs

(*Posterior means*)

Data: $\{Y_i, T_i, \mathbf{X}_i\}_{i=1}^n, \tau \in (0,1)$

Result: $\widehat{QTE}^{dr}(\tau)$

Fit treatment assignment model on $\{T_i, \mathbf{X}_i\}_{i=1}^n$ and obtain B posterior samples $\{\pi^{(b)}(\mathbf{X})\}_{b=1}^B$.

for $t = 0, 1$ **do**

  Fit outcome model on $\{Y_i, \mathbf{X}_i\}_{i:T_i=t}$ and obtain B posterior samples $\{G^{(b)}(y \mid t, \mathbf{X})\}_{b=1}^B$.

end

Derive posterior mean from B posterior samples

  $\hat{\pi}(\mathbf{X}) = \frac{1}{B}\sum_{b=1}^B \pi^{(b)}(\mathbf{X})$ and $\hat{G}(y \mid t, \mathbf{X}) = \frac{1}{B}\sum_{b=1}^B G^{(b)}(y \mid t, \mathbf{X})$. ▷ (2.9)

Solve $\hat{q}_1^{dr}(\tau), \hat{q}_0^{dr}(\tau)$ based on $\hat{\pi}(\mathbf{X})$ and $\hat{G}(y \mid t, \mathbf{X})$. ▷ (2.7) and (2.8)

Calculate $\hat{\Delta}_\tau \coloneqq \widehat{QTE}^{dr}(\tau) = \hat{q}_1^{dr}(\tau) - \hat{q}_0^{dr}(\tau)$. ▷ (2.6)

2.5 Simulation Study

We assess the finite-sample performance of our proposed approach, Bayesian Analogue of Doubly Robust (BADR) estimation, in two simulations with details described below. For each simulation, we specify the distribution of covariates, the treatment assignment mechanism and the distribution of potential outcomes. The first simulation focuses on a linear setting with varying covariate dimensionality to sample size ratio (p/N). In the second simulation, we consider a nonlinear setting and further examine the double robustness of our proposed estimators. Both of these data designs imply that assignment to the treatment is not completely random, but satisfies the *selection-on-observables* assumption. From a theoretical perspective, estimation of treatment effects that fails to account for the selection problem will inevitably produce inconsistent estimates. We regard this approach as a benchmark and consider the `Naive` estimator, which is an estimator of simple differences between empirical quantiles of treated and control groups, without any correction for selection bias.

We develop two versions of the estimators which represent our proposed methodology – Bayesian Doubly Robust estimator (`BDR`) and an extension that adds shrinkage priors (`BDRS`). Specifically, the former employs the original Bayesian Quantile Regression while the latter incorporates the Adaptive Lasso in order to account for sparsity and uncertainty in the outcome model. Both estimators fit the propensity score using a logit BART model in the first step. Furthermore, we also compare our proposed method with existing estimators. The Bayesian nonparametric counterpart (`BNP`) is a fully Bayesian approach developed in Xu et al. (2018), where the propensity score is estimated using a logit BART, then the conditional distribution of the potential outcome given a BART posterior sample of the propensity score in each treatment group is estimated separately using a Dirichlet process mixture of multivariate normals. We additionally compare three frequentist methods – the Localized Debiased Machine Learning (`LDML`) method introduced in Kallus et al. (2024), the Targeted Maximum Likelihood Estimation (`TMLE`) method proposed in Díaz (2017), and Firpo's Inverse Probability Weighted (`FIPW`) method developed in Firpo (2007). Among them, `LDML` and `TMLE` are two estimators that can leverage a variety of machine learning methods. Particularly, in our simulation exercise, Random Forest is incorporated into `LDML` and Lasso is integrated into `TMLE`. Implementation details of these methods can be found in Appendix A.5.1.

In each simulation design, we generate 100 synthetic datasets. For each simulated dataset,

we calculate quantile causal effects for 5 quantile levels, $\tau \in \{0.10, 0.25, 0.50, 0.75, 0.90\}$, and their 95% credible (or confidence) intervals (CIs). We compare all the different approaches in terms of average bias, Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE).

2.5.1 Simulation Design 1 (SD1)

We first consider a linear setting in which the mean of potential outcomes is a linear combination of covariates. We draw 40-dimensional covariates $\mathbf{X}$ ($p = 40$) from the independent standard normal distributions and allow different sample sizes $N \in \{100, 500, 1000\}$ of the dataset. Accordingly, we could evaluate the estimation procedure across varying feature dimensionality (i.e., p/N ratio). The exact form of the true model used to generate synthetic data is as follows:

$$
\begin{aligned}
T \mid \mathbf{X} &\sim \mathcal{B}ern\left(\pi(\mathbf{X})\right), \\
Y^{(0)} \mid \mathbf{X} &\sim \mathcal{N}\left(\mu(\mathbf{X}), 2.5^2\right), \\
Y^{(1)} \mid \mathbf{X} &\sim \mathcal{N}\left(1 + \mu(\mathbf{X}), 3.75^2\right), \\
Y &= T \times Y^{(1)} + (1 - T) \times Y^{(0)}; \\
\text{where} \quad \pi(\mathbf{X}) &= \left\{1 + \exp\left[-(X_1 + X_2 + X_3)\right]\right\}^{-1}, \\
\mu(\mathbf{X}) &= X_1 + X_2 + X_4 + X_5.
\end{aligned}
\tag{2.16}
$$

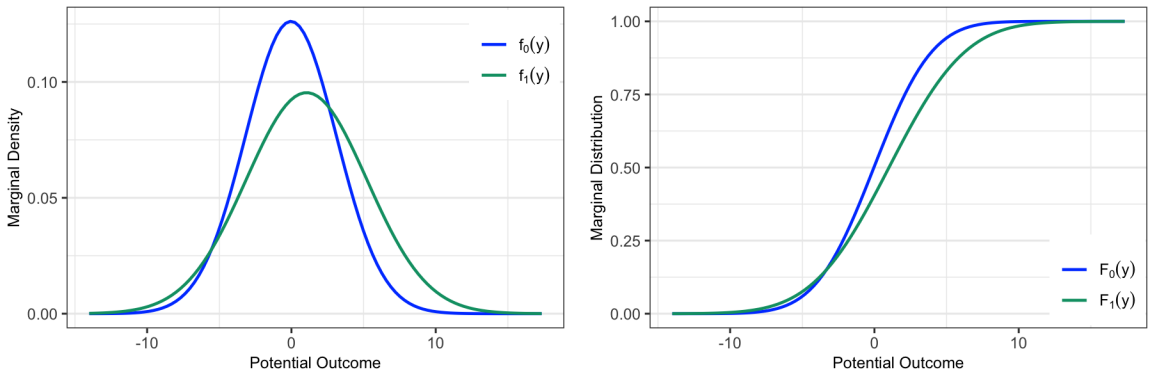

Figure 2.2: True marginal densities and marginal distributions of the treated and untreated potential outcomes in SD1. This design emulates a thought experiment relevant to policy evaluation literature. Hypothetically assigning the entire population either to treatment or to control induces a change in both location and shape of the outcome distribution.

Under this specification, the unconditional distribution of potential outcomes are $Y^{(0)} \sim \mathcal{N}(0, 10.25)$ and $Y^{(1)} \sim \mathcal{N}(1, 18.0625)$. Figure 2.2 provides a visual illustration of the corresponding marginal densities and marginal distributions. As a result, the population

quantile treatment effects can be computed analytically. In particular, the true 10^{th}, 25^{th}, 50^{th}, 75^{th} and 90^{th} QTEs are $\Delta_{0.10} = (-4.447) - (-4.103) = -0.344$, $\Delta_{0.25} = (-1.866) - (-2.159) = 0.293$, $\Delta_{0.5} = 1 - 0 = 1$, $\Delta_{0.75} = 3.866 - 2.159 = 1.707$, and $\Delta_{0.90} = 6.447 - 4.103 = 2.344$, respectively.

Table 2.1: Comparison of point estimates for QTEs across 100 replicates $(N = 1000, p = 40)$

	Percentiles				
	10th	25th	50th	75th	90th
True QTEs	**-0.34**	**0.29**	**1.00**	**1.71**	**2.34**
BDR	-0.37	0.30	0.95	1.64	2.30
	(-1.20, 0.46)	(-0.38, 0.97)	(0.37, 1.53)	(1.00, 2.27)	(1.50, 3.11)
BDRS	-0.34	0.31	0.95	1.65	2.34
	(-1.16, 0.48)	(-0.34, 0.96)	(0.38, 1.53)	(1.03, 2.27)	(1.56, 3.12)
BNP	0.92	1.55	2.24	2.93	3.59
	(0.24, 1.58)	(1.01, 2.09)	(1.74, 2.74)	(2.39, 3.48)	(2.91, 4.25)
LDML	0.29	0.96	1.63	2.35	3.04
	(-0.74, 1.31)	(-0.08, 2.01)	(0.25, 3.01)	(-0.00, 4.70)	(-1.74, 7.82)
TMLE	-0.38	0.39	1.07	1.75	2.32
	(-1.63, 0.86)	(-0.44, 1.22)	(0.36, 1.78)	(0.94, 2.56)	(1.15, 3.49)
FIPW	-0.38	0.27	0.92	1.64	2.25
	(-1.71, 0.96)	(-0.93, 1.47)	(-0.18, 2.02)	(0.43, 2.85)	(0.88, 3.63)
Naive	0.94	1.58	2.25	2.97	3.65
	(0.14, 1.74)	(0.97, 2.19)	(1.67, 2.84)	(2.35, 3.59)	(2.86, 4.43)

Notes: 95% CIs in parentheses correspond to 95% confidence intervals in Frequentist approach or 95% posterior credible intervals in Bayesian approach. To estimate these 95% CIs, LDML and FIPW use analytical standard errors, whereas others rely on the bootstrap method.

Table 2.1 presents point estimates for the quantile treatment effects, along with the average lower and upper bounds of the corresponding 95% CIs across 100 simulated datasets. These computations are based on a sample size of $N = 1000$ and $p = 40$. It is clear that the `Naive` method exhibits substantial bias in its point estimates. This can be attributed to the absence of adjustment for confounders in $\mathbf{X}$, resulting in poor performance as expected. In comparison, all other methods considered in our current setting outperform the `Naive` method in terms of both bias and coverage, proving their effectiveness in correcting selection bias to some extent.

Our proposed estimators, `BDR` and `BDRS`, yield point estimates closest to the true values of QTEs. Notably, incorporating a shrinkage prior, as in `BDRS`, further enhances the performance

of BDR, particularly when the object of interest is extreme tails (i.e., 10^{th} and 90^{th} percentiles).

Despite sharing a probabilistic approach, the Bayesian nonparametric estimator, BNP, demonstrates differences from our proposed estimators. While the nonparametric method produces point estimates slightly better than Naive method, they are still far from the truth. Moreover, the 95% credible intervals associated with BNP fail to cover the true values of QTEs at any percentile. It aligns with the observation that BNP exhibits the smallest CI widths among all surveyed methods, posing challenges in achieving satisfactory coverage rates. It is worth noting that because both BDR and BNP use BART-logit to model the treatment assignment in the first stage, their distinct performance illustrates the role of modelling the conditional distribution of potential outcomes given confounders. Intuitively, BNP avoids directly modelling the conditional distribution of potential outcomes given confounders. Instead, it is grounded in the balancing property of the propensity score (Rosenbaum and Rubin, 1983) to model the conditional distribution of the outcome given the propensity score alone. While this approach involves estimating a less complex distribution due to having only one binary regressor (i.e., the estimated propensity score) in the second stage, it becomes skeptical in the case of misspecified treatment assignment. According to Monte Carlo results in the original paper by Xu et al. (2018), the inclusion of non-confounders in the treatment assignment equation entails less precise estimations of QTEs, thereby compromising the performance of the BNP method. Aside from efficiency loss, finite-sample bias is also a notable drawback of methods targeting a set of variables that best predict treatment assignment without accounting for how these variables are related to the outcome, as widely discussed in the context of average treatment effect (Belloni et al., 2014; Zigler and Dominici, 2014, etc.).

Among frequentist approaches, TMLE and FIPW perform reasonably well in terms of bias, although they do not surpass our proposed estimators. Whilst the bootstrapped standard errors of TMLE are smaller than the estimated asymptotic standard errors of FIPW[7], both methods provide corresponding 95% confidence intervals that contain the truth at any percentile. In contrast, LDML yields point estimates that are less favourable compared to TMLE and FIPW. However, its asymptotically calibrated confidence intervals still effectively capture the true QTEs, despite having the widest spans across all quantile levels.

Boxplots in Figure 2.3 offer more insight on the sampling distributions of the difference between the true and estimated quantities which produced by all estimators across 100

[7]Firpo (2007) also recommends bootstrapping as possibly a good alternative to analytical standard errors estimation in FIPW.

simulated datasets. When $N = 1000$, BDR and BDRS showcase nearly zero median (or mean) bias as well as small variation, outperforming other methods. Their strong performance persists even in smaller sample size of $N = 500$. Interestingly, the advantage of BDRS, which is developed by adopting a hierarchical shrinkage prior, becomes prominent when $N = 100$. While the performance of BDR exhibits instability in the presence of high-dimensional covariates, BDRS handles such settings more effectively, as evidenced by the remarkably reduced box widths observed for all QTEs of interest.

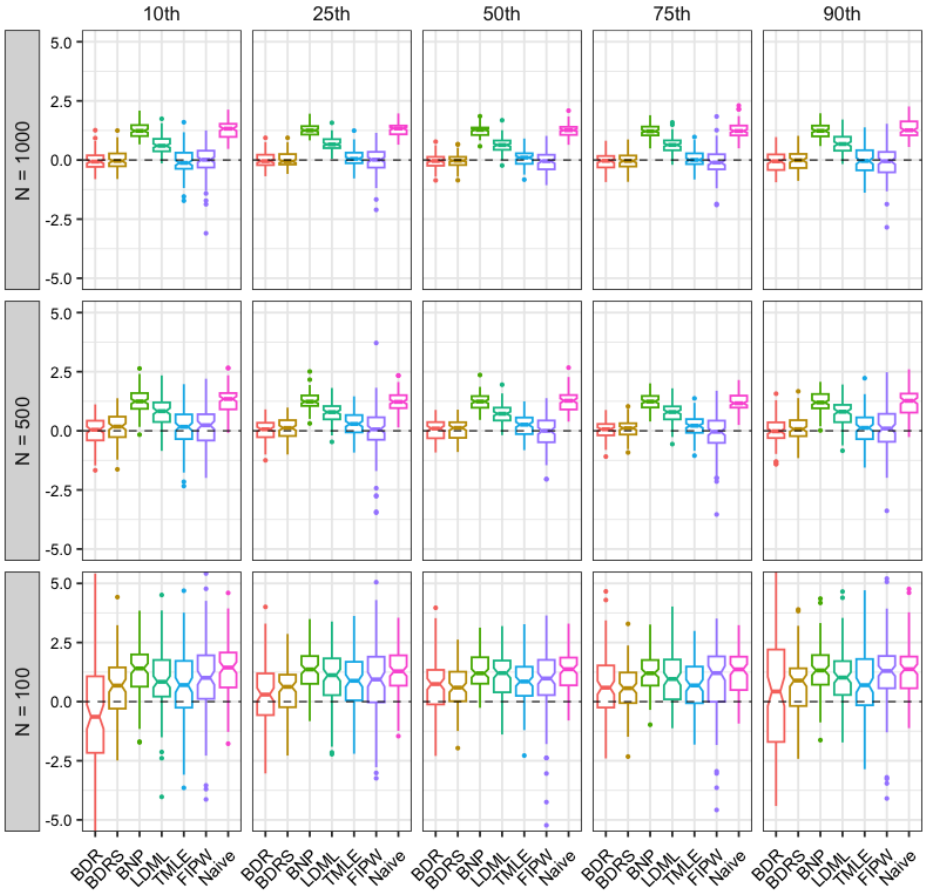

Figure 2.3: Sampling distributions of the difference between the true and estimated quantities for 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th} QTEs across 100 replicates. The dashed line indicates zero difference.

Table 2.2 numerically validates our above findings on the pattern of mean bias. Both BDR and BDRS alternately secure the top rank, exhibiting the smallest average bias across all computed percentiles. While the challenge of pronounced average bias is inherent in the high-dimensional setting ($N = 100$ and $p = 40$), as the p/N ratio decreases, the average bias diminishes relatively fast for QTEs estimated by methods BDR, BDRS, TMLE, and FIPW. Additionally, LDML also exhibits a declining trend in average bias, albeit at a slower rate. This phenomenon is not observed with the Naive and BNP estimators.

Table 2.2: Simulation Results for SD1, Average Bias

		Estimation Methods						
Percentiles	N	BDR	BDRS	BNP	LDML	TMLE	FIPW	Naive
10th	1000	-0.022	**0.001**	1.261	0.63	-0.041	-0.034	1.282
	500	**0.008**	0.153	1.261	0.794	0.121	0.102	1.269
	100	-0.659	**0.56**	1.267	0.928	0.724	0.901	1.398
25th	1000	**0.003**	0.017	1.261	0.669	0.1	-0.025	1.288
	500	0.047	0.103	1.245	0.764	0.266	**-0.035**	1.237
	100	**0.265**	0.49	1.32	1.03	0.793	0.794	1.296
50th	1000	-0.049	**-0.045**	1.24	0.632	0.071	-0.08	1.25
	500	**0.035**	0.053	1.209	0.718	0.215	-0.054	1.241
	100	0.696	**0.595**	1.284	1.073	0.85	0.852	1.295
75th	1000	-0.071	-0.057	1.228	0.641	**0.045**	-0.068	1.266
	500	**0.022**	0.071	1.192	0.743	0.172	-0.111	1.226
	100	0.809	**0.574**	1.24	0.941	0.673	0.822	1.179
90th	1000	-0.039	**-0.006**	1.247	0.696	-0.024	-0.089	1.302
	500	**-0.008**	0.096	1.21	0.736	0.094	0.07	1.206
	100	**0.564**	0.712	1.293	1.036	0.825	1.093	1.314

Notes: This table displays the average bias across 100 replicates of different estimation methods. The rows contain results for various percentile levels and for various sample size N.

With respect to the relative Mean Absolute Error (MAE), as presented in Table 2.3, our proposed approach outperforms all competitors at high percentiles 50^{th}, 75^{th}, and 90^{th}. Meanwhile, `BDRS` performs better than `BDR` in the majority of cases, especially in high-dimensional scenarios. In addition, line plots of raw MAE in Figure 2.4 illustrate a downward trend for both `BDR` and `BDRS` across all quantile levels as the p/N ratio decreases.

In conclusion, `BDR` and `BDRS` demonstrate similarly excellent performance in moderate dimensionality, thereby facilitating robustness checks in practical use. `BDRS` even provides extra merit thanks to its adaptation to high-dimensional settings. It is trivial that the Bayesian Adaptive Lasso in `BDRS` is only one option among a wide range of shrinkage priors which can be incorporated into our proposed framework. Thus, the results from this simulation exercise imply the great potential of our methodology in flexibly handling high dimensions when estimating quantile treatment effects.

Table 2.3: Simulation Results for SD1, Relative MAE

		Estimation Methods					
Percentiles	N	BDR	BDRS	BNP	LDML	TMLE	FIPW
10th	1000	1.67	1.645	**0.99**	1.169	1.741	1.741
	500	1.571	1.462	**0.964**	1.078	1.547	1.634
	100	2.449	1.194	**0.945**	1.123	1.272	1.346
25th	1000	0.999	0.996	0.969	**0.862**	0.978	1.077
	500	1.135	1.117	1.017	**0.947**	1.07	1.391
	100	1.182	0.986	**0.93**	1	0.992	1.107
50th	1000	**0.628**	**0.628**	0.994	0.739	0.64	0.688
	500	0.673	**0.666**	0.976	0.778	0.676	0.694
	100	0.876	**0.81**	0.964	0.943	0.876	1.012
75th	1000	**0.477**	0.479	0.981	0.704	0.513	0.523
	500	**0.547**	0.56	0.981	0.776	0.592	0.604
	100	0.927	**0.771**	1.004	0.912	0.81	0.992
90th	1000	**0.519**	0.529	0.979	0.771	0.539	0.534
	500	**0.572**	0.596	1.002	0.819	0.623	0.669
	100	1.109	**0.822**	0.988	0.919	0.851	1.06

Notes: This table displays the relative Mean Absolute Error (MAE) of different estimation methods across 100 replicates. The rows contain results for various percentile levels and for various sample size N. The relative MAE is the MAE in comparison with the Naive method as the benchmark, where $MAE = R^{-1}\sum_{r=1}^{R}|\hat{\alpha}_r - \alpha|$ and $R = 100$.

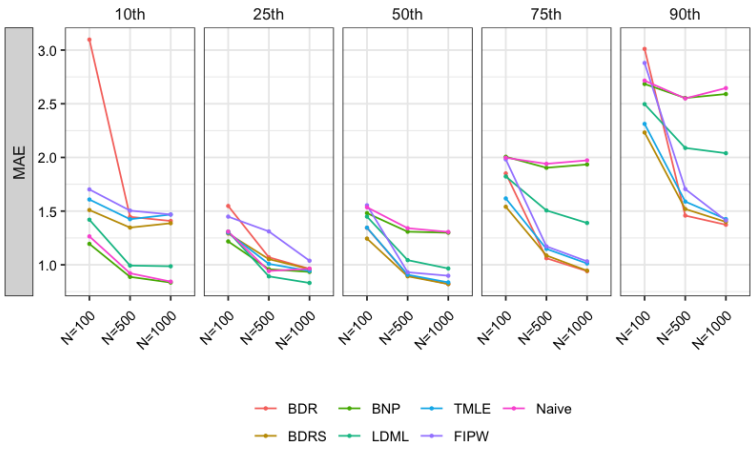

Figure 2.4: Line plots of raw MAE for 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th} QTEs estimated based on 100 replicates.

2.5.2 Simulation Design 2 (SD2)

In this simulation design, we explore a setting with nonlinearities, where the mean of potential outcomes involves polynomial functions of covariates. We draw covariates $\mathbf{X}$ with $p = 5$ from the independent standard normal distributions. The sample size is fixed at $N = 1000$ for the remainder of the exercise given small sample sizes are likely to be inadequate to explore nonlinearities. The true model for data generation takes the following form:

$$\begin{aligned} T \mid \mathbf{X} &\sim \mathcal{B}ern\left(\pi(\mathbf{X})\right), \\ Y^{(0)} \mid \mathbf{X} &\sim \mathcal{N}\left(\mu(\mathbf{X}), 1^2\right), \\ Y^{(1)} \mid \mathbf{X} &\sim \mathcal{N}\left(1 + \mu(\mathbf{X}), 1.5^2\right), \\ Y &= T \times Y^{(1)} + (1 - T) \times Y^{(0)}; \\ \text{where} \quad \pi(\mathbf{X}) &= \{1 + \exp\left[-(-0.6X_1 + 0.8X_2 + 1.2X_3)\right]\}^{-1}, \\ \mu(\mathbf{X}) &= -X_1 + X_2^2 + 1.5X_3X_4 + 1.5X_5^3. \end{aligned} \tag{2.17}$$

Unlike the first simulation study, the true unconditional density and the true quantiles of the potential outcomes for this simulation are not analytically achievable. However, the true unconditional quantiles can be derived approximately from a large sample. At sample size of 10^7, the approximate values for true 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th} QTEs are $\Delta_{0.10} = 0.705$, $\Delta_{0.25} = 0.794$, $\Delta_{0.5} = 1.022$, $\Delta_{0.75} = 1.205$, and $\Delta_{0.90} = 1.205$, respectively.

We introduce two simpler variants of our proposed framework in this simulation exercise. The first variant consists of Bayesian Outcome Modelling without and with shrinkage priors, represented by `BOM` and `BOMS` estimators, respectively. It could be regarded as an outcome-regression-based approach that omits the treatment assignment model fitted in the initial step of the BADR framework. Instead, it focuses solely on estimating the conditional distribution by using multiple Bayesian quantile regressions in the outcome model of each treatment group. Shrinkage priors, akin to the doubly robust approach, can be readily incorporated. In particular, the `BOMS` estimator considers the Adaptive Lasso prior. The second variant is Bayesian Propensity Score Analysis (`BPSA`), a treatment-assignment-based approach. Specifically, it involves fitting the treatment assignment using a logit BART model. Subsequently, it employs multiple Bayesian quantile regressions to model the conditional distribution of the outcome given the posterior mean of the propensity score in each treatment group. Further details on the implementation can be found in Appendix A.5.2.

We evaluate the performance of the methods with respect to two distinct modelling strategies: linear and nonlinear specification. For the linear specification, we use 5 raw covariates $X_1, \ldots, X_5$. For the nonlinear specification, we expand the covariates $\mathbf{X}$ to a 55-dimensional space by incorporating full cubic polynomials along with interaction terms. `BDR`, `BOM`, and `BNP` are excluded as competitors in the second specification since they are less suitable for high-dimensional contexts.

SD2a. Linear Specification

Table 2.4 illustrates simulation results when 5-dimensional covariates are employed as control variables. Overall, `BDR` and `BDRS` outperform all competing methods, achieving the smallest average bias at 25^{th} and 75^{th} percentiles. Frequentist methods including `TMLE`, `LDML`, and `FIPW`, individually rank first once at 10^{th}, 50^{th}, and 90^{th}, respectively. However, each of them is less superior to our proposed estimators in at least three of the five quantile levels of interest. The Bayesian nonparametric method, `BNP`, continues to register the lowest rank, offering only a marginal reduction in bias compared to the benchmark.

Table 2.4: Simulation Results for SD2a, Average Bias and Relative MAE

	Bias					MAE				
	10th	25th	50th	75th	90th	10th	25th	50th	75th	90th
Linear specification										
BDR	0.055	0.002	-0.032	-0.008	0.016	1.349	1.049	0.621	0.597	0.781
BDRS	0.061	0.002	-0.030	-0.006	0.024	1.331	1.056	0.624	0.595	0.790
BOM	0.040	0.092	-0.013	-0.094	-0.053	0.926	0.808	0.505	0.559	0.828
BOMS	0.059	0.102	-0.002	-0.076	-0.025	0.924	0.801	0.507	0.558	0.823
BPSA	0.078	-0.042	-0.044	0.061	0.112	0.752	0.945	0.493	0.555	0.725
BNP	0.335	0.368	0.343	0.328	0.284	1.114	1.005	0.861	0.939	1.095
LDML	0.009	0.010	-0.009	0.019	0.163	0.997	0.963	0.584	0.620	0.832
TMLE	0.000	-0.005	-0.022	-0.041	0.025	1.284	1.181	0.657	0.683	1.156
FIPW	0.019	-0.039	-0.034	-0.010	0.014	1.576	1.218	0.671	0.717	1.240
No covariates										
Naive	0.468	0.434	0.430	0.377	0.312	1.000	1.000	1.000	1.000	1.000

Notes: This table displays the average bias and the relative Mean Absolute Error (MAE) of different estimation methods across 100 replicates. The relative MAE is the MAE in comparison with the Naive method as the benchmark, where $MAE = R^{-1}\sum_{r=1}^{R}|\hat{\alpha}_r - \alpha|$ and $R = 100$.

Further investigation can unveil the mechanics of our analogue doubly robust estimators. It

is essential to note that, by considering only 5 raw covariates $\mathbf{X}$, there is a misspecification in the functional form of covariates in the outcome equation.

Bayesian Outcome Modelling estimators, `BOM` and `BOMS`, inherit the advantages of Bayesian Quantile Regression and shrinkage priors, as same as our primary approach. Nonetheless, since these estimators ignore the treatment assignment equation, they exhibit significantly higher average bias than doubly robust estimators, irrespective of whether penalisation in covariate space is introduced or not. In contrast, by fitting the propensity score, doubly robust estimators gain another protective layer against misspecification of the outcome. A similar rationale applies to the favourable performance of `TMLE`, which is originally a doubly robust estimator from frequentist viewpoints. Other methods, `LDML` and `FIPW`, do not utilise both the whole conditional cumulative distribution function and the propensity score function as inputs in the doubly robust estimation procedure. However, their reliance on the treatment assignment equation from the outset makes them less affected by the misspecification of the outcome, resulting in reasonably good performance.

Bayesian Propensity Score Analysis estimator, `BPSA`, exhibits lower average bias than both `BOM` and `BOMS` when estimating 25^{th} and 75^{th} QTEs. Nevertheless, its performance is dominated by both `BDR` and `BDRS` in terms of average bias across all evaluated quantile levels. Despite sharing the first stage with doubly robust estimators when fitting the propensity score by a logit BART model, `BPSA` then uses posterior samples of propensity score rather than 5-dimensional covariates as control variables to estimate the conditional distribution of potential outcomes. Its inferiority compared to `BDR` underscores the doubly robust approach, suggesting that using the estimated propensity score alone is less favourable, especially when the treatment assignment and potential outcome equations contain different sets of control variables. This observation also aligns with the poor performance of `BNP` and reinforces our conclusion from the first simulation exercise.

SD2b. Nonlinear Specification

Table 2.5 presents simulation results when the 55-dimensional expansion of covariates is utilised as control variables. It can be seen that the performance of `BDRS` is noticeably improved, particularly in the extreme tails. `BDRS` outperforms all frequentist methods in terms of both average bias and MAE, across most quantile levels except for two instances when it ranks second after `TMLE`. This finding again highlights the superiority of `BDRS` in

high dimensions. Continuing our previous discussion on the double robustness of BDRS, when considering this basis expansion of covariates, the treatment assignment equation is misspecified to some extent. Because the logit link is maintained across Bayesian methods, the use of high-order polynomials induces a nonlinear functional form of $\mathbf{X}$, whereas the true model involves only a linear combination of X_1, X_2, and X_3. BPSA produces larger average bias than BOMS across almost all quantile levels, other than 25^{th} QTEs, and remains persistently dominated by BDRS.

Table 2.5: Simulation Results for SD2b, Average Bias and Relative MAE

	Bias					MAE				
	10th	25th	50th	75th	90th	10th	25th	50th	75th	90th
Nonlinear specification										
BDRS	-0.014	0.015	0.019	0.027	0.011	0.724	0.859	0.503	0.527	0.550
BOMS	0.041	0.027	0.020	0.015	0.001	0.547	0.709	0.463	0.455	0.441
BPSA	0.140	-0.019	-0.029	0.077	0.099	0.744	0.913	0.494	0.562	0.724
LDML	0.111	0.052	0.047	0.062	0.182	0.880	0.922	0.663	0.683	0.820
TMLE	-0.020	0.009	0.023	0.023	0.064	0.728	0.863	0.500	0.533	0.667
FIPW	0.034	0.120	0.090	-0.050	-0.051	1.577	1.290	0.873	0.926	1.298
No covariates										
Naive	0.468	0.434	0.430	0.377	0.312	1.000	1.000	1.000	1.000	1.000

Notes: This table displays the average bias and the relative Mean Absolute Error (MAE) of different estimation methods across 100 replicates. The relative MAE is the MAE in comparison with the Naive method as the benchmark, where $MAE = R^{-1}\sum_{r=1}^{R}|\hat{\alpha}_r - \alpha|$ and $R = 100$.

In summary, our proposed doubly robust estimators (BDR and BDRS) consistently surpass at least one among outcome-regression-based estimators (BOM and BOMS) or treatment-assignment-based estimator (BPSA) regarding the average bias, when either the outcome equation or treatment equation is misspecified. By flexibly incorporating shrinkage priors, BDRS outperforms its Bayesian nonparametric counterpart and all frequentist competitors in high-dimensional settings. This result demonstrates that our proposed framework features not only adaptability to complexity but also robustness to misspecification.

2.6 Empirical Illustration

2.6.1 Overview

To demonstrate the applicability and usefulness of our proposed method, we revisit the microcredit study by Crépon et al. (2015), which was derived from a randomised experiment conducted in Morocco. The dataset enables us to examine the potential of our approach in two distinct contexts. In the first setting, we employ the *random treatment assignment* available in the original research to investigate the effect of microcredit availability on household borrowing activities, such as the total amount of loans. Our second setting deviates from randomisation – we instead use *observational data* while assuming *selection-on-observables* to evaluate the welfare impact of household loans.

The evaluation was conducted across 162 Moroccan villages that were paired based on their observable similarities. The intervention was microcredit availability, which was randomly assigned to one village within each pair. These designated villages constituted the treated group, whereas the remaining villages formed the control group. In particular, a microfinance institution was established in the treated villages between 2006 and 2007. In 2009, a follow-up study surveyed 5551 households in both treated and control villages.

The expansion of microcredit, or access to loans in general, can have potentially heterogeneous effects on household welfare for several reasons. First, households have diverse loan take-up behaviours. They may differentially select into borrowing activities based on their characteristics, leading to varying outcomes. Those who do not take up loans may end up worse off due to effects on wages or the displacement of informal lending in a dynamic general equilibrium (Kaboski and Townsend, 2011; Morduch, 1999). Second, among borrowers, the effects may vary due to differences in the efficiency of loan use and uneven investment opportunities. Indeed, certain households may not benefit from loans if the requirements for investment purposes are restrictive or the term to maturity is too short (Banerjee, 2013). Additionally, multiple microlenders in a community can engage in exploitative lending practices and "overlending" to households who cannot feasibly repay the loan (Ahmad, 2003; Schicks, 2013). This can result in high-productivity borrowers benefiting from the positive impact, whereas the most vulnerable borrowers are systematically harmed by the saturation of credit markets. In summary, there are potential winners and losers to financial market expansion. Even if disadvantaged groups are small, social welfare

consequences could be substantial, particularly if economic inequality across households is exacerbated (Meager, 2022).

The average treatment effect (ATE), which is most commonly utilised in empirical research, cannot reveal this heterogeneity. Even though loan access might have no impact on average, it could still have significant positive or negative effects on different types of households. This policy implication is particularly critical for developing countries. To gain a more comprehensive understanding of causal effects, it is worthwhile to estimate unconditional quantile treatment effects (QTEs), which offer a valid measurement that goes beyond the ATE for the entire population. Therefore, our proposed framework is well suited for this empirical context.

In contrast to the original paper and previous studies that typically rely on randomised controlled trial (RCT) design and ad hoc selection of baseline covariates, our approach offers more flexible specifications and data-driven estimation. This enables us to conduct new analyses using either data from randomised experiments or observational data, as demonstrated in sections 2.6.2 and 2.6.3, respectively.

Specifically, our general strategy is to initially create a large set of covariates by combining village pair dummies and full cubic polynomials along with interaction terms of household observed characteristics. Once collinear columns are removed, this set serves as the baseline specification of $\mathbf{X}$ and can be readily integrated into our Bayesian Analogue of Doubly Robust (BADR) estimation framework. Given the high dimensionality of this empirical issue, we opt for the Bayesian Doubly Robust estimator with Adaptive Lasso (`BDRS`) due to its proven merits in our prior simulation study. To compare our results with the benchmark, we also include the Naive estimator (`Naive`) in our analysis.

2.6.2 Impact of Microcredit Availability on Loan Amount

We begin with the context of random treatment assignment, where our objective of interest is the effect of microcredit availability on the total amount of loans at the household level. To examine the balance between the treated and control groups, we select pre-treatment covariates which are observed characteristics for each household, including head age, education of the head, number of adults, total number of members in a household, indicators for households doing animal husbandry, doing other non-agricultural activities, and whether household spouse responded to the survey. Table 2.6 reports the mean values of these covariates in

addition to the outcome and treatment variables, both for the whole sample and for each of the treated and control groups.

Table 2.6: Summary Statistics of Households

	Treated	Control	Treated – Control		
Treatment: Microcredit Availability (RCT)	Mean (sd)	Mean (sd)	Diff.Mean		t-statistic
Outcome variable					
Total amount of loans (in MAD)	2350.44 (10865.84)	1547.75 (7381.73)	802.69	*	2.54
Baseline covariates					
Head age	49.18 (15.83)	48.14 (15.85)	1.05	.	1.95
Head with no education	0.67 (0.47)	0.68 (0.47)	-0.01		-0.89
Number of members	5.70 (2.54)	5.64 (2.44)	0.06		0.71
Number of adults	3.81 (1.99)	3.76 (1.91)	0.05		0.83
Number of members aged 6-16	1.22 (1.29)	1.25 (1.26)	-0.03		-0.72
Declared animal husbandry activities	0.60 (0.49)	0.55 (0.50)	0.05	**	2.73
Declared non-agricultural activities	0.17 (0.37)	0.21 (0.41)	-0.04	**	-3.15
Spouse of head responded	0.09 (0.29)	0.07 (0.26)	0.02	*	2.27
Member responded	0.05 (0.22)	0.05 (0.21)	0.00		0.62

Data sources: Moroccan household survey (Crépon et al., 2015).

Although the randomisation of microcredit availability and the absence of confounding factors leading to self-selection into treatment is plausible, there are slightly imbalances in covariates across the two groups. Regarding the unconditional means, the households' total loan amount for the treated group (2350.44) significantly exceeds that of the control group (1547.75). The potential heterogeneity of microcredit motivates us to further investigate this positive average treatment effect using quantile analysis.

The results of Quantile Treatment Effects (QTEs), as estimated by the `Naive` and `BDRS` methods, are depicted in Figure 2.5. According to the findings, microcredit expansion has a precise zero effect below the 75^{th} percentile of the distribution of total loan amounts, but exhibits positive effects above this threshold. In particular, at the 90^{th} percentile, the positive effect is statistically significant (2300), contributing to the decomposition of the average treatment effect (802.69). Compared with naive estimates, `BDRS` produces similar results, only higher at the 90^{th} percentile; however, the difference is insignificant. The result is robust after adjusting for the influence of covariate imbalance on the outcome.

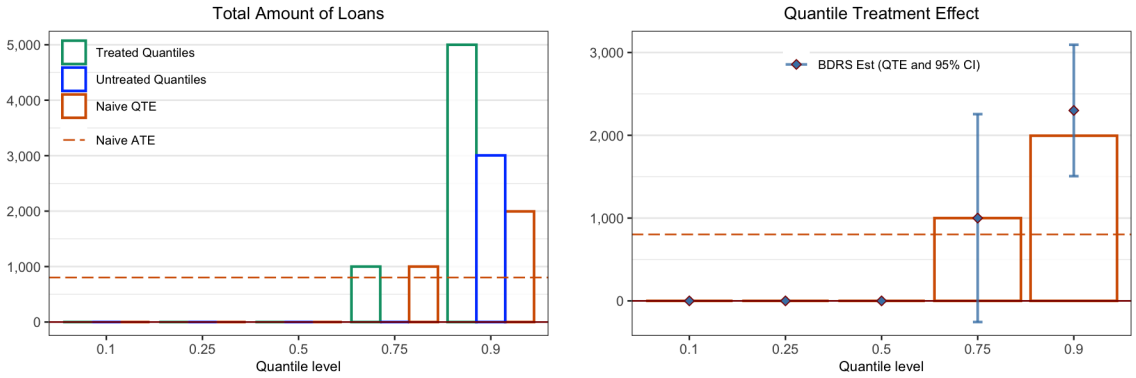

Figure 2.5: Quantile Treatment Effects (QTEs) of microcredit availability on households' total loan amount. The graph on the left demonstrates Naive estimation results. Red bar plots represent naive QTEs, which are differentials between empirical quantiles of treated group (in green) and control group (in blue). Red dashed line indicates naive Average Treatment Effects (ATE), which is simple mean difference between these two groups. Results obtained using BDRS method, QTE point estimates and corresponding 95% CI at five quantile levels based on 100 bootstrap replications, are plotted as error bars in the right-hand graph.

Using the same dataset, findings in Chernozhukov et al. (2017) and Jacob (2021) also document the heterogeneity of the microcredit availability on total loan amount; however, their estimand is conditional ATE, different from this paper (QTEs).

2.6.3 Impact of Loan Access on Household Outcomes

Our second objective is to explore the causal impact of access to loans on household welfare, with a focus on the distribution of consumption and business outcomes, including total consumption, consumption of temptation goods, total output, and total profit. Unconditional QTEs provide deeper insights into the potential heterogeneity of causal effects across the distribution of each outcome interest, as well as the resulting change in household inequality. The binary treatment we consider is the actual borrowing status recorded at the household level. Table 2.7 indicates that the difference in mean between two groups of households (borrowers and non-borrowers) is highly statistically significant regarding consumption, but not for business outcomes. However, there are two caveats to these naive ATE estimates.

Table 2.7: Summary Statistics of Household Outcomes.

	Borrowers		Non-borrowers		Borrowers – Non-borrowers		
Outcome variables	Mean	St.Dev.	Mean	St.Dev.	Diff.Mean		t-statistic
(in MAD)							
Total Consumption	3268.62	(2956.01)	2863.49	(1792.97)	405.13	***	3.82
Temptation Goods	312.33	(229.91)	270.31	(219.33)	42.01	***	4.73
Total Output	32672.06	(85071.58)	30885.38	(85939.63)	1786.68		0.54
Total Profit	10081.86	(37986.07)	8409.95	(45277.88)	1671.91		1.07

Data sources: Moroccan household survey (Crépon et al., 2015). *Definition:* Total Consumption is monthly total consumption (in MAD); Temptation Goods is monthly expenditure on temptation and entertainement(in MAD); Total Output is sum of agricultural, livestock, and non-agricultural business production over the 12 months prior to the survey (in MAD); Total Profit is total profit of self-employment activities over the 12 months prior to the survey (in MAD).

Firstly, all outcome variables in this empirical setting exhibit heavy tails and large variability, as illustrated in histograms in Figure 2.6. This is another motivation for quantile analysis since estimation results for a set of quantiles would be less susceptible to the influence of outliers than results for the mean.

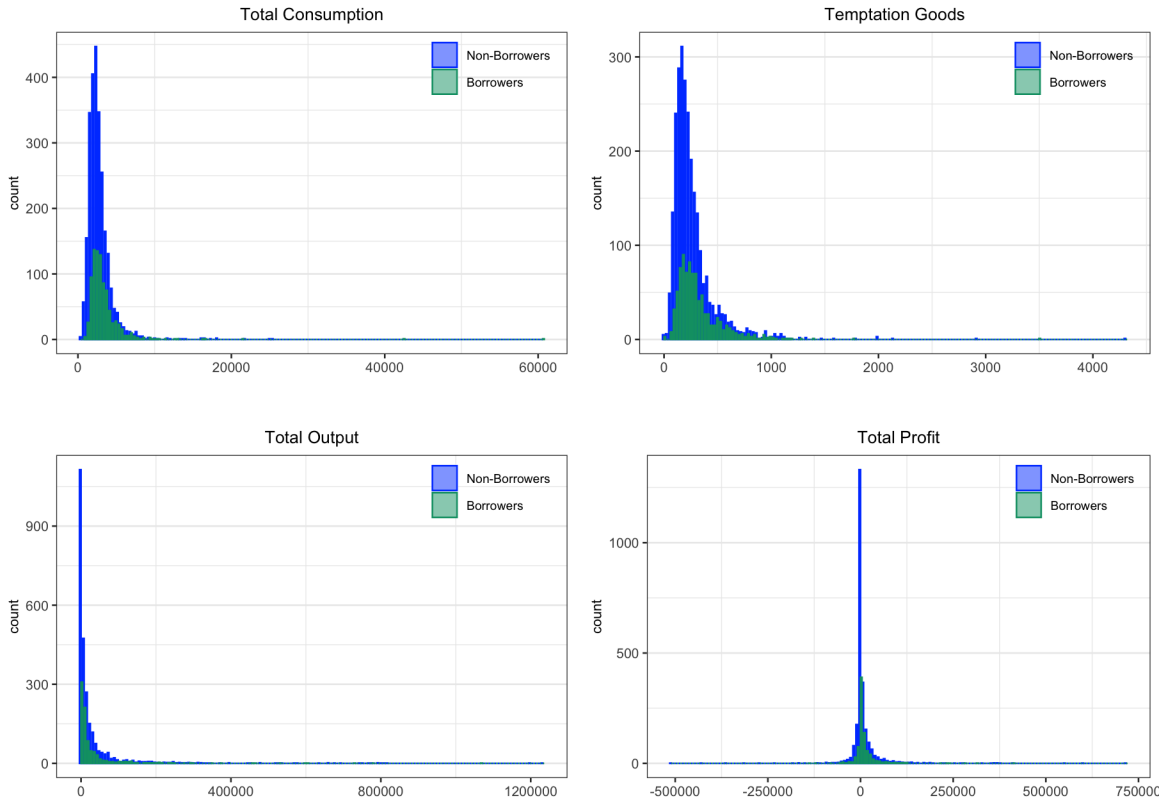

Figure 2.6: Histograms of various consumption and business outcomes of borrowing households (in green) and nonborrowing households (in blue). These graphs display raw data without any truncation applied.

Secondly, the treatment variable – borrowing pattern observed in the dataset – is no longer randomly assigned among households in the present context. This is confirmed by the imbalances between these two groups regarding the mean values of the observed characteristics, as shown in Table 2.8. Specifically, borrowing households tend to have larger average household sizes. They are also more inclined to engage in non-agricultural self-employment activities and reside in villages where microcredit is available. The discrepancy observed is more than what would be expected by pure chance. Therefore, to identify causal effects using non-experimental data, we pursue the *selection-on-observables* assumption. That means, conditional on observed covariates, unmeasured factors that influence household loan access are independent of household outcomes.

Table 2.8: Covariate Balance between Borrowers and Non-borrowers.

	Borrowers	Non-borrowers	Borrowers – Non-borrowers		
Control variables	Mean (sd)	Mean (sd)	Diff.Mean		t-statistic
Head age	49.01 (15.62)	48.53 (15.93)	0.49		0.79
Head with no education	0.68 (0.47)	0.68 (0.47)	0.00		0.05
Number of members	6.06 (2.46)	5.54 (2.48)	0.52	***	5.36
Number of adults	4.02 (2.01)	3.71 (1.92)	0.31	***	3.99
Number of members aged 6-16	1.36 (1.30)	1.19 (1.27)	0.16	**	3.23
Declared animal husbandry activities	0.59 (0.49)	0.57 (0.50)	0.02		1.23
Declared non-agricultural activities	0.23 (0.42)	0.18 (0.38)	0.05	**	3.17
Spouse of head responded	0.05 (0.23)	0.09 (0.29)	-0.04	***	-3.96
Member responded	0.05 (0.21)	0.05 (0.22)	0.00		-0.36
Microcredit availability	0.55 (0.50)	0.47 (0.50)	0.07	***	3.73

Data sources: Moroccan household survey (Crépon et al., 2015).

Whilst a violation of randomisation may threaten the performance of the naive estimator, the `BDRS` estimator serves as a debiasing device, as illustrated in our simulation using synthetic data. Table 2.9 presents the results for key outcome variables related to household consumption and business. Unlike the first setting, the `BDRS` estimates differ considerably from the naive estimates of QTEs because selection bias is accounted for in our proposed approach. Overall, the point estimates at extreme tails (10^{th} and 90^{th} percentiles) are fairly imprecise, as indicated by large credible intervals compared to the other quantile levels. Interestingly, the causal effect in the upper tail remains significantly positive in most cases.

Table 2.9: Quantile Treatment Effects of Loan Access on Household Outcomes.

		BDRS			**Naive**
Outcomes	Percentiles	QTEs	Upper bound	Lower bound	QTEs
Total Consumption	10th	20.093	1469.207	-1429.020	232.795
	25th	9.242	173.511	-155.027	173.456
	50th	79.949	229.587	-69.690	229.753
	75th	132.22	273.974	-9.534	286.680
	90th	237.699	543.355	-67.956	685.442
Temptation Goods	10th	-8.69	65.660	-83.040	17.380
	25th	13.035	29.871	-3.801	21.725
	50th	**30.415**	45.962	14.868	43.450
	75th	**47.795**	79.225	16.365	60.830
	90th	**78.21**	129.145	27.275	78.210
Total Output	10th	0	13146.948	-13146.948	0.000
	25th	-330	19.769	-679.769	1093.446
	50th	50	1385.992	-1285.992	1787.500
	75th	1666	6933.205	-3601.205	2771.616
	90th	**27360**	52964.198	1755.802	2744.044
Total Profit	10th	**-5500**	-1183.536	-9816.464	-1142.697
	25th	-945	158.825	-2048.825	-241.876
	50th	**561**	1117.727	4.273	979.125
	75th	1780.769	4273.915	-712.377	549.373
	90th	**8954.377**	16664.233	1244.520	-1086.350

Notes: Upper bound and Lower bound for BDRS method are associated with the estimates of 95% CI based on 100 bootstrap replications.

Regarding *total consumption*, although all naive estimates of ATE and QTE are positive, estimation results obtained using the `BDRS` method reveal notably lower effects across all quantile levels. The effects of loan access are most pronounced at the 75^{th} and 90^{th} percentiles of the consumption distribution; however, they are insignificantly positive. While the naive method overestimates the effect of loan access compared to the `BDRS` estimator, the upward bias suggests a possible selection-on-gain pattern. Households inclined to borrow to support their consumption are more likely to gain higher total consumption when they have financial access.

Further examination of the impact of borrowing on *temptation consumption* shows a similar upward bias in the `Naive` method relative to the `BDRS` method. The effect is slightly negative

at the lowest percentile (10^{th}) yet clearly insignificant. By contrast, significant positive effects are observed at the median and higher percentiles. This seems inconsistent with other works that have found a statistically significant reduction in nonessential expenditures. However, these studies used different treatment variables and designs compared to this paper.

With respect to household business outcomes, there is a statistically significant increase in *total output*, concentrated only at the highest quantile level. For the rest of the community, specifically those below the 90th percentile, no systematic change appears to be taking place. Consequently, the majority of the total output distribution remains unchanged with or without universal access to loans.

Interestingly, there is notable evidence of heterogeneous effects on *total profit*. The effect on the median household estimated by the `BDRS` method is quite close to the naive ATE estimate, which is moderately positive. In general, access to loans has a favourable impact on households' profit by shifting the center of distribution towards the right. However, the impacts exhibited at extreme tails are more dramatic, with a negative effect at the lowest percentile (10^{th}) and a positive effect at the highest percentile (90^{th}), and both are statistically significant. If the rank invariance assumption is invoked, the rightward expansion of the upper tail means that high-profit households gain benefits, while the leftward expansion of the lower tail means that low-profit households experience loss when loans are accessible to everyone, compared to the opposite counterfactual scenario. While this assumption might be difficult to defend given the complexity and nonlinearities inherent to the financial market, interpretations about the shape change of distribution of household total profit remain valid. There do exist both winners and losers, even when we cannot identify the specific households that belong to each group. The outcome distribution disperses wider leading to the exacerbation of inequality across households.

Taken together, the estimated QTE patterns collectively indicate that broadening financial access is likely to result in an ex-post rise in economic inequality across households. Specifically, the increase in total output and consumption at the household level is solely attributable to the right tail of distributions expanding rightward, suggesting that certain households are likely to experience an improvement in their economic circumstances without incurring any systematic losses from others. Further investigation of total profit, however, reveals a more nuanced picture. While the overall impact of access to loans on total profit is positive, indicating a shift towards higher profits for many households, there is evidence

of extreme heterogeneity. The effect is asymmetric as certain households may experience negative effects on their profits.

Although the treatment variable and identification strategies employed in this setting differ from those used in Crépon et al. (2015), the findings converge in several respects. In the original paper, both the reduced-form quantile regressions and instrumental variable (IV) estimates suggest substantial heterogeneity in the profitability of microfinance investments and emphasise the detrimental effects on certain households. Specifically, their reduced-form quantile analysis measures Intention-to-Treat (ITT) effects because the treatment variable is microcredit availability at the village level rather than actual borrowing at the household level. Additionally, the IV estimates in this study reveal changes in the unconditional distribution of total profit for those who take up microcredit (i.e., compliers only). These results are only valid when randomisation holds. In contrast, the findings of this paper have broader implications for the understanding of economic inequality, as we focus on the entire population of households utilising non-experimental data and a *selection-on-observables* assumption.

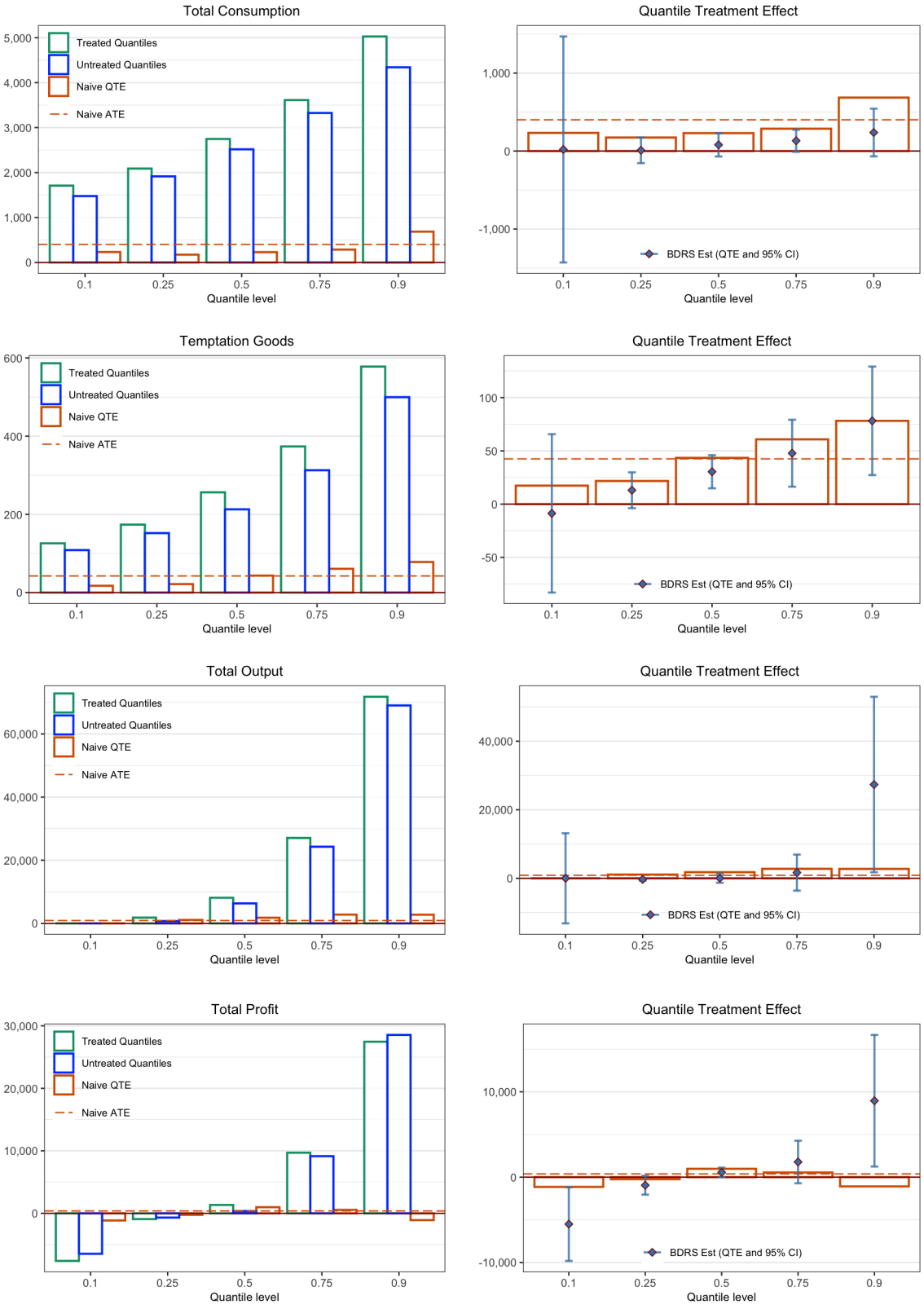

Figure 2.7: Quantile Treatment Effects (QTEs) of loan access on various household outcomes. Graph on the left demonstrates Naive estimation results. Red bar plots represent naive QTEs, which are differentials between empirical quantiles among borrowing households (in green) and nonborrowing households (in blue). Red dashed line indicates naive Average Treatment Effect (ATE), which is simple mean difference between these two groups. Results obtained using BDRS method, QTE point estimates and corresponding 95% CI at five quantile levels based on 100 bootstrap replications, are plotted as error bars in the right-hand graphs.

2.7 Concluding Remarks

The goal of this paper was to address the challenges associated with estimating unconditional Quantile Treatment Effects (QTEs) in observational studies and to make a contribution to the burgeoning econometric literature on QTEs as well as causal machine learning. We introduced a novel approach, Bayesian Analogue of Doubly Robust (BADR) estimation, which accounts for potentially high-dimensional covariates. The framework features a highly flexible Bayesian modelling scheme that showcases favourable frequentist properties in finite samples, even in the presence of high dimensions or model misspecifications, which has not been explored in previous literature. This approach, while not fully Bayesian in nature, offers a straightforward and versatile implementation for integrating probabilistic machine learning techniques into causal analysis on quantiles, with precise estimation and reliable uncertainty quantification. These attributes are particularly advantageous in complex, high-dimensional settings.

The performance of the proposed method was assessed through a simulation study in two different settings. The first simulation focused on a linear setting with varying feature dimensionality, whereas the second simulation considered a nonlinear setting and examined the double robustness of the proposed estimators. Through a comparison with both the naive approach and existing popular estimators, the simulation results consistently indicated a substantial improvement in bias reduction for QTE estimates when using the new method. This finding demonstrates that our proposed framework features not only the ability to adapt to high dimensions and complexity, but also robustness to misspecification.

The empirical illustration of estimating QTEs of financial access on household outcomes showed the potential benefits of using causal inference on quantiles to help characterise the heterogeneity or distributional impact of interventions, which is often appealing to researchers and easily conveyed to policymakers and stakeholders. Our proposed approach makes this possible even in the absence of experimental data. We found strong evidence for an overall positive effect yet heterogeneous across different points of outcome distributions. An ex-post rise in economic inequality among households is likely to occur, primarily driven by significant improvements in consumption and business outcomes at the top quantiles. However, certain households may experience adverse effects on their total profit.

An interesting extension of this framework to be explored in the future would be estimating QTEs when the *selection-on-observable* assumption is violated, that is, when there exists

unmeasured confounding that drives endogenous selection into treatment. Another aspect is improving the bootstrap inference scheme for doubly robust estimators. Although this objective could be achieved effectively in the BADR framework for average treatment effect, it is computationally demanding when applied to quantiles.

Chapter 3

Bayesian Causal Inference in the Presence of Endogenous Selection into Treatment and Spillovers

3.1 Introduction

Spillovers, often referred to as interference, despite being of great significance in economic settings, complicate conventional approaches in causal inference and have received insufficient attention. This phenomenon occurs when the outcome of a unit is influenced not only by its own treatment but also by the treatments applied to other units within a network or spatial domain (see, e.g., Forastiere et al., 2021; Giffin et al., 2022.). This poses a challenge to the Stable Unit Treatment Value Assumption (SUTVA), which stipulates that the potential outcome for any unit remains unaffected by treatment assignments to other units. Understanding both direct and indirect effects is essential for evaluating programs in situations where spillovers exist. Direct effects quantify how an individual's own treatment alters their outcome, while indirect effects quantify how their peers' treatments impact their outcome. Our first motivation comes from the presence of network interference through exogenous social networks. For example, an After-School Program (ASP) aimed at enhancing personal or social skills in youth can lead to spillover effects via knowledge sharing or behavioural influence, given that participants interact with non-program peers. Another important setting is the presence of spatial interference of place-based policies, such as the Opportunity Zone (OZ) program, a U.S. tax incentive program designed to stimulate private investment in economically disadvantaged communities. Spatial interference arises when a census tract receives the program, possibly affecting non-designated neighbouring tracts. This can occur through increased property values and economic activity in Opportunity Zones (OZs), which may spill over into adjacent areas, or by attracting businesses and investors to relocate to designated OZs, influencing surrounding regions. Policymakers may be concerned about the potential for indirect effects to counteract the direct effects of such programs.

In the presence of spillovers, the outcome of an individual is a function of the entire vector of treatments allocated to the population. Consequently, there are an extremely large number of possible potential outcomes, hindering the use of the conventional causal inference framework (Rubin, 1986, 1974). As a result, researchers typically resort to alternative experimental designs, specific assumptions, or additional information. The existing literature can be divided into two primary approaches. The first approach relies on the *partial interference* assumption which requires that units are a priori partitioned into disjoint treatment clusters, and spillovers occur only within clusters and not between them (see, e.g., DiTraglia et al., 2023; Hudgens and Halloran, 2008; Manski, 2013; Sobel, 2006.). Nevertheless, this requirement is only plausible in limited settings when the data naturally segregates a significant distance. In other circumstances, such as when bordering counties in two states influence each other, it may be impractical to apply this assumption. A more recent approach has attempted to relax the assumption of *partial interference*, thereby allowing for spillovers of quite general forms without the notion of cluster. This departure requires extra knowledge of the interference structure, which is increasingly obtainable from the deluge of network or spatial data, and formulates *exposure mappings*[1] through which interference affects individual outcomes. Early studies in econometrics and statistics, which employed this methodology, focused on randomised experiments on social networks, including works by Toulis and Kao (2013), Aronow and Samii (2017), Leung (2020), Yuan and Altenburger (2022), and others. Additionally, there is an active line of research in observational studies where the network is observed by the researcher, with works by Laan and Sofrygin (2017), Forastiere et al. (2021), Forastiere et al. (2022), Ogburn et al. (2022), Sanchez-Becerra (2022), Leung and Loupos (2022), among others. To address identification, these studies establish analogs of the unconfoundedness and support conditions for the applicable settings. Despite the technical advancements, unconfoundedness states that the treatment would be as good as randomly assigned once conditioned on observables. This assumption is not always justifiable in real-world applications, where selection on unobservables is the norm rather than the exception. However, there has been little exploration of this violation. In the context of After-School Programs (ASPs), self-selection is a significant issue. As participation is voluntary, young individuals who enroll in and attend ASPs may differ in numerous ways from those who do not participate. With respect to the OZ program, endogenous selection into treatment may occur in the assignment process of the program. The selection of which communities

[1] For a more comprehensive survey, we refer interested readers to Huber (2023).

to be designated as Opportunity Zones was based on a combination of factors, including a recommendation from the governor of the state and the support of the local government. If the recommendation and support are based on factors related to the economic performance of the community, such as high levels of economic activity or growth potential, this may result in the OZ program only being available to the most promising communities, thereby exaggerating its benefits.

In this paper, we provide methods that use observational network or spatial data to identify and estimate direct and indirect causal effects in the presence of both endogenous selection into treatment and spillovers. This scenario causes the violation of the unconfoundedness and SUTVA assumption. We propose a new econometric framework that nests the Generalised Roy model and accommodates spillovers in a form of neighbourhood treatment (exposure mapping). In this way, we explicitly model endogenous selection into treatment and allow for heterogeneous effects across individuals, making it economically interpretable. For estimation and inference, we develop further Bayesian Data Augmentation algorithms that enable more efficient computation when models involve latent data structures and the maximization of the likelihood function is challenging numerically. We suggest an extension of our method using a Bayesian semiparametric approach that relaxes distributional assumptions.

To assess the performance of our proposed approach in finite samples, we conduct simulations using synthetic data and an empirical Monte Carlo study that utilised friendship networks and covariates from the National Longitudinal Study of Adolescent Health (Add Health). The Bayesian estimators demonstrate strong performance in terms of bias, root mean squared error (RMSE), and coverage rate. Furthermore, the inclusion of neighbourhood treatment terms is found to be plausible, regardless of the presence of interference in the true data-generating process. In contrast, neglecting neighbourhood treatment leads to a larger bias and a lower coverage rate, even when the causal estimand is the direct treatment effect.

Finally, in a realistic setting, we apply the framework to evaluate the effects of the Opportunity Zones (OZ) program, a place-based policy offering tax incentives to promote economic development in distressed communities in the United States. We model the selection process of Qualified Opportunity Zones (QOZs) by state governors and estimate the program's impact on housing unit growth in census tracts. Endogenous selection into treatment is present. Our findings reveal a selection-on-gains pattern, where treatment effects vary across unmeasured tract heterogeneity. Governors tend to designate census tracts based

on characteristics that drive expected future growth, such as the level of prior investment or development. Although these factors are unobservable in the data, governors appear to obtain this information from their local political networks. In fact, the results also indicate that governors are more likely to choose tracts represented by members of their political party, which supports our use of political affiliation as an instrumental variable. With regard to targeted areas, both direct and indirect effects on the treated tracts (QOZs) are positive. However, eligible but unselected tracts (non-QOZs) remain a disadvantaged group that does not experience any positive spillover effects. Additionally, unobserved differences between QOZs and non-QOZs make it unlikely that positive treatment effects on QOZs would be replicated if non-QOZs were granted investment tax credit. In fact, the strongly negative average direct treatment effect on the untreated (ADTUT) even predicts that non-QOZs would likely suffer adverse consequences in that case. Consequently, extending the OZ program to communities that do not currently receive tax credit would not be effective.

Beyond the related literature on violation of the SUTVA, which has been mentioned earlier, our paper also broadly fits the econometrics literature on estimation of treatment-response and selection models under non-random treatment assignment. The proposed framework is closely connected to the canonical Generalised Roy Model (GRM) and its extensions (see, e.g., Abbring and Heckman, 2007; Eisenhauer et al., 2015; Heckman and Vytlacil, 2007, 2005), which form a cornerstone of the literature in econometric causality and structural policy analysis[2]. Specifically, our model inherits GRM's key advantages in modelling heterogeneity and self-selection into treatment based on unobserved gains from treatment, enabling a richer characterisation of causal estimands beyond mean treatment effects. In terms of identification and estimation, the full covariance structure of disturbance terms plays a pivotal yet technically challenging role. First, correlations of unobservables in selection and outcome equations need to be considered to tackle selection bias. It is well known that the variance of the disturbance term in the selection equation is unidentified and therefore requires normalization. Second, when jointly modeling multiple potential outcomes–which is essential for learning distributions of treatment effects–the cross-regime correlation between treated and untreated outcomes becomes another unidentified parameter, since each individual is observed in only one regime. To address these challenges, our Bayesian estimation strategy builds on the previous approaches to Gaussian selection model Poirier and Tobias (2003), which place an Inverse Wishart prior over the full covariance matrix. Our approach, however, differs in how

[2]See Heckman and Pinto (2022; 2024) for recent comprehensive discussions.

we handle the unidentified variance of the selection-equation disturbance, which complicates posterior sampling. Rather than fixing this parameter throughout–as suggested in Nobile (2000)–we employ parameter expansion techniques to construct efficient Gibbs samplers that accommodate the normalization constraint while improving convergence rates. In this regard, our paper also connects to the line of work on Bayesian approaches to sample selection and latent-index models (Ding, 2014; Doğan and Taşpinar, 2018; Imai and Van Dyk, 2005; Jiao and Dyk, 2015). Nevertheless, to our knowledge, no existing studies in this body of research jointly model endogenous selection and spillovers, which is the central contribution of this paper.

The remainder of this paper is structured as follows. In Section 3.2, we formally present a causal framework in the presence of endogenous selection into treatment and spillovers and define causal estimands of interest along with key identification assumptions. In Section 3.3, we propose Bayesian Data Augmentation algorithms to estimate the model and conduct inference. Next, we evaluate the performance of our method using simulations in Section 3.4 and use the proposed approach to investigate the causal impact of the American Opportunity Zones (OZ) program on economic outcomes in Section 3.5. Finally, we conclude the paper in Section 3.6 with brief final remarks on the method and policy recommendations based on our results.

3.2 Causal Framework in the Presence of Endogenous Selection into Treatment and Spillovers

3.2.1 General Model Setup

We consider a general setting for n agents ($i = 1, \ldots, n$) which involves treatment and outcome processes.

Treatment process

- Let D_i be the observed binary treatment decision, which takes the value of 1 if the unit receives the treatment and 0 otherwise. This could be regarded as *individual treatment*. From a choice-theoretical perspective, this selection process can be expressed as an individual decision-making problem relied on cost-benefit assessment. Denote D_i^* the net benefit to the individual of enrolling the program (/the net desire for receipt of treatment), which depends on observable characteristics (Z_i,X_i) and unobservable

factors (U_i). Z_i represents the availability of some exclusion restriction in choice equation that does not appear in outcome equations.

$$\begin{aligned} &D_i^* = \mu(Z_i, X_i) - U_i; \\ &D_i = 1 \text{ if } D_i^* \geq 0; \quad D_i = 0 \text{ otherwise.} \end{aligned} \tag{3.1}$$

Here, U_i is assumed to be a continuous random variable with a strictly increasing distribution function F_U. Define $V_i = F_U(U_i)$, then it has uniformly distribution and indicates different quantile level of U_i. Let $\nu(Z_i, X_i) = F_U(\mu(Z_i, X_i))$, which is the mean scale utility function in discrete choice theory, we can thereby rewriting:

$$D_i = \mathbb{1}\{\mu(Z_i, X_i) \geq U_i\} = \mathbb{1}\{\nu(Z_i, X_i) \geq V_i\}. \tag{3.2}$$

Since V_i enters the selection equation with a negative sign, it embodies characteristics that make individuals less likely to receive treatment yet being unmeasurable by the researcher. Put differently, V_i can be interpreted as the unobservable resistance to the treatment.

- To capture spillovers, for each agent i, we formulate a *neighbourhood treatment* term $\bar{D}_{\mathcal{N}i}$, which is a summary measure of the treatment status of all agent j other than i. It is possible to calculate $\bar{D}_{\mathcal{N}i}$ based on the knowledge of w_{ij}, which represents the connection between units i and j. We can exploit this information from an available adjacency matrix, which defines spillovers structure in the case of either spatial interference or network interference.

$$\bar{D}_{\mathcal{N}i} = \sum_{j=1, j\neq i}^{n} w_{ij} D_j, \qquad \sum_{j=1, j\neq i} w_{ij} = 1. \tag{3.3}$$

Outcome process

- $Y_i^{(1)}$ and $Y_i^{(0)}$ are treated and untreated potential outcomes, respectively. In particular, $Y_i^{(1)}$ is the potential outcome when agent i enrolls the program while $Y_i^{(0)}$ is the potential outcome when agent i does not. In both cases, potential outcome depends on individual characteristics X_i and an idiosyncratic component ($\epsilon_i^{(1)}$ or $\epsilon_i^{(0)}$). We also extend the standard formula of potential outcomes to allow for possible impact of the

neighbourhood treatment term $\bar{D}_{Ni}$.

$$\begin{aligned} Y_i^{(1)} &= \mu_1\left(\bar{D}_{\mathcal{N}i}, X_i\right) + \epsilon_i^{(1)} \text{ and} \\ Y_i^{(0)} &= \mu_0\left(\bar{D}_{\mathcal{N}i}, X_i\right) + \epsilon_i^{(0)}, \end{aligned} \tag{3.4}$$

where $\mu_1\left(\bar{D}_{\mathcal{N}i}, X_i\right) = \mathbb{E}\left[Y_i^{(1)} \mid \bar{D}_{\mathcal{N}i}, X_i\right]$ and $\mu_0\left(\bar{D}_{\mathcal{N}i}, X_i\right) = \mathbb{E}\left[Y_i^{(0)} \mid \bar{D}_{\mathcal{N}i}, X_i\right]$ are mean response functions.

- Y_i is the revealed outcome, which equals treated potential outcome when i is treated ($D_i = 1$) and equals untreated potential outcome when i is untreated ($D_i = 0$)

$$Y_i = D_i Y_i^{(1)} + (1 - D_i) Y_i^{(0)}. \tag{3.5}$$

Taken together, we present a full model specification as follows

$$\begin{aligned} D_i &= \mathbb{1}\{\nu(Z_i, X_i) > V_i\}, \\ \bar{D}_{\mathcal{N}i} &= \sum_{j=1, j\neq i}^{n} w_{ij} D_j, \quad \sum_{j=1, j\neq i} w_{ij} = 1, \\ Y_i^{(1)} &= \mu_1\left(\bar{D}_{\mathcal{N}i}, X_i\right) + \epsilon_i^{(1)}, \\ Y_i^{(0)} &= \mu_0\left(\bar{D}_{\mathcal{N}i}, X_i\right) + \epsilon_i^{(0)}, \\ Y_i &= D_i Y_i^{(1)} + (1 - D_i) Y_i^{(0)}. \end{aligned} \tag{3.6}$$

Example 1. (Causal Inference in the presence of endogenous selection into treatment and spatial interference of place-based policies)

In the context of the Opportunity Zone (OZ) program, there are n census tracts, indexed by $i = 1, \ldots, n$. D_i denotes the participation indicator in the OZ program, while Y_i represents the housing price or new development as the outcome variable. X_i is a vector of exogenous covariates, such as poverty rate, median earnings, and employment rate, among others. Z_i is a vector of instrumental variables that directly determine D_i but only indirectly affect Y_i through D_i. For instance, political alignment between the census tract and the state governor is likely to be correlated with treatment decisions. The unobservable cost V_i would reflects factors such as the level of economic growth potential of the community. For $d \in \{0, 1\}$, $Y_i^{(d)}$ represents the potential outcome when $D_i = d$, and $\epsilon_i^{(d)}$ is the corresponding unobservable determinant of $Y_i^{(d)}$. We may suspect the presence of endogeneity in the OZ program assignment process since

the unobservable component V_i and $(\epsilon_i^{(1)}, \epsilon_i^{(0)})$ may share factors related to local economic performance. It is plausible to imagine that when the OZ program is introduced in community i, the surrounding areas that are not designated as OZs may also be affected due to spatial interference (spillovers). For instance, this program may lead to the displacement of low-income residents and unskilled workers as property values, living costs, and job requirements increase in the designated zone. Since displaced people may have to seek employment opportunities in other areas, the effects on the local economic outcome Y_j in census tract i's neighbouring areas j may be considerable. To measure the neighbourhood treatment term $\bar{D}_{\mathcal{N}i}$, we could employ w_{ij} from a spatial weights matrix.

Example 2. (Causal Inference in the presence of endogenous selection into treatment and network interference via a known exogenous social network)

In the case of the After-School Program (ASP), there are n students, labeled as $i = 1, \ldots, n$. Individual treatment D_i is an index of participation in the program, while Y_i measures a desired outcome, such as academic performance, social-emotional learning (SEL)-related outcomes, and involvement in juvenile crime and violence. The exogenous variables X_i consist of individual characteristics including gender, grade, and race, among others. The instrumental variable Z_i could be in the form of cost-shifters, for example, the distance to the program location, which varies from living directly next to the program to living far away. As participation in the program is voluntary, unmeasured resistance to treatment V_i is likely to be correlated with unobservable components $(\epsilon_i^{(1)}, \epsilon_i^{(0)})$ in the outcome equations. To account for network interference (spillovers) resulting from peer effects on students' outcomes, we compute the neighbourhood treatment $\bar{D}_{\mathcal{N}i}$ as the proportion of the number of treated friends among all friends. Information on network link w_{ij} can be extracted using the adjacency matrix of the students' friendship network.

3.2.2 Identification and Causal Estimands

3.2.2.1 Assumptions

For identification, we make the following assumptions in our setting.

Assumption 1. $\{Y_i, D_i\}_{i=1}^n$ is generated according to a parametric model specified below

$$\begin{aligned} D_i &= \mathbb{1}\{Z_i\alpha + X_i\beta^{(D)} + \epsilon_i^{(D)} > 0\}, \\ \bar{D}_{\mathcal{N}i} &= \sum_{j=1, j\neq i}^{n} w_{ij} D_j, \qquad \sum_{j=1, j\neq i} w_{ij} = 1, \\ Y_i^{(1)} &= \delta^{(1)}\bar{D}_{\mathcal{N}i} + X_i\beta^{(1)} + \epsilon_i^{(1)}, \\ Y_i^{(0)} &= \delta^{(0)}\bar{D}_{\mathcal{N}i} + X_i\beta^{(0)} + \epsilon_i^{(0)}, \\ Y_i &= D_i Y_i^{(1)} + (1 - D_i) Y_i^{(0)}. \end{aligned} \tag{3.7}$$

This specification implies the linearity (in parameters) of mean response functions $\mu_1\left(\bar{D}_{\mathcal{N}i}, X_i\right)$, $\mu_0\left(\bar{D}_{\mathcal{N}i}, X_i\right)$, and $\nu(Z_i, X_i)$. It also allows the outcome Y_i to depend on the (binary) individual treatment D_i, (multi-valued) neighbourhood treatment $\bar{D}_{\mathcal{N}i}$, and their interaction.[3]

Assumption 2. $\left(\epsilon_i^{(D)}, \epsilon_i^{(1)}, \epsilon_i^{(0)}\right) \perp\!\!\!\perp (X_i, Z_i)$, where $\perp\!\!\!\perp$ denotes statistical independence.

This is a stronger independence assumption compared to *exclusion restriction* in conventional IV literature, which requires Z to be exogenous with respect to the outcome processes after conditioning on observed covariates X. An increasing number of empirical studies adopt this assumption in order to overcome the limitations associated with instrumental variation (see, e.g., Brinch et al., 2017; Carneiro et al., 2011; Cornelissen et al., 2018).

Assumption 3. For $i = 1, \ldots, n$

$$\epsilon_i = \begin{bmatrix} \epsilon_i^{(D)} \\ \epsilon_i^{(1)} \\ \epsilon_i^{(0)} \end{bmatrix} \overset{ind}{\sim} \sum_{g=1}^{G} \pi_g \mathcal{N}\left(0, \Sigma_g\right) \tag{3.8}$$

$$\text{where} \quad \sum_{g=1}^{G} \pi_g = 1; \quad \Sigma_g = \begin{bmatrix} 1 & \sigma_{1D,g} & \sigma_{0D,g} \\ & \sigma_{1,g}^2 & \sigma_{10,g} \\ & & \sigma_{0,g}^2 \end{bmatrix}$$

The variance parameter of the disturbance term in the selection equation for the binary

[3] In this way, it connects to *Correlated Random Coefficients (CRC) Model* for potential outcome functions

$$Y_i(D_i, \bar{D}_{\mathcal{N}i}) = \gamma_i D_i + \delta_i \bar{D}_{\mathcal{N}i} + \eta_i D_i \bar{D}_{\mathcal{N}i} + \epsilon_i^{(Y)}$$

treatment indicator D is normalised to unity: $\sigma^2_{D^*} = 1$; because it is only identified up to scale. Without this condition, multiple values for the model parameters give rise to the same value for the likelihood function (see, e.g., Doğan and Taşpinar, 2018).

Assumption 4. For $i = 1, \ldots, n$

$$\begin{aligned} Y_i^{(1)}(\bar{d}_{\mathcal{N}}) &\perp\!\!\!\perp \bar{D}_{\mathcal{N}i} \mid X_i \quad \forall \bar{d}_{\mathcal{N}} \in [0,1], \\ Y_i^{(0)}(\bar{d}_{\mathcal{N}}) &\perp\!\!\!\perp \bar{D}_{\mathcal{N}i} \mid X_i \quad \forall \bar{d}_{\mathcal{N}} \in [0,1]. \end{aligned} \tag{3.9}$$

This assumption is crucial for identifying the causal indirect effects; it could be regarded as *weak unconfoundedness* for the continuous exposure, analogous to continuous-treatment settings, with conditional independence being required to hold for each value of the treatment (Hirano and Imbens, 2004). Unconfoundedness of neighbourhood treatment rules out the presence of unmeasured confounding variables.[4]

Proposition 1. Suppose Assumptions 1 – 4 hold. Denote $\sigma_{1D} = \mathrm{Cov}\left(\epsilon_i^{(1)}, \epsilon_i^{(D)}\right)$ and $\sigma_{0D} = \mathrm{Cov}\left(\epsilon_i^{(0)}, \epsilon_i^{(D)}\right)$. Then, parameters $\delta^{(1)}, \delta^{(0)}, \beta^{(1)}, \beta^{(0)}, \sigma_{1D}, \sigma_{0D}$ are identified.

Proof. See Appendix B.1.

Remark. The proposed model nests the Generalised Roy model (Heckman and Vytlacil, 2005) by maintaining its latent-index selection and correlated potential outcomes, but extends the framework to allow outcomes to depend on the neighbours' treatments through $\bar{D}_{\mathcal{N}i}$ with potentially different slopes when treated versus untreated, $\delta^{(1)}$ and $\delta^{(0)}$. Our resulting model thereby captures both endogenous selection into treatment and network/spatial interference. When neighbourhood exposures are absent ($\bar{D}_{\mathcal{N}i} = 0 \; \forall i$) or exert no effects $\delta^{(0)} = \delta^{(1)} = 0$, the model collapses to the canonical Generalised Roy framework. The identification of the parameters $(\sigma_{1D}, \sigma_{0D})$ in Proposition 1 follows a same principle as in the Heckman selection models (Heckman, 1979): the exclusion of Z_i from the outcome equation and the joint distributional assumption on $\left(\epsilon_i^{(D)}, \epsilon_i^{(1)}, \epsilon_i^{(0)}\right)$ imply that the conditional expectations of outcome errors vary with the inverse Mills ratio derived from the selection index $\nu(X_i, Z_i)$. Our model, however, first partials out the neighbourhood exposure $\bar{D}_{\mathcal{N}i}$ before applying this logic, thus unifying selection correction and spillover identification under Assumptions 1 – 4.

[4] This assumption implicitly requires that units' enrolling decision $\{D_i\}_{i=1}^n$ are independent.

3.2.2.2 Causal Estimands

Our object of interest includes the direct and indirect causal effects of the binary treatment D.

Indirect Causal Effects

Average Indirect Effect (AIE) is the average effect of exogenously increasing an individual's neighbourhood treatment $\bar{D}_{\mathcal{N}i}$ from $\bar{d}_{\mathcal{N}}$ to $\bar{d}_{\mathcal{N}} + \Delta$ while holding the individual's own treatment status D_i fixed at d. In particular, this results in two types of Average Indirect Effects (spillover effects) - either on the treated or on the untreated as follows

$$\begin{aligned} \text{AIE}^{(1)}(\bar{d}_{\mathcal{N}}, \Delta) &\coloneqq \mathbb{E}\left[Y_i^{(1)}(\bar{d}_{\mathcal{N}} + \Delta) - Y_i^{(1)}(\bar{d}_{\mathcal{N}})\right]; \\ \text{AIE}^{(0)}(\bar{d}_{\mathcal{N}}, \Delta) &\coloneqq \mathbb{E}\left[Y_i^{(0)}(\bar{d}_{\mathcal{N}} + \Delta) - Y_i^{(0)}(\bar{d}_{\mathcal{N}})\right]. \end{aligned} \tag{3.10}$$

Suppose assumptions A1-A4 hold, we can identify *Average Partial Indirect Effects* (APIE), which measure the average partial effects of changing the neighbourhood treatment $\bar{D}_{\mathcal{N}}$ on the treated and on the untreated

$$\text{APIE}^{(1)} = \delta^{(1)} \text{ and } \text{APIE}^{(0)} = \delta^{(0)}. \tag{3.11}$$

Direct Causal Effects

Average Direct Treatment Effect (ADTE), is the average effect of exogenously changing an individual's own treatment D_i from 0 to 1 while holding their neighbourhood treatment fixed at $\bar{d}_{\mathcal{N}}$

$$\text{ADTE}(\bar{d}_{\mathcal{N}}) \coloneqq \mathbb{E}\left[Y_i^{(1)} - Y_i^{(0)} \mid \bar{D}_{\mathcal{N}i} = \bar{d}_{\mathcal{N}}, X_i\right], \tag{3.12}$$

where the expectation is taken over all individuals in the population. According to assumptions A1-A4, we obtain

$$\text{ADTE}\left(\bar{d}_{\mathcal{N}}\right) = \left(\delta^{(1)} - \delta^{(0)}\right)\bar{d}_{\mathcal{N}} + \left(\beta^{(1)} - \beta^{(0)}\right)\mathbb{E}\left[X_i\right]. \tag{3.13}$$

Evaluated at mean values of the covariates X, ADTE would exhibit heterogeneity in treatment effects due to $\bar{d}_{\mathcal{N}}$ if $\delta^{(1)} - \delta^{(0)} \neq 0$. Furthermore, $\delta^{(1)} - \delta^{(0)}$ implies the patterns of interaction effects between individual treatment and neighbourhood treatment: Positive interaction ($\delta^{(1)} -$

$\delta^{(0)} > 0$) means the treatment is more valuable when more of neighbours are treated. In contrast, negative interaction ($\delta^{(1)} - \delta^{(0)} < 0$) means the treatment is more valuable when less of neighbours are treated.

Figure 3.1 illustrates a hypothetical example of Opportunity Zone (OZ) program, where outcome variable Y indicates new development in census tract. Both average potential outcomes have positive slopes ($\delta^{(1)} > 0$ and $\delta^{(0)} > 0$), corresponds to positive average partial indirect effects defined in (3.11). This means that new development in a census tract *increases* if more of its neighbouring areas are designated as OZs. As $\delta^{(1)} > \delta^{(0)}$, the spillover is even more beneficial if the census tract is treated, implying a positive interaction effect of the policy. The average direct treatment effect defined in (3.13) and represented by the gap between two lines is positive for all $\bar{d}_{\mathcal{N}}$ and increases as $\bar{d}_{\mathcal{N}}$ does: Receiving investment tax incentive is more valuable to a census tract when more of its neighbouring areas obtain them.

Marginal Direct Treatment Effect (MDTE), is the average direct causal effect of D on Y for individuals with unobserved heterogeneity $V = v$ and observed neighbourhood treatment $\bar{D}_{\mathcal{N}} = \bar{d}_{\mathcal{N}}$

$$\text{MDTE}\left(\bar{d}_{\mathcal{N}}, v\right) \coloneqq \mathbb{E}\left[Y_i^{(1)} - Y_i^{(0)} \mid \bar{D}_{\mathcal{N}i} = \bar{d}_{\mathcal{N}}, V_i = v, X_i\right]. \tag{3.14}$$

Evaluated at mean values of the covariates X, MDTE is a function of not only $\bar{d}_{\mathcal{N}}$ but also unmeasured resistance to the treatment v. This goes beyond average direct treatment effect to uncover direct treatment effects at different points in the distribution of unobserved heterogeneity, providing insights into how effects vary among agents who are marginally indifferent to receiving the treatment. According to assumptions A1-A4,

$$\begin{aligned}
&\text{MDTE}\left(\bar{d}_{\mathcal{N}}, v\right) \\
&= \left(\delta^{(1)} - \delta^{(0)}\right)\bar{d}_{\mathcal{N}} + \mathbb{E}\left[\epsilon_i^{(1)} - \epsilon_i^{(0)} \mid V_i = v\right] + \left(\beta^{(1)} - \beta^{(0)}\right)\mathbb{E}\left[X_i\right] \\
&= \left(\delta^{(1)} - \delta^{(0)}\right)\bar{d}_{\mathcal{N}} + \mathbb{E}\left[\epsilon_i^{(1)} - \epsilon_i^{(0)} \mid \epsilon_i^{(D)} = -F^{-1}(v)\right] + \left(\beta^{(1)} - \beta^{(0)}\right)\mathbb{E}\left[X_i\right].
\end{aligned} \tag{3.15}$$

Without loss of generality, we consider $\mathbb{E}\left[\epsilon_i^{(1)} \mid \epsilon_i^{(D)} = \varepsilon\right]$. From (3.8), we know that

$$\epsilon_i = \begin{bmatrix} \epsilon_i^{(D)} \\ \epsilon_i^{(1)} \end{bmatrix} \overset{ind}{\sim} \sum_{g=1}^{G} \pi_g N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \sigma_{1D,g} \\ \sigma_{1D,g} & \sigma_{1,g}^2 \end{bmatrix}\right).$$

Hence, $\epsilon_i^{(D)} \sim N(0,1)$ and

$$\begin{aligned}\mathbb{E}\left[\epsilon_i^{(1)} \mid \epsilon_i^{(D)} = \varepsilon\right] &= \sum_{g=1}^{G} \pi_g \frac{f_g(\varepsilon)}{f(\varepsilon)} \mathbb{E}_g\left[\epsilon_i^{(1)} \mid \epsilon_i^{(D)} = \varepsilon\right] \\ &= \sum_{g=1}^{G} \pi_g \frac{\phi(\varepsilon)}{\phi(\varepsilon)} \frac{\sigma_{1D,g}}{\sqrt{1}} \epsilon = \sum_{g=1}^{G} \pi_g \sigma_{1D,g} \varepsilon.\end{aligned} \tag{3.16}$$

Therefore,

$$\text{MDTE}\left(\bar{d}_{\mathcal{N}}, v\right) = \left(\delta^{(1)} - \delta^{(0)}\right) \bar{d}_{\mathcal{N}} - \sum_{g=1}^{G} \pi_g \frac{\sigma_{1D,g} - \sigma_{0D,g}}{\sqrt{1}} \Phi^{-1}(v) + \left(\beta^{(1)} - \beta^{(0)}\right) \mathbb{E}\left[X_i\right]. \tag{3.17}$$

In a special case when $G = 1$ (i.e. trivariate normal distribution), the function becomes

$$\text{MDTE}\left(\bar{d}_{\mathcal{N}}, v\right) = \left(\delta^{(1)} - \delta^{(0)}\right) \bar{d}_{\mathcal{N}} + \frac{\sigma_{1D} - \sigma_{0D}}{\sqrt{1}} \Phi^{-1}(v) + \left(\beta^{(1)} - \beta^{(0)}\right) \mathbb{E}\left[X_i\right]. \tag{3.18}$$

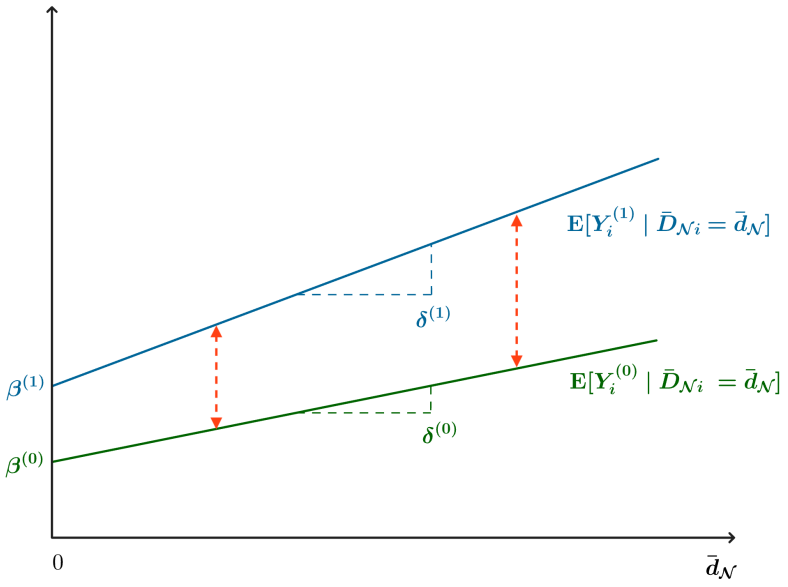

Figure 3.1: Illustration of average potential outcome when being treated (the top line) and when being untreated (the bottom line), as functions of the neighbourhood treatment term $\bar{d}_{\mathcal{N}}$. The corresponding slopes, $\delta^{(1)}$ and $\delta^{(0)}$, are the average partial indirect effect when being treated and when being untreated, respectively. The gap between the two lines is the average direct treatment effect.

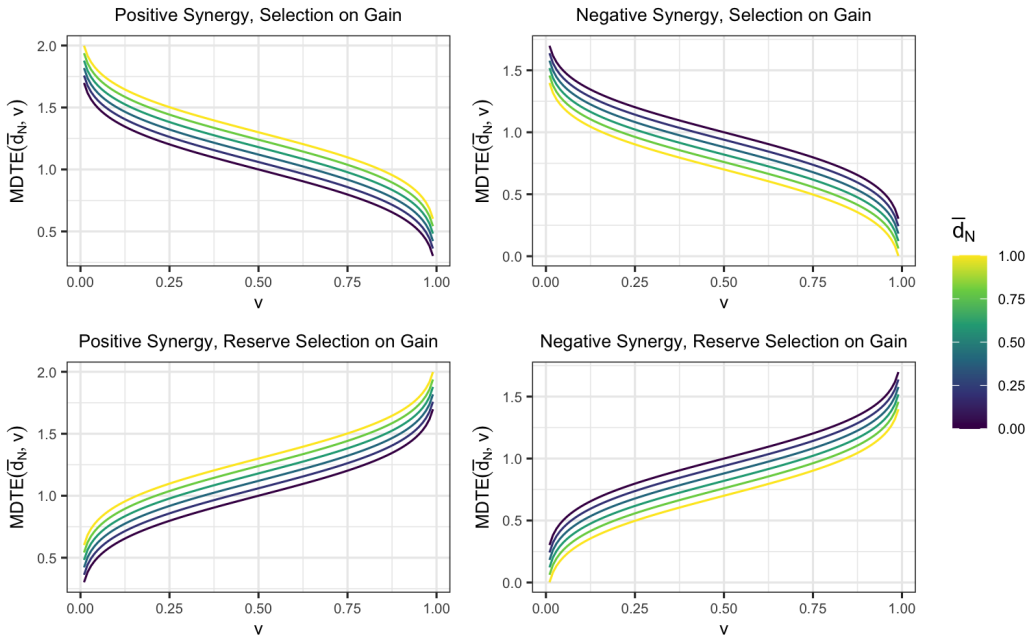

Figure 3.2: Illustration of Marginal Direct Treatment Effects with respect to unmeasured resistance to the treatment v and neighbourhood treatment $\bar{d}_{\mathcal{N}}$. The slope of MDTE curves reveals the patterns of selection into treatment: a rising MDTE curve which exhibits an upward sloping shape ($\sigma_{1D} - \sigma_{0D} > 0$) indicates a pattern of reverse selection on gains in unobserved characteristics; a falling MDTE curve which exhibits a downward sloping shape ($\sigma_{1D} - \sigma_{0D} < 0$) indicates a pattern of positive selection on gains in unobserved characteristics; a flat MDTE curve ($\sigma_{1D} = \sigma_{0D}$) implies there is no selection based on unobserved gains. The dispersion of MDTE curves characterises the patterns of interaction effects, reflecting in $\delta^{(1)} - \delta^{(0)}$ similar to ADTE.

3.3 Bayesian Estimation and Inference

3.3.1 Bayesian Data Augmentation

In this section, we develop Bayesian estimation methods for the model parameters $\delta^{(1)}, \delta^{(0)}, \beta^{(1)}, \beta^{(0)}, \sigma_{1D}$, and σ_{0D}, followed by the causal estimands in Section 3.2.2. Inference about these causal estimands is based on their respective posterior distributions, which accounts for the uncertainty in parameter estimation.

Rewrite $\mathbf{P}_i = [Z_i, X_i]^\top$, $\mathbf{Q}_i = [\bar{D}_{\mathcal{N}i}, X_i]^\top$, $\gamma = \left[\alpha, \beta^{(D)}\right]^\top$, $\kappa_1 = \left[\delta^{(1)}, \beta^{(1)}\right]^\top$, and $\kappa_0 = \left[\delta^{(0)}, \beta^{(0)}\right]^\top$. Then, the model (3.7) can be more compactly written as the following standard form of Generalised Roy model:

$$\begin{aligned} D_i^* &= \mathbf{P}_i^\top \gamma + \epsilon_i^{(D)}, \\ Y_i^{(1)} &= \mathbf{Q}_i^\top \kappa_1 + \epsilon_i^{(1)}, \\ Y_i^{(0)} &= \mathbf{Q}_i^\top \kappa_0 + \epsilon_i^{(0)}, \end{aligned} \tag{3.19}$$

$$\begin{aligned} D_i &= \mathbb{1}\{D_i^* > 0\}, \\ Y_i &= D_i Y_i^{(1)} + (1 - D_i) Y_i^{(0)}. \end{aligned}$$

As a starting point, we consider a fully parametric approach by imposing the trivariate normal distribution assumption on the error terms

$$\epsilon_i = \begin{bmatrix} \epsilon_i^{(D)} & \epsilon_i^{(1)} & \epsilon_i^{(0)} \end{bmatrix}^\top \overset{iid}{\sim} \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}) \tag{3.20}$$

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_D^2 = 1 & \sigma_{1D} & \sigma_{0D} \\ & \sigma_1^2 & \sigma_{10} \\ & & \sigma_0^2 \end{bmatrix} = \begin{bmatrix} 1 & \rho_{1D}\sigma_1 & \rho_{0D}\sigma_0 \\ & \sigma_1^2 & \rho_{10}\sigma_1\sigma_0 \\ & & \sigma_0^2 \end{bmatrix}$$

where $\rho_{10} \coloneqq \text{corr}(\epsilon_1, \epsilon_0)$, $\rho_{1D} \coloneqq \text{corr}(\epsilon_1, \epsilon_D)$, $\rho_{0D} \coloneqq \text{corr}(\epsilon_0, \epsilon_D)$.

A challenge posed in model (3.19) is that the latent utility (D_i^*) and the potential outcomes ($Y_i^{(1)}$ and $Y_i^{(0)}$), for $i = 1, \ldots, n$, are unobserved. Instead, we can only observe the treatment status D_i and the revealed outcome Y_i of each individual. These missing data perspectives necessitate suitable techniques to estimate all model parameters. We therefore consider

the Bayesian data augmentation approach which involves two main stages implemented iteratively: In *imputation stage*, we augment the posterior with the latent utility (D_i^*) and the missing potential outcome (Y_i^{miss}), for $i = 1, \ldots, n$. In *posterior stage*, conditional on these unobserved latent quantities, the regression coefficients and covariance matrix can be sampled in a straightforward manner.

To work with augmented outcome data, for each individual $i = 1, \ldots, n$, we let

$$\mathbf{L}_i^* = \begin{bmatrix} D_i^* \\ Y_i^{(1)} \\ Y_i^{(0)} \end{bmatrix} = \begin{bmatrix} D_i^* \\ D_i Y_i + (1 - D_i) Y_i^{miss} \\ D_i Y_i^{miss} + (1 - D_i) Y_i \end{bmatrix}, \quad \mathbf{R}_i = \begin{bmatrix} \mathbf{P}_i^\top & 0_{1\times k_q} & 0_{1\times k_q} \\ 0_{1\times k_p} & \mathbf{Q}_i^\top & 0_{1\times k_q} \\ 0_{1\times k_p} & 0_{1\times k_q} & \mathbf{Q}_i^\top \end{bmatrix}, \tag{3.21}$$

where k_p and k_q denote the numbers of column vectors in $\mathbf{P}$ and $\mathbf{Q}$, respectively.

We also stack each equation independently

$$\mathbf{L}_{3n\times 1}^* = \begin{bmatrix} \mathbf{D}^* \\ \mathbf{Y}^{(1)} \\ \mathbf{Y}^{(0)} \end{bmatrix} = \begin{bmatrix} \mathbf{D}^* \\ \mathbf{D} \circ \mathbf{Y} + (1 - \mathbf{D}) \circ \mathbf{Y}^{miss} \\ \mathbf{D} \circ Y^{miss} + (1 - \mathbf{D}) \circ \mathbf{Y} \end{bmatrix}, \mathbf{R}_{3n\times k} = \begin{bmatrix} \mathbf{P} & 0_{n\times k_q} & 0_{n\times k_q} \\ 0_{n\times k_p} & \mathbf{Q} & 0_{n\times k_q} \\ 0_{n\times k_p} & 0_{n\times k_q} & \mathbf{Q} \end{bmatrix}, \epsilon = \begin{bmatrix} \epsilon_D \\ \epsilon_1 \\ \epsilon_0 \end{bmatrix}, \tag{3.22}$$

where $k = k_p + 2k_q$.

Denote $\theta = (\gamma^\top, \beta_1^\top, \beta_0^\top)^\top$. Then

$$\begin{aligned} &\mathbf{L}_i^* = \mathbf{R}_i \theta + \epsilon_i \quad \text{for } i = 1, \ldots, n; \\ &\mathbf{L}^* = \mathbf{R}\theta + \epsilon; \quad \text{and} \quad \mathbb{E}[\epsilon^\top \epsilon] = \mathbf{\Omega} = \mathbf{\Sigma} \otimes \mathbf{I}_n. \end{aligned} \tag{3.23}$$

Likelihood function

The augmented likelihood (also known as complete-data likelihood), which includes latent quantities, factorises as

$$p(\mathbf{Y}, \mathbf{D}, \mathbf{L}^* \mid \theta, \mathbf{\Sigma}) = p(\mathbf{Y}, \mathbf{D} \mid \theta, \mathbf{\Sigma}, \mathbf{L}^*) \cdot p(\mathbf{L}^* \mid \theta, \mathbf{\Sigma}), \tag{3.24}$$

where the first term encapsulates the observed data given latent variables and the second term represents how latent variables relate to parameters in our structural model.

The second term is straightforwardly derived from the representation of latent variables in (3.23) and the multivariate normal assumption (3.20)

$$\begin{aligned} p(\mathbf{L}^* \mid \theta, \mathbf{\Sigma}) &= (2\pi)^{-n} \left|\mathbf{\Sigma}\right|^{-n/2} \times \exp\left\{-\frac{1}{2}(\mathbf{L}^* - \mathbf{R}\theta)^\top \mathbf{\Omega}^{-1}(\mathbf{L}^* - \mathbf{R}\theta)\right\} \\ &= (2\pi)^{-n} \left|\mathbf{\Sigma}\right|^{-n/2} \times \prod_{i=1}^{n} \exp\left\{-\frac{1}{2}(\mathbf{L}_i^* - \mathbf{R}_i\theta)^\top \mathbf{\Sigma}^{-1}(\mathbf{L}_i^* - \mathbf{R}_i\theta)\right\}, \end{aligned} \tag{3.25}$$

Since knowing $(Y_i^{(1)}, Y_i^{(0)})$ and the sign of D_i^* perfectly determines the value of D_i and Y_i, we can derive the first term in (3.24) as follows

$$\begin{aligned} &p(\mathbf{Y}, \mathbf{D} \mid \theta, \mathbf{\Sigma}, \mathbf{L}^*) \\ &= \prod_{i=1}^{n} \left[\mathbb{1}(D_i^* > 0)\mathbb{1}(D_i = 1)\mathbb{1}(Y_i = Y_i^{(1)}) + \mathbb{1}(D_i^* \leq 0)\mathbb{1}(D_i = 0)\mathbb{1}(Y_i = Y_i^{(0)})\right]. \\ &= \prod_{\{i:D_i=1\}} p\left(Y_i^{(1)}, D_i^* > 0\right) \prod_{\{i:D_i=0\}} p\left(Y_i^{(0)}, D_i^* \leq 0\right) \\ &= \prod_{\{i:D_i=1\}} \int_0^\infty p\left(Y_i^{(1)}, D_i^*\right) dD_i^* \prod_{\{i:D_i=0\}} \int_{-\infty}^0 p\left(Y_i^{(0)}, D_i^*\right) dD_i^* \\ &= \prod_{\{i:D_i=1\}} \int_0^\infty p\left(D_i^* \mid Y_i^{(1)}\right) p\left(Y_i^{(1)}\right) dD_i^* \prod_{\{i:D_i=0\}} \int_{-\infty}^0 p\left(D_i^* \mid Y_i^{(0)}\right) p\left(Y_i^{(0)}\right) dD_i^* \\ &= \prod_{\{i:D_i=1\}} \Phi\left(\frac{u_{Di} + \rho_{1D}u_{1i}}{(1-\rho_{1D}^2)^{-1/2}}\right) \frac{1}{\sigma_1}\phi(u_{1i}) \prod_{\{i:D_i=0\}} \left[1 - \Phi\left(\frac{u_{Di} + \rho_{0D}u_{0i}}{(1-\rho_{0D}^2)^{-1/2}}\right)\right] \frac{1}{\sigma_0}\phi(u_{0i}), \end{aligned} \tag{3.26}$$

where

$$u_{Di} \coloneqq \mathbf{P}_i^\top \gamma; \quad u_{1i} \coloneqq \frac{Y_i^{(1)} - \mathbf{Q}_i^\top \kappa_1}{\sigma_1}; \quad u_{1i} \coloneqq \frac{Y_i^{(0)} - \mathbf{Q}_i^\top \kappa_0}{\sigma_0}.$$

Prior specification

To carry out full Bayesian inference, a suitable prior specification is necessary. Let $p(\theta, \mathbf{\Sigma})$ be the joint prior distribution function of regression coefficients θ and the covariance matrix $\mathbf{\Sigma}$. We consider a multivariate Normal prior for the regression coefficients

$$\theta \sim \mathcal{N}_{k_p+2k_q}(\mu_{\theta_o}, \mathbf{V}_{\theta_o}). \tag{3.27}$$

Placing a prior over the full covariance matrix $\mathbf{\Sigma}$ mitigate the issue of unidentified correlation parameters between two potential outcomes[5]. In addition, because the $(1,1)^{\text{th}}$

[5]Unidentified cross-regime correlation parameter satisfies $\underline{\rho}_{10} \leq \rho_{10} \leq \overline{\rho}_{10}$; where $\underline{\rho}_{10} = \underline{\rho}_{10}(\rho_{1D}, \rho_{0D}) = \rho_{1D}\rho_{0D} - [(1-\rho_{1D}^2)(1-\rho_{0D}^2)]^{1/2}$ and $\overline{\rho}_{10} = \overline{\rho}_{10}(\rho_{1D}, \rho_{0D}) = \rho_{1D}\rho_{0D} + [(1-\rho_{1D}^2)(1-\rho_{0D}^2)]^{1/2}$ (Li et

element Σ_{11} equals to unity, a natural choice of prior for the covariance matrix $\mathbf{\Sigma}$ is the Inverse Wishart prior along with this normalisation constraint

$$\mathbf{\Sigma} \sim \mathcal{W}^{-1}(\mathbf{\Psi}_o, \nu_o)\mathbb{1}_{(\Sigma_{11}=1)}. \tag{3.28}$$

However, this restriction makes it difficult to sample $\mathbf{\Sigma}$ from its posterior distribution directly[6]. We therefore facilitate the computation by employing the parameter expansion technique. Specifically, we reparameterize the model by introducing an expansion parameter to improve the convergence rate of the resulting Markov chains. A general two-step procedure can be described as follows. First, the model is transformed by the expansion parameter in such a way that the transformed model has an unconstrained covariance matrix and the computational complications are circumvented for the posterior analysis. Second, an Inverse Wishart distribution is assigned as a prior to the unconstrained covariance matrix. At this step, the priors for the expansion parameter and the constrained covariance matrix are also determined to complete a Gibbs sampler.

To begin, we let a positive scalar parameter τ be the expansion parameter and define

$$\tilde{\epsilon}_i = \begin{bmatrix} \tau & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \times (\mathbf{L}_i^* - \mathbf{R}_i\theta). \tag{3.29}$$

Given the normality assumption, we have $\epsilon^* \mid \theta, \tilde{\mathbf{\Sigma}} \sim N\left(\mathbf{0}_{3\times 3}, \tilde{\mathbf{\Sigma}}\right)$, where $\tilde{\mathbf{\Sigma}}$ is the unconstrained covariance matrix of $(\epsilon_{Di}, \epsilon_{1i}, \epsilon_{0i})^\top$ as follows

$$\tilde{\mathbf{\Sigma}}_{3\times 3} = \begin{bmatrix} \tau & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \times \mathbf{\Sigma}_{3\times 3} \times \begin{bmatrix} \tau & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \tau^2 & \rho_{1D}\tau\sigma_1 & \rho_{1D}\tau\sigma_0 \\ \rho_{1D}\tau\sigma_1 & \sigma_1^2 & \rho_{10}\sigma_1\sigma_0 \\ \rho_{1D}\tau\sigma_0 & \rho_{10}\sigma_1\sigma_0 & \sigma_0^2 \end{bmatrix}. \tag{3.30}$$

al., 2004).

[6]Covariance matrix $\mathbf{\Sigma}$: $\mathbf{\Sigma} \mid \Theta_{-\mathbf{\Sigma}}, \mathbf{Y}, \mathbf{D} \sim \mathcal{W}^{-1}\left(\mathbf{M} + \mathbf{\Psi}_o, n + \nu_o,\right) \mathbb{1}_{\{\Sigma_{11}=1\}}$, with

$$\mathbf{M} = \begin{bmatrix} \epsilon_D^\top \epsilon_D & \epsilon_D^\top \epsilon_1 & \epsilon_D^\top \epsilon_0 \\ \epsilon_1^\top \epsilon_D & \epsilon_1^\top \epsilon_1 & \epsilon_1^\top \epsilon_0 \\ \epsilon_0^\top \epsilon_D & \epsilon_0^\top \epsilon_1 & \epsilon_0^\top \epsilon_0 \end{bmatrix}; \quad \epsilon_D = \mathbf{D}^* - \mathbf{P}\gamma; \quad \epsilon_1 = \mathbf{Y}^{(1)} - \mathbf{Q}\kappa_1; \quad \epsilon_0 = \mathbf{Y}^{(0)} - \mathbf{Q}\kappa_0.$$

Then, the Inverse Wishart distribution can be assigned as a prior for $\tilde{\mathbf{\Sigma}}$

$$\tilde{\mathbf{\Sigma}} \sim \mathcal{W}^{-1}(\mathbf{\Psi}_o, \nu_o), \tag{3.31}$$

which implies the following prior for τ^2 and $\mathbf{\Sigma}$

$$p(\tau^2 \mid \mathbf{\Sigma}) \propto \left[\left(\frac{1 + 2\rho_{10}\rho_{1D}\rho_{0D} - \rho_{10}^2 - \rho_{1D}^2 - \rho_{0D}^2}{1 - \rho_{10}^2} \right) \chi^2_{(\nu_o+4)} \right]^{-1}, \tag{3.32}$$

$$p(\mathbf{\Sigma}) \propto |\mathbf{\Sigma}|^{-(\nu_o+4)/2} \exp \left\{ -\frac{1}{2\tau^2} \times \frac{1 - \rho_{10}^2}{1 + 2\rho_{10}\rho_{1D}\rho_{0D} - \rho_{10}^2 - \rho_{1D}^2 - \rho_{0D}^2} \right\}. \tag{3.33}$$

Details of the derivations for (3.32) and (3.33) can be found in Appendix B.2.1. The advantage of this approach is documented in the literature on sample selection models (Ding, 2014; Doğan and Taşpinar, 2018), suggesting that the algorithm improves the convergence rate of resulting Markov chains. Alternatively, we also adopt the marginal data augmentation technique similar to Imai and Van Dyk (2005)[7] for the multinomial probit model, which incorporates the expansion parameter differently. We present our corresponding algorithm in Appendix B.2.2. In the Gaussian Generalised Roy setting, the original approach by Li et al. (2004) and Poirier and Tobias (2003) is fixing the unidentified parameter during the posterior analysis, as per Nobile's (2000) suggestion. Among these algorithms, we found the first approach achieved the best performance during initial simulation exercises, while others might encounter numerical issues. The rest of this paper focuses on the first algorithm to discuss posterior analysis, Monte Carlo experiments with synthetic data, and the empirical application.

3.3.2 Computation

The augmented join posterior including latent quantities $\mathbf{L}^*$ can be expressed as

$$p(\theta, \mathbf{\Sigma}, \mathbf{L}^* \mid \mathbf{Y}, \mathbf{D}) \propto p(\theta, \mathbf{\Sigma}) \cdot p(\mathbf{L}^* \mid \theta, \mathbf{\Sigma}) \cdot p(\mathbf{Y}, \mathbf{D} \mid \theta, \mathbf{\Sigma}, \mathbf{L}^*). \tag{3.34}$$

Accordingly, we can compute conditional posterior distribution for latent quantities and model parameters as detailed below.

First, we can impute the latent quantities based on their conditional posterior distribution given the observed data and the parameters.

[7] There is a correction in Jiao and Dyk (2015).

(a) Missing potential outcome $\mathbf{Y}^{miss}$

$$Y_i^{miss} \mid \Theta_{-Y_i^{miss}}, \mathbf{Y}, \mathbf{D} \overset{ind}{\sim} \mathcal{N}\left((1-D_i)\mu_{1i} + D_i\mu_{0i}, (1-D_i)V_{1i} + D_iV_{0i}\right), \quad (3.35)$$

where

$$\mu_{1i} = \mathbf{Q}_i^\top \kappa_1 + (D_i^* - \mathbf{P}_i^\top \gamma)\left[\frac{\sigma_0^2\sigma_{1D} - \sigma_{10}\sigma_{0D}}{\sigma_0^2 - \sigma_{0D}^2}\right] + (Y_i - \mathbf{Q}_i^\top \kappa_0)\left[\frac{\sigma_{10} - \sigma_{0D}\sigma_{1D}}{\sigma_0^2 - \sigma_{0D}^2}\right],$$
$$\mu_{0i} = \mathbf{Q}_i^\top \kappa_0 + (D_i^* - \mathbf{P}_i^\top \gamma)\left[\frac{\sigma_1^2\sigma_{0D} - \sigma_{10}\sigma_{1D}}{\sigma_1^2 - \sigma_{1D}^2}\right] + (Y_i - \mathbf{Q}_i^\top \kappa_1)\left[\frac{\sigma_{10} - \sigma_{0D}\sigma_{1D}}{\sigma_1^2 - \sigma_{1D}^2}\right],$$
$$V_{1i} = \sigma_1^2 - \frac{\sigma_{1D}^2\sigma_0^2 - 2\sigma_{10}\sigma_{0D}\sigma_{1D} + \sigma_{10}^2}{\sigma_0^2 - \sigma_{0D}^2}, \quad V_{0i} = \sigma_0^2 - \frac{\sigma_{0D}^2\sigma_1^2 - 2\sigma_{10}\sigma_{0D}\sigma_{1D} + \sigma_{10}^2}{\sigma_1^2 - \sigma_{1D}^2}.$$

(b) Latent utility D^*

$$D_i^* \mid \Theta_{-D_i^*}, \mathbf{Y}, \mathbf{D} \overset{ind}{\sim} \begin{cases} \mathcal{TN}_{(0,+\infty)}(\mu_{Di}, V_{Di}) & \text{if } D_i = 1 \\ \mathcal{TN}_{(-\infty,0]}(\mu_{Di}, V_{Di}) & \text{if } D_i = 0 \end{cases}, \quad (3.36)$$

where $\mathcal{TN}$ denotes a truncated normal distribution and

$$\mu_{Di} = \mathbf{P}_i^\top \gamma + \left[D_iY_i + (1-D_i)Y_i^{miss} - \mathbf{Q}_i^\top \kappa_1\right]\left[\frac{\sigma_0^2\sigma_{1D} - \sigma_{10}\sigma_{0D}}{\sigma_1^2\sigma_0^2 - \sigma_{10}^2}\right],$$
$$+ \left[D_iY_i^{miss} + (1-D_i)Y_i - \mathbf{Q}_i^\top \kappa_0\right]\left[\frac{\sigma_1^2\sigma_{0D} - \sigma_{10}\sigma_{1D}}{\sigma_1^2\sigma_0^2 - \sigma_{10}^2}\right],$$
$$V_{Di} = 1 - \frac{\sigma_{1D}^2\sigma_0^2 - 2\sigma_{10}\sigma_{0D}\sigma_{1D} + \sigma_1^2\sigma_{0D}^2}{\sigma_1^2\sigma_0^2 - \sigma_{10}^2}.$$

After imputing the missing data, we can infer the posterior distribution of remaining parameters, conditioning on the complete data.

(c) Regression coefficients θ

$$\theta \mid \Theta_{-\theta}, \mathbf{Y}, \mathbf{D} \sim \mathcal{N}(\mu_\theta, \mathbf{V}_\theta), \quad (3.37)$$

where

$$\mu_\theta = \mathbf{V}_\theta[\mathbf{R}^\top \mathbf{\Omega}^{-1}\mathbf{L}^* + \mathbf{V}_{\theta_o}^{-1}\mu_{\theta_o},$$
$$\mathbf{V}_\theta = [\mathbf{R}^\top \mathbf{\Omega}^{-1}\mathbf{R} + \mathbf{V}_{\theta_o}^{-1}]^{-1}.$$

(d) Parameter expansion entails conjugate posterior distributions to estimate covariance

matrix $\boldsymbol{\Sigma}$

Conditional posterior of $\tilde{\boldsymbol{\Sigma}}$ is of the form

$$p(\tilde{\boldsymbol{\Sigma}} \mid \theta, \tilde{\epsilon}, \mathbf{L}^*, \mathbf{Y}, \mathbf{D}) \propto \left|\tilde{\boldsymbol{\Sigma}}\right|^{-(\nu_o+4)/2} \exp\left\{-\frac{1}{2}\operatorname{tr}(\boldsymbol{\Psi}_o\tilde{\boldsymbol{\Sigma}}^{-1})\right\} \times \left|\tilde{\boldsymbol{\Sigma}}\right|^{-n/2} \exp\left\{-\frac{1}{2}\tilde{\epsilon}^\top\tilde{\boldsymbol{\Omega}}^{-1}\tilde{\epsilon}\right\}$$
$$\propto \left|\tilde{\boldsymbol{\Sigma}}\right|^{-(\nu_o+4)/2} \exp\left\{-\frac{1}{2}\operatorname{tr}(\boldsymbol{\Psi}_o\tilde{\boldsymbol{\Sigma}}^{-1})\right\} \times \left|\tilde{\boldsymbol{\Sigma}}\right|^{-n/2} \exp\left\{-\frac{1}{2}\operatorname{tr}(\tilde{\mathbf{M}}\tilde{\boldsymbol{\Sigma}}^{-1})\right\}, \tag{3.38}$$

with

$$\tilde{\mathbf{M}} = \begin{bmatrix} \tau^2\epsilon_D^\top\epsilon_D & \tau\epsilon_D^\top\epsilon_1 & \tau\epsilon_D^\top\epsilon_0 \\ \tau\epsilon_1^\top\epsilon_D & \epsilon_1^\top\epsilon_1 & \epsilon_1^\top\epsilon_0 \\ \tau\epsilon_0^\top\epsilon_D & \epsilon_0^\top\epsilon_1 & \epsilon_0^\top\epsilon_0 \end{bmatrix}; \quad \epsilon_D = \mathbf{D}^* - \mathbf{P}\gamma; \quad \epsilon_1 = \mathbf{Y}_1 - \mathbf{Q}\kappa_1; \quad \epsilon_0 = \mathbf{Y}_0 - \mathbf{Q}\kappa_0;$$

thus,

$$\tilde{\boldsymbol{\Sigma}} \mid \theta, \tilde{\epsilon}, \mathbf{L}^*, \mathbf{Y}, \mathbf{D} \sim \mathcal{W}^{-1}(\tilde{\mathbf{M}} + \boldsymbol{\Psi}_o, n + \nu_o).$$

To marginalize over τ^2, we draw $\tilde{\boldsymbol{\Sigma}}$ from $\mathcal{W}^{-1}(\tilde{\mathbf{M}} + \boldsymbol{\Psi}_o, n + \nu_o)$ and set $\tau^2 = \tilde{\boldsymbol{\Sigma}}_{11}$.

To recover $\boldsymbol{\Sigma}$, we set

$$\boldsymbol{\Sigma} = \begin{bmatrix} 1/\tau & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \times \tilde{\boldsymbol{\Sigma}} \times \begin{bmatrix} 1/\tau & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{3.39}$$

The details of the implementation are presented in Algorithm 3.1.

Algorithm 3.1: Markov chain Monte Carlo (MCMC) Sampler I

Procedure

Step 0: Initialize parameters $s = 0, \theta^{[0]}, \mathbf{\Sigma}^{[0]}$ for MCMC-chains

while $s < S$ **do**

Step 1: Impute $\mathbf{L}^{*[s+1]}$ via $p(L^* \mid \mathbf{Y}, \mathbf{D}, \theta^{[s]}, \mathbf{\Sigma}^{[s]})$.

sampling $\mathbf{Y}^{miss}$ as in (3.35) and $\mathbf{D}^*$ as in (3.36),

setting $\mathbf{L}^{*[s+1]} = [\mathbf{D}^* \quad \mathbf{Y}^{(1)} \quad \mathbf{Y}^{(0)}]^\top$.

Step 2: Update $\theta^{[s+1]}$ from $p(\theta \mid \mathbf{Y}, \mathbf{D}, \mathbf{L}^{*[s+1]}, \mathbf{\Sigma}^{[s]})$ as in (3.37).

Step 3: Update $\mathbf{\Sigma}^{[s+1]}$ via $p(\mathbf{\Sigma} \mid \mathbf{Y}, \mathbf{D}, \theta^{[s+1]}, \mathbf{L}^{*[s+1]})$ by

(a) sampling $(\tau^2)^\star$ from $p(\tau^2 \mid \mathbf{\Sigma}^{[s]})$

$$(\tau^2)^\star \sim \left[\left(\frac{1 + 2\rho_{10}^{[s]}\rho_{1D}^{[s]}\rho_{0D}^{[s]} - (\rho_{10}^{[s]})^2 - (\rho_{1D}^{[s]})^2 - (\rho_{0D}^{[s]})^2}{1 - (\rho_{10}^{[s]})^2} \right) \chi^2_{(\nu_o+4)} \right]^{-1},$$

(b) calculating $(\epsilon^*)^\star$ via the transformation,

(c) sampling $\tilde{\mathbf{\Sigma}}^\star$ from $p(\tilde{\mathbf{\Sigma}} \mid \mathbf{Y}, \mathbf{D}, \theta^{[s+1]}, \mathbf{L}^{*[s+1]}, (\epsilon^*)^\star)$ as in (3.38),

(d) setting $(\tau^2)^\star = \tilde{\Sigma}^\star_{11}$,

(e) recovering $\mathbf{\Sigma}^{[s+1]}$ according to (3.39).

return $\mathbf{L}^{*[s+1]}, \theta^{[s+1]}, \mathbf{\Sigma}^{[s+1]}$

$s \leftarrow s + 1$

end while

end procedure

3.4 Simulation Study

To evaluate the performance of the proposed framework and Bayesian estimation algorithms, we design a simulation study with details described below. We focus on the primary setting in which both selection-on-unobservables and spillovers are present, as well as several scenarios with misspecifications. In addition to synthetic data, we also conduct an empirical Monte Carlo study which employs the Add Health friendship network data in Appendix B.4.

3.4.1 Data Generating Processes

We draw 5 exogenous variables $\tilde{X}_k$ $(k = 1, \ldots, 5)$ independently from the standard normal distribution and set $X = \left[\iota_n^\top, \tilde{X}^\top\right]^\top$. The instrumental variable Z is also generated from the same distribution. Regarding network/spatial structure, we design the weight matrix W based on the interaction scenario described in Liu and Lee (2010). W is a block diagonal matrices where each block represents the interaction structure of a group. Let the total sample involve $G = 30$ groups where the g^{th} group has the groups size m_g. We

allow m_g to vary across G groups by randomly assigning a value from the set of integers $\{\lfloor n/G \rfloor - 2, \lfloor n/G \rfloor - 1, \ldots, \lfloor n/G \rfloor + 3\}$ to each group size and adjust m_G such that $\sum_{g=1}^{G} m_g = n$. The weight matrix W_g for the g^{th} group is generated in two steps. First, for the i^{th} row of W_g ($i = 1, \ldots, m_g$), an integer value τ_{ig} is uniformly drawn from the set of integer values $\{1, 2, 3, 4\}$. Then, if $i + \tau_{ig} \leq m_g$, we set the $(i+1)^{\text{th}}, \ldots, (\tau_{gi}+1)^{\text{th}}$ elements of the i^{th} of W_g to be ones and the rest elements in that row to be zeros; otherwise, the first $(\tau_{ig} + i - m_g)$ entries of the i^{th} row are set to ones and the rest elements in that row are set to zeros. Finally, we set $W \coloneqq \text{diag}(W_1^\top + W_1, \ldots, W_g^\top + W_g)$ and transform to a row-normalised or doubly stochastic matrix.

The general data generating process (DGP) is based on model (3.7), which corresponds to[8]:

$$\begin{aligned} D_i &= \mathbb{1}\{Z_i\alpha + X_i\beta^{(D)} + \epsilon_i^{(D)} > 0\}, \text{ for } i = 1, \ldots, n \\ \bar{D}_{\mathcal{N}} &= WD, \\ Y^{(1)} &= \delta^{(1)}\bar{D}_{\mathcal{N}} + X\beta^{(1)} + \epsilon^{(1)}, \\ Y^{(0)} &= \delta^{(0)}\bar{D}_{\mathcal{N}} + X\beta^{(0)} + \epsilon^{(0)}, \\ Y &= D \circ Y^{(1)} + (1 - D) \circ Y^{(0)}, \end{aligned} \tag{3.40}$$

where the associated regression coefficients of covariates X in the selection equation and two potential outcome equations are $\beta^{(D)} = \left[0, \tilde{\beta}_{1\times5}^{(D)}\right]^\top$, $\beta^{(1)} = \left[2, \tilde{\beta}_{1\times5}^{(1)}\right]^\top$, $\beta^{(0)} = \left[1, \tilde{\beta}_{1\times5}^{(0)}\right]^\top$. Throughout, true values of the intercepts are fixed in all three equations, and other coefficients are generated from the independent uniform distributions $\mathcal{U}_{[-1,1]}$. The coefficient $\alpha = 1.5$ controls the strength of the instrument Z. The presence of spillovers is implied by $\delta^{(1)} = 1.5$ and $\delta^{(0)} = 0.5$. We specify a multivariate normal distribution of the error term $\epsilon_i = \left[\epsilon_i^{(D)}, \epsilon_i^{(1)}, \epsilon_i^{(0)}\right]^\top$ for $i = 1, \ldots, n$ as below

$$\epsilon_i = \left[\epsilon_i^{(D)}, \epsilon_i^{(1)}, \epsilon_i^{(0)}\right]^\top \overset{iid}{\sim} \mathcal{N}(0, \Sigma); \quad \Sigma = \begin{bmatrix} 1 & 0.9 & 0.7 \\ & 1 & 0.6 \\ & & 1 \end{bmatrix};$$

i.e., $\sigma_D^2 = \sigma_1^2 = \sigma_0^2 = 1$ and $(\rho_{1D}, \rho_{0D}, \rho_{10}) = (0.9, 0.7, 0.6)$, which accommodates positive correlation of unobservables.

We allow different sample sizes $n \in \{500; 1,000; 5,000\}$ of the dataset. For each of

[8]The operator $\circ$ denotes the Hadamard product (also known as the element-wise product).

the generated data sets, we specify two versions of models to be estimated by using the proposed Bayesian MCMC sampler - Algorithm 3.1. First, Gaussian Generalised Roy model without spillovers (`GGRM-noSI`) serves as the benchmark model, without neighbourhood term ($\bar{D}_{\mathcal{N}}$) and with a normal distribution of the error term. Second, Gaussian Generalised Roy model with spillovers (`GGRM-SI`) is the full model with neighbourhood term ($\bar{D}_{\mathcal{N}}$) and a normal distribution of the error term. We run each MCMC algorithm for $11,000$ iterations, with the first $1,000$ draws are discarded as a burn-in period. Throughout our simulation study, the parameters for the prior distributions are chosen as follows: $\mu_{\theta_o} = \mathbf{0}_{21}; \mathbf{V}_{\theta_o} = 10^2 * \mathbf{I}_{21\times 21}; \mathbf{\Psi}_o = \mathbf{I}_{3\times 3}; \nu_o = 4$. The number of independent replicates is $R = 1,000$.

3.4.2 Simulation Results

For each simulated dataset, given the posterior distribution of each model parameter resulted from MCMC draws, we derive the posterior mean for a point estimate and compute the corresponding 95% credible interval. We thereby calculate across the $1,000$ replicates the average bias and the root mean square error (RMSE) of the point estimates, followed by the coverage rate of the 95% credible intervals. The simulation results are presented in Table 3.1, where the true values of the DGP parameters are also listed.

Table 3.1: Simulation Results for Model Parameters

			Quantities of Interest				Other Parameters								
Model	Metric	n	$\delta^{(1)}$	$\delta^{(0)}$	$\delta^{(1)}-\delta^{(0)}$	$\sigma_{1D}-\sigma_{0D}$	α	$\beta^{(D)}$	$\beta_1^{(1)}$	$\beta_1^{(0)}$	σ_1^2	σ_0^2	ρ_{1D}	ρ_{0D}	ρ_{10}
		True Value	1.500	0.500	1.000	0.200	1.500	0.000	2.000	1.000	1.000	1.000	0.900	0.700	0.600
GGRM-noSI	Bias	500	-1.500	-0.500	-1.000	-0.035	0.073	-0.016	0.762	0.243	0.134	0.015	-0.093	-0.012	0.113
		1,000	-1.500	-0.500	-1.000	-0.021	0.033	-0.008	0.756	0.248	0.129	0.016	-0.075	-0.009	0.123
		5,000	-1.500	-0.500	-1.000	-0.004	0.005	-0.001	0.751	0.250	0.127	0.016	-0.055	-0.005	0.138
	RMSE	500	1.500	0.500	1.000	0.134	0.157	0.079	0.768	0.259	0.176	0.100	0.106	0.086	0.132
		1,000	1.500	0.500	1.000	0.093	0.099	0.056	0.759	0.256	0.151	0.076	0.084	0.062	0.134
		5,000	1.500	0.500	1.000	0.044	0.037	0.024	0.752	0.252	0.131	0.036	0.058	0.029	0.148
	Coverage	500	0.000	0.000	0.000	0.947	0.881	0.942	0.000	0.221	0.807	0.956	0.660	0.957	0.987
		1,000	0.000	0.000	0.000	0.947	0.901	0.945	0.000	0.030	0.649	0.934	0.487	0.967	0.984
		5,000	0.000	0.000	0.000	0.947	0.934	0.957	0.000	0.000	0.050	0.914	0.052	0.944	0.847
GGRM-SI	Bias	500	-0.001	-0.024	0.023	-0.050	0.074	-0.024	0.022	0.004	0.004	0.002	-0.053	-0.004	0.105
		1,000	0.000	0.007	-0.007	-0.028	0.035	-0.013	0.011	-0.006	0.003	0.002	-0.028	-0.001	0.110
		5,000	0.000	0.000	0.000	-0.006	0.005	-0.002	0.002	0.000	0.002	0.003	-0.005	0.000	0.112
	RMSE	500	0.226	0.244	0.334	0.133	0.154	0.080	0.138	0.150	0.102	0.099	0.067	0.086	0.128
		1,000	0.144	0.171	0.225	0.089	0.097	0.056	0.089	0.106	0.068	0.073	0.039	0.061	0.124
		5,000	0.066	0.077	0.100	0.040	0.035	0.024	0.041	0.048	0.029	0.032	0.014	0.029	0.119
	Coverage	500	0.954	0.946	0.954	0.934	0.870	0.946	0.955	0.958	0.968	0.954	0.904	0.951	0.973
		1,000	0.968	0.950	0.953	0.947	0.903	0.946	0.961	0.951	0.975	0.941	0.940	0.958	0.960
		5,000	0.960	0.951	0.960	0.957	0.927	0.950	0.963	0.954	0.972	0.954	0.956	0.944	0.879

Notes: This table displays the average bias (Bias), the Root Mean Squared Error (RMSE), and the coverage rate (Coverage) across $R = 1,000$ replicates; where Bias $= R^{-1}\sum_{r=1}^{R}(\hat{\alpha}_r - \alpha)$, RMSE $= \sqrt{R^{-1}\sum_{r=1}^{R}(\hat{\alpha}_r - \alpha)^2}$, and Coverage $= R^{-1}\sum_{r=1}^{R}\mathbb{1}\{\alpha \in \widehat{CI}_{0.95,r}\}$. The rows contain results for models with/without spatial interference and for various sample size n.

It can be seen that under the correct specification, estimating Gaussian Generalised Roy model with spillovers (`GGRM-SI`) using our proposed Bayesian Data Augmentation algorithm can successfully recover the true parameter values. Average bias and RMSE decrease as n increases; the empirical coverage rate is always close to the nominal level (95%). On the other hand, when the existing spillover phenomenon is not taken into consideration (i.e., estimating `GGRM-noSI` model which does not include neighbourhood treatment term $\bar{D}_{\mathcal{N}}$), not only $\delta^{(1)}$ and $\delta^{(0)}$ are clearly ignored but estimates of other relevant parameters ($\beta^{(1)}, \beta^{(0)}, \sigma_1^2, \sigma_0^2, \rho_{1D}, \rho_{0D}$) are also considerably affected – with larger average bias and RMSE, in addition to a lower coverage rate on average.

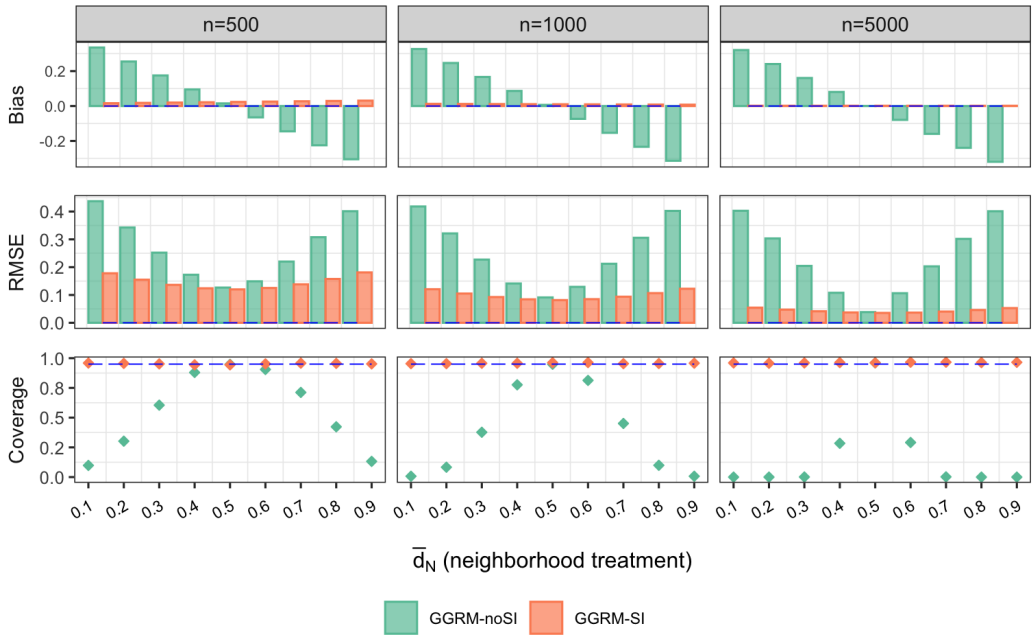

Figure 3.3: Plots of average bias, RMSE, and coverage rate of the estimation for average direct treatment effects. Models without and with spillovers are shown in red and in green, respectively.

Furthermore, we investigate the consequence of neglecting spillovers even when the causal estimand of interest is the direct effect of the treatment. Table 3.2 and Figure 3.3 report the performance of the estimators for Average Direct Treatment Effects, ADTE($\bar{d}_{\mathcal{N}}$), with different levels of the neighbourhood treatment $\bar{d}_{\mathcal{N}} \in \{0.1, 0.2, \ldots, 0.9\}$. Estimating `GGRM-SI` model again achieves good frequentist performance, which is consistent with the relation between ADTE($\bar{d}_{\mathcal{N}}$) and the model parameters in equation (3.13). On the contrary, estimating `GGRM-noSI` leads to a higher magnitude of bias which persists even in the case of a large sample ($n = 5,000$). It is due to the omission of the term $\left(\delta^{(1)} - \delta^{(0)}\right) \bar{d}_{\mathcal{N}}$ when we ignore spillovers. The coverage rate also deteriorates substantially, especially for extreme values of the neighbourhood treatment $\bar{d}_{\mathcal{N}}$. Similar issues is observed for the Marginal Direct

Treatment Effects, MDTE$(\bar{d}_{\mathcal{N}}, v)$. Table 3.3 displays simulation results for this estimand of interest at a fixed value of $\bar{d}_{\mathcal{N}} = 0.5$ and different values of the unmeasured resistance level $v \in \{0.1, 0.2, \ldots, 0.9\}$.

Table 3.2: Simulation Results for Average Direct Treatment Effects

		n = 500			n = 1,000			n = 5,000		
Model	Grid	Bias	RMSE	Coverage	Bias	RMSE	Coverage	Bias	RMSE	Coverage
GGRM-noSI	0.1	0.419	0.437	0.098	0.408	0.418	0.007	0.401	0.403	0.000
	0.2	0.319	0.343	0.302	0.308	0.321	0.083	0.301	0.303	0.000
	0.3	0.219	0.252	0.605	0.208	0.227	0.377	0.201	0.204	0.001
	0.4	0.119	0.173	0.880	0.108	0.141	0.776	0.101	0.108	0.284
	0.5	0.019	0.127	0.948	0.008	0.091	0.949	0.001	0.038	0.961
	0.6	-0.081	0.149	0.905	-0.092	0.129	0.813	-0.099	0.106	0.291
	0.7	-0.181	0.220	0.712	-0.192	0.212	0.451	-0.199	0.203	0.002
	0.8	-0.281	0.308	0.423	-0.292	0.306	0.099	-0.299	0.302	0.000
	0.9	-0.381	0.401	0.132	-0.392	0.402	0.007	-0.399	0.401	0.000
GGRM-SI	0.1	0.020	0.178	0.959	0.016	0.121	0.954	0.002	0.054	0.960
	0.2	0.023	0.155	0.956	0.015	0.105	0.954	0.002	0.047	0.957
	0.3	0.025	0.136	0.952	0.014	0.093	0.957	0.002	0.041	0.960
	0.4	0.027	0.124	0.944	0.014	0.084	0.957	0.002	0.037	0.963
	0.5	0.029	0.120	0.944	0.013	0.082	0.962	0.002	0.036	0.961
	0.6	0.032	0.125	0.952	0.012	0.085	0.963	0.002	0.037	0.966
	0.7	0.034	0.138	0.958	0.011	0.094	0.954	0.002	0.040	0.966
	0.8	0.036	0.158	0.956	0.011	0.107	0.955	0.002	0.046	0.963
	0.9	0.039	0.181	0.952	0.010	0.122	0.958	0.002	0.053	0.965

Notes: This table displays the average bias (Bias), the Root Mean Squared Error (RMSE), and the coverage rate (Coverage) across $R = 1,000$ replicates. The rows contain results for models with/without spatial interference across various sample size n.

Table 3.3: Simulation Results for Marginal Direct Treatment Effects

		n = 500			n = 1,000			n = 5,000		
Model	Grid	Bias	RMSE	Coverage	Bias	RMSE	Coverage	Bias	RMSE	Coverage
GGRM-noSI	0.1	-0.026	0.225	0.330	-0.019	0.160	0.100	-0.004	0.073	0.000
	0.2	-0.011	0.180	0.616	-0.009	0.128	0.404	-0.002	0.057	0.005
	0.3	0.000	0.152	0.823	-0.003	0.109	0.653	-0.001	0.048	0.141
	0.4	0.010	0.135	0.929	0.003	0.097	0.877	0.000	0.042	0.625
	0.5	0.019	0.127	0.948	0.008	0.091	0.949	0.001	0.038	0.961
	0.6	0.028	0.127	0.939	0.014	0.091	0.930	0.002	0.038	0.657
	0.7	0.037	0.137	0.870	0.019	0.098	0.712	0.003	0.042	0.121
	0.8	0.049	0.159	0.734	0.026	0.112	0.484	0.004	0.049	0.003
	0.9	0.064	0.200	0.439	0.035	0.140	0.108	0.006	0.063	0.000
GGRM-SI	0.1	-0.035	0.220	0.956	-0.023	0.153	0.955	-0.005	0.070	0.957
	0.2	-0.013	0.174	0.952	-0.011	0.122	0.955	-0.003	0.055	0.962
	0.3	0.003	0.146	0.950	-0.002	0.102	0.956	-0.001	0.046	0.961
	0.4	0.017	0.128	0.942	0.006	0.089	0.960	0.001	0.039	0.963
	0.5	0.029	0.120	0.944	0.013	0.082	0.962	0.002	0.036	0.961
	0.6	0.042	0.121	0.946	0.020	0.080	0.959	0.004	0.034	0.967
	0.7	0.056	0.131	0.952	0.028	0.085	0.961	0.005	0.036	0.969
	0.8	0.072	0.154	0.955	0.036	0.098	0.955	0.007	0.042	0.965
	0.9	0.094	0.196	0.954	0.049	0.125	0.954	0.009	0.054	0.963

Notes: This table displays the average bias (Bias), the Root Mean Squared Error (RMSE), and the coverage rate (Coverage) across $R = 1,000$ replicates. The rows contain results for models with/without spatial interference across various sample size n.

3.4.3 Other Scenarios

In this section, we slightly depart from the original design by considering other scenarios which include: (i) no spillover; or (ii) non-normality.

Scenario (i). No spillover $\delta^{(1)} = \delta^{(0)} = 0$

We use the same model specification (3.40) except for $\delta^{(1)} = \delta^{(0)} = 0$ to simulate datasets. Table 3.4 indicates that using our proposed Bayesian Data Augmentation algorithm to estimate the Gaussian Generalised Roy model with spillovers (`GGRM-SI`) can recover the true values of model parameters successfully, as well as the quantities of interest. As sample size n increase, the average bias vanishes, the RMSE declines, and the coverage rate of 95% credible interval remains close to the nominal coverage. Compared to the original design in Table 3.1, estimation for the elements of covariance matrix is almost unaffected. These findings suggest the validity of tests for indirect (spillover) effects, patterns of interaction and endogenous selection into treatment, as outlined in Section 3.2. In summary, the inclusion of neighbourhood treatment term to capture potential spillover is plausible, regardless of whether this phenomenon is present in the true data generating process.

Table 3.4: Simulation Results for Scenario (i)

		Quantities of Interest				Other Parameters								
Metric	n	$\delta^{(1)}$	$\delta^{(0)}$	$\delta^{(1)} - \delta^{(0)}$	$\sigma_{1D} - \sigma_{0D}$	α	$\beta^{(D)}$	$\beta_1^{(1)}$	$\beta_1^{(0)}$	σ_1^2	σ_0^2	ρ_{1D}	ρ_{0D}	ρ_{10}
	True Value	0.000	0.000	0.000	0.200	1.500	0.000	2.000	1.000	1.000	1.000	0.900	0.700	0.600
Bias	500	0.000	-0.024	0.023	-0.050	0.075	-0.024	0.022	0.004	0.004	0.002	-0.053	-0.004	0.106
	1,000	0.000	0.007	-0.007	-0.028	0.034	-0.013	0.011	-0.006	0.003	0.002	-0.028	-0.001	0.109
	5,000	0.000	0.000	0.000	-0.006	0.005	-0.002	0.002	0.000	0.002	0.003	-0.005	0.000	0.111
RMSE	500	0.227	0.244	0.335	0.132	0.154	0.080	0.139	0.150	0.101	0.099	0.067	0.086	0.129
	1,000	0.144	0.171	0.226	0.089	0.098	0.056	0.089	0.106	0.069	0.073	0.039	0.061	0.123
	5,000	0.066	0.077	0.100	0.040	0.035	0.024	0.041	0.048	0.029	0.032	0.014	0.029	0.118
Coverage	500	0.952	0.944	0.955	0.937	0.868	0.942	0.958	0.956	0.968	0.954	0.908	0.955	0.977
	1,000	0.970	0.951	0.951	0.946	0.891	0.944	0.962	0.947	0.968	0.940	0.931	0.958	0.969
	5,000	0.957	0.951	0.956	0.953	0.934	0.948	0.962	0.952	0.973	0.952	0.958	0.946	0.894

Notes: This table displays the average bias (Bias), the Root Mean Squared Error (RMSE), and the coverage rate (Coverage) across $R = 1,000$ replicates; where Bias $= R^{-1}\sum_{r=1}^{R}(\hat{\alpha}_r - \alpha)$, RMSE $= \sqrt{R^{-1}\sum_{r=1}^{R}(\hat{\alpha}_r - \alpha)^2}$, and Coverage $= R^{-1}\sum_{r=1}^{R}\mathbb{1}\{\alpha \in \widehat{CI}_{0.95,r}\}$. The rows contain results for various sample size n.

Scenario (ii). Non-normality

Although the assumption of joint normality of the error terms is computationally convenient, it may be inadequate to describe data in some cases. We consider below a finite mixture

of Gaussian distributions of the error terms and examine the influence of this small model misspecification on performance of the proposed approach.

$$\epsilon_i = \left[\epsilon_i^{(D)}, \epsilon_i^{(1)}, \epsilon_i^{(0)}\right]^\top \overset{ind}{\sim} \frac{1}{3}\mathcal{N}\left(0, \Sigma_1\right) + \frac{2}{3}\mathcal{N}\left(0, \Sigma_2\right),$$

where

$$\Sigma_1 = \begin{bmatrix} 7.50 & 6.75 & 5.25 \\ & 7.50 & 4.50 \\ & & 7.50 \end{bmatrix} \quad \text{and} \quad \Sigma_2 = \begin{bmatrix} 0.75 & 0.675 & 0.525 \\ & 0.75 & 0.45 \\ & & 0.75 \end{bmatrix}.$$

It can be seen from Table 3.5 that estimation for parameters related to error terms are affected, which is evident by upward biases and lower coverage rates compared to the original design with joint normality. Estimation for other parameters of interest including $\delta^{(1)}$, $\delta^{(0)}$, and $\delta^{(1)} - \delta^{(0)}$ still exhibits good coverage, although we do observe a slight increase in the RMSE and the bias. This confirms the robustness of the Bayesian estimation approach, which generally exhibits very good frequentist properties. Furthermore, we can introduce greater flexibility into the proposed approach by explicitly consider a finite mixture of normals, with details can be found in Appendix B.3.

Table 3.5: Simulation Results for Scenario (ii)

		Quantities of Interest				Other Parameters								
Metric	n	$\delta^{(1)}$	$\delta^{(0)}$	$\delta^{(1)} - \delta^{(0)}$	$\sigma_{1D} - \sigma_{0D}$	α	$\beta^{(D)}$	$\beta_1^{(1)}$	$\beta_1^{(0)}$	σ_1^2	σ_0^2	ρ_{1D}	ρ_{0D}	ρ_{10}
	True Value	1.500	0.500	1.000	0.200	1.500	0.000	2.000	1.000	1.000	1.000	0.617	0.480	0.600
Bias	500	0.008	0.000	0.008	-0.012	-0.009	-0.005	-0.038	0.037	0.039	0.035	0.031	0.045	-0.003
	1,000	-0.002	-0.001	-0.002	0.005	-0.039	0.001	-0.037	0.035	0.022	0.022	0.057	0.052	0.033
	5,000	0.000	0.002	-0.002	0.025	-0.068	0.002	-0.048	0.032	0.025	0.017	0.076	0.055	0.080
RMSE	500	0.238	0.264	0.352	0.190	0.150	0.076	0.157	0.173	0.196	0.192	0.103	0.142	0.118
	1,000	0.175	0.188	0.260	0.135	0.112	0.059	0.120	0.124	0.133	0.138	0.092	0.108	0.102
	5,000	0.076	0.081	0.110	0.072	0.084	0.024	0.068	0.062	0.062	0.058	0.084	0.072	0.126
Coverage	500	0.967	0.956	0.965	0.949	0.915	0.961	0.959	0.950	0.715	0.726	0.954	0.928	0.999
	1,000	0.945	0.937	0.938	0.945	0.896	0.939	0.925	0.938	0.733	0.699	0.888	0.903	0.999
	5,000	0.962	0.958	0.960	0.904	0.611	0.960	0.843	0.904	0.707	0.719	0.429	0.741	0.939

Notes: This table displays the average bias (Bias), the Root Mean Squared Error (RMSE), and the coverage rate (Coverage) across $R = 1,000$ replicates; where Bias $= R^{-1}\sum_{r=1}^{R}(\hat{\alpha}_r - \alpha)$, RMSE $= \sqrt{R^{-1}\sum_{r=1}^{R}(\hat{\alpha}_r - \alpha)^2}$, and Coverage $= R^{-1}\sum_{r=1}^{R}\mathbb{1}\{\alpha \in \widehat{CI}_{0.95,r}\}$. The rows contain results for various sample size n.

3.5 Empirical Application

To demonstrate the applicability and usefulness of the proposed method, we investigate the effects of the Opportunity Zones (OZ) program, America's largest new place-based policy. This program offers investment tax incentives to a limited number of designated census tracts nationwide to promote economic development in distressed communities. In contrast to previous federal place-based tax policies, which required the U.S. government to determine eligible areas through an application process, state governors held primary authority and significant discretion in selecting their respective states' qualifying zones for the Opportunity Zone program. As the program aimed to encourage investment in low-income and high-poverty neighbourhoods, eligibility for OZ designation was based on the 5-year 2011-2015 American Community Survey (ACS), requiring tracts with poverty rates above 20% or median family incomes below 80% of the area median income. Approximately 40% of the U.S. census tracts, totaling 31,866, were eligible for OZ designation. State governors then had 90-120 days after the law's passage, until March 21, 2018, to nominate a quarter of their eligible tracts for OZ designation. On July 9, 2018, the Treasury (Internal Revenue Service) released a list of 8,764 approved census tracts, hereafter referred to as Qualified Opportunity Zones (QOZs), which included 8,534 low-income communities and 230 contiguous tracts. Figure 3.4 depicts a map of the zones on the U.S. mainland.

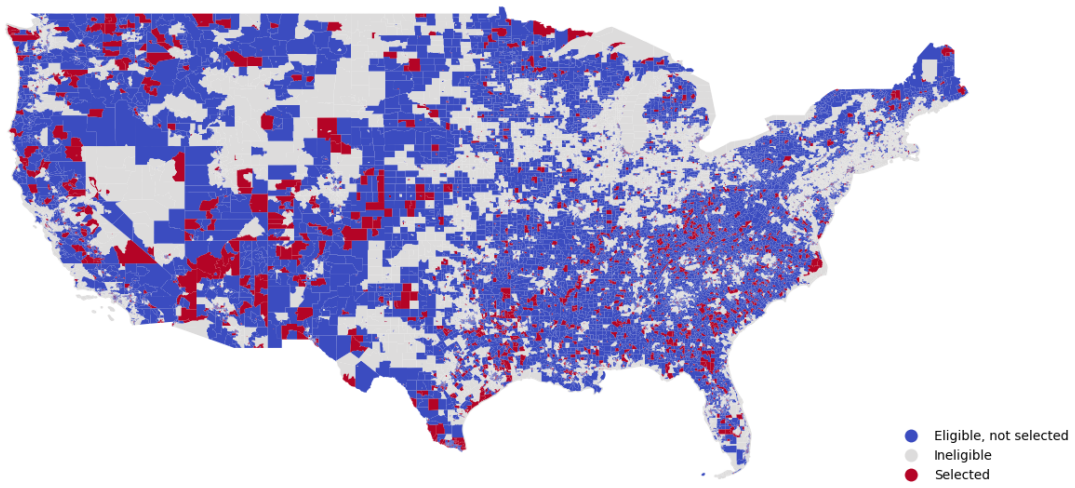

Figure 3.4: Illustration of the treatment assignment under the context of the U.S. Opportunity Zone program. This figure maps the census tracts governors selected as Opportunity Zones (red), as well as the eligible tracts not selected (blue).

It is imperative for policymakers to have a thorough understanding of the impact of interventions on targeted areas, as well as any potential spillover effects, such as beneficial

externalities or reallocations, on neighbouring communities. The economic outcome of interest we examine is the growth of housing units in census tracts, which is closely linked to overall economic growth. Predictions regarding the impacts of the OZ program are diverse[9]. On the one hand, providing tax incentives to investors can make investing in the housing market more financially attractive, thus encouraging the development of new housing projects in OZ areas. This could lead to an increase in housing unit growth. On the other hand, it may be challenging to promote new developments in certain underprivileged communities, particularly those with limited resources and infrastructure. There is also a concern that the OZ status of one location may affect potential outcomes of another, inducing spillovers beyond the direct effects. One possibility is that OZ designations could crowd housing investment into surrounding neighbourhoods through beneficial externalities. Another possibility is that it could reduce housing investment in nearby areas through investors reallocating projects towards tax-advantaged OZs.

3.5.1 Model Specification and Data Source

We consider the following specification that aligns with the proposed framework

$$
\begin{aligned}
QOZ_i &= \mathbb{1}\left\{\alpha Political_i + \beta_0^{(D)} + \sum_{k=1}^{K}\beta_k^{(D)} Demographic_{k,i} + \epsilon_i^{(D)} > 0\right\} \\
\overline{QOZ}_i &= \sum_{j=1,j\neq i} w_{ij} QOZ_j; \quad \sum_{j=1,j\neq i} w_{ij} = 1 \\
\%\Delta Housing_i &= \begin{cases} \delta^{(1)}\overline{QOZ}_i + \beta_0^{(1)} + \sum_{k=1}^{K}\beta_k^{(1)} Demographic_{k,i} + \epsilon_i^{(1)}, & \text{if} \quad QOZ_i = 1 \\ \\ \delta^{(0)}\overline{QOZ}_i + \beta_0^{(0)} + \sum_{k=1}^{K}\beta_k^{(0)} Demographic_{k,i} + \epsilon_i^{(0)}, & \text{if} \quad QOZ_i = 0 \end{cases}
\end{aligned}
\tag{3.41}
$$

$$
\begin{bmatrix} \epsilon_i^{(D)} \\ \epsilon_i^{(1)} \\ \epsilon_i^{(0)} \end{bmatrix} \overset{ind}{\sim} N\left(0, \begin{bmatrix} 1 & \sigma_{1D} & \sigma_{0D} \\ & \sigma_1^2 & \sigma_{10} \\ & & \sigma_0^2 \end{bmatrix}\right)
$$

Under this model specification, individual treatment (QOZ) is an indicator variable equal to one if the tract was selected as an Opportunity Zone from the pool of eligible tracts and zero otherwise. Opportunity Zone details are provided by the Urban Institute, including whether a tract belongs to the $31,866$ eligible tracts and the selected $8,762$ tracts. The neighbourhood treatment ($\overline{QOZ}$) is computed as the proportion of treated neighbours using the spatial

[9] A range of empirical evidences include, but are not limited to, studies by Corinth and Feldman (2024), Freedman et al. (2023), Chen et al. (2023), and Wheeler (2022).

adjacency matrix. Spatial-related information is retrieved from 2010 census tract locations and shapes, available in Census TIGER 2018 shapefiles.

Demographic characteristics ($Demographic$) are incorporated into the selection process to capture criteria that specify eligible census tracts. Some key variables include the poverty rate, median earnings, and employment rate. These variables also potentially affect tract outcomes, thereby being included in outcome equations. We measure them using the American Community Survey (ACS) 2013-2017 5-year estimates.

We employ political affiliation ($Political$) as the instrumental variable in this setting[10]. Regarding the selection process, the governor's party-affiliated census tracts have a higher likelihood of designation as an Opportunity Zone (see, e.g., Alm et al., 2021; Eldar and Garber, 2022; Frank et al., 2022). This can be attributed to two reasons. First, a governor's local political network can provide better information about localities that will benefit the most from the policy, which can influence the selection decision. Second, governors may choose to lend support to politically aligned representatives and constituencies as a means of bolstering their political standing. We measure political affiliation using an indicator equal to one if representative to the state's lower house state of a census tract and the state's governor are members of the same political party. We obtain these data published the month preceding the first Opportunity Zone selection (i.e., on March 1, 2018) from ballotpedia.com and assign the state representatives to each tract using the 2016 State Legislative Block Equivalency Files from the U.S. Census Bureau.

We restrict our analysis to the state of California because of its availability of comprehensive data sources. Our final California data include $3,699$ eligible census tracts, which consist of 727 selected tracts (QOZs) and $2,972$ eligible albeit not selected tracts (Non-QOZs). Details of variable definition, data sources, and a map of Opportunity Zone status of California census tracts are shown in Appendix B.5. Summary statistics and covariate balancing tests between the two subpopulations are reported in Table 3.6. It can be seen that compared to the control group, the selected OZs are less wealthy, less employed, less likely to attain higher education, have more rental units, and are less populated.

[10]That means, political affiliation is the excluded variable which enters the selection equation but does not appear in outcome (housing unit growth) equations. To test this supposition, we incorporated this political affiliation to the outcome equations. We found that it played negligible role in those equations, and its inclusion had little to no effect on the estimates of the remaining parameters.

Table 3.6: Summary Statistics and Balancing Tests

	All tracts (n=3699)	QOZs (n=727)	Non-QOZs (n=2972)	QOZs – Non-QOZs		
Variables	Mean (std)	Mean (std)	Mean (std)	Diff.Mean		t-statistic
Outcome						
Housing Unit Growth	0.03 (0.17)	0.04 (0.14)	0.03 (0.17)	0.02	*	2.57
Observed Characteristics						
Political Affiliation	0.79 (0.41)	0.82 (0.38)	0.78 (0.42)	0.04	**	2.72
Poverty Rate	0.19 (0.09)	0.27 (0.09)	0.17 (0.08)	0.09	***	24.64
Median Earnings	10.17 (0.31)	10.01 (0.26)	10.21 (0.30)	-0.21	***	-18.60
Employment Rate	0.29 (0.07)	0.26 (0.07)	0.29 (0.07)	-0.04	***	-12.87
% White	0.56 (0.21)	0.53 (0.20)	0.56 (0.21)	-0.04	***	-4.35
% Native	0.90 (0.04)	0.89 (0.04)	0.91 (0.04)	-0.02	***	-9.05
% Higher ed.	0.15 (0.09)	0.11 (0.07)	0.16 (0.09)	-0.05	***	-15.23
% Rent	0.57 (0.21)	0.67 (0.19)	0.54 (0.21)	0.13	***	16.47
Population	4509.55 (1613.93)	4305.31 (1476.18)	4559.51 (1642.24)	-254.20	***	-4.07

Notes: This table presents summary statistics at the census tract level in California. All tracts refer to the entire sample of eligible census tracts for Opportunity Zones, which consist of selected tracts (QOZs) and eligible, non-selected tracts (Non-QOZs). Two-sample t-statistics of tests for differences in mean values between two subsamples are reported. The asterisks *, **, and *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

3.5.2 Estimation Results

Our model (3.41) is estimated using the Bayesian Data Augmentation algorithm discussed in previous sections. The estimation results are reported in Table 3.7. With respect to the treatment decision equation, the coefficient estimates are consistent with our expectations. Governors tend to select even more distressed communities, characterised by higher poverty rates and lower median earnings, from the pool of already low-income tracts. Additionally, the instrument appears to play an important role in the selection process, as governors are more likely to choose tracts represented by members of the same political party. Regarding the outcome equations, non-QOZs with lower poverty rates, larger median earnings, and higher employment rates tend to experience greater changes in total housing units on average. These effects are statistically significant at the 5% level. In contrast, for selected census tracts (QOZs), higher poverty rates are associated with more significant changes in total housing units.

Table 3.7: Estimation Results

	(I) No Controls				(II) Additional Controls			
	Mean	Std	LB	UB	Mean	Std	LB	UB
Treatment Decision Equation								
Political Affiliation (α)	0.119	0.043	0.034	0.203	0.116	0.053	0.017	0.221
Intercept ($\beta_0^{(D)}$)	-0.823	0.042	-0.907	-0.742	1.184	1.141	-1.077	3.410
Poverty Rate ($\beta_1^{(D)}$)	–	–	–	–	4.951	0.334	4.307	5.612
Median Income ($\beta_2^{(D)}$)	–	–	–	–	-0.317	0.113	-0.538	-0.095
Employment Rate ($\beta_3^{(D)}$)	–	–	–	–	0.530	0.459	-0.350	1.423
Outcome Equation for QOZs								
Neighbourhood Treatment ($\delta^{(1)}$)	0.034	0.016	0.002	0.066	0.033	0.017	0.001	0.067
Intercept ($\beta_0^{(1)}$)	-0.223	0.028	-0.269	-0.158	-0.573	0.295	-1.164	0.001
Poverty Rate ($\beta_1^{(1)}$)	–	–	–	–	0.857	0.101	0.667	1.054
Median Income ($\beta_2^{(1)}$)	–	–	–	–	0.017	0.030	-0.041	0.074
Employment Rate ($\beta_3^{(1)}$)	–	–	–	–	0.057	0.109	-0.153	0.270
Outcome Equation for Non-QOZs								
Neighbourhood Treatment ($\delta^{(0)}$)	0.016	0.013	-0.010	0.042	0.025	0.014	-0.003	0.052
Intercept ($\beta_0^{(0)}$)	-0.025	0.004	-0.033	-0.017	-0.718	0.152	-1.016	-0.420
Poverty Rate ($\beta_1^{(0)}$)	–	–	–	–	-0.308	0.050	-0.408	-0.212
Median Income ($\beta_2^{(0)}$)	–	–	–	–	0.080	0.015	0.050	0.109
Employment Rate ($\beta_3^{(0)}$)	–	–	–	–	-0.200	0.059	-0.315	-0.084
Correlations and Variances								
σ_1^2	0.049	0.007	0.035	0.062	0.041	0.005	0.033	0.051
σ_0^2	0.038	0.001	0.036	0.041	0.035	0.001	0.033	0.038
ρ_{1D}	0.892	0.046	0.808	0.929	0.868	0.024	0.815	0.907
ρ_{0D}	-0.841	0.011	-0.861	-0.819	-0.811	0.014	-0.835	-0.782
ρ_{10}	-0.752	0.043	-0.800	-0.668	-0.707	0.032	-0.765	-0.640
Criteria								
Log likelihood	-894.939	43.593	-977.288	-801.430	-514.188	41.854	-597.910	-433.570
AICM	1948.754	–	–	–	1102.365	–	–	–
Observations	3699.000	–	–	–	3699.000	–	–	–
Quantities of Interest								
$\delta^{(1)}$	0.034	0.016	0.002	0.066	0.033	0.017	0.001	0.067
$\delta^{(0)}$	0.016	0.013	-0.010	0.042	0.025	0.014	-0.003	0.052
$\Delta_{\delta^{(1)}-\delta^{(0)}}$	0.018	0.021	-0.023	0.058	0.008	0.022	-0.035	0.051
$\Delta_{\sigma_{1D}-\sigma_{0D}}$	0.362	0.021	0.321	0.396	0.329	0.015	0.301	0.358

Notes: This table presents estimation results from Gaussian Generalised Roy model with spillovers. Posterior means, standard deviations, as well as upper bounds and lower bounds of 95% credible intervals are reported. Baseline specification (I) employs no control variables. Specification (II), which uses demographic characteristics as controls, provides robust results of quantities of interest and gains better Log marginal likelihood and Akaike's Information Criterion Monte Carlo (AICM, Raftery (2007)).

Recall that the coefficients of neighbourhood treatment, namely $\delta^{(1)}$ and $\delta^{(0)}$, represent the average partial indirect effects when treated and when untreated, respectively. Although we estimate a significantly positive average indirect effect for treated census tracts, with a point estimate of $\hat{\delta}_1 = 0.033$ and a credible interval of $[0.001, 0.067]$, this effect becomes statistically insignificant for untreated counterparts, with a point estimate of $\hat{\delta}_0 = 0.025$ and a credible interval of $[-0.003, 0.052]$). This suggests that, on average, a greater proportion of

neighbouring census tracts designated as QOZs is typically associated with a higher increase in the growth of housing units in selected tracts. However, the spillover effect is negligible when these tracts are not selected.

Using these estimates, we can also reconstruct corresponding treated and untreated average potential outcome functions, as illustrated in Figure 3.5. We report the average functions as bold lines, which show a moderately upward trend across increasing values of neighbourhood treatment because of the positive estimated slopes. The corresponding (pointwise) 95% credible intervals represented by dashed curves are derived from the posterior samples.

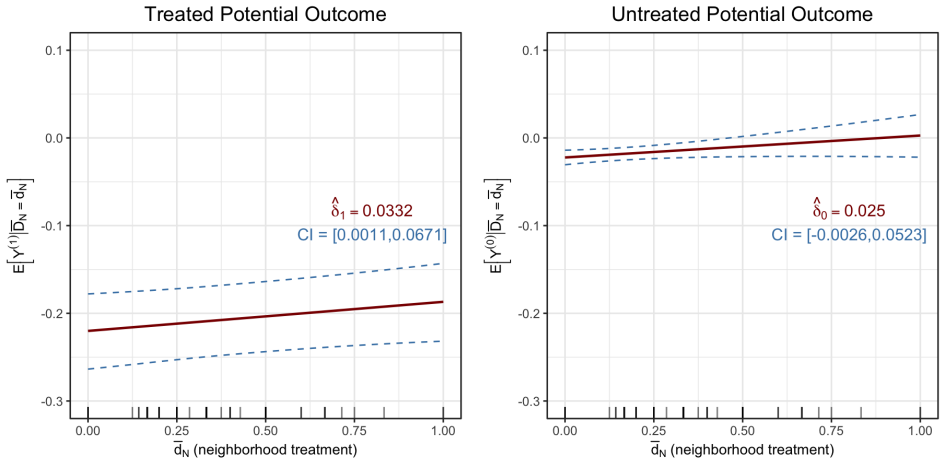

Figure 3.5: The average potential outcome when being treated and when being untreated (right-hand side panel), as a function of the neighbourhood treatment $\bar{d}_{\mathcal{N}}$. The dashed curves represent 95% credible intervals derived from posterior samples. Ticks in the rug plot on the horizontal axis represent empirical distribution of $\bar{d}_{\mathcal{N}}$ based on the real network. Estimation results for specification (II) with control variables are used.

The differential between these two average potential outcomes recovers the average direct treatment effects evaluated at various neighbourhood treatment levels, as depicted in Figure 3.6. The nearly flat curve, of which the estimated slope ($\hat{\Delta}_{\delta(1)-\delta(0)}$) is positive albeit not statistically significant, reveals that the direct effect of being selected for OZ program on census tracts' housing unit growth is almost unaffected by the neighbourhood treatment. Put differently, no significant interaction effect is found. This pattern is consistent with numerical results in Table 3.8. It worth noting that computed 95% credible interval bounds below zero in all cases, implying that the average direct treatment effects of OZ program for a census tract picked at random from the whole population of eligible census tracts are significantly negative. Although there no evidence that these estimated average direct treatment effects are heterogeneous relative to various levels of neighbourhood treatment, the average results may mask positive and negative effects for different subpopulations of census tracts characterised

by other dimensions.

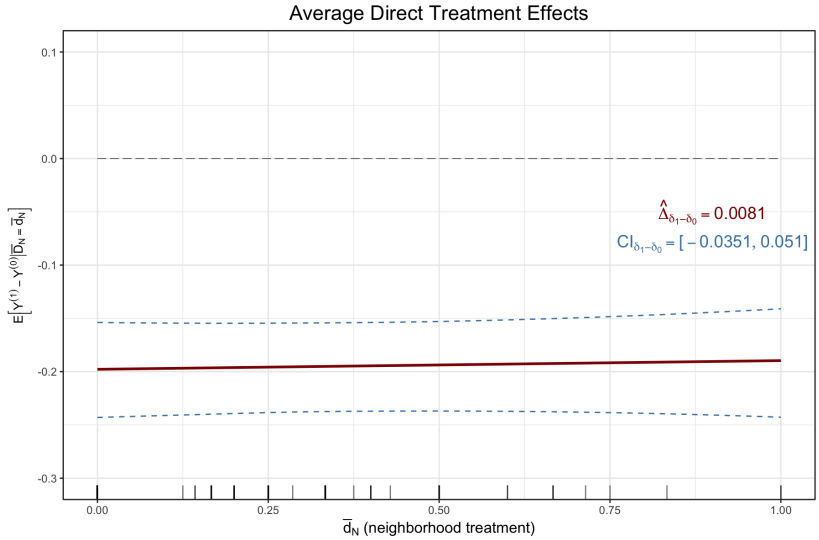

Figure 3.6: The average direct treatment effect as a function of the neighbourhood treatment $\bar{d}_{\mathcal{N}}$. The dashed curves represent 95% credible intervals derived from posterior samples. Ticks in the rug plot on the horizontal axis represent empirical distribution of $\bar{d}_{\mathcal{N}}$ based on the real network. Estimation results for specification (II) with control variables are used.

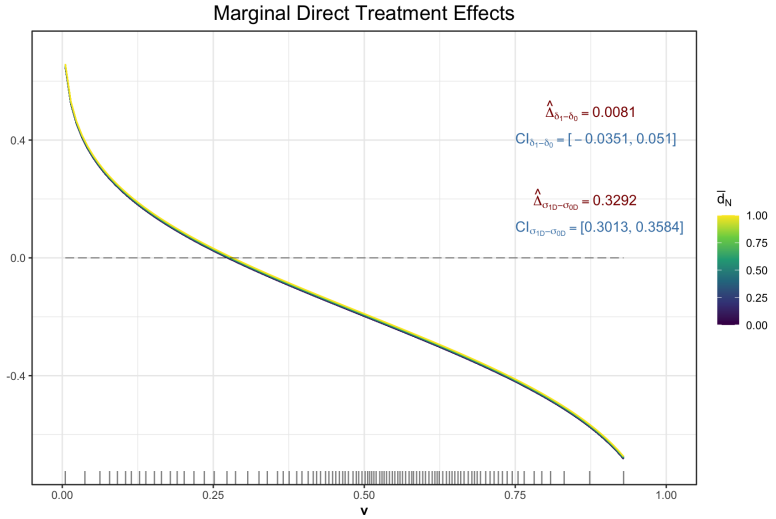

Figure 3.7: The marginal direct treatment effect as a function of the neighbourhood treatment $\bar{d}_{\mathcal{N}}$ and the unmeasured resistance level v, at average values of the covariates. Ticks in the rug plot on the horizontal axis represent empirical distribution of v, obtained from posterior samples. Estimation results for specification (II) with control variables are used.

Table 3.8: Average Direct Treatment Effects

$\bar{d}_{\mathcal{N}}$	Mean (std)	CI95
0.1	-0.1970 (0.0223)	[-0.2413, -0.1546]
0.2	-0.1962 (0.0219)	[-0.2394, -0.1546]
0.3	-0.1954 (0.0218)	[-0.2377, -0.1544]
0.4	-0.1945 (0.0218)	[-0.2373, -0.1538]
0.5	-0.1937 (0.0221)	[-0.2373, -0.1532]
0.6	-0.1929 (0.0225)	[-0.2373, -0.1514]
0.7	-0.1921 (0.0232)	[-0.2379, -0.1493]
0.8	-0.1913 (0.0241)	[-0.2387, -0.1471]
0.9	-0.1905 (0.0251)	[-0.2406, -0.1441]

Notes: This table presents estimation results for Average Direct Treatment Effects, evaluated at nine grid values of the neighbourhood treatment $\bar{d}_{\mathcal{N}}$ and average values of the covariates in our sample. Posterior means, standard deviations, as well as 95% credible intervals for specification (II) with control variables are reported.

Table 3.9: Marginal Direct Treatment Effects

v	Mean (std)	CI95
0.1	0.2258 (0.0127)	[0.2010, 0.2506]
0.2	0.0809 (0.0137)	[0.0538, 0.1075]
0.3	-0.0235 (0.0162)	[-0.0554, 0.0071]
0.4	-0.1127 (0.0190)	[-0.1502, -0.0766]
0.5	-0.1961 (0.0219)	[-0.2393, -0.1546]
0.6	-0.2796 (0.0251)	[-0.3288, -0.2320]
0.7	-0.3688 (0.0286)	[-0.4247, -0.3151]
0.8	-0.4732 (0.0329)	[-0.5376, -0.4115]
0.9	-0.6181 (0.0391)	[-0.6945, -0.5451]

Notes: This table presents estimation results for Marginal Direct Treatment Effects, evaluated at nine grid values of the unmeasured resistance level v and average values of the neighbourhood treatment $\bar{d}_{\mathcal{N}}$ as well as the covariates in our sample. Posterior means, standard deviations, as well as 95% credible intervals for specification (II) with control variables are reported.

Table 3.10: Summary of Direct Treatment Effects

ADTE		ADTT		ADTUT	
Mean (std)	CI95	Mean (std)	CI95	Mean (std)	CI95
-0.196 (0.022)	[-0.239, -0.155]	0.041 (0.015)	[0.012, 0.069]	-0.298 (0.026)	[-0.349, -0.249]

Notes: This table presents estimation results for the average direct treatment effect (ADTE), the average direct treatment effect on the treated (ADTT) and on the untreated (ADTUT) from Gaussian Generalised Roy model with spillovers. Posterior means, standard deviations, as well as 95% credible intervals for specification (II) with control variables are reported.

From estimation results for the elements of covariance matrix in Table 3.7, it is evident that the correlations (ρ_{1D} and ρ_{0D}) between unobserved component in treatment decision equation and each of two potential outcome equations are significantly different from zero. This reveals endogenous selection into treatment, which should be taken into account to accurately characterise the causal effects of the OZ program.

Rewrite (3.41) using $V_i = -\Phi\left(\epsilon_i^{(D)}\right)$ Then, V_i represents unobserved heterogeneity of census tracts, which contributes to the selection process. Because higher values of V_i imply a lower propensity for treatment, V_i can be interpreted as an unmeasured resistance to treatment. Examining further marginal direct treatment effects (MDTE) allows for the estimation of treatment effects at different points in the distribution of unobserved heterogeneity V_i, providing insights into how treatment effects vary among census tracts that are marginally indifferent to receiving the treatment.

Figure 3.7 depicts MDTE curves evaluated at mean values of covariates X in our sample. Because of the null interaction effect documented above, these curves corresponding to various values of $\bar{d}_{\mathcal{N}}$ are nearly identical. However, the MDTE curves decrease with increasing values of the unmeasured resistance to treatment v, revealing a pattern of selection-on-gains. This indicates that "low-resistance" tracts (i.e., very low v), which are most likely to be selected by governors, appear to benefit the most from tax credit. This selection-on-gains pattern, implied by the positive slope of the MDTE curve ($\Delta_{\sigma_{1D}-\sigma_{0D}}$), is statistically significant at the 5% level. The figure referenced in 3.7 depicts mean value curves of MDTE evaluated at different values of $\bar{d}_{\mathcal{N}}$. These curves, which represent the null interaction effect, are nearly identical. However, the MDTE curves decrease with increasing values of the unmeasured treatment resistance v, revealing a selection-on-gains pattern. This indicates that "low-resistance" tracts, which are most likely to be selected by governors and have very low v, benefit the most from tax credits. The slope of the MDTE curve ($\Delta_{\sigma_{1D}-\sigma_{0D}}$) implies this selection-on-gains pattern and is statistically significant at the 5% level.

Interestingly, estimation results from Table 3.9 further suggest that for the 20 percent of census tracts most likely to be selected ($v \leq 0.2$), the direct effects are positively significant. In contrast, high-resistance census tracts that are less likely to be selected by governors exhibit negatively significant direct effects. These tracts constitute 60 percent of the entire eligible population ($v \geq 0.4$). With a realized treatment rate in our data of approximately 20 percent, our further interest is examining the treatment effects on the treated and untreated. By deriving

proper weights and aggregating over the MDTE curve, a range of causal estimands can be computed, as summarized in Table 3.10. The average direct treatment effect (ADTE) suggests that OZ tax credit reduces the growth of housing units by 19.6 percentage points. Although this result is statistically significant, it is generally insufficient for policy analysis and obscures important treatment effect heterogeneity. Specifically, the average direct treatment effect on the treated (ADTT), which places most weight on low-v communities, shows a substantial increase of 4.1 percentage points in housing unit growth for tracts that are ultimately selected as QOZs on average. On the other hand, the average direct treatment effect on the untreated (ADTUT), which places most weight on high-v communities, shows a substantial decrease of 29.8 percentage points for the non-QOZs on average. These findings provide insights into the OZ program's selection mechanism and effectiveness. Selection-on-gains appears to be an empirically significant phenomenon that influences the governor's designation process. Governors prioritize census tracts with characteristics that have potential for ex-post gains, which could generate a positive impact. This goal seems to have been achieved, as indicated by the positive direct treatment effect on the treated tracts (QOZs). Additionally, these findings can be reconciled with our previous discussion on the role of political affiliation in governors' choices. One possible explanation for this is the information advantage that local political networks provide governors, enabling them to better understand which communities will benefit the most from the policy. However, eligible but unselected tracts (non-QOZs) remain a disadvantaged group. Although there is no evidence to support the initial concern about a displacement effect, there are no beneficial externality spillovers into non-QOZs either. Additionally, because of unobserved differences between QOZs and non-QOZs, it is unlikely that positive treatment effects on QOZs would be replicated if non-QOZs were granted investment tax credit. In fact, the strongly negative ADTUT even predicts an adverse impact on non-QOZs in this hypothetical scenario. Therefore, expanding the OZ program to communities that currently do not receive tax credit would not be effective.

These findings have important implications for the design of an optimal policy. Considering the diverse impacts of the OZ program, it is vital to assess its overall efficacy and determine whether its benefits are fairly distributed among communities. Given the role of unmeasured tract heterogeneity in determining the OZ program's effectiveness, an optimal design should adopt a more refined approach to selection mechanisms. In order to address the disadvantages faced by non-QOZs, policymakers should explore alternative strategies to support these communities, rather than relying solely on tax incentives.

3.6 Concluding Remarks

In this paper, we developed methods that use observational network or spatial data to identify and estimate direct and indirect causal effects in the presence of endogenous selection into treatment and spillovers. This scenario arises due to the violation of the unconfoundedness and SUTVA assumptions, which are commonly assumed in the causal inference literature yet implausible in many economic contexts. Our proposed framework nests the Generalised Roy model and captures spillovers in the form of exposure to neighbours' treatment. The advantages of this approach are explicitly modelling selection process and allowing for heterogeneous effects across individuals. Although the model is not fully structural, it is economically interpretable. For estimation and inference, we improved the Bayesian data augmentation algorithms to enable more efficient computation and greater flexibility. All sources of variability are accounted for in the posterior distribution of the target causal estimands. Both the simulation with synthetic data and an empirical Monte Carlo study that uses friendship networks and covariates from Add Health data indicate that our Bayesian estimators perform well in terms of bias, RMSE, and coverage rate.

We also applied the proposed method to evaluate the Opportunity Zones (OZ) program, which offers tax incentives to boost economic development in distressed U.S. communities. We modeled the selection mechanism of designating QOZs by state governors and estimated the program's impact on housing unit growth in census tracts, accounting for endogenous selection into treatment. Our findings indicate a selection-on-gains pattern, where treatment effects vary with unmeasured tract heterogeneity. Both the direct and indirect effects on treated tracts (QOZs) are positive, while unselected tracts (non-QOZs) remain disadvantaged without positive spillover effects. The unobserved differences between QOZs and non-QOZs suggest that replicating positive treatment effects in non-QOZs is unlikely. Moreover, our model predicts adverse outcomes for non-QOZs if they were granted investment tax credits. Therefore, extending the OZ program to currently uncredited communities would be ineffective.

Extending the methods discussed earlier, a Bayesian semiparametric approach can be implemented to incorporate heterogeneous indirect effects and relax the distributional assumption for disturbance terms. We outline an extension that adopts a finite mixture model in Appendix B.3. A more general approach could be built on the work of Conley et al. (2008), which utilises the Dirichlet process (DP) to enable parameters to be observation-specific

and generated from a common distribution. We leave this direction for future research. Furthermore, variable selection techniques based on spike and slab priors could be employed in our Bayesian estimation procedure to account for uncertainty regarding control variables in both potential outcome equations and the selection equation or to account for uncertainty regarding the structure of the adjacency matrix that represents the network or spatial connection.

Chapter 4

Modelling Interactions with Heterogeneous Effects and Endogenous Network Formation

4.1 Introduction

Modelling network interactions is a significant topic in economics to understand how individuals' actions are shaped by those of their peers, with applications ranging from education and labour markets to industrial organisation and beyond. With the growing availability of detailed data that document connections among individuals, the spatial autoregression model (SAR) has been extensively used to empirically analyse these interactions and estimate peer effects in networks.

The foundational theoretical aspects related to identifying and estimating the standard SAR model as a network interaction model were established by Bramoullé et al. (2009), as well as by Lee et al. (2010), and Lin (2010). These studies are predicated on the assumption that the network is exogenous. Nonetheless, this assumption has been challenged, as unobservable characteristics often simultaneously influence individual activity outcomes and the formation of networks, thereby introducing selection bias into the estimates of interaction effects. This recognition has led to a new wave of research focusing on integrating models of network interaction and formation to structurally address the endogeneity issue. Notable contributions in this domain include subsequent studies by Goldsmith-Pinkham and Imbens (2013), Hsieh and Lee (2016), Qu and Lee (2015), Johnsson and Moon (2021), and Auerbach (2022), which employ various estimation techniques. These developments are naturally linked to evolving literature on the econometrics of network formation (see, e.g., Chandrasekhar, 2016; Graham, 2015), and researchers are increasingly leveraging recent insights from this literature to expand the SAR framework to account for network endogeneity.

Despite these advancements, the literature is still missing due to an implicit assumption

embedded in the conventional specification of SAR. This typically imposes strict homogeneity in the interaction effects, requiring that all individuals are influenced in exactly the same way by the average outcome of their peers. This assumption is implausible in many context when some individuals are clearly more susceptible to peers or exert greater influence than others. Recent surveys (see, e.g., Bramoullé et al., 2020; Kline and Tamer, 2020) have acknowledged the importance of accounting for individual heterogeneity in endogenous interaction effects. Although the economic theory in these cases is well-examined (Jackson and Zenou, 2015), econometric models addressing heterogeneous interaction effects remain underdeveloped. A few exceptions have attempted to incorporate certain degree of pre-stipulated heterogeneity, such as (Arduini et al., 2020a, 2020b), who generalised the standard model to allow for two specific types with between and within-type interactions. Building upon this framework, Beugnot et al. (2019) investigate the gender-heterogeneous peer effects on performance within social networks. Their study reveals that men and women experience distinct peer effects, and individuals may also be influenced differently by male and female peers. Rather than pre-specifying heterogeneity based on observable characteristics of individuals, Masten (2018) introduces the heterogeneity by studying identification for the linear-in-means model with random coefficients. This model accommodates pair-specific interaction effects that purely random and not driven by observable factors. The author proves that the marginal distributions of the coefficients are point identified, provided there exists an instrument with continuous variation over a large support. In a similar vein, Peng (2019) extends SAR model by assuming varying rates of influence among individuals and imposed sparsity on the number of influential individuals. Nonetheless, all of these studies essentially require the network to be exogenously given, which compromises the efforts made in research addressing endogenous network formation. This assumption is subject to considerable skepticism among economists, raising the question of whether relaxing it has major implications for estimates of heterogeneous network interaction effects in these studies. To the best of our knowledge, simultaneously accommodating heterogeneous network interaction effects and accounting for endogenous network formation is a gap in the literature of network/social interactions due to the intertwined challenges.

This paper contributes to the literature by proposing a new framework that extends standard SAR model to incorporate both heterogeneous network interaction effects and endogenous network formation. We develop the Selection-corrected Heterogeneous Spatial Autoregressive (hereafter referred to as SCHSAR) model, a unified framework to achieve these two objectives.

At its core, the SCHSAR model features a mixture structure in which individuals belong to finite latent types, each governed by distinct parameters when they interact with peers in a network. This specification captures rich heterogeneity in network interaction effects, allowing individual outcomes to vary in both the magnitude and direction of peer influence. To address the endogeneity inherent in network formation, we explicitly model dyadic links as the realisation of decision-making processes influenced by observable attributes and unobservable individual-specific factors. By jointly incorporating these unobserved traits as latent variables into the network formation and individual outcome equations, which resemble a two-stage game, the SCHSAR framework corrects for potential selection bias arising from endogenous links. The estimation is performed using a Bayesian data augmentation method, which enables efficient computation and inference despite complex latent structures. Overall, SCHSAR offers a flexible and robust approach to draw credible conclusions about network interaction effects in settings where both individual heterogeneity and network endogeneity are present.

Our methodology builds on two strands of the literature in econometrics, and our contributions are two-fold.

First, our strategy to correct selection bias stemming from endogenous network formation follows the seminal works of Goldsmith-Pinkham and Imbens (2013) and Hsieh and Lee (2016). These studies introduce Bayesian joint modelling approach that uses latent variables to control for unobserved factors influencing both network formation and outcome. Our SCHSAR framework goes beyond these foundational contributions in several ways. Most notably, we relax the assumption of homogeneous peer effects by incorporating a finite mixture structure that enables heterogeneous responses to peers. Broadly speaking, our model nests Goldsmith-Pinkham and Imbens (2013) as a special case when the number of mixture components is reduced to one. Additionally, unlike prior work that focuses solely on unobserved homophily, our framework also allows unobserved degree heterogeneity to drive network formation. This again provides a richer form of individual heterogeneity, facilitating a more robust analysis. Lastly, we adopt a probit instead of logit link function for modelling network formation via introducing normally distributed dyadic shocks. We gain another latent-variable representation that integrates naturally into the hierarchical Bayesian structure of our model, thereby simplifying posterior computation and improving convergence.

Second, to accommodate the heterogeneity in the network interaction effects, our model draws inspiration from Cornwall and Parent (2017), who introduce a finite mixture approach

to spatial econometric modelling. Their Spatial Autoregressive Mixture (henceforth, SAR-M) model marks an innovative step toward integrating spatial dependence and heterogeneity, and connects to a broader literature on heterogeneous spatial models (e.g., Aquaro et al., 2021; LeSage and Chih, 2018, 2016). However, SAR-M and related models share a common and significant limitation: the assumption that the spatial weights matrix (spatial version of the network adjacency matrix) is exogenously given and fixed. This restricts their empirical applicability, particularly in settings where spatial weights are constructed from economic or behavioural variables (for example, GDP, income, or trade flows). In such cases, the endogeneity of the spatial weights matrix becomes a first-order concern, as emphasised in critical works (see Han and Lee, 2016; Kelejian and Piras, 2014; Qu and Lee, 2015). Our SCHSAR framework addresses this limitation by endogenizing the network formation process, allowing the adjacency (or spatial weight) matrix to be driven by unobserved factors. In doing so, our model advances Cornwall and Parent's (2017) SAR-M model in a cross-sectional setup, and the insights may be extended to heterogeneous models utilising panel data, such as in LeSage and Chih (2018).

By bridging recent advancements in two distinct lines of research, SCHSAR provides a unified and flexible framework that can be readily applied to both social networks and spatial data contexts.

To estimate the SCHSAR model, we develop a Bayesian Markov chain Monte Carlo (MCMC) algorithm that offers several methodological and practical advantages. *First*, we leverage the Bayesian data augmentation technique to sample parameters alongside latent variables from the joint posterior distribution. This approach provides a tractable solution to the computational challenges typically faced in maximum likelihood estimation, particularly those involving high-dimensional integrals arising from latent structures. *Second*, the proposed algorithm simultaneously addresses heterogeneity and endogeneity, the two central features of the SCHSAR framework, by utilising latent variables in a unified model. It also enables the seamless integration of model constraints. This unified Bayesian approach facilitates one-step inference across all unknown quantities, providing principled uncertainty quantification. *Third*, the algorithm yields posterior predictive distributions for the individual-level latent variables, including the unobserved random effects and latent type indicators. These posterior draws allow for a data-driven probabilistic assignment of individuals to latent types, enabling the model to uncover clusters of heterogeneous interaction effects through a finite mixture

structure. Moreover, this facilitates an examination of how latent heterogeneity correlates with observable characteristics, enriching the interpretability of the model. *Finally*, the Bayesian framework naturally allows for inference on policy-relevant quantities, such as total spillin and spillout effects, which are highly nonlinear functions of the model parameters. These effects can be computed directly from posterior samples as a by-product of estimation. The computational tractability and convergence of our Bayesian estimation method is supported by simulation evidence. The simulation study also demonstrates favourable frequentist properties, showing that the SCHSAR model delivers valid inference and improved performance relative to more naive approaches that either ignore endogenous network formation or assume homogeneous interaction effects.

Having developed valid tools for estimation, we apply the proposed SCHSAR framework to study the formation of a technological collaboration network among U.S. firms and heterogeneous interaction effects on firm performance, with a focus on their R&D efforts. The empirical analysis confirms significant positive, yet heterogeneous, network interaction effects on corporate R&D investments, even after controlling for selection bias due to endogenous network formation. We find substantial heterogeneity in firm behaviour and uncover notable transmitters and absorbers in response to exogenous R&D policy shocks. This framework facilitates the quantification of firm-level direct, spillin, and spillout effects, thereby offering valuable insights for the design of evidence-based and targeted policy interventions.

The remainder of this paper is organised as follows. In Section 4.2, we formally present the Selection-corrected Heterogeneous Spatial Autoregressive (SCHSAR) framework, accounting for heterogeneous interaction effects and endogenous network formation. In Section 4.3, we develop Bayesian MCMC algorithms to estimate the model and conduct inferences. Next, we evaluate the performance of our method using simulations in Section 4.4 and apply the proposed approach to study U.S. firms' technological collaboration network in Section 4.5. Finally, we conclude the paper in Section 4.6 with brief remarks on the method and policy recommendations based on our results.

4.2 Selection-corrected Heterogeneous Spatial Autoregressive Model

4.2.1 Network Interaction with Heterogeneous Effects

Suppose there are N individuals, indexed by $i \in \{1, \ldots, N\}$ in a network.

Let $\mathbf{W} \coloneqq [w_{ij}]$ be an $N \times N$ adjacency matrix that characterises the network, where w_{ij} equals one if individuals i and j are connected, and zero otherwise. The diagonal entries in $\mathbf{W}$ are always zero.

Let $\mathbf{Y} \coloneqq [Y_1, \ldots, Y_N]^\top$ be an N-dimensional vector of the outcomes of interest, defined for all individuals. Let $\mathbf{X}$ be an $N \times K$ matrix of exogenous covariates that represents individual characteristics, with k-th column denoted as X^k.

The standard spatial autoregressive (SAR) model for studying interactions in networks is specified as

$$Y_i = \lambda \sum_{j=1}^{N} w_{ij} Y_j + X_i \beta + u_i; \tag{4.1}$$

where λ captures the endogenous peer effect (a.k.a. network interaction effect), where an individual's choice of activity level (outcome) may depend on those of their peers; $\beta = [\beta^1, \ldots, \beta^K]^\top$ captures the influence from the individuals' exogenous characteristics on the outcome; and $\mathbf{u} = [u_1, \ldots, u_N]^\top$ is a vector of stochastic errors whose elements are i.i.d. with zero mean and constant variance σ_u^2.

The network interaction model (4.1) has been widely studied (Bramoullé et al., 2009; Lee et al., 2010; Lin, 2010, among many others) and is also referred to as the linear-in-means network model. Despite many available variants in the literature, these models typically assume that the interaction effect parameter λ is constant and common to all individuals. This means that all individuals are influenced in the same way by a summary of their peers' outcomes. However, assuming homogeneous effects is restrictive when some individuals may be highly susceptible to peer effects, while others remain relatively independent in their behaviours.

To accommodate this heterogeneity, we assume that our sample is representative of a population composed of a finite number of latent types of individuals, indexed by $g \in \mathcal{G} \coloneqq \{1, \ldots, G\}$. Each type g is characterised by its own set of parameters $\{\lambda_g, \beta_g, \sigma_{ug}^2\}$, which are similar for all individuals of the same type but vary across types. The type is assigned independently among individuals, following a multinomial distribution with probability

$\boldsymbol{\pi} \coloneqq [\pi_1, \ldots, \pi_G]$, where $0 < \pi_g \leq 1$ and $\sum_{g=1}^{G} \pi_g = 1$. We denote z_{ig} as the individual-type indicators, where $z_{ig} = 1$ means i belongs to type g, and $z_{ig} = 0$ otherwise. Combine $z_i \coloneqq [z_{i1}, z_{i2}, \ldots, z_{iG}]$ and stack them in an $N \times G$ matrix of allocation, $\mathbf{z} \coloneqq [z_1^\top, \ldots, z_N^\top]^\top$. These indicators are stochastic and unobserved, and the probability that the latent type of i is g is $\mathbb{P}(z_{ig} = 1) = \pi_g$. We denote $\mathcal{I}_g$ as the set of type-g individuals and N_g as the corresponding cardinality. $\mathcal{I}_1, \ldots, \mathcal{I}_G$ are thus sets that partition $\{1, \ldots, N\}$, and $\sum_{g=1}^{G} N_g = N$. In this way, all parameters in the original model (4.1) become type-specific as follows

$$Y_i = \lambda_g \sum_{j=1}^{N} w_{ij} Y_j + X_i \beta_g + u_i; \quad u_i \sim \mathcal{N}(0, \sigma_{ug}^2); \tag{4.2}$$

for each individual $i \in \mathcal{I}_g$.

Collect type-specific parameters into G-dimensional vectors and define N-dimensional vectors of individual-specific parameters listed below

$$\begin{aligned}
&\check{\lambda} \coloneqq [\lambda_1, \ldots, \lambda_G]^\top, \quad \tilde{\lambda} \coloneqq \mathbf{z}\check{\lambda}; \\
&\check{\beta}^k \coloneqq [\beta_1^k, \ldots, \beta_G^k]^\top, \quad \tilde{\beta}^k \coloneqq \mathbf{z}\check{\beta}^k, \quad k = 1, \ldots, K; \\
&\check{\sigma}_u^2 \coloneqq [\sigma_{u1}^2, \ldots, \sigma_{uG}^2]^\top, \quad \tilde{\sigma}_u^2 \coloneqq \mathbf{z}\check{\sigma}_u^2.
\end{aligned}$$

Accordingly, we obtain a heterogeneous version of the spatial autoregressive model (HSAR)

$$Y_i = \tilde{\lambda}_i \sum_{j=1}^{N} w_{ij} Y_j + X_i \tilde{\beta}_i + u_i; \quad u_i \sim \mathcal{N}(0, \tilde{\sigma}_{u,i}^2); \tag{4.3}$$

where $\tilde{\lambda}_i$, $\tilde{\beta}_i$, and $\tilde{\sigma}_{u,i}^2$ are individual-specific parameters for individual i, which depend on what type he/she is assigned and parameters of that type.

Remark 1. The HSAR specification offers greater flexibility in modelling network interaction effects than the standard SAR. This relaxes the assumption that peer effects must be uniformly positive or negative across all individuals. Some may be positively influenced, adopting behaviours or attitudes from their peers, whereas others may react negatively. Furthermore, the HSAR specification accounts for varying intensities of peer influence, as represented by the magnitude of $|\tilde{\lambda}_i|$. It is important to note that, when introducing various latent types, the model permits unrestricted interactions both within and between types in the network. Individuals who tend to share the same type, by definition, exhibit similar "receiving rates," but this is not strictly tied to their network positions or any predetermined characteristics (*ex-ante*).

We will demonstrate later that, by placing priors for both type indicators ($\mathbf{z}$) and assignment probabilities ($\boldsymbol{\pi}$), their updated posteriors produces a data-driven probabilistic assignment of individuals to types (groups). For now, we note how our setup with latent types directly leads to a finite mixture approach. Define $\overline{Y}_{(i)} \coloneqq \sum_{j=1}^{N} w_{ij} Y_j$ and $\tilde{Y}_i \coloneqq Y_i - \tilde{\lambda}_i \overline{Y}_{(i)}$, for $i = 1, \ldots, N$. Knowing the latent individual-type indicator $z_{ig} = 1$ informs us: $\tilde{Y}_i \sim \mathcal{N}\left(X_i \beta_g, \sigma^2_{ug}\right)$. Hence, marginalising over the type indicator z_i, we obtain the mixture distribution

$$p\left(\tilde{Y}_i \mid X_i, \{\lambda_g, \beta_g, \sigma^2_{ug}, \pi_g\}_{g=1}^{G}\right) \propto \sum_{g=1}^{G} \pi_g \mathcal{N}(X_i \beta_g, \sigma^2_{ug}), \qquad \text{where } \sum_{g=1}^{G} \pi_g = 1. \tag{4.4}$$

Let $\mathcal{L} \coloneqq \text{diag}(\tilde{\lambda})$ be an $N \times N$ diagonal matrix of which each diagonal element $\tilde{\lambda}_i = \lambda_g$ if individual i belongs to g-type group. Given that the outcomes for all individuals $i = 1, \ldots, N$ obey equations analogous to (4.3), the system of equations can be more compactly written in matrix notation as

$$\mathbf{Y} = \mathcal{L}\mathbf{W}\mathbf{Y} + \sum_{k=1}^{K} \tilde{\beta}^k \circ X^k + \mathbf{u}, \tag{4.5}$$

where $\circ$ denotes the Hadamard (element-wise) multiplication. This can be further expressed in the reduced form below

$$\mathbf{Y} = [\mathbf{I}_N - \mathcal{L}\mathbf{W}]^{-1} \left(\sum_{k=1}^{K} \tilde{\beta}^k \circ X^k + \mathbf{u} \right). \tag{4.6}$$

By definition, $\tilde{\mathbf{Y}} = [\mathbf{I}_N - \mathcal{L}\mathbf{W}]\, \mathbf{Y}$. Conditional on all individual-type indicators, we get the joint density

$$\begin{aligned} p\left(\tilde{\mathbf{Y}} \mid \mathbf{z}, \mathbf{W}, \mathbf{X}, \{\lambda_g, \beta_g, \sigma^2_{ug}, \pi_g\}_{g=1}^{G}\right) &= \prod_{i=1}^{N} \prod_{g=1}^{G} [\phi(X_i \beta_g, \sigma^2_{ug})]^{z_{ig}} \\ &= \prod_{g=1}^{G} (2\pi\sigma^2_{ug})^{-N_g/2} \exp\left\{ \sum_{i=1}^{N} -\frac{z_{ig}(\tilde{Y}_i - X_i\beta_g)^2}{2\sigma^2_{ug}} \right\}, \end{aligned} \tag{4.7}$$

where $N_g = \sum_{i=1}^{N} z_{ig}$, by definition, represents the number of type-g individuals.

Conditional on $(\mathbf{W}, \mathbf{X})$, the augmented likelihood function for the model parameters is associated with the joint distribution of the observed outcome and latent type indicators $(\mathbf{Y}, \mathbf{z})$

and of the form

$$
\begin{aligned}
& p\left(\mathbf{Y}, \mathbf{z} \mid \mathbf{W}, \mathbf{X}, \{\lambda_g, \beta_g, \sigma_{ug}^2, \pi_g\}_{g=1}^G\right) \\
& \quad = p\left(\mathbf{Y} \mid \mathbf{z}, \mathbf{W}, \mathbf{X}, \{\lambda_g, \beta_g, \sigma_{ug}^2\}_{g=1}^G\right) p\left(\mathbf{z} \mid \boldsymbol{\pi}\right) \\
& \quad = |\mathbf{I}_N - \mathcal{L}\mathbf{W}| \prod_{g=1}^G (2\pi\sigma_{ug}^2)^{-N_g/2} \exp\left\{\sum_{i=1}^N -\frac{z_{gi}(Y_i - \lambda_g \overline{Y}_{(i)} - X_i\beta_g)^2}{2\sigma_{ug}^2}\right\} \pi_g^{N_g}.
\end{aligned} \tag{4.8}
$$

Remark 2. For simplicity, we present here a local-aggregate model specification (i.e., the adjacency matrix $\mathbf{W}$ is not row-normalised), but the HSAR is also applicable to a local-average model (i.e., the matrix $\mathbf{W}$ is row-normalised such that each row sums to unity).[1] A pertinent issue concerns the constraints on $\{\lambda_g\}_{g=1}^G$ in order that $\mathbf{I}_N - \mathcal{L}\mathbf{W}$ is invertible. In accordance with the suggestion by Kelejian and Prucha (2010), we restrict the value of all λ_g to the interval $(-1/\tau^*, 1/\tau^*)$, where $\tau^* := \min\left\{\max_{1\le i\le N} \sum_{j=1}^N |w_{ij}|, \max_{1\le j\le N} \sum_{i=1}^N |w_{ij}|\right\}$. When $\mathbf{W}$ is a row-normalised matrix, the condition $|\lambda_g| < 1$ for all $g = 1, \ldots, G$ is sufficient. We impose this assumption throughout the paper and on the estimation procedure for λ_g.

In the spatial econometrics context, Cornwall and Parent (2017) introduces a similar modelling approach, known as the spatial autoregressive mixture (SAR-M) model. They show that the true parameters λ_g can be recovered well. However, the spatial weight matrix in their paper is treated as strictly exogenous, a common yet strong assumption that is unlikely to hold within our network interaction framework. Indeed, if unobserved factors influence both the network links used to construct w_{ij} and individual outcomes Y_i, this introduces endogeneity into the matrix $\mathbf{W}$ in (4.5). We let a_i be a scalar random variable capturing such unobservables which determine the outcome Y_i, and collect $\mathbf{a} := [a_1, \ldots, a_N]^\top$. If the generation of $\mathbf{W}$ also involves $\mathbf{a}$, it would be important to model network formation more explicitly to avoid potential endogeneity. In the following subsection, we examine $\mathbf{W}$ through the lens of a network formation process, which naturally frames our setup as a two-stage game: networks are formed in the first stage, and actions (outcomes) are determined in the second stage given the network structure. Afterward, we will return to the HSAR model to address challenges arising from the potential endogeneity issue.

[1] Aggregate and average models differ in terms of their behavioural foundations, which entails different interpretations for the interaction effect parameter λ (Liu et al., 2014).

4.2.2 Strategic Network Formation

Network formation is modelled through a strategic choice framework, where individuals decide to form a link based on perceived utility gain. We consider each dyad (i, j) composed of individuals i and j, with $i \neq j$. The respective marginal utilities[2] individuals i and j receive from forming a link are

$$U_{ij}(\epsilon^*_{ij}) = \psi_{ij} + \epsilon^*_{ij} \quad \text{and} \quad U_{ji}(\epsilon^*_{ji}) = \psi_{ji} + \epsilon^*_{ji}, \tag{4.9}$$

where ψ_{ij} and ψ_{ji} are score functions that depend on individual attributes of i and j, and $(\epsilon^*_{ij}, \epsilon^*_{ji})$ is a pair of idiosyncratic shocks that captures the uncertainty in their connection decision.

We assume the score function to be symmetric and deterministic across dyads. In addition, while each dyad can fully observe their characteristics when making decisions, there exists a component unobservable to the researcher. Specifically,

$$\psi_{ij} = \psi_{ji} = C_{ij}^\top \gamma + f(a_i, a_j). \tag{4.10}$$

In equation (4.10), C_{ij} is an L-dimensional vector of dyad-specific regressors derived from the observed exogenous characteristics of individuals i and j. For example, $C^l_{ij} = |c^l_i - c^l_j|$ reflects homophily in observables, one of the key features in the network formation literature to acknowledge that individuals prefer linking to similar others. Furthermore, the unobserved component, $f(a_i, a_j)$, is a function of individual unobserved characteristics. Here, we have introduced a_i as a scalar summary of individual i's unobserved characteristics, which is also regarded as an individual random effect. We assume a_i are independent across individuals with a commonly used parametric distribution $p(a_i|\theta_a)$. The specification of $f(a_i, a_j)$ depends on how a_i drives link preferences, and thus could be in the form of either unobserved degree heterogeneity or unobserved homophily. The choice is determined by which feature is more likely to be predominant.

When unobserved homophily is exhibited, the larger the difference in unobservables between individuals i and j, the less likely they are to connect. To capture homophily on unobserved dyad attributes, we use the conventional specification in the literature (see, e.g.,

[2] That is, the difference in utility between two options, linking or not.

Goldsmith-Pinkham and Imbens, 2013; Hsieh and Lee, 2016).

$$f(a_i, a_j) = -|a_i - a_j|. \tag{4.11}$$

Apart from homophily, Graham (2017) highlighted that individual heterogeneity in the number of links (a.k.a. degree), is another common feature of social networks. This is due to the fact that, the degree varies among individuals in many social networks – a few individuals serve as prominent "hubs" with numerous links, whereas others only own a few links. In addressing unobserved degree heterogeneity, we build on studies that accommodate this feature in network formation models (e.g., Ding et al., 2023; Dzemski, 2019) and incorporate individual-specific heterogeneity in an additive manner

$$f(a_i, a_j) = a_i + a_j. \tag{4.12}$$

The unobserved individual characteristic a_i can be interpreted as social capital, which enhances the likelihood of forming a link. Depending on the context, these characteristics may include trustworthiness, socioeconomic status, or charisma. These intangible factors also tend to affect individual activity outcomes, as evidenced in previous studies combining network formation and interaction (Han et al., 2021; Johnsson and Moon, 2021; Weng and Parent, 2023, among others).

Remark 3. To facilitate the discussion of the identification and estimation procedure that will be addressed later, we assume $a_i \sim \mathcal{N}(0, \sigma_a^2)$ in the context of unobserved degree heterogeneity, and $a_i \sim \mathcal{B}ern(p)$ when considering unobserved homophily.

A link between individuals i and j is formed if and only if it improves the average utility of i and j given by

$$w_{ij}^* := \frac{U_{ij}(\epsilon_{ij}^*) + U_{ji}(\epsilon_{ji}^*)}{2} = \psi_{ij} + \frac{\epsilon_{ij}^* + \epsilon_{ji}^*}{2} \geq 0. \tag{4.13}$$

Thus, binary link indicators representing the network are realisations of these latent average utilities

$$w_{ij} = w_{ji} = \begin{cases} 1 & \text{if } w_{ij}^* \geq 0, \\ 0 & \text{if } w_{ij}^* < 0. \end{cases} \qquad \text{for } i = 1, \ldots, N; j = 1, \ldots, i-1. \tag{4.14}$$

By defining the average of dyad-level shocks $\epsilon_{ij} \coloneqq \frac{\epsilon^*_{ij}+\epsilon^*_{ji}}{2}$, we can write the network formation model succinctly as follows

$$\begin{aligned} w^*_{ij} &= C_{ij}^\top \gamma + f(a_i, a_j) + \epsilon_{ij}, \\ w_{ij} &= w_{ji} = \mathbb{1}(w^*_{ij} \geq 0), \quad \text{for } i = 1, \ldots, N; j = 1, \ldots, i-1; \end{aligned} \tag{4.15}$$

where $\mathbb{1}(\cdot)$ is the indicator function.

Define $A_{ij} \coloneqq f(a_i, a_j)$ for $i = 1, \ldots, N$ and $j = 1, \ldots, N$ with $i \neq j$. Substituting into the representation of latent dyadic utilities w^*_{ij}, we obtain

$$w^*_{ij} = C_{ij}^\top \gamma + A_{ij} + \epsilon_{ij}. \tag{4.16}$$

Following the ways the elements of the adjacency matrix $\mathbf{W} = [w_{ij}]$ are indexed, we can stack across row i for a given j as detailed below

$$w^*_{-j} = C_{-j}\gamma + A_{-j} + \epsilon_{-j}, \quad j = 1, \ldots, N; \tag{4.17}$$

where w^*_{-j}, A_{-j}, and ϵ are $(N-1)$-dimensional vectors, C_{-j} is a $(N-1) \times L^2$ matrix

$$w^*_{-j} = \begin{bmatrix} w^*_{1j} \\ \vdots \\ w^*_{j-1,j} \\ w^*_{j+1,j} \\ \vdots \\ w^*_{Nj} \end{bmatrix}, C_{-j} = \begin{bmatrix} C_{1j}^\top \\ \vdots \\ C_{j-1,j}^\top \\ C_{j+1,j}^\top \\ \vdots \\ C_{Nj}^\top \end{bmatrix}, A_{-j} = \begin{bmatrix} A_{1j} \\ \vdots \\ A_{j-1,j} \\ A_{j+1,j} \\ \vdots \\ A_{Nj} \end{bmatrix}, \epsilon_{-j} = \begin{bmatrix} \epsilon_{1j} \\ \vdots \\ \epsilon_{j-1,j} \\ \epsilon_{j+1,j} \\ \vdots \\ \epsilon_{Nj} \end{bmatrix}.$$

We can then stack across the index j to obtain

$$\mathbf{w}^* = \mathbf{C}\gamma + \mathbf{A} + \boldsymbol{\epsilon}, \tag{4.18}$$

where $\mathbf{w}^*$, $\mathbf{A}$, and $\boldsymbol{\epsilon}$ are $N(N-1)$-dimensional vectors and $\mathbf{C}$ is a $N(N-1) \times L$ matrix

$$\mathbf{w}^* = \begin{bmatrix} w^*_{-1} \\ \vdots \\ w^*_{-N} \end{bmatrix}, \mathbf{C} = \begin{bmatrix} C_{-1} \\ \vdots \\ C_{-N} \end{bmatrix}, \mathbf{A} = \begin{bmatrix} A_{-1} \\ \vdots \\ A_{-N} \end{bmatrix}, \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_{-1} \\ \vdots \\ \epsilon_{-N} \end{bmatrix}.$$

As the uncertainty in the linking decision comes solely through ϵ_{ij} in (4.15), we lastly assume that this idiosyncratic error is i.i.d across dyads with a standard normal distribution such that

$$p(\boldsymbol{\epsilon}) = \prod_{i=1}^{N} \prod_{j<i} \phi(\epsilon_{ij}), \tag{4.19}$$

where $\phi(\cdot)$ denotes the standard normal density. This assumption is equivalent to setting the distributions of the original ϵ^*_{ij} as $\mathcal{N}(0, 2)$ independently across dyads. Fixing the variance of shocks, which is unity due to our normalisation, can be seen as an identifying restriction. This is standard in models with binary dependent variables (see, e.g., Cameron and Trivedi, 2005; Chan et al., 2019). Indeed, as we only observe whether a specific link was formed or not, all scalings of the idiosyncratic shocks will be observationally equivalent. Furthermore, while the distribution of ϵ_{ij} is often specified as logistic in the network formation literature, the advantages of (4.19) are demonstrated in Ding et al. (2023). They advocate that this assumption leads to a normal likelihood for the latent variables, which facilitates the incorporation of other elements into their sampling scheme in a Bayesian context.

Conditional on $(\mathbf{C}, \mathbf{a})$, the augmented likelihood[3] function for parameters in the network formation model in (4.15) is associated with the joint distribution of the observed network and latent link utilities $(\mathbf{W}, \mathbf{w}^*)$ and can be expressed as

$$\begin{aligned} & p\Big(\mathbf{W}, \mathbf{w}^* \mid \mathbf{C}, \mathbf{a}, \gamma\Big) \\ & \quad = p\Big(\mathbf{W} \mid \mathbf{w}^*\Big) \cdot p(\mathbf{w}^* \mid \mathbf{C}, \mathbf{a}, \gamma\Big) \\ & \quad = \prod_{i=1}^{N} \prod_{j<i} \Big[\mathbb{1}(w^*_{ij} \geq 0)\mathbb{1}(w_{ij} = 1) + \mathbb{1}(w^*_{ij} < 0)\mathbb{1}(w_{ij} = 0)\Big]\, \phi\Big(w^*_{ij} \mid \psi_{ij}, 1\Big), \end{aligned} \tag{4.20}$$

where $\phi\Big(w^*_{ij} \mid \psi_{ij}, 1\Big) = \phi\Big(w^*_{ij} - C_{ij}^{\top}\gamma - f(a_i, a_j)\Big)$.

This expression immediately implies that the conditional posterior distributions of preference parameters (γ) will only depend on the networks through values of $\mathbf{w}^*$.

[3] Without introducing $\mathbf{w}$ as latent data, the likelihood of observing network $\mathbf{W}$ is given by: $p(\mathbf{W} \mid \mathbf{C}, \mathbf{a}, \gamma) = \prod_{i=1}^{N} \prod_{j<i} [1 - \Phi(\psi_{ij})]^{1-w_{ij}} [\Phi(\psi_{ij})]^{w_{ij}}$. When used in a Bayesian setting, it does not yield conjugate full conditionals, making sampling from the posterior distribution of parameters computationally expensive and less efficient.

4.2.3 SCHSAR – A Joint Modelling Approach

4.2.3.1 Potential Endogeneity

We now reconsider the HSAR model in (4.2), accounting for individual unobserved heterogeneity, represented as $\mathbf{a} = [a_1, \ldots, a_N]^\top$. To expand this idea, we assume that $\mathbf{a}$ results from a component of individual characteristics in $\mathbf{X}$ that is unobservable to the researcher. Isolating $\mathbf{a}$ from $\mathbf{X}$ yields the following version of the model

$$Y_i = \lambda_g \sum_{j=1}^{N} w_{ij} Y_j + X_i \beta_g + a_i \kappa_g + u_i; \quad u_i \sim \mathcal{N}(0, \sigma_{ug}^2), \tag{4.21}$$

for each individual $i \in \mathcal{I}_g$, where $g = 1, \ldots, G$. Similar to β_g^k, κ_g is a type-specific coefficient that captures the magnitude and direction of the effect of a_i on the individual outcome. We define $\kappa \coloneqq [\kappa_1, \ldots, \kappa_G]^\top$ and $\tilde{\kappa} \coloneqq \mathbf{z}\kappa$.

If one ignores the unobserved factor $\mathbf{a}$, the simplified model to be estimated becomes

$$Y_i = \lambda_g \sum_{j=1}^{N} w_{ij} Y_j + X_i \beta_g + v_i, \quad \text{for } i \in \mathcal{I}_g, \tag{4.22}$$

where the error term v_i mistakenly includes the omitted variable a_i: $v_i = a_i \kappa_g + u_i$. Along with the network formation in equation (4.15), this misspecification induces correlation between the peer outcome $\bar{Y}_{(i)} = \sum_{j=1}^{N} w_{ij} Y_j$ and the error term, thus creating omitted variable bias (selection bias), even when observations are correctly classified into type groups. If κ_g were all zero, a_i and v_i would be uncorrelated. In this situation, the influences of the endogeneity of $\mathbf{W}$ on outcomes would be absent and the HSAR model can be estimated by treating $\mathbf{W}$ as exogenously given. Otherwise, the selection bias issue should be addressed.

Therefore, we propose the Selection-corrected Heterogeneous Spatial Autoregressive (SCHSAR) specification which takes into account potential endogenous selection within a network

$$\mathbf{Y} = \mathcal{L}\mathbf{W}\mathbf{Y} + \sum_{k=1}^{K} \tilde{\beta}^k \circ X^k + \tilde{\kappa} \circ \mathbf{a} + \mathbf{u}, \tag{4.23}$$

By explicitly introducing $\mathbf{a}$, the SCHSAR outcome equation in (4.23) is a further generalisation of the HSAR outcome equation in (4.5). Conditional on $(\mathbf{W}, \mathbf{X}, \mathbf{a})$, the augmented likelihood

function for parameters in the outcome equation becomes

$$p\left(\mathbf{Y}, \mathbf{z} \mid \mathbf{W}, \mathbf{X}, \mathbf{a}, \{\lambda_g, \beta_g, \kappa_g, \sigma_{ug}^2, \pi_g\}_{g=1}^G\right) = |\mathbf{I}_N - \mathcal{L}\mathbf{W}| \prod_{g=1}^G (2\pi\sigma_{ug}^2)^{-N_g/2} \exp\left\{\sum_{i=1}^N -\frac{z_{gi}(Y_i - \lambda_g \overline{Y}_{(i)} - X_i\beta_g - a_i\kappa_g)^2}{2\sigma_{ug}^2}\right\} \pi_g^{N_g}. \tag{4.24}$$

Putting everything together, the complete-data likelihood function for all model parameters is associated with the joint distribution of observed and latent data $(\mathbf{W}, \mathbf{w}^*, \mathbf{Y}, \mathbf{z}, \mathbf{a})$ and can be written as

$$\begin{aligned} &p\left(\mathbf{W}, \mathbf{w}^*, \mathbf{Y}, \mathbf{z}, \mathbf{a} \mid \mathbf{C}, \mathbf{X}, \theta_a, \gamma, \{\lambda_g, \beta_g, \kappa_g, \sigma_{ug}^2, \pi_g\}_{g=1}^G\right) \\ &= p\left(\mathbf{W}, \mathbf{w}^*, \mathbf{Y}, \mathbf{z} \mid \mathbf{C}, \mathbf{X}, \mathbf{a}, \gamma, \{\lambda_g, \beta_g, \kappa_g, \sigma_{ug}^2, \pi_g\}_{g=1}^G\right) \cdot p\left(\mathbf{a} \mid \theta_a\right) \\ &= p\left(\mathbf{W}, \mathbf{w}^* \mid \mathbf{C}, \mathbf{a}, \gamma\right) \cdot p\left(\mathbf{Y}, \mathbf{z} \mid \mathbf{W}, \mathbf{X}, \mathbf{a}, \{\lambda_g, \beta_g, \kappa_g, \sigma_{ug}^2, \pi_g\}_{g=1}^G\right) \cdot p\left(\mathbf{a} \mid \theta_a\right), \end{aligned} \tag{4.25}$$

The likelihood function factorises into a part associated with $(\mathbf{W}, \mathbf{w}^*, \mathbf{Y}, \mathbf{z})$ conditional on $\mathbf{a}$, and another part where $\mathbf{a}$ conditional on the parameters. The first part further factorises into separate contributions from the network and the outcome, as derived in (4.20) and (4.24), respectively. To obtain the marginal likelihood function in terms of the observed data $(\mathbf{W}, \mathbf{Y})$[4], we need to integrate out all latent variables $(\mathbf{w}^*, \mathbf{z}, \mathbf{a})$ from the joint likelihood in (4.25). Maximum likelihood estimation of this likelihood is impeded by the challenges associated with high-dimensional integration. Instead, we employ the Bayesian approach for estimation and inference, which is particularly advantageous in this setting. First, Bayesian data augmentation technique (Albert and Chib, 1993; Tanner and Wong, 1987) can be adopted to sample parameters together with latent variables from the joint posterior distribution. This circumvents the need to work with the marginal posterior distribution that lacks a closed-form expression. Second, the full Bayesian framework facilitates inference without relying on asymptotic approximations. By utilising the posterior sample, we can directly perform inference on various functions of the model parameters, such as direct and indirect effects. It is noteworthy that these effects are computed as own- and cross-partial derivatives for the response of outcome Y to changes in the explanatory variable X^k, which exhibit a nonlinear relationship with the model parameters.

Remark 4. Our Bayesian joint modelling approach to correct for selection on unobservables

[4]That is, the observed-data likelihood.

aligns with the classical control function method based on correlated unobservables (see, i.e, Heckman, 1979; Heckman and Robb Jr, 1985; Navarro, 2010), which has been extended to broader settings of social interaction models (see Blume et al., 2015; Johnsson and Moon, 2021 for recent treatments). The key mechanism of the control function approach is to identify and recover a proxy or function of the unobserved variable that drives endogeneity, using its statistical relationship with observed data. This allows us to condition on this function – termed a "control function" – to eliminate bias from omitted variables. In our context, the source of network endogeneity is unobserved individual heterogeneity, a_i, which concurrently influences both the network formation process and the individual outcomes. If a_i were observed, it would be straightforward to include it directly in the estimation and resolve the endogeneity issue. Since a_i is unobserved, we instead specify a joint Bayesian model where a_i enters explicitly as a latent variable in both the network formation equation and the outcome equation. Through this structure, the posterior distribution of a_i is updated based on the observed links and outcomes, effectively recovering a probabilistic representation of the latent heterogeneity. By incorporating a_i into the system, we convert a structural endogeneity problem into one of missing data, solvable through the complete-data likelihood. Model parameters can thus be estimated properly within a coherent inferential framework. Notably, this approach does not require external instruments – especially valuable when valid instruments are weak or unavailable. Moreover, the Bayesian approach is particularly advantageous in our setting, where heterogeneous network interaction effects are modelled via latent types in a finite mixture structure, further complicating the applicability of traditional instrumental variable methods. Broadly speaking, our SCHSAR approach inherits the Bayesian recipe by Goldsmith-Pinkham and Imbens (2013) and Hsieh and Lee (2016), but nests their models as a special case when the number of mixture components G is set to one.

4.2.3.2 Identification

We briefly outline our identification strategy for the proposed model as follows. First, the network model is semiparametrically identified, i.e., parameters in the deterministic components as well as distributions of disturbances (including individual unobserved heterogeneity, $\mathbf{a}$, and the idiosyncratic error, $\boldsymbol{\epsilon}$) are identified. The network formation equation in (4.15) implies a single-index equation

$$\mathbb{E}(w_{ij}|\mathbf{C}) = \mathbb{P}(w_{ij} = 1|\mathbf{C}) = \mathbb{P}(C_{ij}\gamma + f(a_i, a_j) + \epsilon_{ij} \geq 0|\mathbf{C}) = 1 - F_{\eta_{ij}}(-\xi_{ij}), \quad (4.26)$$

where $\xi_{ij} \coloneqq C_{ij}\gamma$ represents the deterministic component and $\eta_{ij} \coloneqq f(a_i, a_j) + \epsilon_{ij}$ represents the stochastic component which has the distribution function $F_{\eta_{ij}}(\cdot)$. Ichimura (1993) shows that, even when $F_{\eta_{ij}}(\cdot)$ is unknown, parameters in the linear index ξ_{ij} are identified up to a scale. With further parametric assumptions and constraints on η_{ij}, parameters in ξ_{ij} are identified, and hence ξ_{ij} can be determined. As ξ_{ij} is identified, the distribution function $F_{\eta_{ij}}(\cdot)$ can also be identified (estimated) from the data by a nonparametric kernel regression with ξ_{ij} as the regressor. Given that η_{ij} is continuous, if we assume ξ_{ij} can take values that cover the support of the probability density function $f_{\eta_{ij}}(\cdot)$, the moments of η_{ij} can also be estimated from the data. Accordingly, we can obtain $p(\mathbf{a}|\theta_a)$ in both cases of unobserved degree heterogeneity[5] and unobserved homophily.[6]

Concerning the identification of unknown parameters in the SCHSAR outcome equation in (4.23), we consider first $G = 1$.

$$\mathbf{Y} = \lambda\mathbf{W}\mathbf{Y} + \mathbf{X}\beta + \mathbf{a}\kappa + \mathbf{u}. \tag{4.27}$$

Deriving an expectation conditional on $\mathbf{W}$ of both sides, we obtain

$$\mathbb{E}(\mathbf{Y}|\mathbf{W}) = \lambda\mathbb{E}(\mathbf{W}\mathbf{Y}|\mathbf{W}) + \mathbb{E}(\mathbf{X}|\mathbf{W})\beta + \mathbb{E}(\mathbf{a}|\mathbf{W})\kappa, \tag{4.28}$$

where $\mathbb{E}(\mathbf{Y}|\mathbf{W})$, $\mathbb{E}(\mathbf{W}\mathbf{Y}|\mathbf{W})$. and $\mathbb{E}(\mathbf{X}|\mathbf{W})$ can be identified from the data[7]. Although $\mathbf{a}$ is unobserved, we can identify

$$\mathbb{E}(\mathbf{a}|\mathbf{W}) = \int_a \mathbf{a}p(\mathbf{a}|\mathbf{W})d\mathbf{a} = \int_a \mathbf{a}\frac{p(\mathbf{a})p(\mathbf{W}|\mathbf{a})}{p(\mathbf{W})}d\mathbf{a}, \tag{4.29}$$

provided that (i) $p(\mathbf{W})$ and parameters in $p(\mathbf{W}|\mathbf{a})$ are identifiable from the network data; and (ii) $p(\mathbf{a})$ is specified. Noting that $p(\mathbf{W}|\mathbf{a})$ is invariant between 1 and 0 under unobserved homophily, we restrict κ to be positive in this case for identification. Now, we let $\mathbf{\Omega} = [\mathbb{E}(\mathbf{W}\mathbf{Y}|\mathbf{W}), \mathbb{E}(\mathbf{X}|\mathbf{W}), \mathbb{E}(\mathbf{a}|\mathbf{W})]$. The necessary condition that $\mathbf{\Omega}^\top\mathbf{\Omega}$ has a full rank will identify the parameters in (4.28), including κ. Hence, there is no rotational indeterminacy issue on $\mathbf{a}$ and κ. This identification strategy suggests the possibility of a two-step estimation approach similar to Heckman's (1979) correction for sample selection.

[5] $\eta_{ij} = a_i + a_j + \epsilon_{ij}$, where $a_i \sim \mathcal{N}(0, \sigma_a^2)$; thus, $\mathbb{V}(\eta_{ij}) = \mathbb{V}(a_i) + \mathbb{V}(a_j) + \mathbb{V}(\epsilon_{ij}) = 2\sigma_a^2 + 1$.

[6] $\eta_{ij} = -|a_i - a_j| + \epsilon_{ij}$, where $a_i \sim \mathcal{B}ern(p)$; thus, $\mathbb{V}(\eta_{ij}) = \mathbb{V}(|a_i - a_j|) + \mathbb{V}(\epsilon_{ij}) = [2p(1-p) - 4p^2(1-p)^2] + 1$.

[7] Exclusion restriction condition for the selection models: in our specification the dyad-specific regressors in the network formation model are naturally excluded from the outcome model (Hsieh and Lee, 2016).

However, this two-step procedure might be difficult to implement in practice, because the calculation of $\mathbb{E}(\mathbf{a}|\mathbf{W})$ involves a high-dimensional integration. A standard method to replace integration in $\mathbb{E}(\mathbf{a}|\mathbf{W})$ is to use the sampling average of $\mathbf{a}$ from the target density $p(\mathbf{a}|\mathbf{W})$ via simulation. As mentioned earlier, our Bayesian approach coupled with posterior simulators offer computational tractability and straightforward inference.

To extend our identification arguments to the proposed framework when the number of latent groups $G > 1$, we could build on insight from Frühwirth-Schnatter (2006) on mixtures of Gaussian regression models. A trivial identification issue can occur due to the invariability of the likelihood to permutations of the labels. We avoid these problems by assuming that all types have pairwise different sizes and restricting the feasible parameter space: $1 > \pi_g > \pi_{g'} > 0$ for all $1 \leq g < g' \leq G$.

4.3 Bayesian Estimation

We implement Bayesian inference to simultaneously handle two key elements of the proposed SCHSAR framework – heterogeneity and endogeneity – with latent variables playing a crucial role, while allowing for seamless incorporation of model constraints. The procedure begins with a prior probabilistic belief about unknown parameters, collected into $\boldsymbol{\theta}$, and systematically updates this belief using observed sample data, denoted by $\boldsymbol{\mathcal{D}}$. Formally, Bayes' rule gives us the posterior distribution as follows

$$p(\boldsymbol{\theta} \mid \boldsymbol{\mathcal{D}}) \propto p(\boldsymbol{\mathcal{D}} \mid \boldsymbol{\theta}) \times p(\boldsymbol{\theta}),$$

where $p(\boldsymbol{\mathcal{D}} \mid \boldsymbol{\theta})$ is the likelihood function, and $p(\boldsymbol{\theta})$ is the prior distribution.

Throughout, we define $\boldsymbol{\theta}$ to encompass all model parameters – $\{\lambda_g, \beta_g, \kappa_g, \sigma^2_{u,g}, \pi_g\}_{g=1}^G$, γ, θ_a^2 – augmented with latent variables comprising individual random effects $\mathbf{a}$, latent individual-type indicators $\mathbf{z}$, and latent network link utilities $\mathbf{w}^*$. The observed outcomes, network data, and exogenous characteristics are compiled into $\boldsymbol{\mathcal{D}} = (\mathbf{W}, \mathbf{Y}, \mathbf{C}, \mathbf{X})$. This unified Bayesian approach is computational tractability and enables one-step inference across all unknown quantities, offering principled uncertainty quantification.

Having specified the likelihood function characterising our econometric model in (4.25), we proceed with providing a prior specification for posterior inference. For the ease of exposition, we initially focus on the case involving unobserved degree heterogeneity, where

the individual random effects influence network formation according to the linear additive form $f(a_i, a_j) = a_i + a_j$. Unobserved homophily, characterised by $f(a_i, a_j) = -|a_i - a_j|$, can be accommodated similarly within our Bayesian framework, albeit with minor modifications. This extension is elaborated at the end of this section.

4.3.1 Prior Specification

We specify prior distributions for all unknown quantities in our model. For computational convenience, we employ the conjugate priors commonly used in the Bayesian literature.

Regarding the individual random effect a_i, we assume a normal distribution with prior mean of zero and the variance σ_a^2, where the variance itself follows an Inverse Gamma distribution. This hyperparameter will be updated during our estimation procedure

$$a_i \mid \sigma_a^2 \overset{iid}{\sim} \mathcal{N}(0, \sigma_a^2), \tag{4.30}$$

$$\sigma_a^2 \sim \mathcal{IG}(\underline{s}_a, \underline{r}_a) \Leftrightarrow p(\sigma_a^2) = \frac{\underline{r}_a^{\underline{s}_a}}{\Gamma(\underline{s}_a)} (\sigma_a^2)^{-(\underline{s}_a+1)} \exp\left(-\frac{\underline{r}_a}{\sigma_a^2}\right). \tag{4.31}$$

For the network formation equation, we consider a multivariate normal prior for parameters γ

$$\gamma \sim \mathcal{N}_L(\underline{\nu}_\gamma, \underline{\Sigma}_\gamma). \tag{4.32}$$

For the outcome equation, we assume standard priors independently for β_g, κ_g, and σ_{ug}^2. Specifically, for each $g = 1, \ldots, G$,

$$\beta_g \sim \mathcal{N}_K(\underline{\nu}_\beta, \underline{\Sigma}_\beta), \tag{4.33}$$

$$\kappa_g \sim \mathcal{N}(\underline{\nu}_\kappa, \underline{\Sigma}_\kappa), \tag{4.34}$$

$$\sigma_{ug}^2 \sim \mathcal{IG}(\underline{s}_u, \underline{r}_u) \Leftrightarrow p(\sigma_{ug}^2) \propto \frac{\underline{r}_u^{\underline{s}_u}}{\Gamma(\underline{s}_u)} (\sigma_{ug}^2)^{-(\underline{s}_u+1)} \exp\left(-\frac{\underline{r}_u}{\sigma_{ug}^2}\right). \tag{4.35}$$

We also assume $\lambda_1, \ldots, \lambda_G$ are independent a priori. The prior for each parameter λ_g is assumed to follow a Beta distribution centered on zero, $\mathcal{B}(d, d)$, which is introduced by LeSage and Parent (2007) to represent an alternative to the uniform prior on the interval $(-1, 1)$.

$$\lambda_g \sim \mathcal{B}(d, d) \Leftrightarrow p(\lambda_g) = \frac{1}{\text{Beta}(d, d)} \frac{(1+\lambda_g)^{d-1}(1-\lambda_g)^{d-1}}{2^{2d-1}}, \tag{4.36}$$

where $\text{Beta}(d,d) = \int_0^1 t^{d-1}(1-t)^{d-1}dt$ is the Beta function. Specifically, values for hyperparameter d_o close to unity induce a relatively uninformative prior that places zero prior weight on end points of the interval for λ_g. In our setting, we initialize $\underline{d} = 1.01$. Recall that in our model setup, the individual type indicator vector z_i follows a Multinomial distribution that depends on a vector of assignment probabilities $\boldsymbol{\pi} = [\pi_1, \ldots, \pi_G]^\top$. We complete by assigning a Dirichlet prior on the distribution of $\boldsymbol{\pi}$[8]

$$z_i \mid \boldsymbol{\pi} \overset{iid}{\sim} \mathcal{M}ult(1,\pi) \Leftrightarrow p(\mathbf{z} \mid \boldsymbol{\pi}) = \prod_{i=1}^{N} p(z_i \mid \boldsymbol{\pi}) = \prod_{i=1}^{N}\prod_{g=1}^{G} \pi_g^{z_{gi}}, \tag{4.37}$$

$$\boldsymbol{\pi} \sim \mathcal{D}ir(\underline{\alpha}_1, \underline{\alpha}_2, \ldots, \underline{\alpha}_G) \Leftrightarrow p(\pi) \propto \pi_1^{\underline{\alpha}_1 - 1}\pi_2^{\underline{\alpha}_2 - 1}\ldots\pi_G^{\underline{\alpha}_G - 1}. \tag{4.38}$$

In our implementation for the simulation study and the empirical application in subsequent sections, we set the prior hyperparameters to standard non-informative values as follows: $\underline{\nu}_\gamma = \mathbf{0}_L$, $\underline{\Sigma}_\gamma = 10^4\mathbf{I}_L$, $\underline{\nu}_\beta = \mathbf{0}_K$, $\underline{\Sigma}_\beta = 10^4\mathbf{I}_K$, $\underline{\nu}_\kappa = 0$, $\underline{\Sigma}_\kappa = 10^4$, $\underline{r}_a = \underline{s}_a \approx 0$, $\underline{r}_u = \underline{s}_u \approx 0$, and $\underline{\alpha}_1 = \ldots = \underline{\alpha}_G = 1/G$.

4.3.2 Posterior Analysis

From Bayes' theorem, the augmented joint posterior density of interest can be expressed as

$$p(\boldsymbol{\theta} \mid \boldsymbol{\mathcal{D}}) \propto p\Big(\mathbf{W}, \mathbf{w}^*, \mathbf{Y}, \mathbf{z}, \mathbf{a} \mid \mathbf{C}, \mathbf{X}, \sigma_a, \gamma, \{\lambda_g, \beta_g, \kappa_g, \sigma_{u,g}^2, \pi_g\}_{g=1}^G\Big) \cdot p(\lambda, \beta, \kappa, \sigma_u^2, \boldsymbol{\pi}, \gamma, \sigma_a^2). \tag{4.39}$$

The first term on the right-hand side is the likelihood function defined previously in (4.25), which factorises further into

$$\begin{aligned} &p\Big(\mathbf{W}, \mathbf{w}^* \mid \mathbf{C}, \mathbf{a}, \gamma\Big) \cdot p\Big(\mathbf{Y}, \mathbf{z} \mid \mathbf{W}, \mathbf{X}, \mathbf{a}, \{\lambda_g, \beta_g, \kappa_g, \sigma_{u,g}^2, \pi_g\}_{g=1}^G\Big) \cdot p\Big(\mathbf{a} \mid \sigma_a^2\Big) \\ &= p\Big(\mathbf{W} \mid \mathbf{w}^*\Big) \cdot p(\mathbf{w}^* \mid \mathbf{C}, \mathbf{a}, \gamma\Big) \cdot p\Big(\mathbf{Y} \mid \mathbf{z}, \mathbf{W}, \mathbf{X}, \sigma_a, \gamma, \{\lambda_g, \beta_g, \kappa_g, \sigma_{u,g}^2, \pi_g\}\Big) \cdot p\Big(\mathbf{z} \mid \boldsymbol{\pi}\Big) \cdot p\Big(\mathbf{a} \mid \sigma_a^2\Big), \end{aligned} \tag{4.40}$$

[8] We requires an explicit specification regarding the number of types (i.e., G, the number of mixture components). This could be done by employing model selection criteria along with diagnostic plots.

and the second term in (4.39) gives the joint prior of the model parameters and factorises into independent priors as follows

$$p(\lambda, \beta, \kappa, \sigma_u^2, \boldsymbol{\pi}, \gamma, \sigma_a^2) = p(\lambda) \cdot p(\beta) \cdot p(\kappa) \cdot p(\sigma_u^2) \cdot p(\boldsymbol{\pi}) \cdot p(\gamma) \cdot p(\sigma_a^2). \tag{4.41}$$

With this representation, we estimate the model using the data-augmented Markov chain Monte Carlo (MCMC) approach. Specifically, our MCMC scheme treats latent variables $(\mathbf{w}^*, \mathbf{z}, \mathbf{a})$ as parameters to be estimated, and samples unknowns from the joint posterior distribution by cycling through three blocks of conditional distributions. The first block of conditionals is used for updating the individual random effects $\{a_i\}_{i=1}^N$ and their variance σ_a^2. The second block of conditionals is for other parameters and latent network utilities associated with the network formation equation. The third block of conditionals is for those associated with the outcome equation, including latent mixture indicators. These blocks respectively are

$$\begin{aligned} \mathbf{a}, \sigma_a^2 \mid \underline{s}_a, \underline{r}_a, \mathbf{w}^*, \gamma, \mathbf{z}, \{\lambda_g, \beta_g, \kappa_g, \sigma_{ug}^2, \pi_g\}_{g=1}^G, \boldsymbol{\mathcal{D}}; \\ \mathbf{w}^*, \gamma \mid \underline{\nu}_\gamma, \underline{\Sigma}_\gamma, \mathbf{a}, \boldsymbol{\mathcal{D}}; \\ \mathbf{z}, \{\lambda_g, \beta_g, \kappa_g, \sigma_{ug}^2, \pi_g\}_{g=1}^G \mid \underline{d}, \underline{\nu}_\beta, \underline{\Sigma}_\beta, \underline{\nu}_\kappa, \underline{\Sigma}_\kappa, \underline{s}_u, \underline{r}_u, \mathbf{a}, \boldsymbol{\mathcal{D}}. \end{aligned} \tag{4.42}$$

For notational simplicity, let $\boldsymbol{\theta}_{-\theta_1}$ denote the set of all parameters (including latent variables) in $\boldsymbol{\theta}$ excluding the component θ_1.

4.3.2.1 Conditional posteriors for $\mathbf{a}$, σ_a^2, $\mathbf{w}^*$, γ (random-effect and network blocks)

To facilitate efficient sampling, we introduce the following quantities relevant to the network formation equation

$$\begin{aligned} &\mathbf{w} = [w_{-1}^\top, \ldots, w_{-N}^\top]^\top, \quad w_{-j} = [w_{1j}, \ldots, w_{j-1,j}, w_{j+1,j}, \ldots, w_{Nj}]^\top, \quad j = 1, \ldots, J; \\ &\mathbf{F} = [F_{-1}, \ldots, F_{-N}]^\top, \quad F_{-j} = [e_1^\top, \ldots, e_{j-1}^\top, e_{j+1}^\top, \ldots, e_N^\top]^\top, \quad j = 1, \ldots, J; \\ &\mathbf{E} = \mathbf{I}_N \otimes \iota_{N-1}; \\ &\mathbf{H} = \mathbf{F} + \mathbf{E}, \end{aligned}$$

where $\otimes$ denotes the Kronecker product. Then, we can rewrite $\mathbf{A} = \mathbf{H}\mathbf{a}$ and replace into equation (4.15)

$$\mathbf{w}^* = \mathbf{C}\gamma + \mathbf{H}\mathbf{a} + \boldsymbol{\epsilon}. \tag{4.43}$$

Consequently, the full conditional posterior distributions for parameters and latent variables in the random-effects and network formation blocks can be written explicitly as follows

$$\mathbf{a} \mid \boldsymbol{\theta}_{-a_i}, \boldsymbol{\mathcal{D}} \sim \mathcal{N}\left(\overline{\nu}_a, \overline{\Sigma}_a\right), \tag{4.44}$$

$$\sigma_a^2 \mid \boldsymbol{\theta}_{-\sigma_a^2}, \boldsymbol{\mathcal{D}} \sim \mathcal{IG}(\overline{s}_a, \overline{r}_a), \tag{4.45}$$

$$\mathbf{w}^* \mid \boldsymbol{\theta}_{-\mathbf{w}^*}, \boldsymbol{\mathcal{D}} \sim TM\mathcal{N}_{\mathcal{S}}\left(\mathbf{C}\gamma + \mathbf{H}\mathbf{a}, \mathbf{I}_{N(N-1)}\right), \tag{4.46}$$

$$\gamma \mid \boldsymbol{\theta}_{-\gamma}, \boldsymbol{\mathcal{D}} \sim \mathcal{N}_L(\overline{\nu}_\gamma, \overline{\Sigma}_\gamma), \tag{4.47}$$

Here, $TM\mathcal{N}_{\mathcal{S}}$ denotes a multivariate normal distribution truncated to the region $\mathcal{S}$ implied by the binary adjacency vector $\mathbf{w}$ above.[9]. The posterior hyperparameters governing other conditional distributions are defined as

$$\begin{aligned}
\overline{\Sigma}_a &\coloneqq \left[\sigma_a^{-2}\mathbf{I}_N + \mathbf{H}^\top\mathbf{H} + \tilde{\kappa}\tilde{\kappa}^\top \circ \operatorname{diag}(\tilde{\sigma}_u^{-2})\right]^{-1}, \\
\overline{\nu}_a &\coloneqq \overline{\Sigma}_a\left[\mathbf{H}^\top(\mathbf{w}^* - \mathbf{C}\gamma) + \tilde{\kappa} \circ \tilde{\sigma}_u^{-2} \circ \left(\tilde{\mathbf{Y}} - \sum_{k=1}^{K} X^k \circ \tilde{\beta}^k\right)\right], \\
\overline{s}_a &\coloneqq \underline{s}_a + \frac{N}{2}, \\
\overline{r}_a &\coloneqq \underline{r}_a + \frac{\mathbf{a}^\top\mathbf{a}}{2}, \\
\overline{\Sigma}_\gamma &\coloneqq \left[\underline{\Sigma}_\gamma^{-1} + \mathbf{C}^\top\mathbf{C}\right]^{-1}, \\
\overline{\nu}_\gamma &\coloneqq \overline{\Sigma}_\gamma\left[\underline{\Sigma}_\gamma^{-1}\underline{\nu}_\gamma + \mathbf{C}^\top(\mathbf{w}^* - \mathbf{H}\mathbf{a})\right].
\end{aligned}$$

4.3.2.2 Conditional posteriors for z, $\{\lambda_g, \beta_g, \kappa_g, \sigma_{ug}^2, \pi_g\}_{g=1}^G$ (outcome block)

At each iteration, given the latent group assignments $\mathbf{z}$, the data are partitioned into G mixture components. For each group $g = 1 \ldots, G$, let $\mathcal{I}_g$ index observations assigned to group g and $N_g \coloneqq \sum_{i=1}^N z_{ig}$ denote the group size. We define $\tilde{Y}_g \coloneqq \{\tilde{Y}_i\}_{i\in\mathcal{I}_g}$, $X_g \coloneqq \{X_i\}_{i\in\mathcal{I}_g}$ and $a_g \coloneqq \{a_i\}_{i\in\mathcal{I}_g}$. The conditional posteriors of group-specific parameters including λ_g, β_g, κ_g, and σ_{ug}^2, can be characterised accordingly. Conversely, conditional on the parameters of the mixture, allocation z_i follows an independent Multinomial distribution with the classification weights (i.e., the probability that each observation i belongs to a given group g) can be calculated from the predictive densities. Finally, the conditional posterior for the G-dimensional vector of component probability π follows a Dirichlet distribution.

[9]Specifically, the truncated region is $[0, +\infty)$ for elements corresponding to observed network links ($w_{ij} = 1$), and $(-\infty, 0)$ otherwise ($w_{ij} = 0$). Sampling from this truncated multivariate normal distribution can be implemented efficiently using the algorithm proposed by Botev (2017).

The full conditional posteriors for the parameters are given as follows

$$p(\lambda_g \mid \boldsymbol{\theta}_{-\lambda_g}, \boldsymbol{\mathcal{D}}) \propto p(\lambda_g) \cdot |\mathbf{I}_N - \mathcal{L}\mathbf{W}| \cdot (2\pi\sigma_{ug}^2)^{-N_g/2} \tag{4.48}$$
$$\times \prod_{g=1}^{G} \exp\left\{-\frac{1}{2\sigma_{ug}^2}(\tilde{Y}_g - X_g\beta_g - a_g\kappa_g)^\top(\tilde{Y}_g - X_g\beta_g - a_g\kappa_g)\right\}, \quad g = 1, \ldots, G,$$

which is the only one that does not belong to a known class of distributions, whereas the conditional posteriors for other parameters follow standard conjugate results

$$\beta_g \mid \boldsymbol{\theta}_{-\beta_g}, \boldsymbol{\mathcal{D}} \sim \mathcal{N}_K\left(\overline{\nu}_{\beta g}, \overline{\Sigma}_{\beta g}\right), \quad g = 1, \ldots, G, \tag{4.49}$$
$$\kappa_g \mid \boldsymbol{\theta}_{-\kappa_g}, \boldsymbol{\mathcal{D}} \sim \mathcal{N}\left(\overline{\nu}_{\kappa g}, \overline{\Sigma}_{\kappa g}\right), \quad g = 1, \ldots, G, \tag{4.50}$$
$$\sigma_{ug}^2 \mid \boldsymbol{\theta}_{-\sigma_{ug}^2}, \boldsymbol{\mathcal{D}} \sim \mathcal{IG}\left(\overline{s}_{ug}, \overline{r}_{ug}\right), \quad g = 1, \ldots, G, \tag{4.51}$$
$$z_i \mid \boldsymbol{\theta}_{-z_i}, \boldsymbol{\mathcal{D}} \sim \mathcal{M}ult\left(1, [\omega_{i1}, \omega_{i2}, \ldots, \omega_{iG}]\right), \quad i = 1, \ldots, N, \tag{4.52}$$
$$\boldsymbol{\pi} \mid \boldsymbol{\theta}_{-\boldsymbol{\pi}}, \boldsymbol{\mathcal{D}} \sim \mathcal{D}ir(\overline{\alpha}_1, \overline{\alpha}_2, \ldots, \overline{\alpha}_G), \tag{4.53}$$

with posterior hyperparameters (for $i = 1, \ldots, N$ and $g = 1, \ldots, G$) are

$$\begin{aligned}
\overline{\Sigma}_{\beta g} &\coloneqq \left(\underline{\Sigma}_\beta^{-1} + \sigma_{ug}^{-2} X_g^\top X_g\right)^{-1}, \\
\overline{\nu}_{\beta g} &\coloneqq \overline{\Sigma}_{\beta g}\left[\underline{\Sigma}_\beta^{-1}\underline{\nu}_\beta + \sigma_{ug}^{-2} X_g^\top(\tilde{Y}_g - a_g\kappa_g)\right], \\
\overline{\Sigma}_{\kappa g} &\coloneqq \left(\underline{\Sigma}_\kappa^{-1} + \sigma_{u,g}^{-2} a_g^\top a_g\right)^{-1}, \\
\overline{\nu}_{\kappa g} &\coloneqq \overline{\Sigma}_{\kappa g}\left[\underline{\Sigma}_\kappa^{-1}\underline{\nu}_\kappa + a_g^\top(\tilde{Y}_g - X_g\beta_g)\right], \\
\overline{s}_{ug} &\coloneqq \underline{s}_u + \frac{N_g}{2}, \\
\overline{r}_{ug} &\coloneqq \underline{r}_u + \frac{1}{2}\left(\tilde{Y}_g - X_g\beta_g - a_g\kappa_g\right)^\top\left(\tilde{Y}_g - X_g\beta_g - a_g\kappa_g\right), \\
\omega_{ig} &\coloneqq \Pr\left(z_{ig} = 1 \mid \boldsymbol{\theta}_{-z_{ig}}, \boldsymbol{\mathcal{D}}\right) = \frac{\pi_g q_{ig}}{\sum_{g=1}^{G} \pi_g q_{ig}}, \\
q_{ig} &\coloneqq (2\pi\sigma_{ug}^2)^{-1/2} \exp\left\{-\frac{1}{2\sigma_{ug}^2}(Y_i - \lambda_g \overline{Y}_{(i)} - X_i\beta_g - a_i\kappa_g)^2\right\}, \\
\overline{\alpha}_g &\coloneqq \underline{\alpha}_g + N_g.
\end{aligned}$$

4.3.3 Markov Chain Monte Carlo (MCMC) Algorithm

As in the preceding posterior analysis, with conjugate priors, $\{a_i\}$, $\{w_{ij}^*\}$, γ, $\{z_{ig}\}$, $\{\beta_g\}$, $\{\kappa_g\}$, $\{\sigma_{ug}^2\}$, $\{\pi_g\}$ can be sampled straightforwardly via Gibbs sampling. By contrast, the conditional posterior distribution of $\{\lambda_g\}$ shown in (4.48) does not conform to a standard

form. To address this issue, a Metropolis-Hastings (M-H) step with an important twist is incorporated into the procedure, resulting in Metropolis-within-Gibbs sampling. Specifically, at the s-th iteration, for each group $g = 1, \ldots, G$, sampling λ_g from $p(\lambda_g \mid \boldsymbol{\theta}_{-\lambda_g}, \boldsymbol{\mathcal{D}})$ involves two main steps:

First, we generate a candidate value for λ_g by perturbing the current value using a proposal distribution:
Propose $\lambda_g^\star = \lambda_g^{(s-1)} + \mathcal{N}(0, \tau_g)$, where the scaling parameter (i.e., proposal increment shape) τ_g is tuned according to Adaptive Scaling Metropolis (ASM) algorithm (Andrieu and Thoms, 2008; Vihola, 2022), with details can be found in Appendix C.1.1. Compared with the standard random-walk Metropolis sampler, the key idea is to implement an automatic adjustment of τ_g based on monitoring the acceptance rates and stepsizes during the MCMC sampling procedure. As a result, this adaptation can learn from the historical MCMC draws (accepted draws) to make the proposal distribution better suited to the target distribution, thereby improving the efficiency and convergence of the algorithm.

Second, we compute the acceptance rate using the ratio of the posterior densities at the proposed and current values:
Let $\mathcal{L}_{\cdot|g}^\star = \text{diag}\left(\mathbf{z}[\lambda_1, \ldots, \lambda_{g-1}, \lambda_g^\star, \lambda_{g+1}, \ldots, \lambda_G]^\top\right)$, $\tilde{\mathbf{Y}}_g^\star = \left[\mathbf{I}_N - \tilde{\mathcal{L}}_{\cdot|g}^\star \mathbf{W}\right] \mathbf{Y} \circ z_g$; and
$\mathcal{L}_{\cdot|g}^{(s-1)} = \text{diag}\left(\mathbf{z}[\lambda_1, \ldots, \lambda_{g-1}, \lambda_g^{(s-1)}, \lambda_{g+1}, \ldots, \lambda_G]^\top\right)$,$\tilde{\mathbf{Y}}_g^{(s-1)} = \left[\mathbf{I}_N - \tilde{\mathcal{L}}_{\cdot|g}^{(s-1)} \mathbf{W}\right] \mathbf{Y} \circ z_g$.
Also,

$$u_{\cdot|g}^\star = \tilde{\mathbf{Y}}_g^\star - X_g\beta_g - a_g\kappa_g$$
$$u_{\cdot|g}^{(s-1)} = \tilde{\mathbf{Y}}_g^{(s-1)} - X_g\beta_g - a_g\kappa_g$$

Then, recall that $p(\lambda_g)$ is the density function of the Beta prior distribution defined in (4.36), with the acceptance rate

$$\alpha(\lambda_g^\star, \lambda_g^{(s-1)}) := \min\left\{ \frac{|\mathbf{I}_N - \tilde{\mathcal{L}}_{\cdot|g}^\star \mathbf{W}| \exp\left[-u_{\cdot|g}^{\star\top} u_{\cdot|g}^\star / (2\sigma_{ug}^2)\right]}{|\mathbf{I}_N - \tilde{\mathcal{L}}_{\cdot|g}^{(s-1)} \mathbf{W}| \exp\left[-u_{\cdot|g}^{(s-1)\top} u_{\cdot|g}^{(s-1)} / (2\sigma_{ug}^2)\right]} \times \frac{p(\lambda_g^\star)}{p(\lambda_g^{(s-1)})}, 1 \right\}, \tag{4.54}$$

update $\lambda_g^{(s)} = \lambda_g^\star$, else set $\lambda_g^{(s)} = \lambda_g^{(s-1)}$.

The full implementation of MCMC sampling scheme for the SCHSAR model with unobserved degree heterogeneity is outlined in Algorithm 4.1.

Algorithm 4.1: MCMC Sampler for SCHSAR Model
(*Unobserved Degree Heterogeneity*)

Procedure

Set initial values for all model parameters (augmented with latent variables) in $\boldsymbol{\theta}$.
Step 1: Update individual random effects $\mathbf{a}$ by sampling
  (a1) $\mathbf{a}$ from $\mathcal{N}\left(\overline{\nu}_a, \overline{\Sigma}_a\right)$, given $\boldsymbol{\theta}_{-\mathbf{a}}, \boldsymbol{\mathcal{D}}$; ▷ (4.44)
  (a2) σ_a^2 from $\mathcal{IG}(\overline{s}_a, \overline{r}_a)$, given $\mathbf{a}$. ▷ (4.45)
Step 2: Conditional on $\mathbf{a}$, update parameters in the network formation equation by sampling
  (b1) $\mathbf{w}^*$ from $TM\mathcal{N}_{\mathcal{S}}\left(\mathbf{C}\gamma + \mathbf{H}\mathbf{a}, \mathbf{I}_{N(N-1)}\right)$, given $\boldsymbol{\theta}_{-\mathbf{w}^*}, \boldsymbol{\mathcal{D}}$; ▷ (4.46)
  (b2) γ from $\mathcal{N}_L(\overline{\nu}_\gamma, \overline{\Sigma}_\gamma)$, given $\boldsymbol{\theta}_{-\gamma}, \boldsymbol{\mathcal{D}}$. ▷ (4.47)
Step 3: Conditional on $\mathbf{a}$ and the allocations $\mathbf{z}$, update parameters in the outcome equation by sampling
  (c1) $\boldsymbol{\pi}$ from $\mathcal{D}ir(\overline{\alpha}_1, \overline{\alpha}_2, \ldots, \overline{\alpha}_G)$, given $\boldsymbol{\theta}_{-\boldsymbol{\pi}}, \boldsymbol{\mathcal{D}}$; ▷ (4.53)
  (c2) each β_g, for $g = 1, \ldots, G$, from $\mathcal{N}_K\left(\overline{\nu}_{\beta g}, \overline{\Sigma}_{\beta g}\right)$, given $\boldsymbol{\theta}_{-\beta g}, \boldsymbol{\mathcal{D}}$; ▷ (4.49)
  (c3) each κ_g, for $g = 1, \ldots, G$, from $\mathcal{N}\left(\overline{\nu}_{\kappa g}, \overline{\Sigma}_{\kappa g}\right)$, given $\boldsymbol{\theta}_{-\kappa_g}, \boldsymbol{\mathcal{D}}$; ▷ (4.50)
  (c4) each σ_{ug}^2, for $g = 1, \ldots, G$, from $\mathcal{IG}\left(\overline{s}_{ug}, \overline{r}_{ug}\right)$, given $\boldsymbol{\theta}_{-\sigma_{ug}^2}, \boldsymbol{\mathcal{D}}$; ▷ (4.51)
  (c5) each λ_g, for $g = 1, \ldots, G$, via ASM algorithm, given $\boldsymbol{\theta}_{-\lambda_g}, \boldsymbol{\mathcal{D}}$. ▷ (4.48)
Step 4: For each $i = 1, \ldots, N$, sample the allocation z_i from $\mathcal{M}ult\left(1, [\omega_{i1}, \omega_{i2}, \ldots, \omega_{iG}]\right)$, given $\boldsymbol{\theta}_{-z_i}, \boldsymbol{\mathcal{D}}$. ▷ (4.52)
Repeat Steps 1-4 using the most recently updated values until convergence.
end procedure

4.3.4 Extension

We develop Bayesian estimation for the SCHSAR framework in the case of unobserved homophily, where the individual random effects influence network formation via the function $f(a_i, a_j) = -|a_i - a_j|$. In this setting, we assume that the unobserved types a_i are binary, modelled as independent Bernoulli random variables with mean p

$$a_i \overset{iid}{\sim} \mathcal{B}ern\,(p) \quad i = 1, \ldots, N. \tag{4.55}$$

Due to the absolute difference structure $|a_i - a_j|$, the signs of κ_g are not fully identified, as the likelihood is invariant under the transformations between $\kappa_g a_i$ and $(-\kappa_g)(-a_i)$. Therefore, we restrict κ_g to be nonnegative for estimation, which can be achieved seamlessly within the MCMC sampling procedure. A careful initialisation for $\mathbf{a}$ based on community detection algorithms described in Appendix C.1.2 further improves computational efficiency.

The conditional posterior for each a_i is

$$a_i \mid \boldsymbol{\theta}_{-a_i}, \boldsymbol{\mathcal{D}} \sim \mathcal{B}ern\left(\bar{p}\right), \quad i = 1, \ldots, N, \tag{4.56}$$

where

$$\begin{aligned}
\bar{p}_i &:= \frac{P_{1,i}}{P_{1,i} + P_{0,i}}, \\
P_{0,i} &:= \underline{p} \times \exp\left\{ -\sum_{g=1}^{G} \frac{z_{gi}(Y_i - \lambda_g \overline{Y}_{(i)} - X_i\beta_g - \kappa_g)^2}{2\sigma_g^2} - \sum_{j \neq i} \frac{(w_{ij}^* - C_{ij}\gamma + |1 - a_j|)^2}{2} \right\}, \\
P_{1,i} &:= (1 - \underline{p}) \times \exp\left\{ -\sum_{g=1}^{G} \frac{z_{gi}(Y_i - \lambda_g \overline{Y}_{(i)} - X_i\beta_g)^2}{2\sigma_g^2} - \sum_{j \neq i} \frac{(w_{ij}^* - C_{ij}\gamma + |a_j|)^2}{2} \right\}.
\end{aligned}$$

The conditional posterior distributions for the other parameters remain the same as derived previously, with adjustments required specifically for $\mathbf{w}^*, \gamma, \kappa_g$. Explicitly, their conditional posteriors become

$$\mathbf{w}^* \mid \boldsymbol{\theta}_{-\mathbf{w}^*}, \boldsymbol{\mathcal{D}} \sim TM\mathcal{N}_{\mathcal{S}}\left(\mathbf{C}\gamma - \mathbf{A}, \mathbf{I}_{N(N-1)}\right), \tag{4.57}$$

$$\gamma \mid \boldsymbol{\theta}_{-\gamma}, \boldsymbol{\mathcal{D}} \sim \mathcal{N}\left(\overline{\nu}_\gamma, \overline{\Sigma}_\gamma\right), \tag{4.58}$$

$$\kappa_g \mid \boldsymbol{\theta}_{-\kappa_g}, \boldsymbol{\mathcal{D}} \sim \mathcal{TN}_{[0,+\infty)}\left(\overline{\nu}_{\kappa g}, \overline{\Sigma}_{\kappa g}\right), \tag{4.59}$$

where $\mathcal{TN}_{[0,+\infty)}$ denotes a truncated normal distribution constrained to $\kappa_g \geq 0$, and

$$\begin{aligned}
\overline{\Sigma}_\gamma &:= \left[\underline{\Sigma}_\gamma^{-1} + \mathbf{C}^\top\mathbf{C}\right]^{-1}, \\
\overline{\nu}_\gamma &:= \overline{\Sigma}_\gamma\left[\underline{\Sigma}_\gamma^{-1}\underline{\nu}_\gamma + \mathbf{C}^\top(\mathbf{w}^* + \mathbf{A}\right], \\
\overline{\Sigma}_\gamma &:= \left[\underline{\Sigma}_\gamma^{-1} + \mathbf{C}^\top\mathbf{C}\right]^{-1}, \\
\overline{\nu}_{\kappa g} &:= \overline{\Sigma}_{\kappa g}\left[\underline{\Sigma}_\kappa^{-1}\underline{\nu}_\kappa + a_g^\top(\tilde{Y}_g - X_g\beta_g)\right], \quad g = 1, \ldots, G.
\end{aligned}$$

The full implementation of MCMC sampling scheme for the SCHSAR model with unobserved homophily is outlined in Algorithm 4.2.

Algorithm 4.2: MCMC Sampler for SCHSAR Model (*Unobserved Homophily*)

Procedure

Set initial values for all model parameters (augmented with latent variables) in $\boldsymbol{\theta}$.

Step 1: Sample unobserved component $\mathbf{a}$ by sampling

(a1) each a_i, for $i = 1, \ldots, N$, from $\mathcal{B}ern\,(\bar{p})$, given $\boldsymbol{\theta}_{-a_i}, \boldsymbol{\mathcal{D}}$. ▷ (4.56)

Step 2: Conditional on $\mathbf{a}$, sample parameters in the network formation equation by sampling

(b1) $\mathbf{w}^*$ from $TM\mathcal{N}_{\mathcal{S}}\left(\mathbf{C}\gamma - \mathbf{A}, \mathbf{I}_{N(N-1)}\right)$, given $\boldsymbol{\theta}_{-\mathbf{w}^*}, \boldsymbol{\mathcal{D}}$; ▷ (4.57)

(b2) γ from $\mathcal{N}_L(\overline{\nu}_\gamma, \overline{\Sigma}_\gamma)$, given $\boldsymbol{\theta}_{-\gamma}, \boldsymbol{\mathcal{D}}$. ▷ (4.58)

Step 3: Conditional on $\mathbf{a}$ and the allocations $\mathbf{z}$, sample parameters in the outcome equation by sampling

(c1) $\boldsymbol{\pi}$ from $\mathcal{D}ir(\overline{\alpha}_1, \overline{\alpha}_2, \ldots, \overline{\alpha}_G)$, given $\boldsymbol{\theta}_{-\boldsymbol{\pi}}, \boldsymbol{\mathcal{D}}$; ▷ (4.53)

(c2) each β_g, for $g = 1, \ldots, G$, from $\mathcal{N}_K\left(\overline{\nu}_{\beta g}, \overline{\overline{\Sigma}}_{\beta g}\right)$, given $\boldsymbol{\theta}_{-\beta_g}, \boldsymbol{\mathcal{D}}$; ▷ (4.49)

(c3) each κ_g, for $g = 1, \ldots, G$, from $\mathcal{TN}_{[0,+\infty)}\left(\overline{\nu}_{\kappa g}, \overline{\overline{\Sigma}}_{\kappa g}\right)$, given $\boldsymbol{\theta}_{-\kappa_g}, \boldsymbol{\mathcal{D}}$; ▷ (4.59)

(c4) each σ^2_{ug}, for $g = 1, \ldots, G$, from $\mathcal{IG}\left(\overline{s}_{ug}, \overline{r}_{ug}\right)$, given $\boldsymbol{\theta}_{-\sigma^2_{ug}}, \boldsymbol{\mathcal{D}}$; ▷ (4.51)

(c5) each λ_g, for $g = 1, \ldots, G$, via ASM algorithm, given $\boldsymbol{\theta}_{-\lambda_g}, \boldsymbol{\mathcal{D}}$. ▷ (4.48)

Step 4: For each $i = 1, \ldots, N$, sample the allocation z_i from $\mathcal{M}ult\,(1, [\omega_{i1}, \omega_{i2}, \ldots, \omega_{iG}])$, given $\boldsymbol{\theta}_{-z_i}, \boldsymbol{\mathcal{D}}$. ▷ (4.52)

Repeat Steps 1-4 using the most recent values until convergence.

end procedure

4.4 Simulation Study

We conduct a Monte Carlo simulation study to analyse the finite sample performance of the proposed Bayesian MCMC algorithm for estimating the joint SCHSAR model of the network formation equation in (4.15) and the outcome equation in (4.23). The simulation exercise also enables us to examine the extent of estimation bias in the network interaction effect parameter λ that arises when network endogeneity is ignored. Additionally, we evaluate the convergence and mixing behaviour of the Markov chain by inspecting diagnostics from representative simulation replications (see Appendix C.2.1).

4.4.1 Simulation Design

The general data-generating process (DGP) is based on the SCHSAR framework in Section 4.2, with the following key elements.

Generation of the individual random effects

Each individual $i = 1, \ldots, N$ is assigned an unobserved individual-specific factor a_i, which plays a central role in both the network formation and outcome equations. The form and distribution of a_i vary across simulation settings (detailed below).

Generation of the network data $\{w_{ij}\}_{i,j=1}^{N}$

We generate observed dyad-specific exogenous variable $\{C_{ij}\}_{i,j=1}^{N}$ by first drawing two random variables, v_1 and v_2, from the uniform distribution $U(0,1)$. We then set $C_{ij} = 1$ if both v_1 and v_2 are below 0.3 or above 0.7, and set $C_{ij} = 0$ otherwise. The corresponding coefficient of C_{ij} is set to $\gamma = 1.5$. We simulate each entry of the adjacency matrix $\mathbf{W}$ based on latent utilities following the network formation equation

$$w_{ij} = w_{ji} = \mathbb{1}\{C_{ij}\gamma + f(a_i, a_j) + \epsilon_{ij} \geq 0\}, \quad \text{for } i = 1, \ldots, N; j = 1, \ldots, i-1, \quad (4.60)$$

with idiosyncratic shock ϵ_{ij} drawn either from a standard normal or logistic distribution, depending on the specification.[10]

Generation of the outcome data $\{Y_i\}_{i=1}^{N}$

Individuals are randomly assigned to one of three latent types (i.e., $G = 3$) in each

[10] In other words, the conditional probability of each w_{ij} is $\mathbb{P}(w_{ij} = 1 \mid \cdot) = \mathcal{H}\left[C_{ij}^{\top}\gamma + f(a_i, a_j)\right]$, where $\mathcal{H}$ is the link function, either probit or logit, to be specified later.

simulation with fixed probabilities $\boldsymbol{\pi} = [0.45, 0.35, 0.2]$. Each individual is endowed with two observed covariates, $\{X_i^1\}_{i=1}^N$ and $\{X_i^2\}_{i=1}^N$, both from $\mathcal{N}(0, 4)$, whereas their corresponding coefficients are type-specific and set to $\beta^1 = [-0.5, 0.5, -1.0]^\top$ and $\beta^2 = [-0.75, 0.8, 1.2]^\top$. The effect of unobserved individual-specific factor a_i is associated with $\kappa = [0.8, 0.6, 0.25]^\top$. The type-specific peer effect parameter of interest is $\lambda = [-0.15, 0.15, 0.3]^\top$. The error term u_i is normally distributed with variance scaled across types by $\sigma_u^2 = c_\sigma \times [1, 0.75, 0.5]$, where c_σ controls the desired signal-to-noise ratio. Combining all components, the outcome variable is generated from the reduced form of the SCHSAR model[11]

$$\mathbf{Y} = [\mathbf{I}_N - \mathcal{L}\mathbf{W}]^{-1} \left(\sum_{k=1}^{K} \tilde{\beta}^k \circ X^k + \tilde{\kappa} \circ \mathbf{a} + \mathbf{u} \right), \tag{4.61}$$

where $\tilde{\beta}^k = \mathbf{z}\beta^k (k \in \{1, 2\})$, and $\tilde{\kappa} = \mathbf{z}\kappa$, with $\mathbf{z}$ being the matrix of latent type indicators.

The sample size is kept constant across the simulations and is reflective of a large sample, where $N = 1000$. Meanwhile, the specification of a_i, ϵ_{ij}, and c_σ are varied in the simulation study to account for different scenarios:

1. The form of the unobserved component in the network formation equation (source of network endogeneity): In the large literature on economic and social network analysis, the latent part associated with the endogeneity between network formation and individual outcomes could present as unobserved degree heterogeneity or unobserved homophily, depending on specific application contexts. Thus, we examine both cases separately.

$$f(a_i, a_j) = a_i + a_j, \text{ where } a_i \sim \mathcal{N}(0, \sigma_a^2) \text{ and } \sigma_a^2 = 2; \tag{4.62}$$

$$\text{or} \quad f(a_i, a_j) = -|a_i - a_j|, \text{ where } a_i \sim \mathcal{B}ern\left(\frac{1}{2}\right). \tag{4.63}$$

2. Possibility of misspecification of the link function: We switch the dyad-shock distribution from standard normal (probit link) to logistic (logit link) to explore the robustness of the estimator under misspecification.

3. The signal-to-noise ratio (SNR) in the outcome equation: We set $c_\sigma \in \{0.01, 0.1, 1\}$ which represents *high*, *medium*, and *low* levels of signal, respectively.

[11] We employ the row-normalised version of $\mathbf{W}$ in this outcome equation to ensure compatibility with interpretations in the empirical application. This can be implemented without loss of generality within our SCHSAR framework.

For each DGP, we generate $N_{sim} = 100$ independent replications. For each generated dataset, we estimate the following two models:

1. Heterogeneous Spatial Autoregressive (HSAR) Model: This benchmark model ignores the endogeneity of network structure by treating $\mathbf{W}$ as exogenous. It is conceptually aligned with the SAR-M model introduced by (Cornwall and Parent, 2017), which allows for heterogeneous interaction effects but assumes a fixed spatial weight matrix.

2. Selection-Corrected Heterogeneous Spatial Autoregressive (SCHSAR) Model: This is the fully flexible model we propose, which jointly models network formation and outcomes while accounting for heterogeneity in network interaction effects and endogeneity in network links through the inclusion of latent variables.

Each MCMC estimation is run for $5,500$ iterations, with the first 500 iterations discarded as burn-in.

4.4.2 Simulation Results

For every simulated dataset, given the posterior distribution of each model parameter resulting from MCMC draws, we derive the posterior mean for a point estimate and compute the corresponding equal-tailed 95% credible interval. We aggregate the results over 100 Monte Carlo replications and evaluate performance through the average bias and the root mean squared error (RMSE) of the point estimates, followed by the coverage rate of the 95% credible intervals. The simulation results are presented in Table 4.1–4.4, where the true values of the DGP parameters of interest are also listed in each table for ease of comparison.

Overall, across all data-generating processes (DGP I–IV) the proposed SCHSAR estimators produce near-unbiasedness and nominal coverage. When the signal-to-noise ratio (SNR) is high or medium, the true parameters are recovered very well as posterior means cluster tightly around the truth. Precision falls in low-signal setting as expected, yet coverage remains close to 0.95, indicating that intervals widen appropriately. By contrast, the benchmark HSAR estimator, which ignores endogenous network formation, displays severe upward or downward bias and virtually zero coverage for all peer-effect coefficients λ. Failing to model the link formation process when the peer effect parameter is of interest, therefore, renders the statistical inference unreliable.

When the latent component is unobserved degree heterogeneity (DGP I and II), SCHSAR

estimators achieve negligible bias for the peer effects (λ) even at low SNR. The variance of the individual random effects, σ_a^2, is recovered quite accurately, indicating that the continuous latent heterogeneity is effectively captured leads to the desirable performance of SCHSAR. In contrast, under unobserved homophily (DGP III and IV), the latent mechanism penalises dissimilar types. SCHSAR estimators perform well but both bias and variability of the peer-effect estimates increase, particularly under low signal. A similar pattern arises for the loadings κ, suggesting additional challenges added when capturing latent variables in this scenario.

DGP II and IV purposely replace the probit link assumed in estimation with a logit link in the true model. The performance of SCHSAR degrades gracefully: RMSE of the peer-effect estimates (λ) roughly increases relative to correctly specified cases but remains modest. SCHSAR estimators also maintain decent coverage, with a slight drop in the case of unobserved homophily yet stay above 0.85. Notably, while estimation for the network-covariate coefficient γ becomes fragile under misspecification, this does not propagate to peer effects, our primary quantities of interest. In short, link-function misspecification is tolerable for peer-effect estimation, indicating SCHSAR is reasonably robust.

In terms of other parameters, SCHSAR estimators for exogenous slope parameters β and group shares w exhibit virtually zero bias and coverage between 0.93 and 0.97 across all designs. These results bolster confidence that the mixture structure and covariate effects are recovered faithfully alongside endogenous peer effect parameters.

In summary, the Monte Carlo findings collectively demonstrate that the SCHSAR framework delivers reliable estimation and inference.

Table 4.1: DGP I: $N = 1000$, Unobserved Degree Heterogeneity

			SCHSAR					HSAR				
SNR	Parameter	True Value	Mean	Std	Bias	RMSE	Coverage	Mean	Std	Bias	RMSE	Coverage
	λ_1	-0.15	-0.151	0.010	-0.001	0.011	0.92	-0.881	0.101	-0.731	0.738	0.00
	λ_2	0.15	0.149	0.009	-0.001	0.010	0.94	-0.502	0.157	-0.652	0.670	0.00
	λ_3	0.30	0.298	0.019	-0.002	0.019	0.96	0.026	0.070	-0.274	0.283	0.00
	ω_1	0.45	0.452	0.015	0.002	0.015	0.98	0.451	0.020	0.001	0.020	0.98
	ω_2	0.35	0.349	0.016	-0.001	0.016	0.99	0.350	0.020	0.000	0.020	0.99
	ω_3	0.20	0.199	0.015	-0.001	0.015	0.93	0.199	0.017	-0.001	0.017	0.91
	β_{11}	-0.50	-0.500	0.003	0.000	0.003	0.93	-0.502	0.029	-0.002	0.029	0.92
	β_{12}	0.50	0.498	0.017	-0.002	0.017	0.86	0.496	0.026	-0.004	0.026	0.92
High	β_{13}	-1.00	-0.966	0.332	0.034	0.333	0.98	-1.000	0.014	0.000	0.014	0.97
	β_{21}	-0.75	-0.750	0.003	0.000	0.003	0.93	-0.749	0.027	0.001	0.027	0.95
	β_{22}	0.80	0.800	0.004	0.000	0.004	0.95	0.801	0.021	0.002	0.021	0.96
	β_{23}	1.20	1.158	0.404	-0.042	0.406	0.95	1.201	0.013	0.001	0.013	0.97
	κ_1	0.80	0.798	0.005	-0.002	0.005	0.90	–	–	–	–	–
	κ_2	0.60	0.598	0.007	-0.002	0.007	0.99	–	–	–	–	–
	κ_3	0.25	0.227	0.227	-0.023	0.228	0.94	–	–	–	–	–
	γ	1.50	1.501	0.010	0.001	0.010	0.80	–	–	–	–	–
	σ_a^2	2.00	2.001	0.083	0.001	0.083	0.99	–	–	–	–	–
	λ_1	-0.15	-0.153	0.033	-0.003	0.033	0.93	-0.876	0.105	-0.726	0.734	0.00
	λ_2	0.15	0.148	0.029	-0.002	0.029	0.95	-0.502	0.162	-0.652	0.672	0.00
	λ_3	0.30	0.300	0.031	0.000	0.031	0.96	0.028	0.081	-0.272	0.284	0.01
	ω_1	0.45	0.452	0.015	0.002	0.016	0.97	0.451	0.020	0.001	0.020	0.97
	ω_2	0.35	0.348	0.016	-0.002	0.016	0.98	0.349	0.021	-0.001	0.021	0.98
	ω_3	0.20	0.200	0.014	0.000	0.014	0.94	0.200	0.017	0.000	0.017	0.92
	β_{11}	-0.50	-0.501	0.009	-0.001	0.009	0.91	-0.503	0.030	-0.003	0.030	0.94
	β_{12}	0.50	0.500	0.010	0.000	0.010	0.90	0.496	0.028	-0.004	0.028	0.93
Medium	β_{13}	-1.00	-0.999	0.008	0.002	0.008	0.97	-0.998	0.017	0.002	0.017	0.96
	β_{21}	-0.75	-0.750	0.008	0.000	0.008	0.94	-0.749	0.029	0.001	0.029	0.94
	β_{22}	0.80	0.800	0.008	0.000	0.008	0.95	0.801	0.023	0.001	0.023	0.96
	β_{23}	1.20	1.200	0.011	0.000	0.011	0.95	1.202	0.016	0.002	0.016	0.96
	κ_1	0.80	0.798	0.013	-0.002	0.013	0.95	–	–	–	–	–
	κ_2	0.60	0.598	0.013	-0.002	0.013	0.95	–	–	–	–	–
	κ_3	0.25	0.249	0.014	-0.001	0.014	0.96	–	–	–	–	–
	γ	1.50	1.501	0.010	0.001	0.010	0.80	–	–	–	–	–
	σ_a^2	2.00	2.002	0.083	0.002	0.083	0.99	–	–	–	–	–
	λ_1	-0.15	-0.158	0.116	-0.008	0.117	0.95	-0.828	0.133	-0.678	0.691	0.00
	λ_2	0.15	0.135	0.117	-0.015	0.117	0.93	-0.492	0.203	-0.642	0.673	0.02
	λ_3	0.30	0.293	0.132	-0.007	0.132	0.95	0.011	0.175	-0.289	0.338	0.50
	ω_1	0.45	0.450	0.018	0.000	0.018	0.97	0.449	0.022	-0.001	0.022	0.99
	ω_2	0.35	0.348	0.022	-0.002	0.022	0.96	0.346	0.026	-0.004	0.027	0.98
	ω_3	0.20	0.203	0.018	0.003	0.018	0.95	0.205	0.022	0.005	0.023	0.93
	β_{11}	-0.50	-0.496	0.101	0.004	0.101	0.91	-0.487	0.141	0.013	0.141	0.95
	β_{12}	0.50	0.491	0.109	-0.009	0.110	0.91	0.468	0.194	-0.032	0.197	0.94
Low	β_{13}	-1.00	-0.990	0.033	0.010	0.035	0.96	-0.977	0.137	0.024	0.139	0.95
	β_{21}	-0.75	-0.735	0.154	0.015	0.155	0.94	-0.720	0.216	0.030	0.218	0.93
	β_{22}	0.80	0.784	0.153	-0.016	0.154	0.95	0.771	0.230	-0.029	0.232	0.92
	β_{23}	1.20	1.202	0.037	0.002	0.037	0.92	1.194	0.076	-0.006	0.076	0.94
	κ_1	0.80	0.798	0.048	-0.002	0.048	0.95	–	–	–	–	–
	κ_2	0.60	0.597	0.051	-0.003	0.051	0.94	–	–	–	–	–
	κ_3	0.25	0.246	0.062	-0.004	0.062	0.97	–	–	–	–	–
	γ	1.50	1.501	0.010	0.001	0.010	0.80	–	–	–	–	–
	σ_a^2	2.00	2.002	0.083	0.002	0.083	0.98	–	–	–	–	–

Notes: This table displays results based on $R = 100$ replicates. The values include the average and standard deviation of the point estimates; the average bias (Bias), the Root Mean Squared Error (RMSE), and the coverage rate (Coverage) across replicates; where Bias $= R^{-1}\sum_{r=1}^{R}(\hat{\alpha}_r - \alpha)$, RMSE $= \sqrt{R^{-1}\sum_{r=1}^{R}(\hat{\alpha}_r - \alpha)^2}$, and Coverage $= R^{-1}\sum_{r=1}^{R}\mathbb{1}\{\alpha \in \widehat{CI}_{0.95,r}\}$.

Table 4.2: DGP II: $N = 1000$, Unobserved Degree Heterogeneity, Link Misspecification

			SCHSAR					HSAR				
SNR	Parameter	True Value	Mean	Std	Bias	RMSE	Coverage	Mean	Std	Bias	RMSE	Coverage
	λ_1	-0.15	-0.150	0.014	0.000	0.014	0.94	-0.878	0.110	-0.728	0.736	0.00
	λ_2	0.15	0.151	0.013	0.001	0.013	0.96	-0.590	0.217	-0.740	0.771	0.00
	λ_3	0.30	0.300	0.012	0.000	0.012	0.98	-0.021	0.114	-0.321	0.341	0.01
	ω_1	0.45	0.452	0.015	0.002	0.015	0.98	0.450	0.020	0.000	0.020	0.98
	ω_2	0.35	0.349	0.015	-0.001	0.015	0.98	0.350	0.021	0.000	0.021	0.99
	ω_3	0.20	0.200	0.014	0.000	0.014	0.91	0.200	0.017	0.000	0.017	0.93
	β_{11}	-0.50	-0.500	0.003	0.000	0.003	0.97	-0.502	0.031	-0.002	0.031	0.94
	β_{12}	0.50	0.500	0.003	0.000	0.004	0.81	0.495	0.028	-0.005	0.028	0.92
High	β_{13}	-1.00	-1.000	0.002	0.000	0.002	0.98	-0.999	0.017	0.001	0.017	0.98
	β_{21}	-0.75	-0.750	0.003	0.000	0.003	0.96	-0.749	0.028	0.001	0.028	0.96
	β_{22}	0.80	0.800	0.003	0.000	0.003	0.96	0.802	0.023	0.002	0.023	0.97
	β_{23}	1.20	1.200	0.003	0.000	0.003	0.91	1.198	0.035	-0.002	0.035	0.98
	κ_1	0.80	1.387	0.000	0.587	0.587	0.00	–	–	–	–	–
	κ_2	0.60	1.041	0.000	0.441	0.441	0.00	–	–	–	–	–
	κ_3	0.25	0.434	0.006	0.184	0.184	0.00	–	–	–	–	–
	γ	1.50	0.864	0.000	-0.636	0.636	0.00	–	–	–	–	–
	σ_a^2	2.00	0.661	0.023	-1.339	1.339	0.00	–	–	–	–	–
	λ_1	-0.15	-0.153	0.042	-0.003	0.042	0.94	-0.872	0.115	-0.722	0.731	0.00
	λ_2	0.15	0.150	0.037	0.000	0.037	0.96	-0.588	0.217	-0.738	0.769	0.00
	λ_3	0.30	0.300	0.039	0.000	0.039	0.97	-0.019	0.128	-0.319	0.344	0.03
	ω_1	0.45	0.452	0.016	0.002	0.016	0.98	0.451	0.020	0.001	0.020	0.99
	ω_2	0.35	0.348	0.016	-0.002	0.016	0.97	0.349	0.021	-0.001	0.021	0.98
	ω_3	0.20	0.200	0.014	0.000	0.014	0.96	0.200	0.018	0.000	0.018	0.94
	β_{11}	-0.50	-0.501	0.009	-0.001	0.009	0.90	-0.503	0.031	-0.003	0.032	0.94
	β_{12}	0.50	0.500	0.010	0.000	0.010	0.86	0.496	0.029	-0.004	0.029	0.94
Medium	β_{13}	-1.00	-0.999	0.008	0.002	0.008	0.96	-0.998	0.017	0.002	0.017	0.97
	β_{21}	-0.75	-0.750	0.008	0.000	0.008	0.95	-0.749	0.030	0.001	0.030	0.94
	β_{22}	0.80	0.799	0.008	-0.001	0.008	0.95	0.801	0.025	0.001	0.025	0.96
	β_{23}	1.20	1.200	0.009	0.000	0.010	0.94	1.202	0.017	0.002	0.017	0.96
	κ_1	0.80	1.390	0.024	0.590	0.591	0.00	–	–	–	–	–
	κ_2	0.60	1.042	0.019	0.442	0.442	0.00	–	–	–	–	–
	κ_3	0.25	0.433	0.024	0.183	0.185	0.00	–	–	–	–	–
	γ	1.50	0.864	0.000	-0.636	0.636	0.00	–	–	–	–	–
	σ_a^2	2.00	0.660	0.023	-1.340	1.340	0.00	–	–	–	–	–
	λ_1	-0.15	-0.160	0.148	-0.011	0.148	0.95	-0.816	0.144	-0.666	0.681	0.00
	λ_2	0.15	0.136	0.147	-0.014	0.147	0.92	-0.554	0.237	-0.704	0.743	0.05
	λ_3	0.30	0.281	0.171	-0.019	0.172	0.94	-0.040	0.226	-0.340	0.408	0.58
	ω_1	0.45	0.450	0.018	0.000	0.018	0.99	0.449	0.022	-0.001	0.022	0.99
	ω_2	0.35	0.348	0.022	-0.002	0.022	0.97	0.346	0.027	-0.004	0.027	0.98
	ω_3	0.20	0.202	0.018	0.002	0.018	0.95	0.205	0.022	0.005	0.023	0.92
	β_{11}	-0.50	-0.496	0.101	0.004	0.101	0.91	-0.486	0.141	0.014	0.142	0.95
	β_{12}	0.50	0.491	0.109	-0.009	0.110	0.90	0.481	0.155	-0.019	0.156	0.95
Low	β_{13}	-1.00	-0.989	0.033	0.011	0.035	0.96	-0.988	0.042	0.012	0.044	0.94
	β_{21}	-0.75	-0.735	0.154	0.015	0.155	0.94	-0.720	0.216	0.030	0.218	0.93
	β_{22}	0.80	0.784	0.153	-0.016	0.154	0.95	0.768	0.227	-0.032	0.229	0.93
	β_{23}	1.20	1.200	0.039	0.000	0.039	0.93	1.201	0.039	0.001	0.039	0.96
	κ_1	0.80	1.391	0.082	0.591	0.596	0.00	–	–	–	–	–
	κ_2	0.60	1.041	0.084	0.441	0.449	0.00	–	–	–	–	–
	κ_3	0.25	0.427	0.105	0.177	0.206	0.59	–	–	–	–	–
	γ	1.50	0.864	0.000	-0.636	0.636	0.00	–	–	–	–	–
	σ_a^2	2.00	0.660	0.023	-1.340	1.341	0.00	–	–	–	–	–

Notes: This table displays results based on $R = 100$ replicates. The values include the average and standard deviation of the point estimates; the average bias (Bias), the Root Mean Squared Error (RMSE), and the coverage rate (Coverage) across replicates; where Bias $= R^{-1}\sum_{r=1}^{R}(\hat{\alpha}_r - \alpha)$, RMSE $= \sqrt{R^{-1}\sum_{r=1}^{R}(\hat{\alpha}_r - \alpha)^2}$, and Coverage $= R^{-1}\sum_{r=1}^{R}\mathbb{1}\{\alpha \in \widehat{CI}_{0.95,r}\}$.

Table 4.3: DGP III: $N = 1000$, Unobserved Homophily

			SCHSAR					HSAR				
SNR	Parameter	True Value	Mean	Std	Bias	RMSE	Coverage	Mean	Std	Bias	RMSE	Coverage
High	λ_1	-0.15	-0.102	0.228	0.048	0.232	0.91	0.973	0.055	1.123	1.124	0.00
	λ_2	0.15	0.184	0.170	0.034	0.172	0.92	0.983	0.030	0.833	0.834	0.00
	λ_3	0.30	0.318	0.108	0.018	0.109	0.90	0.744	0.166	0.444	0.474	0.00
	ω_1	0.45	0.452	0.015	0.002	0.015	0.97	0.454	0.020	0.004	0.020	0.94
	ω_2	0.35	0.348	0.015	-0.002	0.015	0.99	0.347	0.018	-0.003	0.018	0.97
	ω_3	0.20	0.199	0.013	-0.001	0.013	0.94	0.199	0.014	-0.001	0.014	0.94
	β_{11}	-0.50	-0.500	0.003	0.000	0.003	0.95	-0.490	0.071	0.010	0.071	0.92
	β_{12}	0.50	0.500	0.002	0.000	0.002	0.97	0.480	0.142	-0.020	0.142	0.94
	β_{13}	-1.00	-1.000	0.002	0.000	0.002	0.96	-0.991	0.071	0.009	0.071	0.92
	β_{21}	-0.75	-0.750	0.003	0.000	0.003	0.94	-0.716	0.236	0.034	0.238	0.94
	β_{22}	0.80	0.800	0.002	0.000	0.002	0.96	0.769	0.218	-0.031	0.219	0.97
	β_{23}	1.20	1.201	0.003	0.001	0.003	0.93	1.162	0.275	-0.038	0.276	0.91
	κ_1	0.80	0.753	0.224	-0.047	0.228	0.89	–	–	–	–	–
	κ_2	0.60	0.567	0.166	-0.034	0.168	0.91	–	–	–	–	–
	κ_3	0.25	0.235	0.079	-0.015	0.080	0.88	–	–	–	–	–
	γ	1.50	1.501	0.005	0.001	0.005	0.81	–	–	–	–	–
Medium	λ_1	-0.15	-0.105	0.251	0.045	0.254	0.91	0.966	0.052	1.116	1.117	0.00
	λ_2	0.15	0.182	0.195	0.032	0.197	0.94	0.970	0.033	0.820	0.821	0.00
	λ_3	0.30	0.311	0.156	0.011	0.155	0.92	0.736	0.088	0.436	0.444	0.00
	ω_1	0.45	0.452	0.016	0.002	0.016	0.97	0.453	0.018	0.003	0.018	0.97
	ω_2	0.35	0.349	0.017	-0.001	0.017	1.00	0.348	0.017	-0.002	0.017	0.98
	ω_3	0.20	0.199	0.015	-0.001	0.015	0.92	0.198	0.015	-0.002	0.015	0.94
	β_{11}	-0.50	-0.500	0.008	0.000	0.008	0.96	-0.500	0.012	0.000	0.012	0.93
	β_{12}	0.50	0.498	0.014	-0.002	0.014	0.96	0.499	0.013	-0.001	0.013	0.96
	β_{13}	-1.00	-0.999	0.013	0.001	0.013	0.95	-0.999	0.018	0.001	0.018	0.97
	β_{21}	-0.75	-0.750	0.008	0.000	0.008	0.95	-0.749	0.011	0.001	0.011	0.95
	β_{22}	0.80	0.800	0.009	0.000	0.009	0.96	0.800	0.010	0.000	0.010	0.97
	β_{23}	1.20	1.197	0.043	-0.003	0.043	0.92	1.199	0.031	-0.001	0.031	0.92
	κ_1	0.80	0.754	0.232	-0.046	0.235	0.89	–	–	–	–	–
	κ_2	0.60	0.568	0.171	-0.032	0.173	0.94	–	–	–	–	–
	κ_3	0.25	0.237	0.098	-0.013	0.098	0.92	–	–	–	–	–
	γ	1.50	1.501	0.005	0.001	0.005	0.82	–	–	–	–	–
Low	λ_1	-0.15	-0.111	0.308	0.039	0.309	0.95	0.879	0.059	1.029	1.031	0.00
	λ_2	0.15	0.096	0.326	-0.054	0.329	0.94	0.864	0.067	0.714	0.717	0.00
	λ_3	0.30	0.142	0.289	-0.158	0.328	0.94	0.612	0.128	0.312	0.337	0.59
	ω_1	0.45	0.451	0.020	0.001	0.020	0.99	0.452	0.021	0.002	0.021	0.97
	ω_2	0.35	0.347	0.022	-0.003	0.022	0.98	0.347	0.023	-0.003	0.023	0.99
	ω_3	0.20	0.202	0.019	0.002	0.019	0.96	0.201	0.019	0.001	0.019	0.98
	β_{11}	-0.50	-0.488	0.099	0.012	0.100	0.96	-0.498	0.030	0.002	0.030	0.96
	β_{12}	0.50	0.489	0.104	-0.011	0.104	0.94	0.500	0.030	0.000	0.030	0.95
	β_{13}	-1.00	-1.000	0.032	0.000	0.032	0.97	-1.000	0.033	0.000	0.033	0.98
	β_{21}	-0.75	-0.735	0.159	0.015	0.159	0.92	-0.749	0.029	0.001	0.028	0.93
	β_{22}	0.80	0.782	0.157	-0.018	0.157	0.92	0.797	0.031	-0.003	0.031	0.92
	β_{23}	1.20	1.206	0.033	0.006	0.034	0.95	1.206	0.033	0.006	0.034	0.94
	κ_1	0.80	0.754	0.230	-0.046	0.233	0.95	–	–	–	–	–
	κ_2	0.60	0.610	0.210	0.010	0.209	0.95	–	–	–	–	–
	κ_3	0.25	0.308	0.173	0.058	0.182	0.92	–	–	–	–	–
	γ	1.50	1.501	0.005	0.001	0.005	0.84	–	–	–	–	–

Notes: This table displays results based on $R = 100$ replicates. The values include the average and standard deviation of the point estimates; the average bias (Bias), the Root Mean Squared Error (RMSE), and the coverage rate (Coverage) across replicates; where Bias $= R^{-1}\sum_{r=1}^{R}(\hat{\alpha}_r - \alpha)$, RMSE $= \sqrt{R^{-1}\sum_{r=1}^{R}(\hat{\alpha}_r - \alpha)^2}$, and Coverage $= R^{-1}\sum_{r=1}^{R}\mathbb{1}\{\alpha \in \widehat{CI}_{0.95,r}\}$.

Table 4.4: DGP IV: $N = 1000$, Unobserved Homophily, Link Misspecification

			SCHSAR					HSAR				
SNR	Parameter	True Value	Mean	Std	Bias	RMSE	Coverage	Mean	Std	Bias	RMSE	Coverage
High	λ_1	-0.15	-0.066	0.294	0.084	0.305	0.89	0.961	0.067	1.111	1.113	0.00
	λ_2	0.15	0.208	0.220	0.059	0.226	0.89	0.974	0.044	0.824	0.825	0.00
	λ_3	0.30	0.335	0.141	0.035	0.145	0.89	0.723	0.216	0.423	0.474	0.00
	ω_1	0.45	0.452	0.015	0.002	0.015	0.98	0.454	0.021	0.004	0.022	0.95
	ω_2	0.35	0.349	0.016	-0.001	0.016	0.98	0.348	0.017	-0.002	0.017	0.99
	ω_3	0.20	0.198	0.015	-0.002	0.015	0.93	0.198	0.017	-0.002	0.017	0.93
	β_{11}	-0.50	-0.500	0.003	0.000	0.003	0.93	-0.475	0.112	0.025	0.114	0.86
	β_{12}	0.50	0.498	0.021	-0.002	0.021	0.97	0.451	0.219	-0.049	0.224	0.91
	β_{13}	-1.00	-1.000	0.002	0.000	0.002	0.97	-0.978	0.102	0.022	0.103	0.89
	β_{21}	-0.75	-0.750	0.003	0.000	0.003	0.95	-0.663	0.375	0.087	0.383	0.91
	β_{22}	0.80	0.800	0.007	0.000	0.007	0.94	0.722	0.340	-0.078	0.347	0.94
	β_{23}	1.20	1.201	0.003	0.001	0.003	0.93	1.112	0.397	-0.088	0.405	0.90
	κ_1	0.80	0.717	0.302	-0.083	0.312	0.85	–	–	–	–	–
	κ_2	0.60	0.539	0.224	-0.061	0.231	0.89	–	–	–	–	–
	κ_3	0.25	0.236	0.184	-0.014	0.184	0.87	–	–	–	–	–
	γ	1.50	1.144	0.005	-0.356	0.356	0.00	–	–	–	–	–
Medium	λ_1	-0.15	-0.080	0.287	0.070	0.294	0.88	0.956	0.060	1.106	1.107	0.00
	λ_2	0.15	0.194	0.224	0.044	0.227	0.91	0.962	0.039	0.812	0.813	0.00
	λ_3	0.30	0.324	0.158	0.024	0.160	0.91	0.731	0.095	0.431	0.441	0.00
	ω_1	0.45	0.452	0.016	0.002	0.016	0.97	0.453	0.018	0.003	0.018	0.97
	ω_2	0.35	0.349	0.017	-0.001	0.017	1.00	0.348	0.017	-0.002	0.017	0.97
	ω_3	0.20	0.199	0.015	-0.001	0.015	0.92	0.199	0.015	-0.001	0.015	0.93
	β_{11}	-0.50	-0.500	0.008	0.000	0.008	0.95	-0.500	0.013	0.000	0.013	0.94
	β_{12}	0.50	0.498	0.013	-0.002	0.013	0.94	0.500	0.011	0.000	0.011	0.96
	β_{13}	-1.00	-0.998	0.013	0.002	0.013	0.96	-1.001	0.009	-0.001	0.009	0.96
	β_{21}	-0.75	-0.750	0.008	0.000	0.008	0.95	-0.749	0.011	0.001	0.011	0.97
	β_{22}	0.80	0.800	0.009	0.000	0.008	0.94	0.800	0.010	0.000	0.010	0.97
	β_{23}	1.20	1.195	0.041	-0.005	0.041	0.92	1.202	0.010	0.002	0.010	0.93
	κ_1	0.80	0.726	0.292	-0.074	0.299	0.88	–	–	–	–	–
	κ_2	0.60	0.550	0.213	-0.050	0.218	0.92	–	–	–	–	–
	κ_3	0.25	0.225	0.112	-0.025	0.114	0.90	–	–	–	–	–
	γ	1.50	1.144	0.005	-0.356	0.356	0.00	–	–	–	–	–
Low	λ_1	-0.15	-0.053	0.375	0.097	0.385	0.89	0.865	0.064	1.015	1.017	0.00
	λ_2	0.15	0.136	0.365	-0.014	0.364	0.88	0.846	0.076	0.696	0.700	0.00
	λ_3	0.30	0.192	0.300	-0.108	0.318	0.89	0.598	0.135	0.298	0.327	0.66
	ω_1	0.45	0.451	0.020	0.001	0.020	0.98	0.452	0.022	0.002	0.022	0.97
	ω_2	0.35	0.347	0.022	-0.003	0.022	0.98	0.347	0.023	-0.003	0.023	1.00
	ω_3	0.20	0.202	0.019	0.002	0.019	0.97	0.201	0.019	0.001	0.019	0.97
	β_{11}	-0.50	-0.488	0.100	0.012	0.100	0.95	-0.498	0.031	0.002	0.031	0.96
	β_{12}	0.50	0.489	0.104	-0.011	0.104	0.93	0.500	0.030	0.000	0.030	0.96
	β_{13}	-1.00	-1.000	0.032	0.000	0.032	0.97	-0.999	0.034	0.001	0.034	0.97
	β_{21}	-0.75	-0.735	0.158	0.015	0.158	0.92	-0.749	0.029	0.001	0.029	0.94
	β_{22}	0.80	0.782	0.157	-0.018	0.158	0.92	0.797	0.032	-0.002	0.032	0.93
	β_{23}	1.20	1.206	0.034	0.006	0.034	0.95	1.204	0.037	0.004	0.037	0.95
	κ_1	0.80	0.691	0.340	-0.109	0.356	0.88	–	–	–	–	–
	κ_2	0.60	0.565	0.261	-0.035	0.262	0.88	–	–	–	–	–
	κ_3	0.25	0.269	0.187	0.019	0.187	0.89	–	–	–	–	–
	γ	1.50	1.144	0.005	-0.356	0.356	0.00	–	–	–	–	–

Notes: This table displays results based on $R = 100$ replicates. The values include the average and standard deviation of the point estimates; the average bias (Bias), the Root Mean Squared Error (RMSE), and the coverage rate (Coverage) across replicates; where Bias $= R^{-1}\sum_{r=1}^{R}(\hat{\alpha}_r - \alpha)$, RMSE $= \sqrt{R^{-1}\sum_{r=1}^{R}(\hat{\alpha}_r - \alpha)^2}$, and Coverage $= R^{-1}\sum_{r=1}^{R}\mathbb{1}\{\alpha \in \widehat{CI}_{0.95,r}\}$.

4.5 Empirical Analysis

In this section, we apply the proposed methodology in the context of market-for-technology network formation and corporate research and development (R&D) investments in the United States.

4.5.1 Data Summary

We first construct a firm-level panel dataset for the period 1980–2014 using multiple sources, which combines accounting data from US Compustat, patent trades between firms from the USPTO Patent Assignment Dataset (PAD), and R&D tax credit information. More specifically, we utilise the Link Compustat – USPTO Patent Assignment Dataset (PAD)[12], shared publicly by Arqué-Castells and Spulber (2022). These authors effectively matched assignor/assignee names in the PAD to Compustat GVKEYs, almost directly producing a match between Compustat and patent transactions that took place from 1980 to 2014. Information on companies corresponding to these GVKEYs and their annual balance sheets from the S&P North America Annual Compustat is available through Wharton Research Data Services (WRDS). Federal tax information required to build the tax price of R&D is also acquired from Arqué-Castells and Spulber (2022) and linked to firms by year[13]. The broadest possible sample resulting from merging all the data sources includes $3,896$ Compustat firms that interact in the market for technology with at least one other firm in the sample and for which the deal has a known execution date.

To create a firm-level technology-collaboration network of interest, which is essential for our analysis, we apply further filters. We keep innovating firms that have both adopter and provider roles and invest in R&D during the sample periods in which they are available, with non-missing information on relevant variables such as sales, capital, and employment. We create undirected links among firms, represented by the binary variable w_{ij} which indicates the presence or absence of at least one transaction in the technology market between firms i and j. To maintain a manageable sample size, we exploit only the cross-sectional variation between firm dyads and treat the network as time-invariant network. This approach allows for a more stable analysis of relationships between firms during the study period. Firms that

[12]Data source: https://zenodo.org/record/6352358. DISCERN2 provides updated data covering 1980-2021.

[13]To calculate the federal component of the firm-specific tax price of R&D, they take advantage of the dataset produced by Wilson (2009), who calculates the user cost of R&D faced by a representative firm conducting R&D within a given state.

are not connected to the network are excluded to ensure that all entities in the final sample have meaningful technological collaborations. The resulting dataset comprises $1,150$ firms with a total of $5,576$ links, providing a substantial network for examining the formation of technology collaboration and its impact on firm performance and innovation. Figure 4.1 illustrates the network of interactions in the market for technology. Groups 2 (computers and communications) and 3 (drugs and medical) in the National Bureau of Economic Research Patent Data Project (NBER PDP) six-group aggregation of technology fields (see Hall et al., 2001) – on the left and right extremes of the network, respectively, dominate the market for technology. In between, there is a gradation of smaller technology clusters: chemical (Group 1), electrical and electronic (Group 4), mechanical (Group 5), and others (Group 6). The fact that firms heavily cluster by technology field suggests that technological proximity plays an important role in shaping the market-for-technology network.

Figure 4.1: Network of interactions by technological fields. This figure includes 1,150 Compustat firms that play both roles as adopters and providers in the market-for-technology network (patent transactions). All these firms are connected to the network through 5,576 links that accumulated from 1980 to 2014. Each node represents a firm, with node size being proportional to the number of connections of the firm). Nodes are arranged following the Fruchterman-Reingold force-directed algorithm, and coloured by technology class (six groups in the NBER PDP classification: 1 (chemical), 2 (computers and communications), 3 (drugs and medical), 4 (electrical and electronic), 5 (mechanical), and 6 (others). The main technology class of a firm is defined as the modal class of the patents of the firm.

The primary firm outcome of interest in our application is the corporate R&D efforts,

measured by the R&D expenditure-to-sales ratio, often referred to as R&D intensity. The natural logarithm of R&D intensity among firms, as depicted in Figure 4.2, reveals a notable multimodal distribution. Table 4.5 presents the definitions and descriptive statistics for essential variables in our final sample, including various firm-level characteristics and the federal component of the firm-specific tax price of R&D. The latter represents a supply-side shock to corporate R&D induced by government policy.

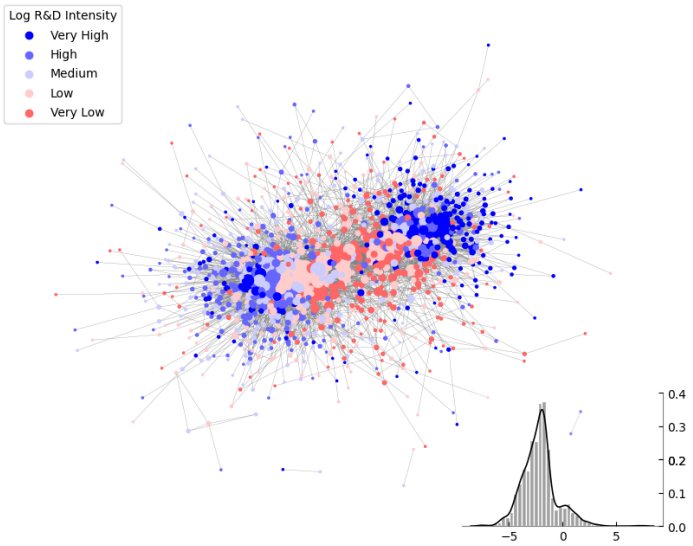

Figure 4.2: Distribution of Research & Development intensity among firms (post transformation) exhibits visible multimodality.

Table 4.5: Data Statistics and Definitions ($N = 1,150$)

Variable	Mean	Med	Std	Min	Max	Description
RDintensity	3.075	0.113	63.804	0.000	1939.069	R&D intensity
RDexpense	252.891	31.579	764.907	0.095	7684.677	Annual R&D investment
Sales	5381.282	304.322	18627.362	0.005	230887.253	Sales
Capital	321.358	13.515	1333.668	0.007	15503.994	Capital expenditure
Employment	14.606	1.244	41.260	0.004	526.483	Employment
EBIT	468.475	16.656	1663.241	-380.792	21915.117	Earnings before interest and taxes
Revenue	4505.899	256.781	15820.807	0.004	197726.303	Revenue
Size	5616.906	359.009	21191.756	0.266	380628.074	Total assets
TaxPrice	0.910	0.910	0.040	0.820	1.019	Federal component of R&D tax price

Notes: This table presents summary statistics for the firm-level dataset, which includes $1,150$ firms in total with no missing data. The values of RDexpense, Sales, Capital, Revenue, and Size are in millions of 2010 dollars.

4.5.2 Model Specification

We employ the general SCHSAR framework to jointly model the formation of a firm collaboration network and their R&D efforts in two sequential stages as follows.

Network Formation: *Endogenous Formation of Technological Linkages*

In the first stage, the network among N firms is formed with links determined by

$$w_{ij} = w_{ji} = \mathbb{1}\{w_{ij}^* \geq 0\} \quad \text{for } i = 1, \ldots, N; j = 1, \ldots, i-1; \tag{4.64}$$

where w_{ij} indicates whether firms i and j have at least one transaction in the market of technology, and w_{ij}^* measures the utility that each of the firms i and j gains from forming a link

$$w_{ij}^* = \gamma_0 + \gamma_1 \cdot \text{sameSIC}_{ij} + \gamma_2 \cdot \text{sameAAclass}_{ij} + a_i + a_j + \epsilon_{ij}, \text{ where } \epsilon_{ij} \overset{iid}{\sim} \mathcal{N}(0,1). \tag{4.65}$$

In particular, the utility depends on predetermined dyad-specific regressors that include: Homophily measures such as whether the two firms are in the same industry (sameSIC_{ij}), whether they are in the same technology class (sameAAclass_{ij}), etc. γ_0 is the fixed cost of maintaining links. a_i and a_j represent firm-specific unobserved degree heterogeneity which is concurrent with their ability to create linkages.

Network Interaction: *Potential Heterogeneous Effects*

Once the network is formed in the first stage, in the second stage, firms choose their actions (e.g., R&D efforts or other firm outcomes) taking the network structure as given. To maximise their quadratic payoff function, a firm's activity intensity follows a best-response function accounting for the choices of the others. Denote Y_i as firm i's logged R&D intensity, which is measured by the natural logarithm of R&D expenditure-Sales ratio. Let X_i^k be the k-th observed firm-specific characteristic. The degree heterogeneity a_i can also affect Y_i through the unobserved part, making the $N \times N$ network adjacency matrix $\mathbf{W} = [w_{ij}]$ potentially endogenous.

In accordance with the SCHSAR framework, the outcome equation for firm i can be

expressed as

$$Y_i = \lambda_g \sum_{j=1}^{N} w_{ij} Y_j + \sum_{k=1}^{K} X_i^k \beta_g^k + \kappa_g a_i + u_i; \quad u_i \sim \mathcal{N}(0, \sigma_{u,g}^2) \tag{4.66}$$

for i belongs to cluster g ($g \in \{1, \ldots, G\}$). $\{\lambda_g\}_{g=1}^G$ represents the network interaction effect and $\{\beta_g^k\}_{g=1}^G$ captures the own influence of the k-th observed firm-specific characteristic. This finite mixture structure flexibly allows for heterogeneity as it nests both the standard SAR ($G = 1$) and the heterogeneous coefficients SAR ($G = N$) specification as special cases. Furthermore, this specification explicitly incorporates $\kappa_g a_i$ as a control function to model the endogeneity between network formation and firm outcomes, thereby correcting the selection bias in the estimation of endogenous network effect λ_g. Using similar notations and derivations to Section 4.2, we can obtain the reduced form below

$$\mathbf{Y} = [\mathbf{I}_N - \mathcal{L}\mathbf{W}]^{-1} \left(\sum_{k=1}^{K} \tilde{\beta}^k \circ X^k + \tilde{\kappa} \circ \mathbf{a} + \mathbf{u} \right). \tag{4.67}$$

This specification not only enables the heterogeneous network interaction effects (λ), but also leads to a much richer interpretation of the effects of explanatory variables on the outcome. Specifically, the marginal effects of a change in the k-th variable vector, $X^k = \left[X_1^k, \ldots, X_N^k\right]^\top$, are given by the following matrix of partial derivatives

$$\begin{aligned}
\frac{\partial Y}{\partial X^{k\top}} &= \begin{bmatrix} \frac{\partial Y_1}{\partial X_1^k} & \frac{\partial Y_1}{\partial X_2^k} & \cdots & \frac{\partial Y_1}{\partial X_N^k} \\ \frac{\partial Y_2}{\partial X_1^k} & \frac{\partial Y_2}{\partial X_2^k} & \cdots & \frac{\partial Y_2}{\partial X_N^k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial Y_N}{\partial X_1^k} & \frac{\partial Y_N}{\partial X_2^k} & \cdots & \frac{\partial Y_N}{\partial X_N^k} \end{bmatrix} \\
&= (\mathbf{I}_N - \mathcal{L}\mathbf{W})^{-1} \times \operatorname{diag}\left(\tilde{\beta}^k\right) \\
&= \left(\mathbf{I}_N - \begin{bmatrix} \tilde{\lambda}_1 & 0 & \cdots & 0 \\ 0 & \tilde{\lambda}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \tilde{\lambda}_N \end{bmatrix} \mathbf{W} \right)^{-1} \times \begin{bmatrix} \tilde{\beta}_1^k & 0 & \cdots & 0 \\ 0 & \tilde{\beta}_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \tilde{\beta}_N^k \end{bmatrix} = [\ell_{ij}]_{i,j=1}^N.
\end{aligned} \tag{4.68}$$

The direct effects show how changes in a firm's own k-th characteristic influence its own outcomes. Conversely, spillin effects represent the cumulative impact of changes in the k-th characteristic of peer firms on a firm's outcomes, while spillout effects demonstrate how

changes in a firm's k-th characteristic affect peers' outcomes. The latter two effects are also known as indirect effects.

We note that, $G = 1$ results in homogeneous models, allowing us to derive standard summaries of the effects at the aggregate level. As defined by LeSage and Pace (2009), the direct effect (DE) is quantified by averaging the diagonal elements of the matrix of partial derivatives, and the indirect effect (IE) is determined by averaging the cumulative sums of off-diagonal row or column-elements. The total effect is simply given by: $TE = DE + IE$.

When $G > 1$, the models become heterogeneous because of the mixture structure. Unlike homogeneous models where scalar summary measures are used, in heterogeneous models, these effects of interest are reported at the observation level for each firm in the sample to capture the parameter heterogeneity. In particular, because of the heterogeneity of λ and β^k, the N-dimensional vectors of firm-level spillin and spillout effects are not equal even when using a doubly-stochastic weight matrix, necessitating the calculation of both types of indirect effects. In particular, they are computed as the cumulative sum of off-diagonal elements in each row and in each column of the matrix of the partial derivatives, respectively. Furthermore, both quantities depend not only on the individual group type but also on the network position. From spatial perspective, a detailed discussion on the interpretation of such heterogeneous models can be found in LeSage and Chih (2016) and Cornwall (2017), which pertains to our proposed framework.

4.5.3 Estimation Results

We estimate the model parameters using the MCMC procedure described in Section 4.3. We run the MCMC algorithm for $50,491$ iterations and drop the first 500 draws for burnin and keep every 10th of the remaining draws to conduct the posterior analysis, that is, we compute the posterior mean (as a point estimate) and posterior variance for each parameter.

4.5.3.1 Homogeneous Models

We begin with homogeneous models for comparison. That is the case when the number of mixture components in the outcome equation reduces to $G = 1$, implying homogeneous network interaction effect.

Table 4.6: Parameter Estimates for the HSAR and SCHSAR Models (when $G = 1$)

		SAR		SCSAR	
		Mean (std)	CI95	Mean (std)	CI95
Network Interaction					
Interaction Effect	λ	0.190 (0.017)	[0.156, 0.223]	0.157 (0.016)	[0.125, 0.189]
Intercept	β_1	1.236 (0.196)	[0.861, 1.623]	2.664 (0.221)	[2.228, 3.098]
logTaxPrice	β_2	-8.332 (0.610)	[-9.527, -7.147]	-8.063 (0.581)	[-9.210, -6.915]
logCapitalExpense	β_3	0.616 (0.038)	[0.543, 0.691]	0.540 (0.036)	[0.469, 0.611]
EBIT	β_4	0.071 (0.018)	[0.035, 0.108]	0.042 (0.018)	[0.008, 0.075]
logEmployment	β_5	-0.025 (0.049)	[-0.120, 0.071]	-0.073 (0.045)	[-0.162, 0.015]
logRevenue	β_6	-0.983 (0.040)	[-1.061, -0.905]	-0.962 (0.038)	[-1.038, -0.887]
Correlation	κ	–	–	0.802 (0.067)	[0.671, 0.933]
Unobserved Heterogeneity	σ_a^2	–	–	3.092 (0.130)	[2.841, 3.358]
Network Formation					
SIC homophily	γ_1	–	–	0.722 (0.017)	[0.689, 0.755]
Tech homophily	γ_2	–	–	0.691 (0.010)	[0.671, 0.711]
Criteria					
Log likelihood		-1484.935 (2.023)		-45471.583 (7812.020)	
AICM		2978.054 (0.279)		122146268.899 (2862451.223)	
Observations		1150		1150	

Notes: This table presents the estimation results for the homogeneous SAR and SCSAR models. MCMC sampling runs a total of $50,491$ iterations, where the first 500 iterations are discarded as burn-in and every 10th draw is retained, yielding $5,000$ effective draws. Posterior means, standard deviations, and 95% equal-tailed intervals (ETI) are computed using MCMC draws.

Table 4.6 reports the estimation results of the homogeneous SAR and SCSAR models. Parameter estimates for the standard SAR specification, presented in the first panel, suggests that a majority of firm-specific characteristics are significant determinants of private R&D efforts. Furthermore, the network interaction effect (λ) is positive and statistically significant, confirming the theoretical conjecture that corporate R&D efforts are positively influenced by their collaborators. Concerning the endogenous adjacent matrix inherent in this benchmark model, the SCSAR model in the second panel takes into account the network formation process among the firms. It can be seen that the homophily indeed matters for firm collaboration in the market for technology, where firms are more likely to form links with those from the same industry or technology cluster. Controlling for firms' unobserved heterogeneity in both network formation and economic outcome corrects for selection bias. In particular, the network interaction effect (λ) decreases considerably from 0.190 to 0.156 when introducing

unobserved heterogeneity. This positive effect remains statistically significant, as evidenced by the corresponding 95% credible intervals.

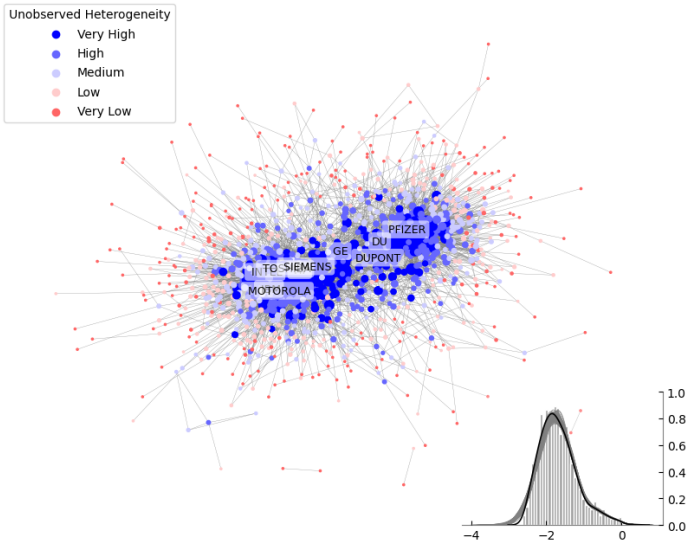

Figure 4.3: Posterior mean of the firms' unobserved degree heterogeneity.

It is worth exploring deeper insight into the role of firms' unobserved heterogeneity, a_i, in the SCSAR model. This serves as the random effect of each firm $i = 1, \ldots, N$ on both network formation and interaction stages. Figure 4.3 presents the posterior summary of $\{a_i\}_{i=1}^{N}$. The distribution of posterior mean estimates is depicted in the histogram and kernel density, characterising latent heterogeneity across firms. In the firm network graph, each node i is coloured according to the associated posterior mean of a_i. Clearly, firms who possess unobserved degree heterogeneity with higher values (darker colours) tend to gain more connections in the network, scattering surround the central area. This finding is also validated numerically in Table C.5 and Figure C.33, which shows strong positive correlation between the a_i and common network centrality measures including degree, betweenness, closeness, and eigenvector centrality. More importantly, a_i is also positively correlated with R&D efforts, as captured by a significant estimate of its coefficient (κ) in the R&D intensity equation. Put differently, unmeasured confounding exists in the form of firm-specific latent advantages that make firms become popular and exert greater R&D efforts simultaneously. To avoid inaccurate inferences, it is essential to adequately model and estimate unobserved heterogeneity that could affect both network formation and outcome.

Table 4.7: Direct, Indirect and Total Effects from Explanatory Variables

	SAR			**SCSAR**		
	Direct	Indirect	Total	Direct	Indirect	Total
logTaxPrice	-8.564	-1.723	-10.287	-8.216	-1.354	-9.570
	(0.621)	(0.197)	(0.749)	(0.589)	(0.171)	(0.693)
	[-9.79, -7.35]	[-2.13, -1.35]	[-11.73, -8.84]	[-9.37, -7.05]	[-1.71, -1.03]	[-10.92, -8.19]
logCapitalExpense	0.633	0.127	0.760	0.550	0.091	0.641
	(0.039)	(0.014)	(0.047)	(0.037)	(0.011)	(0.044)
	[0.56, 0.71]	[0.10, 0.16]	[0.67, 0.86]	[0.48, 0.62]	[0.07, 0.11]	[0.56, 0.73]
EBIT	0.073	0.015	0.088	0.042	0.007	0.049
	(0.019)	(0.004)	(0.023)	(0.018)	(0.003)	(0.021)
	[0.04, 0.11]	[0.01, 0.02]	[0.04, 0.13]	[0.01, 0.08]	[0.00, 0.01]	[0.01, 0.09]
logEmployment	-0.026	-0.005	-0.031	-0.074	-0.012	-0.086
	(0.050)	(0.010)	(0.060)	(0.046)	(0.008)	(0.054)
	[-0.12, 0.07]	[-0.03, 0.01]	[-0.15, 0.09]	[-0.16, 0.02]	[-0.03, 0.00]	[-0.19, 0.02]
logRevenue	-1.011	-0.203	-1.214	-0.980	-0.162	-1.142
	(0.041)	(0.021)	(0.054)	(0.039)	(0.019)	(0.049)
	[-1.09, -0.93]	[-0.25, -0.16]	[-1.32, -1.11]	[-1.06, -0.91]	[-0.20, -0.13]	[-1.24, -1.05]

Notes: This table presents the posterior means, standard deviations, and 95% credible intervals for each effect across firms. The derivation of each effect is based on posterior samples of each parameter.

In addition to the change in the estimated network interaction effect parameter λ, we can further examine how other estimated coefficients of the explanatory variables and their interactions with λ are adjusted by the selection-corrected approach. Owing to the structure of the SAR and SCSAR models, a change of any given explanatory variable in an individual will affect not only the dependent variable of its own (direct effect), but also the dependent variables of the others (indirect effects). Table 4.7 summarizes the average of direct, indirect and total effects for the explanatory variables derived from the estimates of the SAR and SCSAR models. Because a firm's R&D intensity is positively influenced by its peers' R&D efforts through network interaction effects, the government R&D tax price for a specific firm directly affects the firm itself and its peer firms indirectly. On average, both direct and indirect effects are significantly negative, contributing to the average total impact of 1% increase in the tax price the $i-$th firm receives on firms' R&D intensity is significantly negative, with 10.29% and 9.57% declines in the SAR and SCSAR models, respectively. In other words, the peer effect on corporate R&D policies generates a multiplier, defined as the ratio of the total effect to the direct effect. By accounting for the endogenous network structure, the SCSAR model avoids overestimating λ, thereby providing robust evidence for the multiplier although

the magnitude of the ratio is slightly reduced compared to the SAR model.

4.5.3.2 Heterogeneous Models

Table 4.8 presents estimation results of the HSAR and SCHSAR models with two component distributions (i.e., $G = 2$). Under a three-component model (i.e., $G = 3$), the parameter estimates for the third type, including the assignment probability and network effect parameter, are not significantly distinguishable from zero (see Table C.6).

Comparing the SCHSAR model with the homogeneous SCSAR, the estimates for parameters in the network formation equation are almost preserved up to two decimals, as well as those for the random effect variance. Visualising firms' unobserved degree heterogeneity $\{a_i\}_{i=1}^N$ and computing relevant statistics also give us similar pattern as in the previous case. This aligns with our modelling approach when the first stage (network formation) is specified for the entire population and does not depend on G. In terms of the network interaction equation, the correlation coefficient κ remains positive in both types and is significant in the first type. The two groups are ordered based on their weights to the sample population, with the smaller group recording larger network interaction effect ($\hat{\lambda}_2 \approx 1.7\hat{\lambda}_1$) and lower regression coefficient of logTaxPrice ($\hat{\beta}_{22} \approx 4.28\hat{\beta}_{21}$). This interesting distinction indicates that some firms are highly responsive to their peer R&D activities but less susceptible to direct R&D tax price intervention from the government, and vice versa. Perhaps the latter type includes more "self-reliant" innovators (and network peer effects are less important to these firms), while firms of the former type are more influenced by what others are doing (and hence less directly responsive to their own tax changes). For simplicity, we refer to these groups as high-λ (more peer-driven) and low-λ (more self-driven) types. The group labels themselves have no direct interpretation[14], but some attributes of each group provide further insight into the underlying mechanisms driving the heterogeneity.

[14] See Geweke (2007) and Frühwirth-Schnatter (2006) for full discussion.

Table 4.8: Parameter Estimates for the HSAR and SCHSAR Models (when $G = 2$)

		HSAR		SCHSAR	
		1st Type	2nd Type	1st Type	2nd Type
Network Interaction					
Assignment Probability	π	0.588 (0.043) [0.50, 0.67]	0.412 (0.043) [0.33, 0.50]	0.660 (0.041) [0.58, 0.74]	0.340 (0.041) [0.26, 0.42]
Interaction Effect	λ	0.150 (0.026) [0.10, 0.20]	0.226 (0.035) [0.16, 0.29]	0.127 (0.024) [0.08, 0.17]	0.215 (0.038) [0.14, 0.29]
Intercept	β_1	1.315 (0.277) [0.77, 1.86]	-3.171 (0.464) [-4.08, -2.23]	3.153 (0.340) [2.54, 3.79]	-3.077 (0.632) [-4.23, -1.87]
logTaxPrice	β_2	-10.637 (1.035) [-12.66, -8.65]	-2.428 (0.853) [-4.08, -0.77]	-9.525 (0.930) [-11.28, -7.76]	-2.223 (1.041) [-4.23, -0.29]
logCapitalExpense	β_3	0.685 (0.058) [0.57, 0.80]	0.292 (0.069) [0.16, 0.43]	0.589 (0.054) [0.48, 0.69]	0.263 (0.076) [0.12, 0.42]
EBIT	β_4	0.059 (0.025) [0.01, 0.11]	0.052 (0.027) [-0.00, 0.11]	0.045 (0.021) [0.00, 0.09]	0.058 (0.037) [-0.01, 0.13]
logEmployment	β_5	-0.102 (0.081) [-0.27, 0.05]	-0.617 (0.084) [-0.79, -0.45]	-0.086 (0.070) [-0.22, 0.05]	-0.639 (0.097) [-0.82, -0.45]
logRevenue	β_6	-1.099 (0.051) [-1.20, -1.00]	0.109 (0.094) [-0.08, 0.29]	-1.080 (0.055) [-1.17, -0.99]	0.135 (0.112) [-0.07, 0.33]
Correlation	κ	–	–	0.933 (0.108) [0.72, 1.15]	0.083 (0.131) [-0.17, 0.34]
Unobserved Heterogeneity	σ_a^2	–	–	3.093 (0.131) [2.85, 3.37]	
Network Formation					
SIC homophily	γ_1	–	–	0.724 (0.017) [0.69, 0.76]	
Tech homophily	γ_2	–	–	0.692 (0.010) [0.67, 0.71]	
Criteria					
Log likelihood		-1958.68 (20.94)		-44949.63 (9729.57)	
AICM		4794.66 (20.68)		189418920.84 (4440159.24)	
Observations		1150		1150	

Notes: This table presents the estimation results for the HSAR and SCHSAR models with $G = 2$. MCMC sampling runs a total of $50,491$ iterations, where the first 500 iterations discarded as burn-in and every 10th draw is retained, yielding $5,000$ effective draws. Posterior means, standard deviations, and 95% equal-tailed intervals (ETI) are computed using these MCMC draws.

In Table 4.9, we present regression results of the probability that firm i belongs to the high-λ type on multiple firm characteristics. For the linear specification, we use a

continuous dependent variable obtained from the posterior means of z_i in (4.52). For the logit specification, we use the binary type indicators derived from posterior median of of z_i. The findings consistently suggest that the larger firm size (measured by total assets), the higher the probability that it belongs to the high-λ type. Technological class also plays a role in determining the latent types.

Table 4.9: Regressions of Posterior Probability of Inclusion on Firm Characteristics

Dependent Variables (Regresion)	Probability (Linear)		Binary (Logistic)	
	(1)	(2)	(3)	(4)
logTotalAsset	0.016***	0.013***	0.150***	0.129***
	(0.003)	(0.003)	(0.033)	(0.040)
Computers & Communications	0.210***	0.205***	1.109***	1.077***
	(0.021)	(0.021)	(0.273)	(0.275)
Drugs & Medical	0.065***	0.059***	0.682**	0.643**
	(0.022)	(0.022)	(0.292)	(0.295)
Electrical & Electronic	0.191***	0.189***	1.177***	1.156***
	(0.025)	(0.025)	(0.303)	(0.304)
Mechanical	0.135***	0.140***	0.862**	0.887***
	(0.029)	(0.029)	(0.340)	(0.341)
Others	0.035	0.039	0.506	0.527
	(0.030)	(0.030)	(0.372)	(0.373)
Degree Centrality		0.981*		4.751
		(0.501)		(5.061)
Intercept	0.115***	0.132***	-2.955***	-2.843***
	(0.027)	(0.028)	(0.350)	(0.367)
Observations	1150	1150	1150	1150
Adjusted R^2	0.138	0.140		
AIC			1228.424	1229.540

Notes: Standard errors are reported in parentheses. Coefficients marked with *, **, and *** are significant at the 10%, 5%, and 1% levels, respectively.

Turning to the response of corporate R&D efforts in the firm-level network to government tax incentives for R&D, as mentioned earlier in the case of heterogeneous model, our focus of inference are the direct and indirect (spillin and spillout) effect estimates at individual/observational levels (i.e., one value for each firm). For $i = 1, \ldots, N$, the direct effect shows the elasticity response of i-th firm's R&D intensity to its own R&D tax price, given the fact that we employ the logarithm transformation for both quantities. The indirect

spillin effect estimate represents the cumulative impact of changes in peer firms' R&D tax prices on i-th firm's R&D intensity, whereas the indirect spillout effect estimate represents the cumulative impact of changes in R&D tax prices of i-th firm on its peers' R&D intensity. These quantities provide fine-grained heterogeneity across firms, thus enabling a more effective targeted policy.

In a policy scenario in which the government seeks to encourage corporate R&D investments by lowering the costs incurred by firms, we assess the effects of a 1% reduction in firm-specific R&D tax prices using estimation results from the SCHSAR model. To summarize the observational-level effects of interest succinctly, we employ a combination of coloured graphs and histograms to visualise these quantities. Specifically, the histograms depict the empirical frequency of the posterior mean estimates across $1,150$ firms. We then categorize these values into five quintiles and map corresponding colours to the firms in the network graph. Additionally, we rank the firms and highlight those with the highest effects, along with several of their characteristics. The results reveal a rich pattern of firm-level heterogeneity in the estimates. Overall, all firms display values significantly different from zero, as indicated by 95% credible intervals derived from the MCMC draws.

The direct effect estimates presented in Figure 4.4 and Table 4.10 align with the previous analysis, revealing a bimodal distribution that supports the predictions of the two-component mixture SCHSAR model. This bimodality is indicative of mixture-driven heterogeneity rather than sampling noise, with most values being significantly positive and falling between 2 and 10. The distribution suggests two distinct firm types with varying levels of tax price elasticity and dependence on peer network effects for innovation output. More specifically, firms exhibiting the highest direct effects predominantly belong to the low-λ type which has high tax price elasticity (large β_2). This implies that these firms are less reliant on peer network effects and more responsive to their own R&D incentives. Notably, firms in the Drugs & Medical and Chemical classes are heavily represented in this group. Interestingly, the analysis reveals that a firm's centrality or size in the network negatively correlate with its direct effects. For instance, despite its low size percentile, Cellcy Pharmaceuticals exhibits very high direct effects, highlighting that even smaller or less-connected firms can demonstrate substantial R&D responsiveness to tax incentives.

Figure 4.5 and Table 4.11 represent the indirect spillin effects across firms due to a 1% reduction in peers' R&D tax prices. The density plot reveals a less polarized distribution

centred around 1.0-1.3, indicating that most firms experience positive spillin effects, albeit at a lower magnitude than direct effects. This suggests that while peer effects exist and firms benefit from their peers' R&D becoming cheaper, these indirect effects are generally smaller than the firm-specific responses. Large standard deviations relative to the effect size indicate greater estimation uncertainty. This is consistent with the fact that these effects depend on the matrix inverse $(\mathbf{I}_N - \mathcal{L}\mathbf{W})^{-1}$, which propagate uncertainty from both λ and $\mathbf{W}$. The empirical findings illustrate that a firm's network position and network interaction elasticity (λ) are pivotal in determining its responsiveness to peer-based incentives. Firms exhibiting high spillin effects predominantly belong to high-λ type, emphasising the importance of network interaction effects in shaping spillin patterns. From a theoretical standpoint, we refer to these firms as "responsive absorbers" to distinguish them from "self-reliant innovators," who exhibit strong direct effects. In contrast to self-reliant innovators, responsive absorbers are more reliant on external innovations, such as partnerships, and consequently display greater peer dependency and weaker direct responses. Our data indicates that these top-ranked firms span various technology classes, including Chemical, Drugs & Medical, Mechanical, and Computers & Communications, indicating a broader mix compared to the direct effects. Additionally, larger firm size and degree centrality are slightly linked to higher spillin effects. Nonetheless, firms that are not highly central can still be significantly exposed to peer innovation, as demonstrated by firms such as Fmc Corp and Seclone Pharmaceuticals, which exhibit high spillins despite relatively low centrality. The top responsive absorbers, highlighted in blue nodes, are also dispersed throughout the network graph rather than being heavily clustered.

Figure 4.6 and Table 4.12 illustrate the indirect spillout effects from each firm to its peers resulting from a 1% reduction in the firm's own R&D tax price. When firms experience a 1% decrease in their R&D tax prices, their peers' R&D intensities increase significantly, as evidenced by all 95% intervals lying entirely above zero. However, these indirect effects have wider credible intervals than direct effects, indicating greater uncertainty in their estimation. While top spillin firms often belong to the high-λ type, most spillout leaders are classified in the low-λ type. These low-λ firms are more sensitive to direct firm-level R&D incentives (large β_2), confirming that their own influence prevails even when their connected peers are less elastic. Beyond network interaction effects λ, the calculation for spillouts involve firms' own price elasticity and network position, which enhance firms' ability to transmit cost shocks outward through collaboration and co-patenting links. Top spillout firms (in

blue) are prominently centralised in the network graph, highlighting their potential as central hubs capable of significantly affecting broader network activities. In accordance with theory, we refer to them as "influential transmitters," who can magnify strategic complementarities (Bulow et al., 1985; Cooper and John, 1988) by virtue of their extensive linkages, serving as critical "gatekeepers" or "bridges" (Katz and Shapiro, 1986) and transferring technological shocks to many connected firms simultaneously. Firm size complements network centrality in amplifying spillout effects; nearly all top spillout firms are in the 90^{th}–100^{th} percentile of total asset. This joint condition – being large and central – defines the profile of influential transmitters in the network. Additionally, sector concentration is evident, with the Computers & Communications sector accounting for 13 out of the top 20 spillout firms, indicating that technology-intensive classes drive network-wide R&D diffusion.

The empirical patterns observed across all three sets of results provide strong support for the SCHSAR framework, demonstrating its ability to capture the intended network mechanisms. Although direct incentives exhibit stronger effects, the significance of indirect effects is noteworthy. These indirect effects are inherently more complex because of their dependence on heterogeneous network interaction effect λ and network structure $\mathbf{W}$, resulting in larger standard deviations and wider credible intervals than direct effects. Furthermore, our empirical findings distinguish firms that rank high in either indirect spillin or spillout effects, characterising responsive absorbers and influential transmitters based on their specific values of λ, β, and $\mathbf{W}$. This observation aligns with the technical distinction made in the SCHSAR framework, highlighting its flexibility and robustness in modelling complex network interactions.

Taken together, these results underscore the importance of considering both firm-level direct and indirect effects when evaluating the efficacy of government R&D tax incentives. Ignoring the indirect effects inherent in the collaboration network can lead to a significant underestimation of a policy's true impact. The rich heterogeneity observed across firms reveals various R&D behaviours within the market-for-technology network, enabling the identification of self-reliant innovators, responsive absorbers, and influential transmitters, among others. Measuring firm-level total effects – the sums of direct and indirect spillin or spillout effects above – offers the most comprehensive basis for designing policies that generate balanced and synergistic innovation outcomes.

The total spillin effect captures the cumulative impact of a 1% reduction in the R&D

tax price of every firm in the network on a given firm's R&D intensity. This measure is especially relevant under a uniform tax policy in which all firms receive the same marginal incentive from the government. This fully reflects both the firm's own direct responsiveness and the accumulated influence of peer firms' incentive-driven R&D responses. Firms with the highest (or lowest) total spillin effects are therefore best (or least) positioned to benefit from such broad-based policy interventions. The empirical evidence presented in Figure 4.7 and Table 4.13 suggest that own-price elasticity (β_2) plays a more dominant role than network elasticity (λ) in determining top beneficiaries. In other words, firms that rank highest on total spillin effect are often self-reliant innovators with strong internal responsiveness, rather than responsive absorbers primarily driven by peer R&D behaviours. Many of these firms are found in the Drugs & Medical and Chemical sectors.

The total spillout effect, on the contrary, captures the cumulative increase in network-wide R&D intensity resulting from a 1% reduction in a single firm's R&D tax price. This measure is the most informative for designing cost-effective, high-impact policies, especially when the government operates under a constrained budget and must selectively target a subset of firms. It combines both the firm's own R&D response and its downstream influence on peers through the network. High-spillout firms act as policy multipliers; thus, subsidising them can propagate innovation widely across the network at a relatively low cost. Figure 4.8 and Table 4.14 confirm that the top total spillout firms tend to be innovation initiators and influencers, combining strong direct effects with extensive network reach. They are typically highly central in the firm network and are often found in the Computers & Communications, Chemical, or Drugs & Medical sectors. While some are large, capitalized industrial leaders (e.g., GE, Apple, P&G), interestingly, well-connected but much smaller firms also appear on the list. Conditional on sector and centrality, smaller firmsize tends to generate higher total spillout effects. To maximise systemic diffusion from limited resources, policymakers should strategically prioritize firms with high total spillout effects.

Direct Effects for each firm due to a 1% reduction in the firm's own R&D tax price

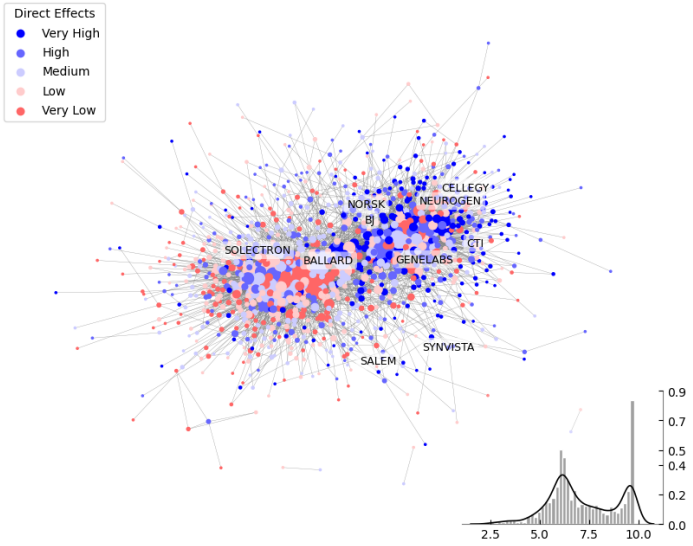

Figure 4.4: Direct effects of a 1% reduction in a firm's own R&D tax price. The histogram shows the distribution of the effects of interest across firms. In the network graph, firms are represented as nodes colored by the effect magnitude.

Table 4.10: The Top 20 Firms with the Highest Direct Effects

						Direct Effects	
Rank	Firm	Field	Centrality	Size	Type	Mean (std)	CI95
1	Norsk Hydro Asa	Chemical	57	93	Low λ	9.770 (0.941)	[7.950, 11.571]
2	Solectron Corp	Computers & Communications	80	82	Low λ	9.744 (0.927)	[7.930, 11.544]
3	Cellegy Pharmaceuticals -Old	Drugs & Medical	37	5	Low λ	9.743 (0.925)	[7.949, 11.535]
4	Synvista Therapeutics Inc	Drugs & Medical	37	12	Low λ	9.742 (0.927)	[7.938, 11.568]
5	Ballard Power Systems Inc	Electrical & Electronic	68	52	Low λ	9.741 (0.979)	[7.926, 11.548]
6	Neurogen Corp	Drugs & Medical	57	26	Low λ	9.741 (0.929)	[7.946, 11.543]
7	Cti Biopharma Corp	Drugs & Medical	48	30	Low λ	9.737 (0.923)	[7.946, 11.571]
8	Bj Services Co	Others	37	76	Low λ	9.737 (0.952)	[7.919, 11.572]
9	Salem Corp	Others	8	21	Low λ	9.730 (0.923)	[7.925, 11.553]
10	Genelabs Technologies Inc	Drugs & Medical	78	11	Low λ	9.730 (0.922)	[7.942, 11.536]
11	Linde Plc	Others	83	89	Low λ	9.730 (0.930)	[7.931, 11.538]
12	Windtree Therapeutics Inc	Drugs & Medical	24	11	Low λ	9.730 (0.922)	[7.928, 11.528]
13	Church & Dwight Inc	Chemical	68	68	Low λ	9.729 (0.998)	[7.905, 11.548]
14	Protein Polymer Technologies	Chemical	37	0	Low λ	9.729 (0.926)	[7.944, 11.531]
15	Aradigm Corp	Drugs & Medical	48	19	Low λ	9.728 (0.995)	[7.907, 11.559]
16	Respirerx Pharmaceuticals	Drugs & Medical	37	1	Low λ	9.717 (0.981)	[7.896, 11.541]
17	Immunogen Inc	Drugs & Medical	84	27	Low λ	9.717 (0.946)	[7.905, 11.529]
18	Sonus Pharmaceuticals Inc	Drugs & Medical	63	11	Low λ	9.717 (0.954)	[7.902, 11.519]
19	Hanson Plc	Chemical	24	93	Low λ	9.716 (0.966)	[7.899, 11.549]
20	Carpenter Technology Corp	Mechanical	57	68	Low λ	9.714 (1.036)	[7.869, 11.564]

Notes: Centrality is the percentile ranking of degree centrality; Size is the percentile ranking of a firm's total asset; Type represents high/low network interaction effect (λ).

Indirect Spillin Effects to each firm from its peers due to a 1% reduction in peers' R&D tax price.

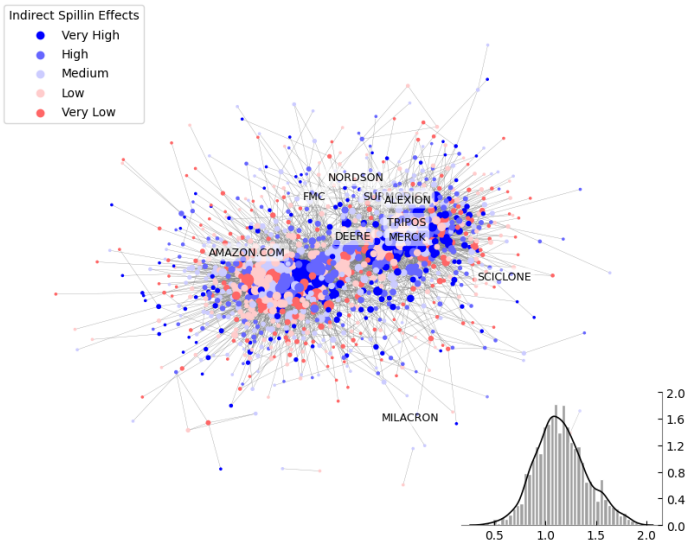

Figure 4.5: Indirect Spillin Effects to each firm from a 1% reduction in peers' R&D tax price. The histogram shows the distribution of the effects of interest across firms. In the network graph, firms are represented as nodes colored by the effect magnitude.

Table 4.11: The Top 20 Firms with the Highest Indirect Spillin Effects

						Indirect Spillin Effects	
Rank	Firm	Field	Centrality	Size	Type	Mean (std)	CI95
1	Fmc Corp	Chemical	24	80	High λ	1.914 (0.568)	[0.898, 3.009]
2	Sciclone Pharmaceuticals Inc	Drugs & Medical	24	23	High λ	1.910 (0.596)	[0.952, 3.104]
3	Alexion Pharmaceuticals Inc	Drugs & Medical	48	63	High λ	1.854 (0.612)	[0.885, 3.110]
4	Milacron Inc	Mechanical	8	64	High λ	1.848 (0.606)	[0.831, 3.029]
5	Amazon.com Inc	Computers & Communications	63	89	High λ	1.842 (0.584)	[0.675, 3.007]
6	Nordson Corp	Chemical	24	59	High λ	1.833 (0.639)	[0.764, 3.111]
7	Merck	Drugs & Medical	98	97	High λ	1.829 (0.584)	[0.864, 3.009]
8	Surmodics Inc	Drugs & Medical	68	29	High λ	1.822 (0.587)	[0.867, 3.000]
9	Deere & Co	Others	63	94	High λ	1.810 (0.719)	[0.287, 3.007]
10	Tripos Inc	Computers & Communications	75	18	High λ	1.799 (0.604)	[0.712, 2.960]
11	Tsi Corp	Drugs & Medical	8	10	High λ	1.794 (0.590)	[0.854, 2.992]
12	Astrazeneca	Drugs & Medical	96	95	High λ	1.792 (0.586)	[0.857, 2.993]
13	Lifecore Biomedical Inc	Drugs & Medical	24	15	High λ	1.791 (0.594)	[0.889, 3.013]
14	Visteon Corp	Mechanical	71	87	High λ	1.783 (0.584)	[0.868, 2.962]
15	Mattson Technology Inc	Electrical & Electronic	8	37	High λ	1.782 (0.656)	[0.541, 3.041]
16	Dot Hill Systems Corp	Computers & Communications	24	31	High λ	1.776 (0.589)	[0.860, 2.979]
17	Johnson Controls Intl Plc	Others	71	89	High λ	1.769 (0.592)	[0.861, 2.954]
18	Ici-Imperial Chem Inds Plc	Chemical	88	92	High λ	1.762 (0.516)	[0.823, 2.803]
19	Rohm And Haas Co	Chemical	84	84	High λ	1.761 (0.588)	[0.842, 2.966]
20	Parlex Corp	Electrical & Electronic	24	19	High λ	1.758 (0.613)	[0.667, 2.955]

Notes: Centrality is the percentile ranking of degree centrality; Size is the percentile ranking of a firm's total asset; Type represents high/low network interaction effect (λ).

Indirect Spillout Effects from each firm to its peers due to a 1% reduction in the firm's own R&D tax price.

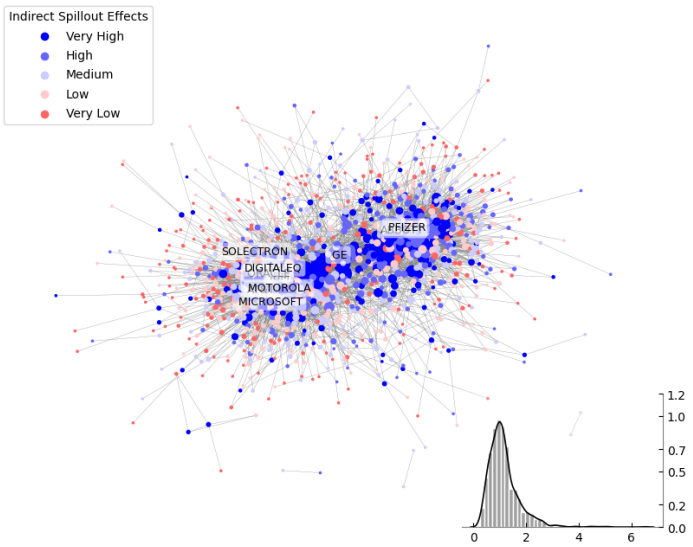

Figure 4.6: Indirect Spillout Effects from each firm to its peers due to a 1% reduction in the firm's own R&D tax price. The histogram shows the distribution of the effects of interest across firms. In the network graph, firms are represented as nodes colored by the effect magnitude.

Table 4.12: The Top 20 Firms with the Highest Indirect Spillout Effects

						Indirect Spillout Effects	
Rank	Firm	Field	Centrality	Size	Type	Mean (std)	CI95
1	Ibm	Computers & Communications	100	99	High λ	6.490 (4.719)	[0.642, 15.169]
2	Motorola	Computers & Communications	100	93	Low λ	5.144 (2.298)	[0.735, 8.860]
3	Oracle	Computers & Communications	99	94	Low λ	4.980 (2.762)	[0.670, 9.781]
4	Ge	Electrical & Electronic	99	100	Low λ	4.734 (1.197)	[1.534, 6.972]
5	Abbott	Drugs & Medical	98	94	Low λ	4.512 (2.805)	[0.464, 8.951]
6	Microsoft	Computers & Communications	99	98	Low λ	4.399 (2.928)	[0.435, 9.913]
7	Digitaleq	Computers & Communications	95	88	Low λ	3.897 (1.961)	[0.430, 7.288]
8	Pfizer	Drugs & Medical	100	98	High λ	3.836 (2.916)	[0.356, 9.307]
9	At&T Corp	Computers & Communications	95	99	Low λ	3.463 (2.016)	[0.410, 7.165]
10	Solectron Corp	Computers & Communications	80	82	Low λ	3.318 (0.864)	[1.813, 5.272]
11	Hp	Computers & Communications	100	97	High λ	3.310 (2.727)	[0.356, 9.346]
12	Intel	Computers & Communications	100	96	Low λ	3.226 (2.106)	[0.297, 7.068]
13	Endo International Plc	Drugs & Medical	68	79	Low λ	3.126 (1.016)	[0.531, 4.859]
14	Apple	Computers & Communications	97	95	Low λ	3.070 (0.849)	[1.630, 5.054]
15	Alcatel	Computers & Communications	99	97	Low λ	3.051 (1.503)	[0.442, 5.612]
16	Emc	Computers & Communications	97	90	Low λ	3.033 (1.898)	[0.304, 6.247]
17	Basf	Chemical	98	98	Low λ	3.024 (0.723)	[1.809, 4.639]
18	Verisign Inc	Computers & Communications	80	80	Low λ	2.763 (1.308)	[0.367, 5.271]
19	Illinois Tw	Mechanical	92	86	Low λ	2.727 (1.010)	[0.478, 4.859]
20	Schering	Drugs & Medical	97	87	Low λ	2.722 (0.977)	[0.468, 4.239]

Notes: Centrality is the percentile ranking of degree centrality; Size is the percentile ranking of a firm's total asset; Type represents high/low network interaction effect (λ).

Total Spillin Effects on each firm due to a 1% reduction in the R&D tax price for all firms

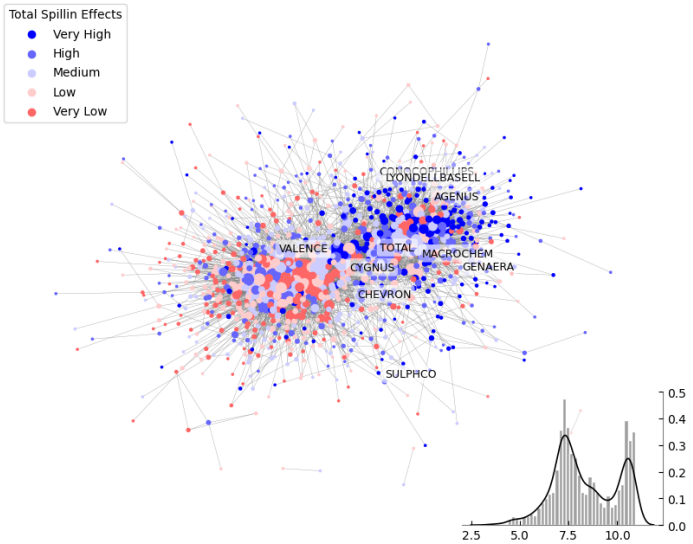

Figure 4.7: Total Spillin Effects on each firm due to a 1% reduction in the R&D tax price for all firms. The histogram shows the distribution of the effects of interest across firms. In the network graph, firms are represented as nodes colored by the effect magnitude.

Table 4.13: The Top 20 Firms with the Highest Total Spillin Effects

						Total Spillin Effects	
Rank	Firm	Field	Centrality	Size	Type	Mean (std)	CI95
1	Cygnus Inc	Drugs & Medical	78	16	Low λ	10.921 (1.026)	[8.911, 12.930]
2	Total Se	Chemical	87	99	Low λ	10.921 (1.026)	[8.911, 12.929]
3	Sulphco Inc	Chemical	8	2	Low λ	10.920 (1.026)	[8.911, 12.929]
4	Agenus Inc	Drugs & Medical	48	23	Low λ	10.920 (1.025)	[8.911, 12.930]
5	Conocophillips	Chemical	24	98	Low λ	10.920 (1.025)	[8.911, 12.922]
6	Genaera Corp	Drugs & Medical	37	10	Low λ	10.920 (1.024)	[8.911, 12.922]
7	Lyondellbasell Industries Nv	Chemical	24	95	Low λ	10.919 (1.027)	[8.905, 12.929]
8	Chevron Corp	Chemical	75	99	Low λ	10.918 (1.025)	[8.909, 12.925]
9	Valence Technology Inc	Electrical & Electronic	37	18	Low λ	10.918 (1.027)	[8.908, 12.922]
10	Macrochem Corp/De	Drugs & Medical	63	2	Low λ	10.916 (1.032)	[8.903, 12.922]
11	Emisphere Technologies Inc	Drugs & Medical	68	14	Low λ	10.915 (1.035)	[8.902, 12.919]
12	Abeona Therapeutics Inc	Drugs & Medical	37	2	Low λ	10.913 (1.021)	[8.903, 12.909]
13	Archer-Daniels-Midland Co	Drugs & Medical	78	92	Low λ	10.913 (1.025)	[8.908, 12.916]
14	Corixa Corp	Drugs & Medical	81	38	Low λ	10.913 (1.041)	[8.898, 12.930]
15	Orbital Atk Inc	Chemical	71	75	Low λ	10.913 (1.047)	[8.898, 12.922]
16	Bp Plc	Chemical	90	99	Low λ	10.913 (1.023)	[8.908, 12.913]
17	Crop Genetics Intl Corp	Others	24	9	Low λ	10.912 (1.027)	[8.902, 12.919]
18	Mobil Corp	Chemical	48	97	Low λ	10.911 (1.042)	[8.898, 12.913]
19	Colgate-Palmolive Co	Chemical	57	86	Low λ	10.910 (1.040)	[8.893, 12.915]
20	Atrix Laboratories Inc	Drugs & Medical	68	21	Low λ	10.910 (1.031)	[8.899, 12.908]

Notes: Centrality is the percentile ranking of degree centrality; Size is the percentile ranking of a firm's total asset; Type represents high/low network interaction effect (λ).

Total Spillout Effects from each firm due to a 1% reduction in the firm's own R&D tax price

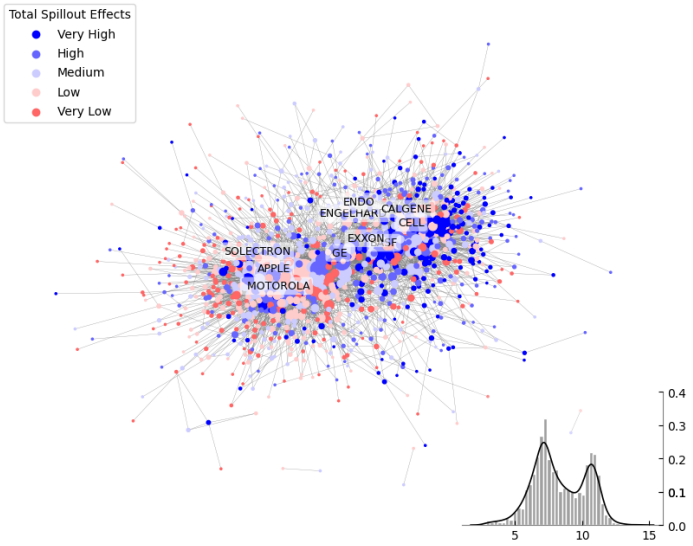

Figure 4.8: Total Spillout Effects from each firm due to a 1% reduction in the firm's own R&D tax price. The histogram shows the distribution of the effects of interest across firms. In the network graph, firms are represented as nodes colored by the effect magnitude.

Table 4.14: The Top 20 Firms with the Highest Total Spillout Effects

						Total Spillout Effects	
Rank	Firm	Field	Centrality	Size	Type	Mean (std)	CI95
1	Ge	Electrical & Electronic	99	100	Low λ	14.167 (2.588)	[4.660, 17.978]
2	Motorola	Computers & Communications	100	93	Low λ	13.158 (5.375)	[2.025, 19.635]
3	Solectron Corp	Computers & Communications	80	82	Low λ	13.061 (1.504)	[10.315, 16.257]
4	Apple	Computers & Communications	97	95	Low λ	12.743 (1.692)	[9.954, 15.976]
5	Basf	Chemical	98	98	Low λ	12.665 (1.486)	[9.989, 15.585]
6	Cell Genesys Inc	Drugs & Medical	85	46	Low λ	12.400 (1.362)	[9.870, 15.055]
7	Engelhard Corp	Chemical	71	75	Low λ	12.254 (1.598)	[9.653, 14.978]
8	Calgene Inc	Drugs & Medical	80	26	Low λ	12.202 (1.429)	[9.733, 14.831]
9	Exxon Mobil Corp	Chemical	87	100	Low λ	12.185 (1.346)	[9.723, 15.029]
10	Endo International Plc	Drugs & Medical	68	79	Low λ	12.167 (3.096)	[2.273, 15.714]
11	Xoma Corp	Drugs & Medical	81	25	Low λ	12.163 (1.445)	[9.612, 15.080]
12	Oracle	Computers & Communications	99	94	Low λ	12.094 (6.259)	[1.682, 20.562]
13	Usg Corp	Others	24	79	Low λ	12.071 (1.290)	[9.621, 14.605]
14	Soligenix Inc	Chemical	24	2	Low λ	12.059 (1.334)	[9.612, 14.676]
15	Ciba	Chemical	95	87	Low λ	11.974 (1.339)	[9.674, 14.316]
16	Ionis Pharma	Chemical	92	47	Low λ	11.914 (2.148)	[3.727, 14.760]
17	P&G	Chemical	96	98	Low λ	11.843 (2.470)	[3.035, 14.976]
18	Bemis Co Inc	Others	57	73	Low λ	11.837 (1.599)	[9.108, 14.579]
19	Texaco Inc	Chemical	84	96	Low λ	11.796 (1.269)	[9.453, 14.456]
20	Ibm	Computers & Communications	100	99	High λ	11.793 (8.407)	[1.206, 25.793]

Notes: Centrality is the percentile ranking of degree centrality; Size is the percentile ranking of a firm's total asset; Type represents high/low network interaction effect (λ).

For comparison with the homogeneous models in the previous section, we again summarize the average direct, indirect, and total effects of 1% reduction in the R&D tax price for all firms. As shown in Table 4.15, one consequence of averaging indirect effects over all observations (firms) is that the average of the row- and column- sums of the derivative matrix are inevitably the same. In fact, this equality is a result of uniform weighting, which is meaningful only for a policy experiment in which every firm receives exactly the same 1% tax-price shock. In other words, this can speak only to a "one-size-fits-all" counterfactual – a critical limitation of homogeneous models (SAR/SCSAR). Compared to the homogeneous benchmarks in Table 4.7,[15] the magnitude of the estimated effects shrinks once latent types are introduced to capture firm heterogeneity (via a mixture structure). The reduction is not a loss of explanatory power but a correction for over-aggregation. Once heterogeneity and endogenous network selection are properly modelled under the SCHSAR framework, the inflated indirect effects of the homogeneous SAR collapse to more credible values while remaining statistically and economically significant. Furthermore, by relaxing the homogeneity restriction, SCHSAR yields firm-specific effect estimates and delivers richer interpretations for targeted policy designs.

Table 4.15: Direct, Indirect and Total Effects from R&D Tax Price.

	HSAR		**SCHSAR**	
	Mean (std)	CI95	Mean (std)	CI95
Direct	7.275 (0.613)	[6.052, 8.482]	7.168 (0.601)	[5.971, 8.296]
Indirect Splillin	1.320 (0.188)	[0.981, 1.707]	1.157 (0.154)	[0.871, 1.473]
Indirect Splillout	1.320 (0.188)	[0.981, 1.707]	1.157 (0.154)	[0.871, 1.473]
Total Splillin	8.595 (0.730)	[7.171, 10.053]	8.325 (0.699)	[6.945, 9.654]
Total Splillout	8.595 (0.730)	[7.171, 10.053]	8.325 (0.699)	[6.945, 9.654]

Notes: This table presents the posterior means, standard deviations, and 95% credible intervals for each effect across firms. The derivation of each effect is based on posterior samples of each parameter.

[15] The first row reports the effects of 1% *increase* in the tax price and therefore carries the opposite sign.

4.6 Concluding Remarks

This paper introduces the Selection-corrected Heterogeneous Spatial Autoregressive (SCHSAR) framework to address two critical challenges in network interaction models: heterogeneous effects and endogenous network formation. The proposed approach extends the standard SAR model by incorporating a finite mixture structure to capture rich heterogeneity in network interaction effects, while simultaneously modelling the network formation process to correct for potential selection bias. Three types of latent variables are integrated into the framework: individual-type indicators associated with mixture components in the outcome equation, dyadic utilities governing link formation in the network equation, and individual-specific random effects accounting for endogeneity. This hierarchical structure, while posing a challenge to the maximum likelihood estimators, is efficiently handled by our Bayesian algorithm based on data augmentation techniques.

The simulation study validates our fully Bayesian approach. It is computationally tractable and able to deliver reliable inference for the SCHSAR model across various data-generating processes. The proposed estimators exhibit near-unbiasedness and nominal coverage, especially when the signal level is favourable and endogenous network formation is mainly driven by unobserved degree heterogeneity. In contrast, reliance on either exogeneity or homogeneity results in significant bias and unreliable inference for interaction effect coefficients when these assumptions are violated.

Our empirical application to a technological collaboration network among U.S. firms provides several key insights. First, technological homophily significantly shapes the network structure, complemented by firm-specific latent advantages that affect both linkage ability and R&D intensity. Second, positive network interaction effects on corporate R&D investments persist after correcting for selection bias due to endogenous network formation. Our heterogeneous model reveals two distinct types of firms, with varying levels of network peer effects and tax price responsiveness. Third, we find substantial firm-level heterogeneity in both the direct and indirect (spillin and spillout) effects of R&D tax incentives. Different firm behaviours are revealed: “self-reliant innovators” with strong direct effects, “responsive absorbers” with high spillin effects, and “influential transmitters” with significant spillout effects. This highlights the role of network position and firm attributes, such as firm size, in driving innovation dynamics. Finally, evaluating firm-level total effects facilitates the design of targeted policy interventions. For example, prioritising firms with high total spillout

effects could be a cost-effective strategy to maximise systemic innovation diffusion across the network.

The SCHSAR framework provides a flexible and robust approach to drawing credible conclusions about network interaction effects by effectively accommodating both heterogeneity and endogeneity. Future research directions include extending the framework to dynamic panel settings, automatically selecting the number of mixture components, and improving unobserved homophily handling. Although we focus on the analysis of firm network data, a wide range of potential applications can be explored using this method.

Chapter 5

Conclusion

5.1 Summary & Implications

This thesis leverages Bayesian techniques to develop methods that tackle challenges in three microeconometric settings with a focus on causality and heterogeneity.

As consistently demonstrated in the three essays, the Bayesian modelling approach, in combination with efficient computational methods, showcases numerous advantages. First, it provides flexible data-driven solutions to high-dimensional problems, nonlinearities, endogeneity, and heterogeneity through sophisticated hierarchical priors. For example, high-dimensional covariates are tackled by introducing regularisation through shrinkage priors (as in quantile regressions), or unmeasured confounding is captured via latent variables (as in the modelling of endogenous treatment selection and network formation). This flexibility allows for effective and simultaneous handling of all key elements within a unified framework, even in settings where traditional methods fail or are computationally infeasible. Second, it enables principled uncertainty quantification through a probabilistic framework that incorporates all sources of variability within the posterior distribution, including parameters, latent variables, and missing data. This property is particularly valuable for drawing credible inferences in small samples and enabling direct inference on nonlinear/nonstandard causal estimands or policy-relevant quantities that would be difficult to obtain through traditional approaches. Ultimately, the Bayesian approach not only provides computational tractability but also fundamental improvements in handling complexity, addressing uncertainty, and enhancing interpretability in modern microeconometric analysis. Across all three essays, proposed methods deliver more accurate inference and substantial bias reduction compared to naive or misspecified approaches.

Apart from methodological insights, the three essays in this thesis provide strong evidence that accounting for both causality and heterogeneity is not merely an academic exercise but a

policy imperative. Without proper causal identification, policies risk being based on spurious relationships that lead to ineffective or harmful interventions. Without heterogeneity analysis, policies miss opportunities for optimisation and may produce unintended distributional consequences. In the first essay, quantile treatment effect (QTE) estimation reveals how universal financial access affects different segments of the outcome distribution, highlighting that while some households benefit substantially, others may experience negative impacts. These distributional insights, which are hidden by average treatment effects, underscore the importance of designing financial interventions that consider potential inequality-amplifying consequences. The second essay shows that accounting for both endogenous selection and spillovers is essential when evaluating place-based policies like the Opportunity Zone (OZ) program. The findings indicate that benefits are concentrated among the selected zones, driven by unobserved tract characteristics, whereas neighbouring areas do not experience positive spillovers, suggesting that simply expanding the program may not yield similar returns. Ignoring this heterogeneity could lead to misguided scaling of such policies. The third essay demonstrates that firms' responses to R&D tax incentives vary significantly with their network positions and latent attributes. Identifying distinct types of firms, such as "influential transmitters", enables more targeted and cost-effective innovation policies, maximising the total benefits across the network.

5.2 Avenues for Future Research

This thesis opens up several promising directions for future research, particularly at the intersection of Bayesian methods and microeconometric analysis. In the first essay, future work could improve bootstrap-based inference for QTE estimators or develop scalable fully Bayesian inference procedures, especially for large datasets. Efficient implementation could also facilitate the examination of the performance of the method with various choices of priors. Additionally, establishing more rigorous theoretical results, such as deriving Bernstein-von Mises theorems for Bayesian estimators of quantile treatment effects in high-dimensional settings, would significantly contribute to the literature. The second and third essays collaboratively demonstrate the capacity of the Bayesian approach to address unmeasured confounding and capture heterogeneity in complex settings. As these essays target different sources of endogeneity, namely endogenous selection into treatment and endogenous network formation, an important extension would be to develop a unified framework that jointly handles multiple causal challenges in the presence of networks. Although the proposed

methods are grounded in parametric assumptions to ensure interpretability and computational feasibility, future research could relax these assumptions to introduce greater flexibility. Altogether, these directions highlight a rich research agenda in Bayesian microeconometrics, which will be crucial for advancing empirical research in economics, especially in contexts where classical tools fall short.

Bibliography

ABBRING, J. H., AND J. J. HECKMAN. (2007): "Econometric evaluation of social programs, part III: Distributional treatment effects, dynamic treatment effects, dynamic discrete choice, and general equilibrium policy evaluation," *Handbook of econometrics*, 6, 5145–5303.

AHMAD, M. M. (2003): "Distant voices: The views of the field workers of NGOs in bangladesh on microcredit," *Geographical Journal*, 169, 65–74.

ALBERT, J. H., AND S. CHIB. (1993): "Bayesian analysis of binary and polychotomous response data," *Journal of the American statistical Association*, 88, 669–79.

ALHAMZAWI, R., AND H. TAHA MOHAMMAD ALI. (2020): "A new gibbs sampler for bayesian lasso," *Communications in Statistics-Simulation and Computation*, 49, 1855–71.

ALHAMZAWI, R., K. YU, AND D. F. BENOIT. (2012): "Bayesian adaptive lasso quantile regression," *Statistical Modelling*, 12, 279–97.

ALM, J., T. DRONYK-TROSPER, AND S. LARKIN. (2021): "In the land of OZ: Designating opportunity zones," *Public Choice*, 188, 503–23.

ANDRIEU, C., AND J. THOMS. (2008): "A tutorial on adaptive MCMC," *Statistics and computing*, 18, 343–73.

ANGRIST, J. D., AND J.-S. PISCHKE. (2009): *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton university press.

ANTONELLI, J., M. CEFALU, N. PALMER, AND D. AGNIEL. (2018): "Doubly robust matching estimators for high dimensional confounding adjustment," *Biometrics*, 74, 1171–79.

ANTONELLI, J., G. PAPADOGEORGOU, AND F. DOMINICI. (2022): "Causal inference in high dimensions: A marriage between bayesian modeling and good frequentist properties," *Biometrics*, 78, 100–114.

ANTONELLI, J., G. PARMIGIANI, AND F. DOMINICI. (2019): "High-dimensional confounding adjustment using continuous spike and slab priors," *Bayesian analysis*, 14, 805.

AQUARO, M., N. BAILEY, AND M. H. PESARAN. (2021): "Estimation and inference for spatial models with heterogeneous coefficients: An application to US house prices," *Journal of Applied Econometrics*, 36, 18–44.

ARDUINI, T., E. PATACCHINI, AND E. RAINONE. (2020a): "Identification and Estimation of Network Models with Heterogeneous Interactions," in *The econometrics of networks*, Emerald Publishing Limited, 3–25.

---. (2020b): "Treatment effects with heterogeneous externalities," *Journal of Business & Economic Statistics*, 38, 826–38.

ARONOW, P. M., AND C. SAMII. (2017): "Estimating average causal effects under general interference, with application to a social network experiment,"

ARQUÉ-CASTELLS, P., AND D. F. SPULBER. (2022): "Measuring the private and social returns to r&d: Unintended spillovers versus technology markets," *Journal of Political Economy*, 130, 1860–1918.

ATCHADÉ, Y. F., AND J. S. ROSENTHAL. (2005): "On adaptive markov chain monte carlo algorithms," *Bernoulli*, 11, 815–28.

ATHEY, S., AND G. W. IMBENS. (2007): "Discrete choice models with multiple unobserved choice characteristics," *International Economic Review*, 48, 1159–92.

ATHEY, S., G. W. IMBENS, AND S. WAGER. (2018): "Approximate residual balancing: Debiased inference of average treatment effects in high dimensions," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 80, 597–623.

AUERBACH, E. (2022): "Identification and estimation of a partially linear regression model using network data," *Econometrica*, 90, 347–65.

BANERJEE, A. V. (2013): "Microcredit under the microscope: What have we learned in the past two decades, and what do we need to know?" *Annu. Rev. Econ.*, 5, 487–519.

BANG, H., AND J. M. ROBINS. (2005): "Doubly robust estimation in missing data and causal inference models," *Biometrics*, 61, 962–73.

BELLONI, A., V. CHERNOZHUKOV, I. FERNÁNDEZ-VAL, AND C. HANSEN. (2017): "Program evaluation and causal inference with high-dimensional data," *Econometrica*, 85, 233–98.

BELLONI, A., V. CHERNOZHUKOV, AND C. HANSEN. (2014): "Inference on treatment effects after selection among high-dimensional controls," *The Review of Economic Studies*, 81,

608–50.

BENOIT, D. F., AND D. VAN DEN POEL. (2017): "bayesQR: A bayesian approach to quantile regression," *Journal of Statistical Software*, 76, 1–32.

BENVENISTE, A., M. MÉTIVIER, AND P. PRIOURET. (2012): *Adaptive Algorithms and Stochastic Approximations*,. Vol. 22 Springer Science & Business Media.

BEUGNOT, J., B. FORTIN, G. LACROIX, AND M. C. VILLEVAL. (2019): "Gender and peer effects on performance in social networks," *European Economic Review*, 113, 207–24.

BHADRA, A., J. DATTA, N. G. POLSON, AND B. WILLARD. (2019): "Lasso meets horseshoe: A survey," *Statistical Science*, 34, 405–27.

BLUME, L. E., W. A. BROCK, S. N. DURLAUF, AND R. JAYARAMAN. (2015): "Linear social interactions models," *Journal of Political Economy*, 123, 444–96.

BOTEV, Z. I. (2017): "The normal law under linear restrictions: Simulation and estimation via minimax tilting," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 79, 125–48.

BRAMOULLÉ, Y., H. DJEBBARI, AND B. FORTIN. (2009): "Identification of peer effects through social networks," *Journal of econometrics*, 150, 41–55.

---. (2020): "Peer effects in networks: A survey," *Annual Review of Economics*, 12, 603–29.

BREUNIG, C., R. LIU, AND Z. YU. (2024): "Semiparametric bayesian difference-in-differences," *arXiv preprint arXiv:2412.04605*,.

---. (2025): "Double robust bayesian inference on average treatment effects," *Econometrica*, 93, 539–68.

BRINCH, C. N., M. MOGSTAD, AND M. WISWALL. (2017): "Beyond LATE with a discrete instrument," *Journal of Political Economy*, 125, 985–1039.

BRYAN, G., S. CHOWDHURY, AND A. M. MOBARAK. (2014): "Underinvestment in a profitable technology: The case of seasonal migration in bangladesh," *Econometrica*, 82, 1671–1748.

BULOW, J. I., J. D. GEANAKOPLOS, AND P. D. KLEMPERER. (1985): "Multimarket oligopoly: Strategic substitutes and complements," *Journal of Political economy*, 93, 488–511.

CALLAWAY, B., AND T. LI. (2019): "Quantile treatment effects in difference in differences models with panel data," *Quantitative Economics*, 10, 1579–1618.

CAMERON, A. COLIN, AND P. K. TRIVEDI. (2005): *Microeconometrics: Methods and Applications*, Cambridge university press.

CAMERON, A. COLIN, AND P. K. TRIVEDI. (2022): *Microeconometrics Using Stata: Volume II — Nonlinear Models and Causal Inference Methods*, 2nd ed. Vol. II College Station, TX: Stata Press.

CARNEIRO, P., J. J. HECKMAN, AND E. J. VYTLACIL. (2011): "Estimating marginal returns to education," *American Economic Review*, 101, 2754–81.

CARVALHO, C. M., N. G. POLSON, AND J. G. SCOTT. (2009): "Handling Sparsity via the Horseshoe," *Artificial intelligence and statistics*, PMLR, 73–80.

---. (2010): "The horseshoe estimator for sparse signals," *Biometrika*, 97, 465–80.

CHAMBERLAIN, G., AND K. HIRANO. (1999): "Predictive distributions based on longitudinal earnings data," *Annales d'Economie et de Statistique*, 211–42.

CHAMBERLAIN, G., AND G. W. IMBENS. (2003): "Nonparametric applications of bayesian inference," *Journal of Business & Economic Statistics*, 21, 12–18.

CHAN, J., G. KOOP, D. J. POIRIER, AND J. L. TOBIAS. (2019): *Bayesian Econometric Methods*,. Vol. 7 Cambridge University Press, 242–44.

CHANDRASEKHAR, A. G. (2016): "Econometrics of network formation,"

CHEN, J., E. GLAESER, AND D. WESSEL. (2023): "JUE insight: The (non-) effect of opportunity zones on housing prices," *Journal of Urban Economics*, 133, 103451.

CHERNOZHUKOV, V., D. CHETVERIKOV, M. DEMIRER, E. DUFLO, C. HANSEN, W. NEWEY, AND J. ROBINS. (2018): "Double/Debiased Machine Learning for Treatment and Structural Parameters,"Oxford University Press Oxford, UK.

CHERNOZHUKOV, V., M. DEMIRER, E. DUFLO, AND I. FERNÁNDEZ-VAL. (2017): "Generic machine learning inference on heterogenous treatment effects in randomized experiments," *arXiv e-prints*, arXiv–1712.

CHERNOZHUKOV, V., I. FERNÁNDEZ-VAL, AND B. MELLY. (2013): "Inference on counterfactual distributions," *Econometrica*, 81, 2205–68.

CHERNOZHUKOV, V., W. K. NEWEY, AND R. SINGH. (2022): "Automatic debiased machine learning of causal and structural effects," *Econometrica*, 90, 967–1027.

CHETTY, R., N. HENDREN, AND L. F. KATZ. (2016): "The effects of exposure to better

neighborhoods on children: New evidence from the moving to opportunity experiment," *American Economic Review*, 106, 855–902.

CHIB, S., AND L. JACOBI. (2007): "Modeling and calculating the effect of treatment at baseline from panel outcomes," *Journal of Econometrics*, 140, 781–801.

CHIPMAN, H. A., E. I. GEORGE, AND R. E. MCCULLOCH. (2012): "BART: Bayesian additive regression trees," *Annals of Applied Statistics*, 6(1), 266-298.

CHIPMAN, H., E. GEORGE, AND R. MCCULLOCH. (2006): "Bayesian ensemble learning," *Advances in neural information processing systems*, 19.

CONLEY, T. G., C. B. HANSEN, R. E. MCCULLOCH, AND P. E. ROSSI. (2008): "A semi-parametric bayesian approach to the instrumental variable problem," *Journal of Econometrics*, 144, 276–305.

COOPER, R., AND A. JOHN. (1988): "Coordinating coordination failures in keynesian models," *The Quarterly Journal of Economics*, 103, 441–63.

CORINTH, K., AND N. FELDMAN. (2024): "Are opportunity zones an effective place-based policy?" *Journal of Economic Perspectives*, 38, 113–36.

CORNELISSEN, T., C. DUSTMANN, A. RAUTE, AND U. SCHÖNBERG. (2018): "Who benefits from universal child care? Estimating marginal returns to early child care attendance," *Journal of Political Economy*, 126, 2356–2409.

CORNWALL, G. J. (2017): "Three Essays on Bayesian Econometric Methods," University of Cincinnati.

CORNWALL, G. J., AND O. PARENT. (2017): "Embracing heterogeneity: The spatial autoregressive mixture model," *Regional Science and Urban Economics*, 64, 148–61.

CRÉPON, B., F. DEVOTO, E. DUFLO, AND W. PARIENTÉ. (2015): "Estimating the impact of microcredit on those who take it up: Evidence from a randomized experiment in morocco," *American Economic Journal: Applied Economics*, 7, 123–50.

DANIEL, R. M. (2014): "Double robustness," *Wiley StatsRef: Statistics Reference Online*, 1–14.

DÍAZ, I. (2017): "Efficient estimation of quantiles in missing data models," *Journal of Statistical Planning and Inference*, 190, 39–51.

DING, C., J. ESTRADA, AND S. MONTOYA-BLANDÓN. (2023): "Bayesian inference of

network formation models with payoff externalities,"

DING, P. (2014): "Bayesian robust inference of sample selection using selection-t models," *Journal of Multivariate Analysis*, 124, 451–64.

DITRAGLIA, F. J., C. GARCIA-JIMENO, R. OKEEFFE-ODONOVAN, AND A. SANCHEZ-BECERRA. (2023): "Identifying causal effects in experiments with spillovers and non-compliance," *Journal of Econometrics*, 235(2), 1589-1624.

DOĞAN, O., AND S. TAŞPINAR. (2018): "Bayesian inference in spatial sample selection models," *Oxford Bulletin of Economics and Statistics*, 80, 90–121.

DOKSUM, K. (1974): "Empirical probability plots and statistical inference for nonlinear models in the two-sample case," *The annals of statistics*, 267–77.

DUFLO, E., P. DUPAS, AND M. KREMER. (2021): The Impact of Free Secondary Education: Experimental Evidence from Ghana, National Bureau of Economic Research.

DZEMSKI, A. (2019): "An empirical model of dyadic link formation in a network with unobserved heterogeneity," *Review of Economics and Statistics*, 101, 763–76.

EISENHAUER, P., J. J. HECKMAN, AND E. VYTLACIL. (2015): "The generalized roy model and the cost-benefit analysis of social programs," *Journal of Political Economy*, 123, 413–43.

ELDAR, O., AND C. GARBER. (2022): "Does government play favorites? Evidence from opportunity zones," *Journal of Law and Economics (forthcoming, 2023)*,.

FARRELL, M. H. (2015): "Robust inference on average treatment effects with possibly more covariates than observations," *Journal of Econometrics*, 189, 1–23.

FIRPO, S. (2007): "Efficient semiparametric estimation of quantile treatment effects," *Econometrica*, 75, 259–76.

FIRPO, S., AND C. PINTO. (2016): "Identification and estimation of distributional impacts of interventions using changes in inequality measures," *Journal of Applied Econometrics*, 31, 457–86.

FORASTIERE, L., E. M. AIROLDI, AND F. MEALLI. (2021): "Identification and estimation of treatment and interference effects in observational studies on networks," *Journal of the American Statistical Association*, 116, 901–18.

FORASTIERE, L., F. MEALLI, A. WU, AND E. M. AIROLDI. (2022): "Estimating causal

effects under network interference with bayesian generalized propensity scores," *Journal of Machine Learning Research*, 23, 1–61.

FRANK, M. M., J. L. HOOPES, AND R. LESTER. (2022): "What determines where opportunity knocks? Political affiliation in the selection of opportunity zones," *Journal of Public Economics*, 206, 104588.

FREEDMAN, M., S. KHANNA, AND D. NEUMARK. (2023): "JUE insight: The impacts of opportunity zones on zone residents," *Journal of Urban Economics*, 133, 103407.

FRÖLICH, M., AND B. MELLY. (2013): "Unconditional quantile treatment effects under endogeneity," *Journal of Business & Economic Statistics*, 31, 346–57.

FRÜHWIRTH-SCHNATTER, S. (2006): *Finite Mixture and Markov Switching Models*, Springer.

GEWEKE, J. (2007): "Interpretation and inference in mixture models: Simple MCMC works," *Computational Statistics & Data Analysis*, 51, 3529–50.

GIFFIN, A., B. REICH, S. YANG, AND A. RAPPOLD. (2022): "Generalized propensity score approach to causal inference with spatial interference," *Biometrics*, 79(3), 2220-2231.

GLEWWE, P., AND P. TODD. (2022): *Impact Evaluation in International Development: Theory, Methods, and Practice*, World Bank Publications.

GOLDSMITH-PINKHAM, P., AND G. W. IMBENS. (2013): "Social networks and the identification of peer effects," *Journal of Business & Economic Statistics*, 31, 253–64.

GRAHAM, B. S. (2015): "Methods of identification in social networks," *Annu. Rev. Econ.*, 7, 465–85.

---. (2017): "An econometric model of network formation with degree heterogeneity," *Econometrica*, 85, 1033–63.

GRAMACY, R. B., AND N. G. POLSON. (2012): "Simulation-based regularized logistic regression," *Bayesian Analysis* 7.3, 567-590.

GUO, B. (2019): "Variational Inference for Quantile Regression,"Washington University in St. Louis.

HAHN, P. R., C. M. CARVALHO, D. PUELZ, AND J. HE. (2018): "Regularization and confounding in linear regression for treatment effect estimation," *Bayesian Analysis*, 13, 163–82.

HAHN, P. R., J. S. MURRAY, AND C. M. CARVALHO. (2020): "Bayesian regression tree

models for causal inference: Regularization, confounding, and heterogeneous effects (with discussion)," *Bayesian Analysis*, 15, 965–1056.

HALL, B. H., A. B. JAFFE, M. TRAJTENBERG, ET AL. (2001): "The NBER patent citations data file: Lessons, insights and methodological tools," *NBER working paper*, 8498, 1–74.

HAN, X., C.-S. HSIEH, AND S. I. KO. (2021): "Spatial modeling approach for dynamic network formation and interactions," *Journal of Business & Economic Statistics*, 39, 120–35.

HAN, X., AND L.-F. LEE. (2016): "Bayesian analysis of spatial panel autoregressive models with time-varying endogenous spatial weight matrices, common factors, and random coefficients," *Journal of Business & Economic Statistics*, 34, 642–60.

HANS, C. (2009): "Bayesian lasso regression," *Biometrika*, 96, 835–45.

HASTIE, T., AND R. TIBSHIRANI. (2000): "Bayesian backfitting (with comments and a rejoinder by the authors," *Statistical Science*, 15, 196–223.

HECKMAN, J. J. (1979): "Sample selection bias as a specification error," *Econometrica: Journal of the econometric society*, 153–61.

---. (2001): "Micro data, heterogeneity, and the evaluation of public policy: Nobel lecture," *Journal of political Economy*, 109, 673–748.

HECKMAN, J. J., AND R. PINTO. (2022): "The econometric model for causal policy analysis," *Annual review of economics*, 14, 893–923.

HECKMAN, J. J., AND R. ROBB JR. (1985): "Alternative methods for evaluating the impact of interventions: An overview," *Journal of econometrics*, 30, 239–67.

HECKMAN, J. J., AND E. J. VYTLACIL. (2005): "Structural equations, treatment effects, and econometric policy evaluation 1," *Econometrica*, 73, 669–738.

---. (2007): "Econometric evaluation of social programs, part II: Using the marginal treatment effect to organize alternative econometric estimators to evaluate social programs, and to forecast their effects in new environments," *Handbook of econometrics*, 6, 4875–5143.

HECKMAN, J., AND R. PINTO. (2024): "Econometric causality: The central role of thought experiments," *Journal of econometrics*, 243, 105719.

HILL, J. L. (2011): "Bayesian nonparametric modeling for causal inference," *Journal of Computational and Graphical Statistics*, 20, 217–40.

HINES, O., O. DUKES, K. DIAZ-ORDAZ, AND S. VANSTEELANDT. (2022): "Demystifying statistical learning based on efficient influence functions," *The American Statistician*, 76, 292–304.

HIRANO, K., AND G. W. IMBENS. (2004): "The propensity score with continuous treatments," *Applied Bayesian modeling and causal inference from incomplete-data perspectives*, 226164, 73–84.

HIRANO, K., G. W. IMBENS, AND G. RIDDER. (2003): "Efficient estimation of average treatment effects using the estimated propensity score," *Econometrica*, 71, 1161–89.

HSIEH, C.-S., AND L. F. LEE. (2016): "A social interactions model with endogenous friendship formation and selectivity," *Journal of Applied Econometrics*, 31, 301–19.

HUBER, M. (2023): *Causal Analysis: Impact Evaluation and Causal Machine Learning with Applications in r*, MIT Press.

HUDGENS, M. G., AND M. E. HALLORAN. (2008): "Toward causal inference with interference," *Journal of the American Statistical Association*, 103, 832–42.

ICHIMURA, H. (1993): "Semiparametric least squares (SLS) and weighted SLS estimation of single-index models," *Journal of econometrics*, 58, 71–120.

IMAI, K., AND D. A. VAN DYK. (2005): "A bayesian analysis of the multinomial probit model using marginal data augmentation," *Journal of econometrics*, 124, 311–34.

IMBENS, G. W. (2022): "Causality in econometrics: Choice vs chance," *Econometrica*, 90, 2541–66.

ISHWARAN, H., AND J. S. RAO. (2005): "Spike and slab variable selection: Frequentist and bayesian strategies," *The Annals of Statistics*, 33, 730–73.

JACKSON, M. O., AND Y. ZENOU. (2015): "Games on Networks," in *Handbook of game theory with economic applications*, Elsevier, 95–163.

JACOB, D. (2021): "CATE meets ML: Conditional average treatment effect and machine learning," *Digital Finance*, 3, 99–148.

JACOBI, L., H. WAGNER, AND S. FRÜHWIRTH-SCHNATTER. (2016): "Bayesian treatment effects models with variable selection for panel outcomes with an application to earnings effects of maternity leave," *Journal of Econometrics*, 193, 234–50.

JIAO, X., AND D. A. VAN DYK. (2015): "A corrected and more efficient suite of MCMC

samplers for the multinomal probit model," *arXiv preprint arXiv:1504.07823*,.

JOHNSSON, I., AND H. R. MOON. (2021): "Estimation of peer effects in endogenous social networks: Control function approach," *Review of Economics and Statistics*, 103, 328–45.

KABOSKI, J. P., AND R. M. TOWNSEND. (2011): "A structural evaluation of a large-scale quasi-experimental microfinance initiative," *Econometrica*, 79, 1357–1406.

KALLUS, N., X. MAO, AND M. UEHARA. (2024): "Localized debiased machine learning: Efficient inference on quantile treatment effects and beyond," *Journal of Machine Learning Research*, 25, 1–59.

KATZ, M. L., AND C. SHAPIRO. (1986): "Technology adoption in the presence of network externalities," *Journal of political economy*, 94, 822–41.

KEHAGIAS, A. (2018): "Community detection toolbox," *MATLAB Central File Exchange*,.

KELEJIAN, H. H., AND G. PIRAS. (2014): "Estimation of spatial models with endogenous weighting matrices, and an application to a demand model for cigarettes," *Regional Science and Urban Economics*, 46, 140–49.

KELEJIAN, H. H., AND I. R. PRUCHA. (2010): "Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances," *Journal of econometrics*, 157, 53–67.

KENNEDY, E. H. (2024): "Semiparametric doubly robust targeted double machine learning: A review," *Handbook of statistical methods for precision medicine*, 207-236.

KLINE, B., AND E. TAMER. (2020): "Econometric Analysis of Models with Social Interactions," in *The econometric analysis of network data*, Elsevier, 149–81.

KOENKER, R., AND G. BASSETT. (1978): "Regression quantiles," *Econometrica: journal of the Econometric Society*, 33–50.

KOROBILIS, D., AND K. SHIMIZU. (2022): "Bayesian approaches to shrinkage and sparse estimation," *Foundations and Trends® in Econometrics*, 11, 230–354.

KOTZ, S., T. KOZUBOWSKI, AND K. PODGORSKI. (2012): *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance*, Springer Science & Business Media.

KOZUMI, H., AND G. KOBAYASHI. (2011): "Gibbs sampling methods for bayesian quantile regression," *Journal of statistical computation and simulation*, 81, 1565–78.

LAAN, M. J. VAN DER, AND O. SOFRYGIN. (2017): "Semi-parametric estimation and inference for the mean outcome of the single time-point intervention in a causally connected population,"

LEE, L., X. LIU, AND X. LIN. (2010): "Specification and estimation of social interaction models with network structures," *The Econometrics Journal*, 13, 145–76.

LEHMANN, E. L. ET AL. (1974): "Statistical methods based on ranks," *Nonparametrics. San Francisco, CA, Holden-Day*,.

LESAGE, J. P., AND Y.-Y. CHIH. (2016): "Interpreting heterogeneous coefficient spatial autoregressive panel models," *Economics Letters*, 142, 1–5.

---. (2018): "A bayesian spatial panel model with heterogeneous coefficients," *Regional Science and Urban Economics*, 72, 58–73.

LESAGE, J. P., AND R. K. PACE. (2009): *Introduction to Spatial Econometrics*, Chapman; Hall/CRC.

LESAGE, J. P., AND O. PARENT. (2007): "Bayesian model averaging for spatial econometric models," *Geographical Analysis*, 39, 241–67.

LEUNG, M. P. (2020): "Treatment and spillover effects under network interference," *Review of Economics and Statistics*, 102, 368–80.

LEUNG, M. P., AND P. LOUPOS. (2022): "Unconfoundedness with network interference," *arXiv preprint arXiv:2211.07823*,.

LI, F., P. DING, AND F. MEALLI. (2023): "Bayesian causal inference: A critical review," *Philosophical Transactions of the Royal Society A*, 381, 20220153.

LI, MINGLIANG, D. J. POIRIER, AND J. L. TOBIAS. (2004): "Do dropouts suffer from dropping out? Estimation and prediction of outcome gains in generalized selection models," *Journal of Applied Econometrics*, 19, 203–25.

LI, MINGLIAN, AND J. TOBIAS. (2011): "Bayesian methods in microeconometrics."

LIM, D., B. PARK, D. NOTT, X. WANG, AND T. CHOI. (2020): "Sparse signal shrinkage and outlier detection in high-dimensional quantile regression with variational bayes," *Statistics and Its Interface*, 13, 237–49.

LIN, X. (2010): "Identifying peer effects in student academic achievement by spatial autoregressive models with group unobservables," *Journal of Labor Economics*, 28,

825–60.

LINERO, A. R., AND J. L. ANTONELLI. (2023): "The how and why of bayesian nonparametric causal inference," *Wiley Interdisciplinary Reviews: Computational Statistics*, 15, e1583.

LIU, X., AND L. LEE. (2010): "GMM estimation of social interaction models with centrality," *Journal of Econometrics*, 159, 99–115.

LIU, X., E. PATACCHINI, AND Y. ZENOU. (2014): "Endogenous peer effects: Local aggregate or local average?" *Journal of economic behavior & organization*, 103, 39–59.

LUO, C., AND M. J. DANIELS. (2021): "BNPqte: A bayesian nonparametric approach to causal inference on quantiles in r," *arXiv preprint arXiv:2106.14599*,.

MAKALIC, E., AND D. F. SCHMIDT. (2015): "A simple sampler for the horseshoe estimator," *IEEE Signal Processing Letters*, 23, 179–82.

MALLICK, H., AND N. YI. (2014): "A new bayesian lasso," *Statistics and its interface*, 7, 571–82.

MANSKI, C. F. (2013): "Identification of treatment response with social interactions," *The Econometrics Journal*, 16, S1–23.

MASTEN, M. A. (2018): "Random coefficients on endogenous variables in simultaneous equations models," *The Review of Economic Studies*, 85, 1193–1250.

MCCULLOCH, R., AND P. E. ROSSI. (1994): "An exact likelihood analysis of the multinomial probit model," *Journal of Econometrics*, 64, 207–40.

MEAGER, R. (2022): "Aggregating distributional treatment effects: A bayesian hierarchical analysis of the microcredit literature," *American Economic Review*, 112, 1818–47.

MENG, X.-L., AND D. A. VAN DYK. (1999): "Seeking efficient data augmentation schemes via conditional and marginal augmentation," *Biometrika*, 86, 301–20.

MITCHELL, T. J., AND J. J. BEAUCHAMP. (1988): "Bayesian variable selection in linear regression," *Journal of the american statistical association*, 83, 1023–32.

MORDUCH, J. (1999): "The microfinance promise," *Journal of economic literature*, 37, 1569–1614.

NAVARRO, S. (2010): "Control Functions," in *Microeconometrics*, Springer, 20–28.

NOBILE, A. (2000): "Comment: Bayesian multinomial probit models with a normalization

constraint," *Journal of Econometrics*, 99, 335–45.

OGBURN, E. L., O. SOFRYGIN, I. DIAZ, AND M. J. VAN DER LAAN. (2022): "Causal inference for social network data," *Journal of the American Statistical Association*, 1–15.

PARK, T., AND G. CASELLA. (2008): "The bayesian lasso," *Journal of the American Statistical Association*, 103, 681–86.

PENG, S. (2019): "Heterogeneous endogenous effects in networks," *arXiv preprint arXiv:1908.00663*,.

POIRIER, D. J., AND J. L. TOBIAS. (2003): "On the predictive distributions of outcome gains in the presence of an unidentified parameter," *Journal of Business & Economic Statistics*, 21, 258–68.

POLSON, N. G., AND J. G. SCOTT. (2010): "Shrink globally, act locally: Sparse bayesian regularization and prediction," *Bayesian statistics*, 9, 105.

POLSON, N. G., AND V. SOKOLOV. (2019): "Bayesian regularization: From tikhonov to horseshoe," *Wiley Interdisciplinary Reviews: Computational Statistics*, 11, e1463.

QU, X., AND L. LEE. (2015): "Estimating a spatial autoregressive model with an endogenous spatial weight matrix," *Journal of Econometrics*, 184, 209–32.

RAFTERY, A. (2007): "Estimating the integrated likelihood via posterior simulation using the harmonic mean identity," *Bayesian Statistics*, 8, 1–45.

RAY, K., AND A. VAN DER VAART. (2020): "Semiparametric bayesian causal inference," *The Annals of Statistics*, 48(5), 2999-3020.

ROBERT, C. P. (1995): "Simulation of truncated normal variables," *Statistics and computing*, 5, 121–25.

ROBERTS, G. O., AND J. S. ROSENTHAL. (2009): "Examples of adaptive MCMC," *Journal of computational and graphical statistics*, 18, 349–67.

ROBINS, J. M., M. A. HERNÁN, AND L. WASSERMAN. (2015): "On bayesian estimation of marginal structural models," *Biometrics*, 71, 296.

ROBINS, J. M., AND Y. RITOV. (1997): "Toward a curse of dimensionality appropriate (CODA) asymptotic theory for semi-parametric models," *Statistics in medicine*, 16, 285–319.

ROSENBAUM, P. R., AND D. B. RUBIN. (1983): "The central role of the propensity score in observational studies for causal effects," *Biometrika*, 70, 41–55.

ROSSI, P. E., G. M. ALLENBY, AND S. MISRA. (2024): *Bayesian Statistics and Marketing*, John Wiley & Sons.

ROTHE, C. (2010): "Nonparametric estimation of distributional policy effects," *Journal of Econometrics*, 155, 56–70.

RUBIN, D. B. (1974): "Estimating causal effects of treatments in randomized and nonrandomized studies." *Journal of educational Psychology*, 66, 688.

---. (1981): "The bayesian bootstrap," *The annals of statistics*, 130–34.

---. (1986): "Statistics and causal inference: Comment: Which ifs have causal answers," *Journal of the American Statistical Association*, 81, 961–62.

SAARELA, O., L. R. BELZILE, AND D. A. STEPHENS. (2016): "A bayesian view of doubly robust causal inference," *Biometrika*, 103, 667–81.

SANCHEZ-BECERRA, A. (2022): "The network propensity score: Spillovers, homophily, and selection into treatment," *arXiv preprint arXiv:2209.14391*,.

SCHARFSTEIN, D. O., A. ROTNITZKY, AND J. M. ROBINS. (1999): "Adjusting for nonignorable drop-out using semiparametric nonresponse models," *Journal of the american statistical association*, 1096–1120.

SCHICKS, J. (2013): "From a Supply Gap to a Demand Gap? The Risk and Consequences of over-Indebting the Underbanked," in *Microfinance in developing countries: Issues, policies and performance evaluation*, Springer, 152–77.

SCHIELE, V., AND H. SCHMITZ. (2016): "Quantile treatment effects of job loss on health," *Journal of Health Economics*, 49, 59–69.

SERFLING, R. J. (2009): *Approximation Theorems of Mathematical Statistics*, John Wiley & Sons.

SHIN, H., AND J. ANTONELLI. (2023): "Improved inference for doubly robust estimators of heterogeneous treatment effects," *Biometrics*, 79(4), 3140-3152.

SOBEL, M. E. (2006): "What do randomized studies of housing mobility demonstrate? Causal inference in the face of interference," *Journal of the American Statistical Association*, 101, 1398–1407.

SPARAPANI, R., C. SPANBAUER, AND R. MCCULLOCH. (2021): "Nonparametric machine learning and efficient computation with bayesian additive regression trees: The BART r

package," *Journal of Statistical Software*, 97, 1–66.

SPERTUS, J. V., AND S.-L. T. NORMAND. (2018): "Bayesian propensity scores for high-dimensional causal inference: A comparison of drug-eluting to bare-metal coronary stents," *Biometrical Journal*, 60, 721–33.

TANNER, M. A., AND W. H. WONG. (1987): "The calculation of posterior distributions by data augmentation," *Journal of the American statistical Association*, 82, 528–40.

TIBSHIRANI, R. J., AND B. EFRON. (1993): "An introduction to the bootstrap," *Monographs on statistics and applied probability*, 57, 1–436.

TOULIS, P., AND E. KAO. (2013): "Estimation of Causal Peer Influence Effects," *International conference on machine learning*, PMLR, 1489–97.

VAN DER LAAN, M. J., S. ROSE, ET AL. (2011): *Targeted Learning: Causal Inference for Observational and Experimental Data,*. Vol. 4 Springer.

VAN DER VAART, A. W. (2000): *Asymptotic Statistics,*. Vol. 3 Cambridge university press.

VAN ERP, S., D. L. OBERSKI, AND J. MULDER. (2019): "Shrinkage priors for bayesian penalized regression," *Journal of Mathematical Psychology*, 89, 31–50.

VIHOLA, M. (2022): "Bayesian Inference with Adaptive Markov Chain Monte Carlo," in *Computational statistics in data science*, ed. by Piegorsch, W. W., R. A. Levine, H. H. Zhang, and T. C. M. Lee. Chichester: John Wiley & Sons.

WENG, H., AND O. PARENT. (2023): "Beyond homophilic dyadic interactions: The impact of network formation on individual outcomes," *Statistics and Computing*, 33, 43.

WHEELER, H. (2022): "Locally optimal place-based policies: Evidence from opportunity zones," *Unpublished working paper. https://hbwheeler. github. io/files/JMP_HW. pdf*,.

WILSON, D. J. (2009): "Beggar thy neighbor? The in-state, out-of-state, and aggregate effects of r&d tax credits," *The Review of Economics and Statistics*, 91, 431–36.

XU, D., M. J. DANIELS, AND A. G. WINTERSTEIN. (2018): "A bayesian nonparametric approach to causal inference on quantiles," *Biometrics*, 74, 986–96.

YIU, A., E. FONG, C. HOLMES, AND J. ROUSSEAU. (2025): "Semiparametric posterior corrections," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, qkaf005.

YUAN, Y., AND K. M. ALTENBURGER. (2022): "A Two-Part Machine Learning Approach to

Characterizing Network Interference in a/b Testing,"SSRN.

ZHANG, Z., Z. CHEN, J. F. TROENDLE, AND J. ZHANG. (2012): "Causal inference on quantiles with an obstetric application," *Biometrics*, 68, 697–706.

ZIGLER, C. M. (2016): "The central role of bayes' theorem for joint estimation of causal effects and propensity scores," *The American Statistician*, 70, 47–54.

ZIGLER, C. M., AND F. DOMINICI. (2014): "Uncertainty in propensity score estimation: Bayesian methods for variable selection and model-averaged causal effects," *Journal of the American Statistical Association*, 109, 95–107.

Appendix A

Appendix for Chapter 2

A.1 Doubly Robust Estimator

To derive the doubly robust estimator for potential quantiles q_t $(t = 0, 1)$ as proposed in section 2.4 in the paper, we adopt the general strategy outlined by Kennedy (2024) and Hines et al. (2022). Without loss of generality, the following discussion focuses on the *τ-quantile of treated potential outcome*, denoted by $q_1(\tau)$. Let $\psi(\mathcal{P}_o)$ represent this estimand of interest, where $\mathcal{P}_o$ is the true joint distribution of observed data $Z_i = \{Y_i, T_i, X_i\}$. The procedure first requires calculation of the estimand's efficient influence function[1]. Next, an estimator based on the efficient influence function is constructed. Finally, the asymptotic properties of the doubly robust estimator are briefly verified.

A.1.1 Deriving Influence Functions

Definition 1. For a given functional $\psi(.)$, the influence function for ψ is the function φ satisfying

$$\frac{\partial \psi(\mathcal{P} + \epsilon(\tilde{\mathcal{P}} - \mathcal{P}))}{\partial \epsilon}\bigg|_{\epsilon=0} = \int \varphi(z; \mathcal{P})\{\tilde{p}(z) - p(z)\}dz, \tag{A.1}$$

and $\int \varphi(z; \mathcal{P})p(z)dz = 0$ for any distribution $\mathcal{P}$ and $\tilde{\mathcal{P}}$ with densities p and $\tilde{p}$. The left-hand side measures the sensitivity of $\psi(\mathcal{P})$ to small changes (slight perturbations) in the underlying distribution $\mathcal{P}$, in the direction of a fixed, deterministic distribution $\tilde{\mathcal{P}}$. This quantity is known as the *Gateaux derivative* (Serfling, 2009).

To simplify the calculation of the efficient influence function, we follow the "point mass contamination" strategy. In particular, we can isolate $\varphi(z; \mathcal{P})$ by setting $\tilde{\mathcal{P}}$ equal to a point mass at single observation $\tilde{z}$, denoted by $\mathbb{1}_{\tilde{z}}(z)$. Equation (A.1) reduces to

[1] Efficiency refers to locally minimax semiparametric efficiency.

$$\frac{\partial \psi(\mathcal{P} + \epsilon(\mathbb{1}_{\tilde{z}} - \mathcal{P}))}{\partial \epsilon} \bigg|_{\epsilon=0} = \varphi(\tilde{z}; \mathcal{P}). \tag{A.2}$$

It should be noted that we focus here on perturbations in the direction parameterised via the one-dimensional mixture model

$$\mathcal{P}_\epsilon = \epsilon \mathbb{1}_{\tilde{z}} + (1 - \epsilon)\mathcal{P}, \quad \epsilon \in [0, 1], \tag{A.3}$$

which is called a parametric submodel. Hence, the efficient influence function at observation $\tilde{z}$ is

$$\varphi\left(\tilde{z}; \mathcal{P}\right) = \frac{d\psi\left(\mathcal{P}_\epsilon\right)}{d\epsilon} \bigg|_{\epsilon=0}. \tag{A.4}$$

Building on this general definition, we can calculate the efficient influence function in the context of estimating quantiles of treated potential outcome.

Theorem 1. *(Efficient Influence Function)*
Denote by $\psi_o := \psi(\mathcal{P}_o)$ the τ-quantile of treated potential outcome under the true joint distribution of observed data. The efficient influence function of ψ_o is equal to

$$\varphi(Z; \mathcal{P}_o) = -\frac{1}{f(\psi_o)} \left\{ \frac{\mathbb{1}\{T = 1\}}{\pi(X)} \left[\mathbb{1}\{Y \leq \psi_o\} - G(\psi_o \mid 1, X; \mathcal{P}_o)\right] + G(\psi_o \mid 1, X; \mathcal{P}_o) - \tau \right\}, \tag{A.5}$$

where $\pi(X; \mathcal{P}_o) = \mathbb{P}\left(T = 1 \mid X; \mathcal{P}_o\right)$ and $G(\psi \mid 1, X; \mathcal{P}_o) = \mathbb{P}\left(Y \leq \psi \mid T = 1, X; \mathcal{P}_o\right)$ are the propensity score and the conditional distribution of treated potential outcome, respectively.

Proof of Theorem 1.

By definition, $\psi_\epsilon = \psi\left(\mathcal{P}_\epsilon\right)$ satisfies

$$\iint \mathbb{1}\left\{y \leq \psi_\epsilon\right\} f_\epsilon\left(y \mid 1, x\right) f_\epsilon\left(x\right) dydx = \tau. \tag{A.6}$$

Denote

$$Q\left(\psi, \epsilon\right) = \iint \mathbb{1}\left\{y \leq \psi_\epsilon\right\} f_\epsilon\left(y \mid 1, x\right) f_\epsilon\left(x\right) dydx - \tau, \tag{A.7}$$

which results in $Q\left(\psi_\epsilon, \epsilon\right) = 0$ and $Q\left(\psi_o, 0\right) = 0$.

Also

$$Q\left(\psi,\epsilon\right)=\int_{-\infty}^{\psi}f_{\epsilon}\left(y\right)dy-\tau, \tag{A.8}$$

hence

$$\left[\frac{\partial Q}{\partial\psi}\right]_{(\psi_o,0)}=f\left(\psi_o\right). \tag{A.9}$$

By the Implicit Function Theorem

$$\left.\frac{d\psi_{\epsilon}}{d\epsilon}\right|_{\epsilon=0}=-\left[\frac{\partial Q}{\partial\psi}\right]_{(\psi_0,0)}^{-1}\times\left[\frac{\partial Q}{\partial\epsilon}\right]_{(\psi_o,0)}=-\frac{1}{f\left(\psi_o\right)}\times\left.\frac{dQ\left(\psi_o,\epsilon\right)}{d\epsilon}\right|_{\epsilon=0} \tag{A.10}$$

By the Chain Rule

$$\begin{aligned}\left.\frac{dQ\left(\psi_o,\epsilon\right)}{d\epsilon}\right|_{\epsilon=0}&=\frac{d}{d\epsilon}\left\{\iint\mathbb{1}\left(y\leq\psi_o\right)\frac{f_{\epsilon}\left(y,1,x\right)f_{\epsilon}\left(x\right)}{f_{\epsilon}\left(1,x\right)}dydx\right\}\Bigg|_{\epsilon=0}\\&=\iint\mathbb{1}\left(y\leq\psi_o\right)\left\{\frac{f_{\epsilon}\left(x\right)}{f_{\epsilon}\left(1,x\right)}\frac{d}{d\epsilon}f_{\epsilon}\left(y,1,x\right)\Bigg|_{\epsilon=0}-\frac{f_{\epsilon}\left(y,1,x\right)f_{\epsilon}\left(x\right)}{f_{\epsilon}\left(1,x\right)^2}\frac{d}{d\epsilon}f_{\epsilon}\left(1,x\right)\Bigg|_{\epsilon=0}\right.\\&\quad\left.+\frac{f_{\epsilon}\left(y,1,x\right)}{f_{\epsilon}\left(1,x\right)}\frac{d}{d\epsilon}f_{\epsilon}\left(x\right)\Bigg|_{\epsilon=0}\right\}dydx\\&=\iint\mathbb{1}\left(y\leq\psi_o\right)\frac{f(y,1,x)f(x)}{f(1,x)}\left(\frac{\mathbb{1}_{\tilde{y},\tilde{t},\tilde{x}}(y,1,x)}{f(y,1,x)}-\frac{\mathbb{1}_{\tilde{t},\tilde{x}}(1,x)}{f(1,x)}+\frac{\mathbb{1}_{\tilde{x}}(x)}{f(x)}-1\right)dydx\\&=\frac{\mathbb{1}_{\tilde{t}}(1)}{\pi(\tilde{x};\mathcal{P}_o)}\left[\mathbb{1}\left(\tilde{y}\leq\psi_o\right)-G\left(\psi_o\mid1,\tilde{x};\mathcal{P}_o\right)\right]+G\left(\psi_o\mid1,\tilde{x};\mathcal{P}_o\right)-\tau\end{aligned} \tag{A.11}$$

Hence

$$\begin{aligned}\varphi\left(\tilde{z};\mathcal{P}_o\right)&=\left.\frac{d\psi\left(\mathcal{P}_{\epsilon}\right)}{d\epsilon}\right|_{\epsilon=0}\\&=-\frac{1}{f\left(\psi_o\right)}\left\{\frac{\mathbb{1}_{\tilde{t}}(1)}{\pi(\tilde{x};\mathcal{P}_o)}\left[\mathbb{1}\left(\tilde{y}\leq\psi_o\right)-G\left(\psi_o\mid1,\tilde{x};\mathcal{P}_o\right)\right]+G\left(\psi_o\mid1,\tilde{x};\mathcal{P}_o\right)-\tau\right\}.\end{aligned} \tag{A.12}$$

A.1.2 Efficient Influence Function-Based Estimator

Let $\mathcal{P}_n$ denote the empirical distribution from a sample of size n. Denote $h_o(X,\psi)=G(\psi\mid 1,X;\mathcal{P}_o)$; and $\pi_o(X)=\pi(X;\mathcal{P}_o)$. Then, $h_o(X,\psi)$ and $\pi_o(X)$ are function-valued nuisance parameters of the estimation problem for the τ-quantile of treated potential outcome, which is our target parameter denoted by $\psi_o\coloneqq\psi(\mathcal{P}_o)$.

The moment condition associated with the efficient influence function in (A.5) to identify the target parameter value ψ_o is

$$\mathbb{E}\left[\varphi(Z; h_o(X, \psi_o), \pi_o(X)), \psi_o\right] = 0, \tag{A.13}$$

in which the moment has zero derivative with respect to nuisances at ψ_o, h_o and π_o. Intuitively, this moment condition satisfies Neyman orthogonality, that is the first-order insensitivity of target parameter value to local perturbations of the values of nuisance parameters. This property is desirable because it helps ensure that the estimation of the parameter of interest remains robust even when there are small errors or uncertainties in the estimation of nuisance parameters. When regularization methods are needed to handle high-dimensional covariates or nonlinearities when estimating nuisance parameters, the use of Neyman orthogonal moment conditions help eliminate the first-order biases stemming from these plugging-in estimators (see e.g., Belloni et al., 2017; Chernozhukov et al., 2018; Kallus et al., 2024).

Therefore, the efficient influence function-based estimator for $\psi(\mathcal{P}_o)$ is defined as a solution to the estimating equation

$$\begin{aligned} &\mathbb{E}_{\mathcal{P}_n}\left[\varphi(Z; \hat{h}(X, \psi), \hat{\pi}(X)), \psi\right] = 0 \\ \Leftrightarrow\ &\frac{1}{n}\sum_{i=1}^{n} \frac{\mathbb{1}\{T_i = 1\}}{\hat{\pi}(X_i)}\left[\mathbb{1}\{Y_i \leq \psi\} - \hat{h}(X_i, \psi)\right] + \hat{h}(X_i, \psi) - \tau = 0. \end{aligned} \tag{A.14}$$

Denote by $\widehat{\psi}^{dr}$ the resulting estimator from (A.14). We will show that $\widehat{\psi}^{dr}$ is a doubly robust estimator, which is consistent provided that either one of the nuisance estimators – $\hat{h}$ or $\hat{\pi}$ – is consistent, but not necessarily both.

Lemma 1. (*Double Robustness of Efficient Influence Function*)
Let $\eta^* = (h^*, \pi^*)$ with either $h^* = h_o$ or $\pi^* = \pi_o$. Then $\mathbb{E}_{\mathcal{P}_o}\left[\varphi(\eta^*, \psi_o) = 0\right]$.

Sketch of Proof for Lemma 1.

- By the law of iterated expectation

$$\begin{aligned} \mathbb{E}_{\mathcal{P}_o}\left[\varphi(\eta^*, \psi)\right] &= -\frac{1}{f(\psi)}\mathbb{E}_{\mathcal{P}_o}\left[\frac{\pi_o}{\pi^*}\left(h_{o,\psi} - h^*_{\psi}\right) + h^*_{\psi} - \tau\right] \\ &= -\frac{1}{f(\psi)}\mathbb{E}_{\mathcal{P}_o}\left[\left(\frac{\pi_o}{\pi^*} - 1\right)\left(h_{o,\psi} - h^*_{\psi}\right) + h_{o,\psi} - \tau\right]. \end{aligned}$$

- When either $\pi^* = \pi_o$ or $h^* = h_o$, substituting ψ with ψ_o leads to $\mathbb{E}_{\mathcal{P}_o}[\varphi(\eta^*, \psi_o)] = 0$, and the lemma follows.

Theorem 2. (*Consistency of the Point Estimator*)
Under Identifying Assumptions 2.2-2.4 and additional Regularity Conditions below, $\widehat{\psi}^{dr}$ is consistent if either nuisance estimator $\hat{h}$ or $\hat{\pi}$ is consistent:

1. The cumulative distribution function F has compact support $[a, b] \subset \mathbb{R}$ and is continuously differentiable on its support with strictly positive derivative f.
2. The class function $\{\varphi(\eta, \psi) : |\psi - \psi_o| < \delta, \|h_\psi - h^*_\psi\| < \delta, \|\pi_\psi - \pi^*_\psi\| < \delta\}$ is Donsker for some $\delta > 0$ and such that $\mathcal{P}_o\{\varphi(\eta, \psi) - \varphi(\eta^*, \psi_o)\}^2 \to 0$ as $(\eta, \psi) \to (\eta^*, \psi_o)$.

Sketch of Proof for Theorem 2.

- By construction of $\widehat{\psi}^{dr}$ in (A.14) we have $\mathbb{E}_{\mathcal{P}_n}\varphi(\hat{\eta}, \widehat{\psi}^{dr}) = 0$, where $\hat{\eta} = (\hat{h}(X, \widehat{\psi}^{dr}), \hat{\pi})$.
- By Lemma 1 we have $\mathbb{E}_{\mathcal{P}_o}[\varphi(\eta^*, \psi_o)] = 0$.
- An application of Theorem 5.9 of Van der Vaart (2000) yields $\widehat{\psi}^{dr} = \psi_o + o_{\mathcal{P}}(1)$, thereby $\widehat{\psi}^{dr}$ is consistent. This completes the proof.

A.1.3 Bernstein–von Mises (BvM) Theorem

Definition 2. *(Asymptotic efficiency)*

A sequence of regular estimators $\hat{\psi}_n = \hat{\psi}_n(Z^{(n)})$ is said to be *asymptotically efficient* at $\mathcal{P}_o$ if

$$\sqrt{n}\Big(\hat{\psi}_n - \psi(\mathcal{P}_o)\Big) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\varphi(Z_i; \mathcal{P}_o) + o_{\mathcal{P}_o}(1), \tag{A.15}$$

where $\varphi(Z; \mathcal{P}_o)$ is the efficient influence function in accordance with (A.5). We define the variance $V_o := \mathbb{E}_{\mathcal{P}_o}[\varphi^2(Z_i; \mathcal{P}_o)]$.

Definition 3. *(Semiparametric BvM)*

Let $\mathcal{L}_\Pi\Big(\sqrt{n}\Big(\psi(\mathcal{P}) - \hat{\psi}_n\Big) \mid Z^{(n)}\Big)$ denote the posterior law of $\sqrt{n}\Big(\psi(\mathcal{P}) - \hat{\psi}_n\Big)$, where $\hat{\psi}_n$ is any sequence of estimators satisfying (A.15). The posterior satisfies the *semiparametric Bernstein–von Mises (BvM) theorem* if

$$d_{\mathrm{BL}}\Big(\mathcal{L}_\Pi\Big(\sqrt{n}\Big(\psi(\mathcal{P}) - \hat{\psi}_n\Big) \mid Z^{(n)}\Big),\ \mathcal{N}\Big(0, V_o\Big)\Big) \underset{\mathcal{P}_o}{\longrightarrow} 0.$$

where d_{BL} is the *bounded Lipschitz distance*. In simpler terms, we can say that the posterior

law of $\sqrt{n}\big(\psi(\mathcal{P}) - \hat{\psi}_n\big)$ "converges weakly to $\mathcal{N}\big(0, V_0\big)$ in probability" (Yiu et al., 2025).

Assumption A1. *(Nuisance convergence rate and complexity)*

There exists a sequence of measurable sets $(\mathcal{H}_n)_n$ of $\mathcal{P}$ satisfying $\Pi(\mathcal{P} \in \mathcal{H}_n \mid Z^{(n)}) \underset{\mathcal{P}_o}{\longrightarrow} 1$ such that:

(a) *(Rates of convergence of π and h)* There exist numbers $\rho_n, \varepsilon_n \to 0$ such that

$$\sup_{\mathcal{P}\in\mathcal{H}_n} \|\pi - \pi_o\|_{\mathcal{P}_o} \le \rho_n, \qquad \sup_{\mathcal{P}\in\mathcal{H}_n} \|h - h_o\|_{\mathcal{P}_o} \le \varepsilon_n, \quad \text{and} \quad \sqrt{n}\,\rho_n\,\varepsilon_n \to 0.$$

where $\|\cdot\|_{\mathcal{P}_o}$ denote $L_2(\mathcal{P}_o)$ norm.

(b) *(Uniform bound)* Overlap and tail bounds hold uniformly on $\mathcal{H}_n$.

(c) *(Donsker class)* The sequences of sets $\{\pi : \mathcal{P} \in \mathcal{H}_n\}$ and $\{h : \mathcal{P} \in \mathcal{H}_n\}$ are both eventually contained in fixed $\mathcal{P}_o$–Donsker classes.

Theorem 3. *(BvM for BADR quantiles)*

Under Identifying Assumptions 2.2-2.4 and Assumption A1, the posterior induced by the Bayesian analogue of doubly robust procedure for $\psi := q_1(\tau)$ (shown in Algorithm 2.1) satisfies the semiparametric Bernstein–von Mises theorem:

$$d_{\mathrm{BL}}\Big(\mathcal{L}_{\Pi}\Big(\sqrt{n}\Big(\psi - \hat{\psi}^{dr}\Big) \mid Z^{(n)}\Big),\, \mathcal{N}\Big(0, V_o\Big)\Big) \underset{\mathcal{P}_o}{\longrightarrow} 0.$$

where $\hat{\psi}^{dr}$ is any efficient DR estimator solving (A.14).

Discussion

- The semiparametric BvM result implies that the constructed posterior for ψ is asymptotically normal, centered at an efficient estimator, and with the variance equal to the semiparametric efficiency bound. This also provides guarantees that the point estimator (posterior mean/median) derived from the proposed procedure is semiparametrically efficient.
- Rate double robustness property: By orthogonality, the first-order sensitivity to nuisance error vanishes. The second-order remainder for $\hat{\psi}^{dr}$ (or any plug-in posterior draw $\psi^{(b)}$)

has a cross-term structure in terms of π and h:

$$\|\pi - \pi_0\| \, \|h - h_0\| \; + \; \|\pi - \pi_0\|^2 \; + \; \|h - h_0\|^2,$$

that allows a faster rate of convergence for one nuisance parameter to compensate for a slower rate for the other. Assumption A1(a) states that π and h must both converge uniformly on $\mathcal{H}_n$ to their respective truths in L_2, and their combined rate of convergence must be faster than $n^{-1/2}$ to ensure $o(n^{-1/2})$ remainder and deliver the BvM limit. For instance, in a symmetric case, both nuisances (or their posteriors) contract faster than $n^{-1/4}$ would be sufficient.

- Misspecification of priors for the nuisances: Let $(\pi^\star, h^\star)$ denote the pseudo-true limits of the nuisance posteriors under misspecified priors. Evaluating the efficient influence function mean at ψ_0 gives

$$\mathbb{E}_{\mathcal{P}_o}[\varphi(Z; \pi^\star, h^\star, \psi_o)] = -\frac{1}{f(\psi_o)} \mathbb{E}_{\mathcal{P}_o}\left[\left(\frac{\pi_o(X)}{\pi^\star(X)} - 1\right)\left(h_{o,\psi_o}(X) - h^\star_{\psi_o}(X)\right)\right],$$

which is zero if *either* $\pi^\star = \pi_o$ *or* $h^\star = h_o$ (i.e. double robustness regarding the misspecification of one nuisance), but is generally nonzero if both nuisance posteriors contract to pseudo-truths $\pi^\star \neq \pi_o$ and $h^\star \neq h_o$. In that case, the posterior concentrates at a pseudo-true quantile $\psi^\dagger \neq \psi_o$ solving the biased moment, and a BvM may still hold but centered at $\psi^\dagger$ rather than ψ_o, yielding asymptotically biased inference.

A.2 Bayesian Additive Regression Tree (BART)

A.2.1 BART Model Specifications

BART is a nonparametric modelling technique that translates decision tree-based ensemble methods to a Bayesian framework. Chipman et al. (2012) present a comprehensive overview of the method. In essence, BART is a *sum-of-trees* model with prior distributions are placed over the parameters including *tree depth*, *splitting variables*, *splitting values*, and *terminal node* estimates.

Consider the regression problem that predicts a continuous Y_i using a p-dimensional vector of predictors $\mathbf{X}_i = (X_{i1}, \ldots, X_{ip})^\top (i = 1, \ldots, N)$, BART model can be expressed as

$$Y_i = f_{\text{BART}}(\mathbf{X}_i) + \epsilon_i, \quad \epsilon \overset{iid}{\sim} \mathcal{N}(0, \sigma^2), \quad f_{\text{BART}}(\mathbf{X}_i) = \sum_{m=1}^{M} f_{\text{tree}}(\mathbf{X}_i; \Gamma_m, \mu_m), \tag{A.16}$$

where $f_{\text{tree}}(\mathbf{X}_i; \Gamma_m, \mu_m)$ is a Bayesian single regression tree; Γ_m is a tree structure that consists of a set of splitting rules and a set of terminal nodes; and $\mu = (\mu_{m,1}, \ldots, \mu_{m,b_m})$ is a vector of parameters associated with b_m terminal nodes of Γ_m, such that $f_{\text{tree}}(\mathbf{X}_i; \Gamma_m, \mu_m) = \mu_{m,l}$ if $\mathbf{X}_i$ is corresponding to the l^{th} terminal node of Γ_m.

The prior of BART is specified for three components:

1. The ensemble structure $\{\Gamma_m\}_{m=1}^M$

   Independent regularization prior is placed on Γ_m. It consists of a Bernoulli distribution with probability

   $$\Pr(\text{split} \mid d) = \alpha(1 + d)^{-\beta}, \quad \alpha \in (0, 1), \quad \beta \in (0, \infty), \tag{A.17}$$

   for splitting a node at tree depth d ($d = \{0, 1, \ldots)$ into two child nodes and two discrete uniform distributions for selecting a split variable and a split value given the selected split variable. This regularization prior helps prevent individuals from becoming too influential, thereby enhancing the overall fit and mitigating the risk of overfitting.

2. The parameters $\{\mu_m\}_{m=1}^M$ associated with the terminal nodes given $\{\Gamma_m\}_{m=1}^M$

   $$\mu_{m,l} \overset{iid}{\sim} \mathcal{N}(0, v) \tag{A.18}$$

3. The error variance σ^2 that is independent with the former two

$$\sigma^2 \sim \mathcal{IG}\left(r, s\right) \tag{A.19}$$

The process of sampling the posterior distribution is carried out using a Metropolis-within-Gibbs MCMC sampler, which can also be regarded as a special case of (generalised) Bayesian backfitting algorithm (Hastie and Tibshirani, 2000), to update each tree iteratively. Estimated outcome is achieved by averaging the posterior samples of $f_{\text{BART}}(\mathbf{X}_i)$ after a burn-in period.

For binary outcome, the continuous BART model above has been extended to probit BART and logit BART, which are specified as follows

$$\mathbb{P}(Y_i = 1 \mid \mathbf{X}_i) = H[f_{\text{BART}}(\mathbf{X}_i)], \tag{A.20}$$

where $f_{\text{BART}}(\mathbf{X}_i)$ is the sum-of-trees function in (A.16) and H is the link function with the probit link for probit BART and the logit link for logit BART. Both of the models maintain the same prior assigned to the ensemble structure and the parameters of the terminal nodes, i.e. $\{\Gamma_m, \mu_m\}_{m=1}^M$, but σ^2 is fixed for for the sake of identifiability.

Probit BART employs data augmentation of Albert and Chib (1993) to adapt the Bayesian backfitting sampler used in continuous BART. This involves introducing a latent variable Y_i^* such that $Y_i = \mathbb{1}\{Y_i^* > 0\}$ for each response variable Y_i. At each iteration of the MCMC algorithm, Y_i^* is imputed by sampling from the full conditional distribution of Y_i^* given Y_i and other parameters, which is essentially a truncated normal distribution. The imputed Y_i^*'s are then modelled using the continuous BART model with σ^2 set to 1, enabling the completion of the MCMC algorithm by performing the Bayesian backfitting algorithm of the continuous BART model on the imputed Y_i^*'s.

Logit BART also introduces latent variable Y_i^* that is instead assumed to follow a logistic distribution, which has a heavier tail than a normal distribution, thus improving estimation for extreme instances of $\mathbb{P}\left(Y_i = 1 \mid \mathbf{X_i}\right)$. These latent Y_i^*'s are sampled using the method described by Gramacy and Polson (2012). Conditional on the imputed Y_i^*'s, the continuous BART model with given heteroskedastic variance σ_i^2's are fitted on Y_i^*'s, where σ_i^2's are obtained through the technique outlined by Robert (1995).

A.2.2 On Implementation

Computational details and implementation using **BART** R package can be found in Sparapani et al. (2021). The actual sampling and computation are carried out in C++ code to maximize computational efficiency. Both options for Probit BART and Logit BART are available in this package, which can be utilised directly to fit the treatment assignment model in our proposed BADR framework.

A.3 Bayesian Shrinkage Priors

A.3.1 Hierarchical Bayes for Linear Regression

Consider the linear regression model

$$Y_i = \mathbf{X}_i\beta + \epsilon_i, \quad \epsilon_i \overset{iid}{\sim} \mathcal{N}(0, \sigma^2) \tag{A.21}$$

Assuming that interest lies in learning about the regression coefficients β, then a simple hierarchical specification takes the form

$$\begin{aligned} Y_i \mid \beta, \sigma^2 &\sim \mathcal{N}(\mathbf{X}_i\beta, \sigma^2), i = 1 \ldots, n \\ \beta_j \mid \tau^2, \sigma^2 &\sim \mathcal{N}(0, \sigma^2\tau^2), i = 1 \ldots, p \\ \tau^2 &\sim F(a, b) \\ \sigma^2 &\sim \frac{1}{\sigma^2} \end{aligned} \tag{A.22}$$

where $F(a, b)$ denotes some distribution function with hyper-parameters a, b. Due to the fact that choice of τ^2 is so crucial for the posterior outcome of β_j, the idea behind this hierarchical specification is to treat the hyper-parameter τ^2 as a random variable and learn about it via Bayes Theorem.

A.3.2 Bayesian Shrinkage Priors

The major goal of shrinkage priors is to shrink small coefficients to zero while maintaining true large coefficients, especially in high-dimensional settings. The possible variation in shrinkage amounts among those priors depends on their specific designs. In particular, the sharper the peak is around zero, the stronger shrinkage for small coefficients. Also, the heavier the tail, the lighter the shrinkage for large coefficients.

Bayesian Lasso

The Bayesian counterpart of the Lasso penalty is Laplace prior, which was first proposed by Park and Casella (2008). The Bayesian Lasso can be obtained as a scale mixture of a Normal density with an Exponential density as below:

$$\begin{aligned}
\beta_j \mid \tau_j^2, \sigma^2 &\sim \mathcal{N}\left(0, \sigma^2\tau_j^2\right), \\
\tau_j^2 \mid \lambda^2 &\sim \mathcal{E}xp\left(\frac{\lambda^2}{2}\right), \text{ for } j = 1, \ldots, p, \\
\lambda &\sim \text{half-Cauchy}(0, 1), \\
\sigma^2 &\sim \frac{1}{\sigma^2}.
\end{aligned} \tag{A.23}$$

Integrating τ_j^2 out leads to Double-exponential[2] or Laplace priors on the regression coefficients, i.e.,

$$\beta_j \mid \lambda, \sigma \sim \text{Double-exponential}\left(0, \frac{\sigma}{\lambda}\right), \text{ for } j = 1, \ldots, p \tag{A.24}$$

Although this version of Bayesian Lasso is the most popular form in literature so far; there are also some alternative formulations suggested by Hans (2009), Mallick and Yi (2014) and Alhamzawi and Taha Mohammad Ali (2020).

In addition to the overall shrinkage parameter λ, the Lasso prior has an additional predictor-specific shrinkage parameter τ_j. Therefore, the Lasso prior is more flexible than the Ridge prior, which only relies on the overall shrinkage parameter.

Horseshoe prior

A novel shrinkage prior in the Bayesian literature is the horseshoe prior Carvalho et al. (2010)[3]. This prior is particularly attractive for sparse signal recovery.

$$\begin{aligned}
\beta_j \mid \tau_j^2 &\sim \mathcal{N}\left(0, \tau_j^2\right) \\
\tau_j \mid \lambda &\sim \text{half-Cauchy}(0, \lambda), \text{ for } j = 1, \ldots, p \\
\lambda \mid \sigma &\sim \text{half-Cauchy}(0, \sigma)
\end{aligned} \tag{A.25}$$

[2]Mathematical representation:

$$\int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2 s_j}} e^{\left(-\frac{\beta_j^2}{2\sigma^2 s_j}\right)} \frac{\lambda^2}{2} e^{-\frac{\lambda}{2s_j}} ds_j = \frac{\lambda}{2\sqrt{\sigma^2}} e^{-\lambda|\beta_j|/\sqrt{\sigma^2}}$$

[3]Note that Carvalho et al. (2010) explicitly include the half-Cauchy prior for λ in their specification, thereby implying a full Bayes approach. This formulation results in a horseshoe prior that is automatically scaled by the error standard deviation σ.

The half-Cauchy prior can be written as a mixture of Inverse Gamma densities[4] (Makalic and Schmidt, 2015), so that the horseshoe prior in Equation (A.25) can be equivalently specified as:

$$\begin{aligned}
\beta_j \mid \tau_j^2 &\sim \mathcal{N}\left(0, \tau_j^2\right) \\
\tau_j^2 \mid \omega &\sim \mathcal{IG}\left(\frac{1}{2}, \frac{1}{\omega}\right) \\
\omega \mid \lambda^2 &\sim \mathcal{IG}\left(\frac{1}{2}, \frac{1}{\lambda^2}\right) \\
\lambda^2 \mid \gamma &\sim \mathcal{IG}\left(\frac{1}{2}, \frac{1}{\gamma}\right) \\
\gamma \mid \sigma^2 &\sim \mathcal{IG}\left(\frac{1}{2}, \frac{1}{\sigma^2}\right)
\end{aligned} \tag{A.26}$$

An expression for the marginal prior of the regression coefficients β_j is not analytically tractable, but a tight lower bound Carvalho et al. (2010) can be used instead.

$$-\log p\left(\beta_i \mid \lambda\right) \geq -\log\log\left(1 + \frac{2\lambda^2}{\beta_j^2}\right) \tag{A.27}$$

The key features for the appealing performance of horseshoe prior are its Cauchy-like tails and an asymptote at origin (unique advantage), which make horseshoe adaptive to sparsity and robust to large signals so outperform other shrinkage priors we have discussed.

In the search for intuitive reasons, we consider a common framework of shrinkage rules. Define $\kappa_j = 1/(1+\tau_j^2)$, then κ_j is a random shrinkage coefficient in $[0, 1]$. Under a multivariate normal scale mixture prior (i.e. the general form of all shrinkage priors we are discussing), the posterior mean can be written as a linear function of the observation:

$$\mathbb{E}\left[\beta_j \mid Y_j\right] = \left\{1 - \mathbb{E}\left[\kappa_i \mid Y_j\right]\right\} Y_j \tag{A.28}$$

Hence, $\mathbb{E}\left[\kappa_i \mid Y_j\right]$ implies the amount of weight that the posterior mean for β_j places on 0 once the data have been observed. A shrinkage coefficient κ_j that is close to zero leads to virtually no shrinkage, thus describes signals. A shrinkage coefficient κ_j that is close to one leads to nearly-total shrinkage, thus describes noises. Intuitively speaking, the behavior of a priori $p(\kappa_j)$ near $\kappa_j = 1$ will control the robustness of signal at tail, while near $\kappa_j = 0$ will control the shrinkage of noise toward 0. Because of difference choice of $p(\tau_j)$, each type

[4] If $x^2 \mid z \sim \mathcal{IG}(1/2, 1/z)$ and $z \sim \mathcal{IG}\left(1/2, 1/\alpha^2\right)$ then $x \sim C^+(0, \alpha)$

of shrinkage prior has distinct $p(\kappa_j)$ reflecting its attempt to separate signal and noise. For horseshoe prior, the attempt is even implied in its name, which arises from the fact that for fixed values $\lambda = \sigma = 1$, $p(\kappa_j)$ is similar to a horseshoe-shaped Beta $(1/2, 1/2)$. This prior is symmetric and unbounded at both 0 and 1; thereby, small coefficients (noises) are heavily shrunken towards zero while substantial coefficients (signals) remain large. None of these common shrinkage priors above shares this characteristic. For instance, the Laplace prior, where $p(\kappa_j)$ is bounded at both 0 and 1, tends to over-shrink strong signals yet under-shrink noises. Carvalho et al. (2009), Carvalho et al. (2010) provide more explanation for other priors.

In fact, unlike local shrinkage priors above, the horseshoe prior is a member of a wider class of *global-local shrinkage priors* (Bhadra et al., 2019; Polson and Scott, 2010) because it enables a clear separation between global and local shrinkage effects. Put another way, this class of priors adapt to sparsity by a global shrinkage parameter and recover signals by a local shrinkage parameter.

A.4 Bayesian Quantile Regression (BQR)

A.4.1 Bayesian Quantile Regression

Consider the linear quantile regression model (Koenker and Bassett, 1978) at a given quantile level $\tau \in (0,1)$

$$\mathcal{Q}_\tau \left(Y \mid \mathbf{X} \right) = \mathbf{X}\beta_{(\tau)}, \tag{A.29}$$

The quantile specific coefficient $\beta_{(\tau)}$ can be consistently estimated by

$$\hat{\beta}_{(\tau)} = \underset{\beta}{\text{argmin}} \sum_{i=1}^{n} \rho_\tau \left(Y_i - \mathbf{X}_i\beta \right), \tag{A.30}$$

where $\rho_\tau(u) = u\left(\tau - \mathbb{1}\{u < 0\}\right)$ is the quantile loss function. As this functional form is an asymmetric L_1 loss function proportional to the negative log density of the asymmetric Laplace distribution ($\mathcal{ALD}$), the connection allows researcher to recast the quantile regression as a maximum likelihood problem of the linear model $Y_i = \mathbf{X}_i\beta_{(\tau)} + \epsilon_{i,(\tau)}$ where $\epsilon_{i,(\tau)} \sim \mathcal{ALD}\left(\tau, 0, \sigma_{(\tau)}\right)$[5]. The working likelihood is of the form

$$f\left(Y \mid \mathbf{X}, \beta_{(\tau)}, \sigma_{(\tau)}, \tau\right) = \frac{\tau^n(1-\tau)^n}{\sigma^n_{(\tau)}} \exp\left\{ -\sum_{i=1}^{n} \frac{\rho_\tau \left(Y_i - \mathbf{X}_i\beta \right)}{\sigma_{(\tau)}} \right\}. \tag{A.31}$$

The asymmetric Laplace distribution is known to be expressible as a scale mixture of normals (Kotz et al., 2012), we thus can rewrite $\epsilon_{i,(\tau)}$ as follows

$$\epsilon_{i,(\tau)} = \theta_{(\tau)} z_{i,(\tau)} + \kappa_{(\tau)} \sqrt{\sigma_{(\tau)} z_{i,(\tau)}}\, u_i, \quad \text{with } \theta_{(\tau)} = \frac{1-2\tau}{\tau(1-\tau)} \text{ and } \kappa^2_{(\tau)} = \frac{2}{\tau(1-\tau)},$$

where $z_{i,(\tau)} = \sigma_{(\tau)}\zeta_{i,(\tau)}$ with $\zeta_{i,(\tau)} \sim \mathcal{E}xp(1)$, and $u_i \sim \mathcal{N}(0,1)$.

As a result, the Bayesian quantile regression model has the following representation

$$Y_i = \mathbf{X}_i\beta_{(\tau)} + \theta_{(\tau)} z_{i,(\tau)} + \kappa_{(\tau)} \sqrt{\sigma_{(\tau)} z_{i,(\tau)}}\, u_i, \quad \text{for } i = 1, \ldots, n. \tag{A.32}$$

[5] $\epsilon_{i,(\tau)}$ follows asymmetric Laplace distribution with density

$$f_{\mathcal{ALD}}\left(\epsilon_{(\tau)}\right) = \frac{\tau(1-\tau)}{\sigma_{(\tau)}} \exp\left\{ -\rho_\tau\left(\epsilon_{(\tau)}\right) / \sigma_{(\tau)} \right\}.$$

This leads to the following likelihood function:

$$f\left(Y \mid \mathbf{X}, \beta_{(\tau)}, \sigma_{(\tau)}, \mathbf{z}_{(\tau)}, \tau\right) \propto \exp\left\{-\sum_{i=1}^{n} \frac{\left(Y_i - \mathbf{X}_i\beta_{(\tau)} - \theta_{(\tau)}z_{i,(\tau)}\right)^2}{2\kappa^2_{(\tau)}\sigma_{(\tau)}z_{i,(\tau)}}\right\} \prod_{i=1}^{n} \frac{1}{\sqrt{\sigma_{(\tau)}z_{i,(\tau)}}}. \tag{A.33}$$

We assume the priors as below (Kozumi and Kobayashi, 2011)

$$\beta_{(\tau)} \sim \mathcal{N}\left(0, \mathbf{\Sigma}_{0,(\tau)}\right), \tag{A.34}$$

$$z_{i,(\tau)} \sim \mathcal{E}xp\left(\sigma_{(\tau)}\right) \propto \sigma^{-1}_{(\tau)} \exp\left\{-\sigma^{-1}_{(\tau)}z_{i,(\tau)}\right\} \quad \forall i = 1, \ldots, n, \tag{A.35}$$

$$\sigma_{(\tau)} \sim \mathcal{IG}\left(r_{0,(\tau)}, s_{0,(\tau)}\right) \propto \left(\sigma^{-1}_{(\tau)}\right)^{r_{0,(\tau)}+1} \exp\left\{-s_{0,(\tau)}\sigma^{-1}_{(\tau)}\right\}; \tag{A.36}$$

where for simplicity, $\mathbf{\Sigma}_{0,(\tau)} = \mathbf{D}_{\lambda,(\tau)} = \lambda \times \mathbf{I}_p$ where λ is fixed and known for all τ.

The conditional posteriors are of the form

$$\beta_{(\tau)} \mid \bullet \sim \mathcal{N}_p\left(\mu_{\beta,(\tau)}, \mathbf{\Sigma}_{\beta,(\tau)}\right), \tag{A.37}$$

$$z_{i,(\tau)} \mid \bullet \sim \mathcal{GIG}\left(\frac{1}{2}, a_{i,(\tau)}, b_{i,(\tau)}\right) \propto z^{-\frac{1}{2}}_{i,(\tau)} \exp\left\{-\frac{1}{2}\left(a_{i,(\tau)}z_{i,(\tau)} + b_{i,(\tau)}z^{-1}_{i,(\tau)}\right)\right\}, \tag{A.38}$$

$$\sigma_{(\tau)} \mid \bullet \sim \mathcal{IG}\left(r_{\sigma,(\tau)}, s_{\sigma,(\tau)}\right) \propto \left(\sigma^{-1}_{(\tau)}\right)^{r_{\sigma,(\tau)}+1} \exp\left\{-s_{\sigma,(\tau)}\sigma^{-1}_{(\tau)}\right\}; \tag{A.39}$$

where

$$\mathbf{\Sigma}_{\beta,(\tau)} = \left(\mathbf{X}^\top\mathbf{U}^{-1}\mathbf{X} + \mathbf{\Sigma}^{-1}_{0,(\tau)}\right)^{-1} \text{ and } \mu_{\beta,(\tau)} = \mathbf{\Sigma}_{\beta,(\tau)} \times \mathbf{X}^\top\mathbf{U}^{-1}\left(\mathbf{Y} - \theta_{(\tau)}\mathbf{z}_{(\tau)}\right),$$

$$\mathbf{U} = \left(\sigma_{(\tau)}\kappa^2_{(\tau)}\right) \times \text{diag}\left(\mathbf{z}_{(\tau)}\right), \quad \mathbf{z}_{(\tau)} = \left(z_{1,(\tau)}, \ldots, z_{n,(\tau)}\right)^\top,$$

$$a_{i,(\tau)} = \frac{1}{\sigma_{(\tau)}}\left(2 + \frac{\theta^2_{(\tau)}}{\kappa^2_{(\tau)}}\right) \text{ and } b_{i,(\tau)} = \frac{\left(Y_i - \mathbf{X}_i\beta_{(\tau)}\right)^2}{\sigma_{(\tau)}\kappa^2_{(\tau)}},$$

$$r_{\sigma,(\tau)} = r_{0,(\tau)} + \frac{3n}{2} \text{ and } s_{\sigma,(\tau)} = s_{0,(\tau)} + \sum_{i=1}^{n} \frac{\left(Y_i - \mathbf{X}_i\beta_{(\tau)} - \theta_{(\tau)}z_{i,(\tau)}\right)^2}{2\kappa^2_{(\tau)}z_{i,(\tau)}} + \sum_{i=1}^{n} z_{i,(\tau)};$$

for $i = 1, \ldots, n$.

A.4.2 Bayesian Quantile Regression with the Adaptive Lasso

Bayesian Quantile Regression with the Adaptive Lasso is a Bayesian hierarchical model given by

$$Y_i = \mathbf{X}_i\beta_{(\tau)} + \theta_{(\tau)}z_{i,(\tau)} + \kappa_{(\tau)}\sqrt{\sigma_{(\tau)}z_{i,(\tau)}}u_i; \tag{A.40}$$

$$u_i \sim \mathcal{N}(0,1), \tag{A.41}$$

$$z_{i,(\tau)} \sim \mathcal{E}xp\left(\sigma_{(\tau)}\right) \propto \sigma_{(\tau)}^{-1}\exp\left\{-\sigma_{(\tau)}^{-1}z_{i,(\tau)}\right\} \tag{A.42}$$

$$\beta_{j,(\tau)}, v_{j,(\tau)} \mid \sigma_{(\tau)}, \lambda_{j,(\tau)}^2 \sim \frac{1}{\sqrt{2\pi v_{j,(\tau)}}}\exp\left\{-\frac{\beta_{j,(\tau)}^2}{2v_{j,(\tau)}}\right\}\frac{\sigma_{(\tau)}^{-1}}{2\lambda_{j,(\tau)}^2}\exp\left\{\frac{-\sigma_{(\tau)}^{-1}}{2\lambda_{j,(\tau)}^2}v_{j,(\tau)}\right\}, \tag{A.43}$$

$$\lambda_{j,(\tau)}^2 \sim \mathcal{IG}\left(c_{0,(\tau)}, d_{0,(\tau)}\right) \propto \left(\frac{1}{\lambda_{j,(\tau)}^2}\right)^{c_{0,(\tau)}+1}\exp\left\{-\frac{d_{0,(\tau)}}{\lambda_{j,(\tau)}^2}\right\}, \tag{A.44}$$

$$\sigma_{(\tau)} \sim \mathcal{IG}\left(r_{0,(\tau)}, s_{0,(\tau)}\right) \propto \left(\sigma_{(\tau)}^{-1}\right)^{r_{0,(\tau)}+1}\exp\left\{-s_{0,(\tau)}\sigma_{(\tau)}^{-1}\right\}; \tag{A.45}$$

for $i = 1, \ldots, n$ and $j = 1, \ldots, p$.

The conditional posteriors (Alhamzawi et al., 2012) are of the form

$$z_{i,(\tau)} \mid \bullet \sim \mathcal{GIG}\left(\frac{1}{2}, a_{i,(\tau)}, b_{i,(\tau)}\right) \propto z_{i,(\tau)}^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\left(a_{i,(\tau)}z_{i,(\tau)} + b_{i,(\tau)}z_{i,(\tau)}^{-1}\right)\right\}, \tag{A.46}$$

$$\beta_{j,(\tau)} \mid \bullet \sim \mathcal{N}\left(\mu_{\beta_j,(\tau)}, \Sigma_{\beta_j,(\tau)}\right), \tag{A.47}$$

$$v_{j,(\tau)} \mid \bullet \sim \mathcal{GIG}\left(\frac{1}{2}, \frac{\sigma_{(\tau)}^{-1}}{\lambda_{j,(\tau)}^2}, \beta_{j,(\tau)}^2\right) \propto v_{j,(\tau)}^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\left(\frac{\sigma_{(\tau)}^{-1}}{\lambda_{j,(\tau)}^2}v_{j,(\tau)} + \beta_{j,(\tau)}^2 v_{j,(\tau)}^{-1}\right)\right\}, \tag{A.48}$$

$$\sigma_{(\tau)} \mid \bullet \sim \mathcal{IG}\left(r_{\sigma,(\tau)}, s_{\sigma,(\tau)}\right) \propto \left(\sigma_{(\tau)}^{-1}\right)^{r_{\sigma,(\tau)}+1}\exp\left\{-s_{\sigma,(\tau)}\sigma_{(\tau)}^{-1}\right\}, \tag{A.49}$$

$$\lambda_{j,(\tau)}^2 \mid \bullet \sim \mathcal{IG}\left(c_{0,(\tau)} + 1, d_{0,(\tau)} + \sigma_{(\tau)}^{-1}v_{j,(\tau)}/2\right); \tag{A.50}$$

where

$$a_{i,(\tau)} = \frac{1}{\sigma_(\tau)}\left(2 + \frac{\theta_{(\tau)}^2}{\kappa_{(\tau)}^2}\right) \text{ and } b_{i,(\tau)} = \frac{\left(Y_i - \mathbf{X}_i\beta_{(\tau)}\right)^2}{\sigma_{(\tau)}\kappa_{(\tau)}^2},$$

$$\Sigma_{\beta_j,(\tau)} = \left[\left(\sigma\kappa_{(\tau)}^2\right)^{-1}\sum_{i=1}^{n}x_{ij}^2 z_{i,(\tau)}^{-1} + v_{j,(\tau)}^{-1}\right]^{-1},$$

$$\mu_{\beta_j,(\tau)} = \Sigma_{\beta_j,(\tau)}\left(\sigma\kappa_{(\tau)}^2\right)^{-1}\sum_{i=1}^{n}\left(Y_i - \theta_{(\tau)}z_{i,(\tau)} - \sum_{k=1,k\neq j}^{p}x_{ij}\beta_{j,(\tau)}\right)x_{ij}^2 z_{i,(\tau)}^{-1},$$

$$r_{\sigma,(\tau)} = r_{0,(\tau)} + \frac{3n}{2} + p \text{ and } s_{\sigma,(\tau)} = s_{0,(\tau)} + \sum_{i=1}^{n} \frac{\left(Y_i - \mathbf{X}_i\beta_{(\tau)} - \theta_{(\tau)} z_{i,(\tau)}\right)^2}{2\kappa^2_{(\tau)} z_{i,(\tau)}} + \sum_{i=1}^{n} z_{i,(\tau)} + \sum_{j=1}^{p} \frac{v_{j,(\tau)}}{2\lambda_j^2};$$

for $i = 1, \ldots, n$ and $j = 1, \ldots, p$.

A.4.3 On Implementation

The **bayesQR** R package (Benoit and Van den Poel, 2017) provides the implementation of efficient Gibbs sampler for both Bayesian Quantile Regression and Bayesian Quantile Regression with the Adaptive Lasso outlined above. In addition, the core procedure is programmed in Fortran to speed up the computational time. Therefore, this package can be utilised straightforward to estimate multiple conditional quantiles, which then be used to approximate the condistional distributions of potential outcomes in our proposed BADR framework.

Alternatively, Variational Inference algorithm could be used for Bayesian quantile regression with/without the regularisation (Guo, 2019; Lim et al., 2020), which helps improving the speed of Gibbs sampling while maintaining a comparable accuracy in terms of MSE.

A.5 Implementation of Other Estimators in Simulation Study

A.5.1 Existing Approaches

Firpo's Inverse Probability Weighted method (`FIPW`)

Firpo's Inverse Probability Weighted (`FIPW`) method (Firpo, 2007) involves a two-step estimator. First, the propensity score is estimated nonparametrically as a logistic power series whose degree increases with sample size. In the second step, the quantiles are estimated by minimising an inverse probability weighted check loss function. These weights reflect the fact that the distribution of the covariates differs in the two groups.

Algorithm A.1: FIPW Approach to Estimate QTEs

Data: $\{Y_i, T_i, \mathbf{X}_i\}_{i=1}^n, \tau \in (0,1)$

Result: $\widehat{QTE}(\tau)$

1 *Step(1).* Estimate propensity score $\hat{\pi}(x) = \text{expit}\left(H_K(x)'\hat{p}_K\right)$ where

$$\hat{p}_K := \underset{p \in \mathbb{R}^K}{\text{argmax}} \frac{1}{N} \sum_{i=1}^{N} \{T_i \cdot \log\left(\text{expit}\left(H_K(X_i)'p\right)\right) + (1 - T_i) \cdot \log\left(1 - \text{expit}\left(H_K(X_i)'p\right)\right)\}$$

2 *Step(2).* Derive $\hat{q}_1(\tau)$ and $\hat{q}_0(\tau)$ as the solution to

$$\hat{q}_1(\tau) := \underset{q}{\text{argmin}} \sum_{i=1}^{N} \frac{T_i}{N \cdot \hat{\pi}(\mathbf{X}_i)} \rho_\tau(Y_i - q) \text{ and } \hat{q}_0(\tau) := \underset{q}{\text{argmin}} \sum_{i=1}^{N} \frac{1 - T_i}{N \cdot (1 - \hat{\pi}(\mathbf{X}_i))} \rho_\tau(Y_i - q)$$

where $\rho_\tau(a) = a \cdot (\tau - \mathbb{1}\{a \leq 0\})$ is the check function.

3 Calculate $\widehat{QTE}(\tau) = \hat{q}_1(\tau) - \hat{q}_0(\tau)$.

Targeted Maximum Likelihood Estimation method (TMLE)

The estimation procedure of Targeted Maximum Likelihood Estimation (`TMLE`) method (Díaz, 2017) includes three steps. First, the propensity score and the conditional distribution of the outcome are estimated; second, the quantiles are estimated based on the current cdf of the outcome; and third, the conditional distribution of the outcome is updated based on an exponential submodel for the density of the outcome. The last two steps are iterated until convergence.

Algorithm A.2: TMLE Approach to Estimate QTEs

Data: $\{Y_i, T_i, \mathbf{X}_i\}_{i=1}^n, \tau \in (0,1)$

Result: $\widehat{QTE}(\tau)$

Step (1). Initialize: Obtain initial estimates $\hat{\pi}$ and $\hat{G}$ of π_0 and G_0.

Step (2). Compute $\hat{q}_1(\tau)$: For the current estimate $\hat{G}$, compute

$$\hat{F}(y) = \frac{1}{n}\sum_{i=1}^{n} \hat{G}(y \mid 1, X_i) \text{ and } \hat{q}_1(\tau) = \inf\{y : \hat{F}(y) \geq \tau\}$$

Step (3). Update $\hat{G}$: Let $\hat{g}$ denote the density associated to $\hat{G}$.

(a) Consider the exponential submodel:

$$\hat{g}_\epsilon(y \mid 0, x) = c(\epsilon, \hat{g}) \exp\{\epsilon \hat{H}_{\hat{\eta},\hat{\theta}(z)}\} \hat{g}(y \mid 0, x)$$

where $c(\epsilon, \hat{g})$ is a normalizing constant and

$$\hat{H}_{\hat{\eta},\hat{\theta}(z)} \coloneqq \frac{1}{\hat{\pi}(X)}\{\mathbb{1}_{(-\infty,\hat{\theta}]}(y) - \hat{G}(\hat{\theta} \mid 0, x)\}$$

is the score of the model.

(b) Estimate ϵ

$$\hat{\epsilon} = \text{argmax} \sum_{i=1}^{n} (1 - T_i) \log \hat{g}_\epsilon(Y_i \mid 0, X_i)$$

(c) Calculate $\hat{g}_\epsilon(Y_i \mid 0, X_i)$ as the updated estimator of g

Step (4). Iterate: Let $\hat{g} = \hat{g}_\epsilon$ and iterate steps 2-3 until convergence.

Derive $\hat{q}_0(\tau)$ similarly.

Calculate $\widehat{QTE}(\tau) = \hat{q}_1(\tau) - \hat{q}_0(\tau)$.

Localized Debiased Machine Learning method (LDML)

The Localized Debiased Machine Learning (`LDML`) method (Kallus et al., 2024) is also motivated by the efficient estimation equation, but Inverse Probability Weighted (IPW) estimates are used as rough initial guessed values for $\hat{q}_1$ and $\hat{q}_0$. Then, these values are used to localize the estimation of conditional distributions $\hat{G}(y \mid 0, \mathbf{X})$ and $\hat{G}(y \mid 1, \mathbf{X})$, respectively. This approach aims to refine the IPW estimate while obviating the need to estimate a continuum of continuum nuisances. The main algorithm includes two parts: *three-way-cross-fold nuisance estimation* and *solving the estimating equation*.

Algorithm A.3: LDML Approach to Estimate QTEs

Data: $\{Y_i, T_i, \mathbf{X}_i\}_{i=1}^n, \tau \in (0,1)$

Part(1). Three-way-cross-fold nuisance estimation

Fix integers $K \geq 3$ and $1 \leq K' \leq K-2$.

Randomly permute the data indices and let $\mathcal{D}_k$ be a random even K-fold split of the data.

for $k = 1, \ldots, K$ **do**

(a) Set $\mathcal{H}_{k,1} = \{1, \ldots, K' + \mathbb{1}[k \leq K']\} \setminus \{k\}$ and $\mathcal{H}_{k,2} = \{K' + \mathbb{1}[k \leq K'] + 1, \ldots, K\} \setminus \{k\}$.

(b) Use only the data in $\mathcal{D}_k^{C,1} = \{\mathbf{X}_i : i \in \bigcup_{k' \in \mathcal{H}_{k,1}}\}$ to construct $\hat{q}_{1,init}^{(k)}$.

(c) Use only the data in $\mathcal{D}_k^{C,2} = \{\mathbf{X}_i : i \in \bigcup_{k' \in \mathcal{H}_{k,2}}\}$ to construct $\hat{G}_1^{(k)}(\cdot, \hat{q}_{1,init}^{(k)})$.

(d) Use $\mathcal{D}_k^{C,1} \cap \mathcal{D}_k^{C,2}$ to construct estimator $\hat{\pi}^{(k)}$.

end

Part(2). Solving the average of the estimate in each fold to obtain $\hat{q}_1(\tau)$

$$\frac{1}{N} \sum_{k=1}^{K} \sum_{i \in \mathcal{D}_k} \psi\left(\mathbf{X}_i; q, \hat{G}_1^{(k)}(\mathbf{X}_i, \hat{q}_{1,init}^{(k)}), \hat{\pi}^{(k)}(\mathbf{X}_i)\right) = 0$$

Derive $\hat{q}_0(\tau)$ using two-part procedure similarly.

Calculate $\widehat{QTE}(\tau) = \hat{q}_1(\tau) - \hat{q}_0(\tau)$.

Bayesian nonparametric method (BNP)

Bayesian nonparametric (`BNP`) method (Xu et al., 2018) is a fully Bayesian nonparametric (BNP) approach to estimate QTEs. The estimation procedure includes three steps. First, the propensity score is estimated using a logit BART model. Then, the conditional distribution of the potential outcome given a BART posterior sample of the PS in each treatment group is modelled separately using a Dirichlet process mixture (DPM) of multivariate normals model. Finally, marginalizing the estimated conditional distribution over the population distribution of the confounders using Bayesian bootstrap (Rubin, 1981). Details of implementation using **BNPqte** R package can be found in Luo and Daniels (2021).

Algorithm A.4: BNP Approach to Estimate QTEs

Data: $\{Y_i, T_i, \mathbf{X}_i\}_{i=1}^n, \tau \in (0,1)$

Fit a binary BART model on $\{T_i, X_i\}_{i=1}^n$ and obtain B posterior samples $\{H^{-1}(\pi_i^{\{b\}})\}_{i,b=1}^{n,B}$.

Create a set of grid points of Y values: $(q_1, \ldots, q_S)$.

for $b = 1, \ldots, B$ **do**

  for $t = 0, 1$ **do**

    Fit a DPM of bivariate normals on $\{Y_i, H^{-1}(\pi_i^{\{b\}})\}_{i:T_i=t}$.

    Use Blocked Gibbs sampler to obtain L posterior samples.

    Calculate $\{F^{\{bl\}}(q_s \mid H^{-1}(\pi_i^b), T = t)\}_{i,s,l=1}^{n,S,L}$.

  end

  Sample $\left(u_1^b, \ldots, u_n^b\right)$ from $\mathcal{D}ir(1, \ldots, 1)$

  for $l = 1, \ldots, L$ **do**

    Calculate the CDF of $Y^{(t)}$ as follows:

$$SF_t^{bl}(q_s) = \sum_{i=1}^{n} u_i^b F^{kl}(q_s \mid H^{-1}(\pi_i^b), T = t), \text{ where } 1 \leq s \leq S \text{ and } t \in \{0,1\}$$

    Find a grid point $q_t^{bl}(\tau)$ such that $F_t^{kl}(q_t^{bl}(\tau)) = \tau$ for $t \in \{0,1\}$.

    The τ^{th} quantile from the CDF $F_t^{bl}(.)$ is $q_{t,\tau}^{bl}$ for the group $T = t$.

  end

end

Derive $\hat{q}_1(\tau)$ and $\hat{q}_0(\tau)$ as follows

$$\hat{q}_1(\tau) = \frac{1}{BL}\sum_{b}^{B}\sum_{l=1}^{L} q_1^{bl}(\tau) \quad \text{and} \quad \hat{q}_0(\tau) = \frac{1}{BL}\sum_{b}^{B}\sum_{l=1}^{L} q_0^{bl}(\tau)$$

Calculate $\widehat{QTE}(\tau) = \hat{q}_1(\tau) - \hat{q}_0(\tau)$.

A.5.2 Variants of the Proposed Approach

Bayesian Outcome Modelling (BOM)

Bayesian Outcome Modelling (`BOM`) is an outcome-regression-based approach that omits the treatment assignment model. Instead, it solely focuses on estimating the conditional distribution through multiple Bayesian quantile regressions in the outcome model of each treatment group. Shrinkage priors, akin to the doubly-robust approach, can be readily incorporated.

Algorithm A.5: Bayesian Outcome Modelling to estimate QTE

Data: $\{Y_i, T_i, \mathbf{X}_i\}_{i=1}^n, \tau \in (0,1)$

Result: $\widehat{QTE}^{om}(\tau)$

for $t = 0, 1$ **do**

  Fit outcome model on $\{Y_i, \mathbf{X}_i\}_{i:T_i=t}$ and obtain B posterior samples $\{G^{(b)}(y \mid t, \mathbf{X})\}_{b=1}^B$.

  Derive posterior mean $\hat{G}(y \mid t, \mathbf{X}) = \frac{1}{B}\sum_{b=1}^B G^{(b)}(y \mid t, \mathbf{X})$.

end

Derive $\hat{q}_1^{om}(\tau)$ and $\hat{q}_0^{om}(\tau)$ as the solutions to

$$\frac{1}{n}\sum_{i=1}^n \hat{G}(q_1 \mid 1, \mathbf{X}) = \tau \text{ and } \frac{1}{n}\sum_{i=1}^n \hat{G}(q_0 \mid 0, \mathbf{X}) = \tau$$

Calculate $\hat{\Delta}_\tau \coloneqq \widehat{QTE}^{om}(\tau) = \hat{q}_1^{om}(\tau) - \hat{q}_0^{om}(\tau)$.

Bayesian Propensity Score Analysis (BPSA)

Bayesian Propensity Score Analysis (`BPSA`) is a treatment-assignment-based approach, which involves fitting the treatment assignment and then employs multiple Bayesian quantile regressions to model the conditional distribution of the outcome given the posterior mean of the propensity score in each treatment group.

Algorithm A.6: Bayesian Outcome Modelling to estimate QTE

Data: $\{Y_i, T_i, \mathbf{X}_i\}_{i=1}^n, \tau \in (0,1)$

Result: $\widehat{QTE}^{ps}(\tau)$

Fit treatment assignment model on $\{T_i, \mathbf{X}_i\}_{i=1}^n$ and obtain B posterior samples $\{\pi^{(b)}(\mathbf{X})\}_{b=1}^B$.

Derive posterior mean from B posterior samples $\hat{\pi}(\mathbf{X}) = \frac{1}{B}\sum_{b=1}^B \pi^{(b)}(\mathbf{X})$.

for $t = 0, 1$ **do**

  Fit outcome model on $\{Y_i, \hat{\pi}(\mathbf{X}_i)\}_{i:T_i=t}$ and obtain B posterior samples $\{G^{(b)}(y \mid t, \hat{\pi}(\mathbf{X}))\}_{b=1}^B$.

  Derive posterior mean $\hat{G}(y \mid t, \mathbf{X}) = \frac{1}{B}\sum_{b=1}^B G^{(b)}(y \mid t, \hat{\pi}(\mathbf{X}))$.

end

Derive $\hat{q}_1^{ps}(\tau)$ and $\hat{q}_0^{ps}(\tau)$ as the solutions to

$$\frac{1}{n}\sum_{i=1}^n \hat{G}(q_1 \mid 1, \mathbf{X}) = \tau \text{ and } \frac{1}{n}\sum_{i=1}^n \hat{G}(q_0 \mid 0, \mathbf{X}) = \tau$$

Calculate $\hat{\Delta}_\tau \coloneqq \widehat{QTE}^{ps}(\tau) = \hat{q}_1^{ps}(\tau) - \hat{q}_0^{ps}(\tau)$.

A.6 Additional Simulation Results

Table A.1: Simulation Results for SD1, Relative RMSE

		Estimation Methods					
Percentiles	N	BDR	BDRS	BNP	LDML	TMLE	FIPW
10th	1000	1.71	1.689	**0.991**	1.234	1.771	1.789
	500	1.623	1.54	**0.946**	1.124	1.658	1.774
	100	2.459	1.226	**0.963**	1.193	1.332	1.529
25th	1000	1.052	1.044	0.978	**0.852**	1.022	1.136
	500	1.139	1.117	1.007	**0.92**	1.052	1.468
	100	1.188	0.977	**0.916**	0.995	1.001	1.169
50th	1000	**0.598**	0.599	0.99	0.744	0.623	0.659
	500	0.663	**0.657**	0.978	0.776	0.669	0.712
	100	0.916	**0.819**	0.971	0.944	0.886	1.114
75th	1000	**0.521**	0.525	0.979	0.758	0.568	0.577
	500	**0.586**	0.598	0.983	0.821	0.634	0.65
	100	0.99	**0.799**	0.986	0.92	0.835	1.032
90th	1000	**0.583**	0.593	0.98	0.803	0.605	0.606
	500	**0.635**	0.663	0.991	0.86	0.691	0.728
	100	1.233	**0.853**	0.974	0.928	0.897	1.055

Notes: This table displays the relative Root Mean Squared Error (RMSE) of different estimation methods across 100 replicates. The rows contain results for various percentile levels and for various sample size N. The relative RMSE is the RMSE in comparison with the Naive method as the benchmark, where $RMSE = \sqrt{R^{-1}\sum_{r=1}^{R}(\hat{\alpha}_r - \alpha)^2}$ and $R = 100$.

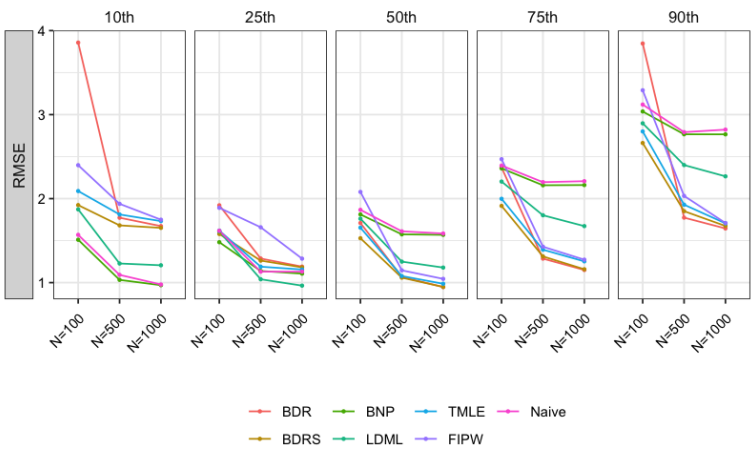

Figure A.1: Line plots of raw RMSE for 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th} QTEs based on 100 replications.

Table A.2: Simulation Results for SD2a, Average Bias and Relative RMSE

	Bias					RMSE				
	10th	25th	50th	75th	90th	10th	25th	50th	75th	90th
Linear										
BDR	0.055	0.002	-0.032	-0.008	0.016	2.036	1.098	0.673	0.637	0.839
BDRS	0.061	0.002	-0.030	-0.006	0.024	2.028	1.098	0.678	0.638	0.846
BOM	0.040	0.092	-0.013	-0.094	-0.053	0.928	0.921	0.562	0.620	0.814
BOMS	0.059	0.102	-0.002	-0.076	-0.025	0.913	0.904	0.554	0.619	0.812
BPSA	0.078	-0.042	-0.044	0.061	0.112	0.735	0.980	0.523	0.595	0.719
BNP	0.335	0.368	0.343	0.328	0.284	1.025	1.042	0.917	0.966	1.038
LDML	0.009	0.010	-0.009	0.019	0.163	0.985	1.029	0.616	0.696	0.819
TMLE	0.000	-0.005	-0.022	-0.041	0.025	1.224	1.229	0.694	0.752	1.180
FIPW	0.019	-0.039	-0.034	-0.010	0.014	1.845	1.286	0.710	0.761	1.285
No covariates										
Naive	0.468	0.434	0.430	0.377	0.312	1.000	1.000	1.000	1.000	1.000

Notes: This table displays the average bias and the relative Root Mean Squared Error (RMSE) of different estimation methods across 100 replicates. The relative RMSE is the RMSE in comparison with the Naive method as the benchmark, where $RMSE = \sqrt{R^{-1}\sum_{r=1}^{R}(\hat{\alpha}_r - \alpha)^2}$ and $R = 100$.

Table A.3: Simulation Results for SD2b, Average Bias and Relative RMSE

	Bias					RMSE				
	10th	25th	50th	75th	90th	10th	25th	50th	75th	90th
Nonlinear										
BDRS	-0.014	0.015	0.019	0.027	0.011	0.680	0.878	0.520	0.570	0.543
BOMS	0.041	0.027	0.020	0.015	0.001	0.519	0.727	0.470	0.490	0.425
BPSA	0.140	-0.019	-0.029	0.077	0.099	0.715	0.955	0.538	0.604	0.705
LDML	0.111	0.052	0.047	0.062	0.182	0.852	1.006	0.722	0.730	0.831
TMLE	-0.020	0.009	0.023	0.023	0.064	0.694	0.898	0.522	0.578	0.669
FIPW	0.034	0.120	0.090	-0.050	-0.051	2.180	2.066	1.981	2.056	1.678
No covariates										
Naive	0.468	0.434	0.430	0.377	0.312	1.000	1.000	1.000	1.000	1.000

Notes: This table displays the average bias and the relative Root Mean Squared Error (RMSE) of different estimation methods across 100 replicates. The relative RMSE is the RMSE in comparison with the Naive method as the benchmark, where $RMSE = \sqrt{R^{-1}\sum_{r=1}^{R}(\hat{\alpha}_r - \alpha)^2}$ and $R = 100$.

A.7 Additional Graphs in Empirical Illustration

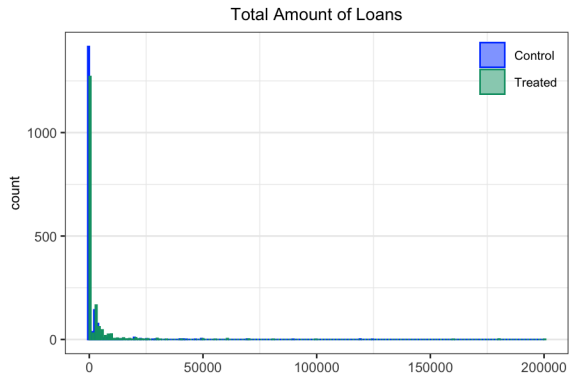

Figure A.2: Histogram of total amount of loans at household level in treated villages and control villages.

Appendix B

Appendix for Chapter 3

B.1 On Identification - Proposition 1

We use the following results in the proof of Proposition 1.

Lemma 1. Let $\epsilon \sim \mathcal{N}(0,1)$ and $D = \mathbb{1}\{\nu + \epsilon > 0\}$. Then

$$\mathbb{E}[\epsilon \mid D = 1] = \lambda(\nu) := \frac{\phi(\nu)}{\Phi(\nu)},$$
$$\mathbb{E}[\epsilon \mid D = 0] = -\lambda(-\nu) := -\frac{\phi(\nu)}{1 - \Phi(\nu)}.$$

where ϕ and Φ are the standard normal pdf and cdf. The function $\lambda(.)$ is called the inverse Mills ratio.

Lemma 2. Under Assumption 3, for each mixture component g,

$$\mathbb{E}[\varepsilon_i^{(1)} \mid \varepsilon_i^{(D)} = \epsilon, g] = \sigma_{1D,g}\epsilon \quad \text{and} \quad \mathbb{E}[\varepsilon_i^{(0)} \mid \varepsilon_i^{(D)} = \epsilon, g] = \sigma_{0D,g}\epsilon.$$

Thus, marginalise over g using the law of total expectation,

$$\mathbb{E}[\varepsilon_i^{(1)} \mid \varepsilon_i^{(D)} = \epsilon] = \sigma_{1D}\epsilon \quad \text{and} \quad \mathbb{E}[\varepsilon_i^{(0)} \mid \varepsilon_i^{(D)} = \epsilon] = \sigma_{0D}\epsilon,$$

where $\sigma_{1D} = \sum_{g=1}^{G} \pi_g \sigma_{1D,g}$ and $\sigma_{0D} = \sum_{g=1}^{G} \pi_g \sigma_{0D,g}$.

Proof of Proposition 1

(i) Identification of $(\delta^{(1)}, \beta^{(1)})$ and $(\delta^{(0)}, \beta^{(0)})$

Condition on treatment status, for $D_i = 1$,

$$\begin{aligned}\mathbb{E}[Y_i \mid D_i = 1, \bar{D}_{\mathcal{N}i}, X_i] &= \mathbb{E}[Y_i^{(1)} \mid D_i = 1, \bar{D}_{\mathcal{N}i}, X_i] \\ &= \delta^{(1)}\bar{D}_{\mathcal{N}i} + X_i\beta^{(1)} + \mathbb{E}[\varepsilon_i^{(1)} \mid D_i = 1, \bar{D}_{\mathcal{N}i}, X_i].\end{aligned}$$

By Assumption 4 and Assumption 2, $\mathbb{E}[\varepsilon_i^{(1)} \mid D_i = 1, \bar{D}_{\mathcal{N}i}, X_i]$ does not depend on $\bar{D}_{\mathcal{N}i}$, so the slope in $\bar{D}_{\mathcal{N}i}$ equals $\delta^{(1)}$, and variation in $(\bar{D}_{\mathcal{N}i}, X_i)$ identifies $(\delta^{(1)}, \beta^{(1)})$. Analogously, for $D_i = 0$ we obtain identification of $(\delta^{(0)}, \beta^{(0)})$.

(ii) Identification of $(\sigma_{1D}, \sigma_{0D})$

By Lemma 1, $\mathbb{E}[\varepsilon_i^{(1)} \mid \varepsilon_i^{(D)} = \epsilon] = \sigma_{1D}\epsilon$ $\mathbb{E}[\varepsilon_i^{(0)} \mid \varepsilon_i^{(D)} = \epsilon] = \sigma_{0D}\epsilon$. Using $D_i = \mathbb{1}\{\nu(X_i, Z_i) + \varepsilon_i^{(D)} > 0\}$ and Lemma 2,

$$\begin{aligned}\mathbb{E}[\varepsilon_i^{(1)} \mid D_i = 1, X_i, Z_i] &= \sigma_{1D}\,\lambda\Big(\nu(X_i, Z_i)\Big), \\ \mathbb{E}[\varepsilon_i^{(0)} \mid D_i = 0, X_i, Z_i] &= -\sigma_{0D}\,\lambda\Big(-\nu(X_i, Z_i)\Big).\end{aligned}$$

From (i), we already have the population identification of $(\delta^{(1)}, \beta^{(1)})$ and $(\delta^{(0)}, \beta^{(0)})$. Form within-arm residuals by netting out the identified linear parts

$$\begin{aligned}&R_i^{(1)} := Y_i - \delta^{(1)}\bar{D}_{\mathcal{N}i} - X_i\beta^{(1)} \quad \text{(for } D_i = 1\text{)}, \\ &R_i^{(0)} := Y_i - \delta^{(0)}\bar{D}_{\mathcal{N}i} - X_i\beta^{(0)} \quad \text{(for } D_i = 0\text{). Then,} \\ &\mathbb{E}[R_i^{(1)} \mid D_i = 1, X_i, Z_i] = \sigma_{1D}\,\lambda\Big(\nu(X_i, Z_i)\Big), \\ &\mathbb{E}[R_i^{(0)} \mid D_i = 0, X_i, Z_i] = -\sigma_{0D}\,\lambda\Big(-\nu(X_i, Z_i)\Big).\end{aligned}$$

By Assumption 1 (exclusion of Z_i from outcomes) and Assumption 2 (independence from disturbances), variation in Z_i moves the selection index $\nu(X_i, Z_i)$ and hence the inverse Mills ratio $\lambda(\nu)$ without directly affecting the outcome equations. With the normalisation $\mathrm{Var}(\varepsilon_i^{(D)}) = 1$ in Assumption 3, the inverse Mills ratio is a known function of ν. Therefore, the slopes of the population regressions of $R_i^{(1)}$ on $\lambda(\nu(X_i, Z_i))$ (within $D_i = 1$) and $R_i^{(0)}$ on $-\lambda(-\nu(X_i, Z_i))$ (within $D_i = 0$) identify σ_{1D} and σ_{0D}, respectively.

This completes the proof.

B.2 Details of Computational Algorithms

B.2.1 Derivations for Algorithm 3.1

Let Σ_{ij} and $\tilde{\Sigma}_{ij}$ denote the (i,j)th element of $\boldsymbol{\Sigma}$ and $\tilde{\boldsymbol{\Sigma}}$, respectively. The Jacobian matrix of the transformation $(\boldsymbol{\Sigma}, \tau^2) \to \tilde{\boldsymbol{\Sigma}}$ is given by

$$\mathcal{J} = \frac{\partial\left(\tilde{\Sigma}_{11}, \tilde{\Sigma}_{12}, \tilde{\Sigma}_{13}, \tilde{\Sigma}_{22}, \tilde{\Sigma}_{23}, \tilde{\Sigma}_{33}\right)}{\partial\left(\Sigma_{12}, \Sigma_{13}, \Sigma_{22}, \Sigma_{23}, \Sigma_{33}, \tau^2\right)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ \tau & 0 & 0 & 0 & 0 & (1/2)\rho_{1D}\sigma_1\tau^{-1} \\ 0 & \tau & 0 & 0 & 0 & (1/2)\rho_{0D}\sigma_0\tau^{-1} \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

If we assign the Inverse Wishart prior $\mathcal{W}^{-1}(\mathbf{I}_3, \nu_o)$ for $\tilde{\boldsymbol{\Sigma}}$ with density

$$p(\tilde{\boldsymbol{\Sigma}}) \propto |\tilde{\boldsymbol{\Sigma}}|^{-(\nu_0+4)/2} \exp\left\{-\frac{1}{2}\operatorname{tr}\left(\mathbf{I}_3\tilde{\boldsymbol{\Sigma}}^{-1}\right)\right\},$$

which is equivalent to the following joint prior for $(\boldsymbol{\Sigma}, \tau^2)$

$$\begin{aligned} p(\boldsymbol{\Sigma}, \tau^2) \propto (\tau^2)^{-(\nu_o+4)/2}|\boldsymbol{\Sigma}|^{-(\nu_o+4)/2} &\exp\left\{\frac{1-\rho_{10}^2}{-2\tau^2\left(1+2\rho_{10}\rho_{1D}\rho_{0D}-\rho_{10}^2-\rho_{1D}^2-\rho_{0D}^2\right)}\right\} \\ &\times \exp\left\{\frac{1-\rho_{0D}^2}{-2\sigma_1^2\left(1+2\rho_{10}\rho_{1D}\rho_{0D}-\rho_{10}^2-\rho_{1D}^2-\rho_{0D}^2\right)}\right\} \\ &\times \exp\left\{\frac{1-\rho_{1D}^2}{-2\sigma_0^2\left(1+2\rho_{10}\rho_{1D}\rho_{0D}-\rho_{10}^2-\rho_{1D}^2-\rho_{0D}^2\right)}\right\} \times |\mathcal{J}| \\ \propto (\tau^2)^{-\left(\frac{\nu_o+4}{2}+1\right)}|\boldsymbol{\Sigma}|^{-(\nu_o+4)/2} &\exp\left\{\frac{1-\rho_{10}^2}{-2\tau^2(1+2\rho_{10}\rho_{1D}\rho_{0D}-\rho_{10}^2-\rho_{1D}^2-\rho_{0D}^2)}\right\} \end{aligned} \tag{B.1}$$

From the joint density of $(\boldsymbol{\Sigma}, \tau^2)$, the prior for τ^2 given $\boldsymbol{\Sigma}$ is

$$p(\tau^2 \mid \boldsymbol{\Sigma}) \sim \left[\left(\frac{1+2\rho_{10}\rho_{1D}\rho_{0D}-\rho_{10}^2-\rho_{1D}^2-\rho_{0D}^2}{1-\rho_{10}^2}\right)\chi^2_{(\nu_o+4)}\right]^{-1}, \tag{B.2}$$

and the prior for $\boldsymbol{\Sigma}$ is

$$p(\boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-(\nu_o+4)/2}\exp\left\{-\frac{1}{2\tau^2} \times \frac{1-\rho_{10}^2}{1+2\rho_{10}\rho_{1D}\rho_{0D}-\rho_{10}^2-\rho_{1D}^2-\rho_{0D}^2}\right\}. \tag{B.3}$$

B.2.2 Alternative Algorithm

This approach is based on the marginal data augmentation method for the multinomial probit model (Imai and Van Dyk, 2005; Jiao and Dyk, 2015). To illustrate the idea, we denote y, θ, and z as generic observed data, unknown parameter of interest, and latent variables, respectively. Marginal data augmentation (MDA) algorithm (Meng and Van Dyk, 1999) introduces τ into the augmented-data model $p(z, y|\theta)$. A MCMC sampler is implemented for the expanded model $\int p(\tilde{z}, y|\theta, \tau)$ which is designed to maintain $p(y|\theta)$ as its marginal distribution:

$$\int p(\tilde{z}, y|\theta, \tau)d\tilde{z} = p(y|\theta)$$

where $\tilde{z} = \mathcal{F}_\tau(z),$ for any given τ, is a one-to-one mapping and $\exists \tau_0 : \mathcal{F}_{\tau_0}(z) = z$. The MDA algorithm iterates

$$(\tilde{z}^{[s+1]}, \tau^\star) \sim p(\tilde{z}, \tau \mid \theta^{[s]}, y),$$
$$(\theta^{[s+1]}, \tau^{[s+1]}) \sim p(\theta, \tau \mid \tilde{z}^{[s+1]}, y).$$

Within the context of our model summarized in (3.23), we introduce a positive scalar parameter τ which serves as the expansion parameter

$$\tau \mathbf{L}^* = \mathbf{R}(\tau\theta) + \tau\epsilon \Leftrightarrow \tilde{\mathbf{L}}^* = \mathbf{R}\tilde{\theta} + \tilde{\epsilon}$$

where $\tilde{\mathbf{L}}^* \coloneqq \tau \mathbf{L}^*$, $\tilde{\theta} \coloneqq \tau\theta$, and $\tilde{\epsilon} \coloneqq \tau\epsilon$. The covariance of $\tilde{\epsilon}$ is $\tilde{\mathbf{\Omega}} = \tau^2\mathbf{\Omega} = \tau^2\mathbf{\Sigma} \otimes \mathbf{I}_n$.

The new unconstrained covariance matrix is $\tilde{\mathbf{\Sigma}} = \tau^2\mathbf{\Sigma}$. Then, the Inverse Wishart distribution can be assigned as a prior for $\tilde{\mathbf{\Sigma}}$

$$\tilde{\mathbf{\Sigma}} \sim \mathcal{W}^{-1}(\mathbf{\Psi}_o, \nu_o),$$

which implies the following joint prior density[1]

$$\begin{aligned} p(\boldsymbol{\Sigma}, \tau^2) &\propto (\tau^2)^{-(\nu_o+4)/2} |\boldsymbol{\Sigma}|^{-(\nu_o+4)/2} \exp\left\{-\frac{1}{2\tau^2}\operatorname{tr}(\boldsymbol{\Psi}_o \boldsymbol{\Sigma}^{-1})\right\} \times |\mathcal{J}| \\ &\propto (\tau^2)^{-(3\nu_o/2+1)} |\boldsymbol{\Sigma}|^{-(\nu_o+4)/2} \exp\left\{-\frac{1}{2\tau^2}\operatorname{tr}(\boldsymbol{\Psi}_o \boldsymbol{\Sigma}^{-1})\right\}, \end{aligned}$$

i.e., $p(\boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-(\nu_o+4)/2} \times [\operatorname{tr}(\boldsymbol{\Psi}_o \boldsymbol{\Sigma}^{-1})]^{-\nu_o}$ and $\tau^2 \mid \boldsymbol{\Sigma} \sim \operatorname{tr}(\boldsymbol{\Psi}_o \boldsymbol{\Sigma}^{-1})/\chi^2_{3\nu_o}$.

The expansion parameter τ^2 should be sample along with $\tilde{\theta}$ and $\tilde{\boldsymbol{\Sigma}}$ to recover θ and $\boldsymbol{\Sigma}$ during the sampling process.

Sampling steps for θ

(i) Sample $\tilde{\theta}$ and τ^2 from

$$p(\tilde{\theta}, \tau^2 \mid \boldsymbol{\Sigma}, \tilde{\mathbf{L}}^*, \mathbf{Y}, \mathbf{D}) = p(\tilde{\theta} \mid \tau^2, \boldsymbol{\Sigma}, \tilde{\mathbf{L}}^*, \mathbf{Y}, \mathbf{D}) \times p(\tau^2 \mid \boldsymbol{\Sigma}, \tilde{\mathbf{L}}^*, \mathbf{Y}, \mathbf{D}).$$

To accomplish this, we obtain a draw of τ^2 from $p(\tau^2 \mid \boldsymbol{\Sigma}, \lambda, \tilde{\mathbf{L}}^*, \mathbf{Y}, \mathbf{D})$ and then use that draw to sample $\tilde{\theta}$ from $p(\tilde{\theta} \mid \tau^2, \boldsymbol{\Sigma}, \tilde{\mathbf{L}}^*, \mathbf{Y}, \mathbf{D})$.

(ii) To marginalize out τ, set $\theta = \tilde{\theta}/\tau^2$.

Sampling steps for $\boldsymbol{\Sigma}$

(i) Sample $\tilde{\boldsymbol{\Sigma}}$ from $p(\tilde{\boldsymbol{\Sigma}} \mid \tau^2, \tilde{\theta}, \lambda, \tilde{L}^*, Y, D)$.

(ii) To marginalize τ, set $\boldsymbol{\Sigma} = \frac{1}{\Sigma_{11}} \times \tilde{\boldsymbol{\Sigma}}$.

Completing posterior analysis, we provide the detailed implementation of the MCMC sampling scheme in Algorithm B.1.

[1]Note that the Jacobian matrix of the transformation $\tau^2 \boldsymbol{\Sigma} = \tilde{\boldsymbol{\Sigma}}$ is given by

$$\mathcal{J} = \frac{\partial\left(\tilde{\Sigma}_{11}, \tilde{\Sigma}_{12}, \tilde{\Sigma}_{13}, \tilde{\Sigma}_{22}, \tilde{\Sigma}_{23}, \tilde{\Sigma}_{33}\right)}{\partial\left(\Sigma_{12}, \Sigma_{13}, \Sigma_{22}, \Sigma_{23}, \Sigma_{33}, \tau^2\right)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ \tau^2 & 0 & 0 & 0 & 0 & \sigma_{1D} \\ 0 & \tau^2 & 0 & 0 & 0 & \sigma_{0D} \\ 0 & 0 & \tau^2 & 0 & 0 & \sigma_1^2 \\ 0 & 0 & 0 & \tau^2 & 0 & \sigma_{10} \\ 0 & 0 & 0 & 0 & \tau^2 & \sigma_0^2 \end{bmatrix}.$$

Algorithm B.1: Markov chain Monte Carlo (MCMC) Sampler II

Procedure

Step 0: Initialize parameters $s = 0, \theta^{[0]}, \mathbf{\Sigma}^{[0]}$ for MCMC-chains

while $s < S$ **do**

Step 1: Update $\left((\tau^2)^\star, (\tilde{\mathbf{L}}^*)^\star\right)$ via $p(\tau^2, \tilde{\mathbf{L}}^* \mid \mathbf{Y}, \mathbf{D}, \theta^{[s]}, \mathbf{\Sigma}^{[s]})$ by

(a) sampling $(\tau^2)^\star$ from $p(\tau^2 \mid \mathbf{\Sigma}^{[s]})$: $(\tau^2)^\star \sim \text{tr}(\Psi_o \mathbf{\Sigma}^{[s]^{-1}})/\chi^2_{3\nu_o}$.

(b) sampling $(\tilde{\mathbf{L}}^*)^\star$ from $p(\tilde{\mathbf{L}}^* \mid \mathbf{Y}, \mathbf{D}, (\tau^2)^\star, \theta^{[s]}), \mathbf{\Sigma}^{[s]})$:

sampling $\mathbf{Y}^{miss}$ as in (1) and $\mathbf{D}^*$ as in (2);

setting $(\tilde{\mathbf{L}}^*)^\star = \tau^\star \mathbf{L}^*$, where $\mathbf{L}^* = [\mathbf{D}^* \quad \mathbf{Y}_1 \quad \mathbf{Y}_0]^\top$.

Step 2: Update $\left((\tau^2)^\star, \theta^{[s+1]}\right)$ via $p(\tau^2, \theta \mid \mathbf{Y}, \mathbf{D}, (\tilde{\mathbf{L}}^*)^\star, \mathbf{\Sigma}^{[s]})$ by

(a) sampling $(\tau^2)^\star$ from $p(\tau^2 \mid \mathbf{Y}, \mathbf{D}, (\tilde{\mathbf{L}}^*)^\star, \mathbf{\Sigma}^{[s]})$

$$(\tau^2)^\star \sim \frac{[(\tilde{\mathbf{L}}^*)^\star - \mathbf{R}\theta_{gls}]^\top \mathbf{\Omega}^{-1}[(\tilde{\mathbf{L}}^*)^\star - \mathbf{R}\theta_{gls}] + \theta_{gls}^\top [V_{\theta_o} + (\mathbf{R}^\top \mathbf{\Omega}^{-1}\mathbf{R})^{-1}]^{-1}\theta_{gls} + \text{tr}(\Psi_o \mathbf{\Sigma}^{[s]^{-1}})}{\chi^2_{3(n+\nu_o)}}$$

where $\theta_{gls} = (\mathbf{R}^\top \mathbf{\Omega}^{-1}\mathbf{R})^{-1}\mathbf{R}^\top \mathbf{\Omega}^{-1}(\tilde{\mathbf{L}}^*)^\star$.

(b) sampling $\tilde{\theta}^\star$ from $p(\tilde{\theta} \mid \mathbf{Y}, \mathbf{D}, (\tilde{\mathbf{L}}^*)^\star, (\tau^2)^\star, \mathbf{\Sigma}^{[s]})$:

$$\tilde{\theta}^\star \sim \mathcal{N}\left(\mu_{\theta_o}, (\tau^2)^\star (\mathbf{V}_{\theta_o}^{-1} + \mathbf{R}^\top \mathbf{\Omega}^{-1}\mathbf{R})\right),$$

where $\mu_\theta = (\mathbf{V}_{\theta_o}^{-1} + \mathbf{R}^\top \mathbf{\Omega}^{-1}\mathbf{R})\mathbf{R}^\top \mathbf{\Omega}^{-1}(\tilde{\mathbf{L}}^*)^\star$;

(c) setting $\theta^{[s+1]} = \tilde{\theta}^\star/\tau^\star$.

Step 3: Update $\left((\tau^2)^{[s+1]}, \Sigma^{[s+1]}\right)$ via $p(\tau^2, \mathbf{\Sigma} \mid \mathbf{Y}, \mathbf{D}, (\tilde{\mathbf{L}}^*)^\star, \theta^{[s+1]})$ by

(a) sampling $\tilde{\mathbf{\Sigma}}^\star$ from $p(\tilde{\mathbf{\Sigma}} \mid \mathbf{Y}, \mathbf{D}, (\tilde{\mathbf{L}}^*)^\star, \theta^{[s+1]}, \tau^\star)$: $\tilde{\mathbf{\Sigma}}^\star \sim \mathcal{W}^{-1}\left(\tilde{\mathbf{M}} + \Psi_o, n + \nu_o\right)$,

where

$$\tilde{\mathbf{M}} = \begin{bmatrix} \tilde{\epsilon}_D^\top \tilde{\epsilon}_D & \tilde{\epsilon}_D^\top \tilde{\epsilon}_1 & \tilde{\epsilon}_D^\top \tilde{\epsilon}_0 \\ \tilde{\epsilon}_1^\top \tilde{\epsilon}_D & \tilde{\epsilon}_1^\top \tilde{\epsilon}_1 & \tilde{\epsilon}_1^\top \tilde{\epsilon}_0 \\ \tilde{\epsilon}_0^\top \tilde{\epsilon}_D & \tilde{\epsilon}_0^\top \tilde{\epsilon}_1 & \tilde{\epsilon}_0^\top \tilde{\epsilon}_0 \end{bmatrix};$$

$$\tilde{\epsilon}_D = (\tilde{\mathbf{D}}^*)^\star - \mathbf{P}\tau^\star \gamma^{[s+1]}; \quad \tilde{\epsilon}_1 = (\tilde{\mathbf{Y}}_1)^\star - \mathbf{Q}\tau^\star \kappa_1^{[s+1]}; \quad \tilde{\epsilon}_0 = (\tilde{\mathbf{Y}}_0)^\star - \mathbf{Q}\tau^\star \kappa_0^{[s+1]};$$

and the first diagonal element of $\tilde{\Sigma}^\star$ must satisfy the following constraint

$$\begin{cases} \mathbf{P}_i^\top \gamma^{[s+1]} \tilde{\Sigma}_{11}^\star + \tilde{\epsilon}_{Di} > 0 & \text{if} \quad D_i = 1 \\ \mathbf{P}_i^\top \gamma^{[s+1]} \tilde{\Sigma}_{11}^\star + \tilde{\epsilon}_{Di} \leq 0 & \text{if} \quad D_i = 0 \end{cases} \quad \text{for} \quad i = 1, \ldots, n$$

(b) setting $\tau^{[s+1]} = \tilde{\Sigma}_{11}^\star, \Sigma^{[s+1]} = \tilde{\Sigma}^\star / \left(\tau^{[s+1]}\right)^2; \mathbf{L}^{*[s+1]} = \left(\tilde{\epsilon} + \tau^{[s+1]}\mathbf{R}\theta^{[s+1]}\right)/\tau^{[s+1]}$.

return $\mathbf{L}^{*[s+1]}, \theta^{[s+1]}, \mathbf{\Sigma}^{[s+1]}$

$s \leftarrow s + 1$

end while

end procedure

B.3 Beyond Normality - Bayesian Semiparametric Approach

A possible extension of the methods described above would be to use a Bayesian semiparametric approach that can accommodate heterogeneous indirect effects and relax the distributional assumption imposed on disturbance terms. We consider a finite mixture of normals

$$\begin{aligned} D_i &= \mathbb{1}\{Z_i\alpha_{\tilde{c}_i} + X_i\beta^{(D)}_{\tilde{c}_i} + \epsilon^{(D)}_i > 0\}, \\ \bar{D}_{\mathcal{N}i} &= \sum_{j=1,j\neq i}^{n} w_{ij}D_j, \quad \sum_{j=1,j\neq i} w_{ij} = 1, \\ Y^{(1)}_i &= \delta^{(1)}_{\tilde{c}_i}\bar{D}_{\mathcal{N}i} + X_i\beta^{(1)}_{\tilde{c}_i} + \epsilon^{(1)}_i, \\ Y^{(0)}_i &= \delta^{(0)}_{\tilde{c}_i}\bar{D}_{\mathcal{N}i} + X_i\beta^{(0)}_{\tilde{c}_i} + \epsilon^{(0)}_i, \\ Y_i &= D_iY^{(1)}_i + (1-D_i)Y^{(0)}_i. \end{aligned} \tag{B.4}$$

where

$$\epsilon_i \coloneqq \begin{bmatrix} \epsilon^{(D)}_i \\ \epsilon^{(1)}_i \\ \epsilon^{(0)}_i \end{bmatrix} \overset{ind}{\sim} \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \sigma_{1D,\tilde{c}_i} & \sigma_{0D,\tilde{c}_i} \\ & \sigma^2_{1,\tilde{c}_i} & \sigma_{10,\tilde{c}_i} \\ & & \sigma^2_{0,\tilde{c}_i} \end{bmatrix} \right) \tag{B.5}$$

and $\Pr(\tilde{c}_i = g) = \pi_g, \quad$ for $g = 1, \ldots, G$ and $\sum_{g=1}^{G}\pi_g = 1$.

Augmenting the model with a set of component label vector, $\{c_i\}_{i=1}^{N}$, where $c_i \coloneqq [c_{i1}, \ldots, c_{iG}]$, and $c_{ig} \coloneqq \mathbb{1}\{\tilde{c}_i = g\}$; i.e., $c_{ig} = 1$ implies that the i^{th} observation is drawn from the g^{th} component of the mixture, and 0 otherwise. Denote $\mathbf{c} \coloneqq [c_1^\top, \ldots, c_N^\top]^\top$.

By transformation, we obtain the equivalent version of (3.23)

$$\begin{aligned} &\mathbf{L}^*_i = \mathbf{R}_i\theta_{\tilde{c}_i} + \epsilon_i \quad \text{for } i = 1, \ldots, n; \\ &\mathbf{L}^* = \mathbf{R}\theta_{\tilde{c}_i} + \epsilon; \quad \text{and} \quad \mathbb{E}[\epsilon^\top\epsilon] = \mathbf{\Omega} = \mathbf{\Sigma} \otimes \mathbf{I}_n. \end{aligned} \tag{B.6}$$

We augment the parameter space with a set of component indicators (labels) $\{c_{gi}\}_{i=1}^{n}$

$$p(\mathbf{L}^*_i \mid \mathbf{c}, \Theta) = \prod_{i=1}^{n}[\phi(\mathbf{L}^*_i; \mathbf{R}_i\theta^1, \mathbf{\Sigma}^1)]^{c_{1i}} \ldots [\phi(\mathbf{L}^*_i; \mathbf{R}_i\theta^G, \mathbf{\Sigma}^G)]^{c_{Gi}}. \tag{B.7}$$

We specify the following priors for component indicators $\{c_i\}_{i=1}^n$ and component probabilities $\boldsymbol{\pi} = [\pi_1, \ldots, \pi_G]^\top$

$$c_i \mid \boldsymbol{\pi} \overset{iid}{\sim} \mathcal{M}ult(1, \boldsymbol{\pi}) \Rightarrow p(\mathbf{c} \mid \boldsymbol{\pi}) = \prod_{i=1}^{n} p(c_i \mid \boldsymbol{\pi}) = \prod_{i=1}^{n} \prod_{g=1}^{G} \pi_g^{c_{gi}}, \tag{B.8}$$

$$\boldsymbol{\pi} \sim \mathcal{D}ir(\omega_{1o}, \ldots, \omega_{Go}) \Rightarrow p(\boldsymbol{\pi}) \propto \pi_1^{\omega_{1o}-1} \ldots \pi_G^{\omega_{Go}-1}. \tag{B.9}$$

Their conditional posteriors are of the form

$$c_i \mid \Theta_{-\boldsymbol{\Sigma}}, \mathbf{Y}, \mathbf{D} \overset{ind}{\sim} \mathcal{M}ult\left(1, \left[\frac{\pi_1 \phi(\mathbf{L}_i^*; \mathbf{R}_i\theta^1, \boldsymbol{\Sigma}^1)}{\sum_{g=1}^G \pi_g \phi(\mathbf{L}_i^*; \mathbf{R}_i\theta^g, \boldsymbol{\Sigma}^g)} \cdots \frac{\pi_G \phi(\mathbf{L}_i^*; \mathbf{R}_i\theta^G, \boldsymbol{\Sigma}^G)}{\sum_{g=1}^G \pi_g \phi(\mathbf{L}_i^*; \mathbf{R}_i\theta^g, \boldsymbol{\Sigma}^g)}\right]^\top\right), \tag{B.10}$$

$$\boldsymbol{\pi} \mid \Theta_{-\boldsymbol{\pi}}, \mathbf{Y}, \mathbf{D} \sim \mathcal{D}ir(n_1 + \omega_{1o}, \ldots, n_G + \omega_{Go}). \tag{B.11}$$

B.4 Empirical Simulation Study with Friendship Network Data

In this section, we design empirical Monte Carlo experiments to illustrate how the proposed framework may be applied and to investigate the finite-sample performance of our Bayesian MCMC algorithms. The aim of an empirical Monte Carlo study is to approximate a real application in economic policy evaluation by taking as many components of the DGP as possible from real data. This simulation study relies on arguably realistic data generation processes (DGPs) based on semi-synthetic data. Specifically, an actual network structure defines the spillover patterns, and real covariates play a role as observed characteristics in both the treatment assignment equation and the outcome equation. However, to analyse the performance of the estimation procedures in different scenarios, we create a hypothetical treatment and generate the rest of the model. In particular, we make use of the Add Health friendship network data and mimic an evaluation of Social-Emotional Learning (SEL)-Focused After-School Programs on youth development.

B.4.1 Add Health Friendship Network Data

We employ in-school friendship network data obtained through a nationally representative longitudinal study of adolescents in grades 7–12 in the US between September 1994 and April 1995. We limit our analysis to the largest community ($n = 2,534$) of 84 surveyed areas.

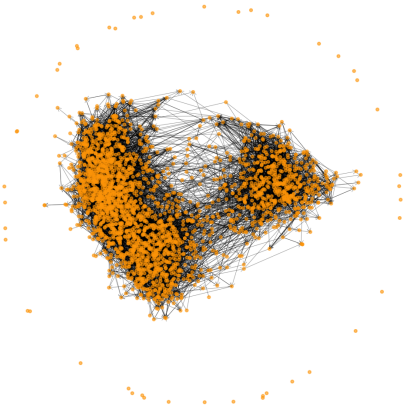

Figure B.1: Friendship network data is demonstrated in a directed node-link graph. Each node represents a student, and network links are measured using student nomination data in the survey (i.e., their best friends, up to five females and up to five males). The network is seperated into two quite distinct clusters, where the within-cluster connections are dense and between-cluster connections are more sparse.

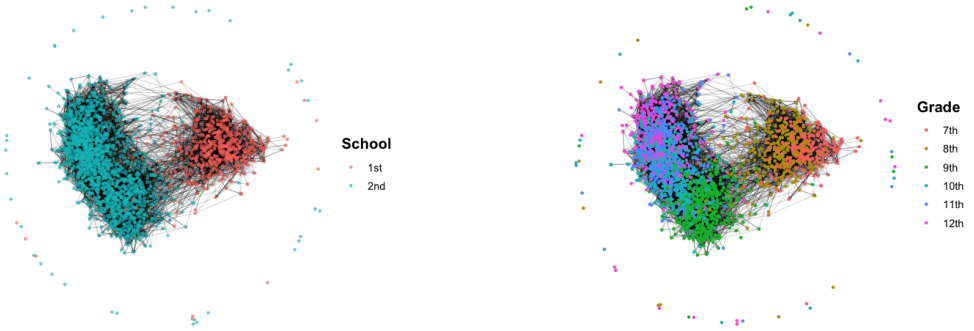

Figure B.2: Two clusters fit a middle school and a high school in the community. The grade is also a telling factor in friendship formation among students. While 9th grade is relatively distinct at the bottom of the graph, 10th, 11th, and 12th graders are harder to distinguish.

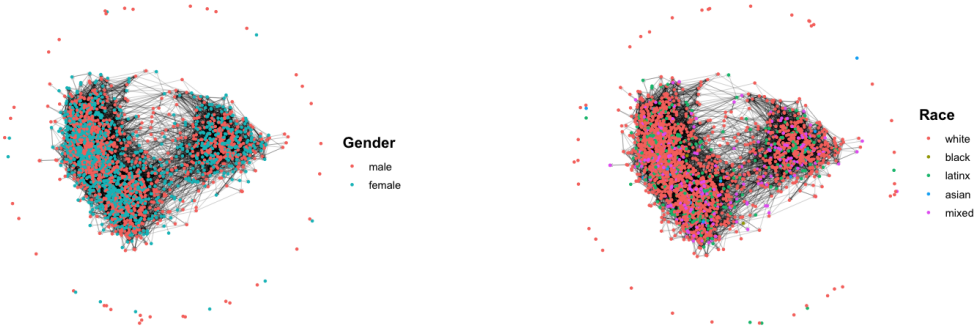

Figure B.3: Gender and race do not play a major role in friendship formation. There is a mixing of gender in all grade levels. Among race groups, White is dominant, however, there is not a clear segregation in friendship according to race.

B.4.2 Data Generating Process

Suppose that our objective is estimating the effect of a social and emotional learning (SEL)-focused after-school program on youth's prosocial development. The treatment assignment mechanism is non-random since participation is inevitably on a voluntary basis. If participants in the program are inherently different from non-participants in ways that are related to the outcomes being measured, it can be difficult to determine whether the program is responsible for any observed changes. Additionally, there is possible spillovers because joining the SEL-focused after-school program may improve youth prosocial behavior for enrolled-students, which in turn can promote prosocial behavior of their friends. The spillovers (positive contagion effects) can occur through several mechanisms. One of them is observational learning, where students may be more likely to exhibit prosocial behaviors by observing their friends' behaviors. Students who participate in the after-school program may develop emotional intelligence, which can improve their ability to understand and regulate emotions in friendship, leading to more positive interactions with their peers. Furthermore, enrolled students may tend to model prosocial behaviors they have learned and demonstrate them in their interactions with their peers. As a result, whether a student enrolls in the program or not, they may still indirectly benefit when there are more people in their friendship network who participate. Moreover, social norms may also play a role, as when students participate in the SEL after-school program, they may develop new social norms that promote prosocial behavior, such as empathy and kindness. These norms may be reinforced among participants who are friends. This indirect effect may favour enrolled students in comparison with those who do not enroll. Taking into account such potential indirect effects would allow for a more accurate estimate of the overall impact of the program on youth prosocial behavior.

Therefore, it is plausible for us to consider the general data generating process according to the model (3.7) proposed in Section 3.2 as follows:

- Let the individual treatment variable D_i be an index of participation in the After-School Program, and let Y_i be a measurement of prosocial behavior. We generate the individual treatment D_i by stylizing enrollment process based on individual choice mechanism. We thereby compute the neighbourhood treatment $\bar{D}_{Ni}$ as the proportion of treated neighbours, the number of "treated" friends among all friends, by using the adjacency matrix of the Add Health friendship network.
- $X = \left[\iota_n^\top, X_{gender}, X_{grade}, X_{race}\right]^\top$ with $X_{gender}, X_{grade}, X_{race}$ are three exogenous

variables standardized from three individual characteristics in Add Health data: $gender$, $grade$, and $race$. The corresponding regression coefficients in the selection equation and two potential outcome equations are $\beta^{(D)} = [0, -0.2, -1, 1]^\top$, $\beta^{(1)} = [2, -0.5, 0.3, 0.2]^\top$, $\beta^{(0)} = [1, 0.3, -0.4, 0.1]^\top$. Here, true values of the intercepts are fixed in all three equations.

- Z is an instrumental variable generated from $\mathcal{N}(0, 1)$ and the regression coefficient α controls the strength of the instrument: $\alpha = 1.5$. In a realistic setting, Z could be cost-shifters (e.g., the distance to the program location and assume that it varies from living directly next to the program location to living very far from the program location).
- We specify both a normal and a finite mixture of normal distribution of the error term $\epsilon_i = \left[\epsilon_i^{(D)}, \epsilon_i^{(1)}, \epsilon_i^{(0)}\right]^\top$. Throughout, we set that

$$\mathbb{E}[\epsilon_i^{(D)}] = \mathbb{E}[\epsilon_i^{(1)}] = \mathbb{E}[\epsilon_i^{(0)}] = 0; \quad \text{for } i = 1, \ldots, n$$

In summary,

$$\begin{aligned}
D_i &= \mathbb{1}\{1.5Z_i - 0.2X_{gender,i} - X_{grade,i} + X_{race,i} + \epsilon_i^{(D)} > 0\} \\
\bar{D}_i &= \sum_{j=1, j\neq i} w_{ij} D_j; \quad \sum_{j=1, j\neq i} w_{ij} = 1 \\
Y_i^{(1)} &= \delta^{(1)} \bar{D}_i + 2 - 0.5X_{gender,i} + 0.3X_{grade,i} + 0.2X_{race,i} + \epsilon_i^{(1)}, \\
Y_i^{(0)} &= \delta^{(1)} \bar{D}_i + 1 + 0.3X_{gender,i} - 0.4X_{grade,i} + 0.1X_{race,i} + \epsilon_i^{(0)}, \\
Y_i &= D_i Y_i^{(1)} + (1 - D_i) Y_i^{(0)}
\end{aligned} \tag{B.12}$$

The simulation study consists of (2×2) scenarios, which are characterized by two following factors

1. The presence of spillovers
   - without spillovers: $\delta^{(1)} = \delta^{(0)} = 0$
   - with spillovers: $\delta^{(1)} = 1.5; \delta^{(0)} = 0.5$
2. The distribution of the error term
   - a normal distribution - for $i = 1, \ldots, n$

$$\epsilon_i = \left[\epsilon_i^{(D)}, \epsilon_i^{(1)}, \epsilon_i^{(0)}\right]^\top \overset{ind}{\sim} \mathcal{N}(0, \Sigma); \quad \Sigma = \begin{bmatrix} 1 & 0.9 & 0.7 \\ & 1 & 0.6 \\ & & 1 \end{bmatrix}$$

i.e. $\sigma_D^2 = \sigma_1^2 = \sigma_0^2 = 1$; $\rho_{1D} = 0.9$; $\rho_{0D} = 0.7$; $\rho_{10} = 0.6$.

- a finite mixture of normal distribution - for $i = 1, \ldots, n$

$$\epsilon_i = \left[\epsilon_i^{(D)}, \epsilon_i^{(1)}, \epsilon_i^{(0)}\right]^\top \overset{ind}{\sim} \frac{1}{3}\mathcal{N}(0, \Sigma_1) + \frac{2}{3}\mathcal{N}(0, \Sigma_2)$$

where

$$\Sigma_1 = \begin{bmatrix} 1 & 1.5435 & 1.2005 \\ & 2.9412 & 1.7647 \\ & & 2.9412 \end{bmatrix} \quad \text{and} \quad \Sigma_2 = \begin{bmatrix} 1 & 0.1543 & 0.1200 \\ & 0.0294 & 0.0176 \\ & & 0.0294 \end{bmatrix}$$

i.e. $\sigma_D^2 = \sigma_1^2 = \sigma_0^2 = 1$; $\rho_{1D} = 0.6174$; $\rho_{0D} = 0.6$; $\rho_{10} = 0.4802$.

For each of the generated data sets, we specify two versions of models to be estimated by using the proposed Bayesian MCMC algorithm. First, Gaussian Generalised Roy model without spillovers (`GGRM-noSI`) serves as the benchmark model, without neighbourhood treatment term ($\bar{D}_{\mathcal{N}}$) and with a normal distribution of the error term. Second, Gaussian Generalised Roy model with spillovers (`GGRM-SI`) is the full model with neighbourhood treatment term ($\bar{D}_{\mathcal{N}}$) and a normal distribution of the error term. We run each MCMC algorithm for $11,000$ iterations, with the first $1,000$ draws are discarded as a burn-in period. Throughout our simulation study, the parameters for the prior distributions are chosen as follows:

$$\mu_{\theta_o} = \mathbf{0}_{15}; \quad \mathbf{V}_{\theta_o} = 10^2 * \mathbf{I}_{15\times 15};$$

$$\mathbf{\Psi}_o = \mathbf{I}_{3\times 3}; \quad \nu_o = 4;$$

The number of replicates in this study is $R = 100$.

Figure B.4: Illustration of the treatment assignment under the context of the after-school program. The node colour represents the individual treatment and the node size represents the neighbourhood treatment (e.g. the proportion of friends who are treated). Students select themselves into the program: half of them enroll while half of them do not. If the friendship network does not matter, students are only exposed to the program via their own treatment status, as demonstrated in the left graph. In contrast, if potential spillovers are considered seriously, neighbourhood treatment acts as an indirect channel through which, the student is also exposed to the program, as shown in the left graph.

B.4.3 Simulation Results

In each replication, we use the posterior mean of MCMC draws as the point estimate for each parameter of interest. We thereby compute across the 100 replicates the average bias and the root mean square error of the point estimates, followed by the coverage rate and the average length of the 95% credible intervals. The simulation results are presented in tables B.1-B.4. The true values of the DGP parameters are also listed in each table. The main findings are summarized as follows: *First,* by estimating the true DGP models, we can successfully recover the true parameter values from our Bayesian MCMC samplers with both Algorithms 3.1 and B.1. Those methods perform well in terms of average bias, root mean square error (RMSE) and coverage rate (close to the nominal level). *Second,* the performance of estimators shows different degrees of deterioration when estimating the misspecified models. When existing spillover phenomenon is not taken into account (i.e., the estimator does not include neighbourhood treatment term $\bar{D}_{\mathcal{N}}$), not only $\delta^{(1)}$ and $\delta^{(0)}$ are clearly ignored but estimating other relevant parameters ($\beta^{(1)}, \beta^{(0)}, \sigma_1^2, \sigma_0^2, \rho_{1D}, \rho_{0D}$) is also considerably affected - with an increase in both the absolute bias and the RMSE, in addition to a wider yet permissive 95% credible interval on average. In contrast, including neighbourhood treatment term when it is not needed is not harmful in general. The performance metrics are almost plausible and insensitive to the inclusion of neighbourhood treatment term no matter whether spillovers are present in the true data generating process or not, especially when Algorithm 3.1 is used. The coverage remains close to the nominal level for all parameters we are interested.

Table B.1: DGP I: without Spillovers, Normal Distribution

Model(Alg)	Metric	α	$\beta^{(D)}$	$\delta^{(1)}$	$\delta^{(0)}$	$\beta_1^{(1)}$	$\beta_1^{(0)}$	σ_1^2	σ_0^2	ρ_{1D}	ρ_{0D}	ρ_{10}
	True Value	1.500	0.000	1.500	0.500	2.000	1.000	1.000	1.000	0.900	0.700	0.600
GGRM-noSI(1)	Mean	1.514	-0.009	0.000	0.000	2.003	0.997	1.002	0.997	0.889	0.698	0.710
	Std	0.053	0.034	0.000	0.000	0.039	0.040	0.042	0.045	0.018	0.042	0.048
	Bias	0.014	-0.009	0.000	0.000	0.003	-0.003	0.002	-0.003	-0.011	-0.002	0.110
	RMSE	0.055	0.035	0.000	0.000	0.039	0.040	0.042	0.045	0.021	0.042	0.120
	Coverage	0.910	0.940	1.000	1.000	0.940	0.950	0.970	0.940	0.950	0.940	0.950
GGRM-SI(1)	Mean	1.514	-0.008	-0.004	0.009	2.004	0.993	1.002	0.997	0.890	0.699	0.705
	Std	0.053	0.034	0.086	0.095	0.055	0.060	0.042	0.045	0.018	0.042	0.048
	Bias	0.014	-0.008	-0.004	0.009	0.004	-0.007	0.002	-0.003	-0.010	-0.002	0.105
	RMSE	0.055	0.035	0.086	0.096	0.055	0.061	0.042	0.045	0.021	0.042	0.115
	Coverage	0.900	0.940	0.930	0.940	0.920	0.960	0.960	0.940	0.950	0.940	0.970
GGRM-noSI(2)	Mean	1.538	-0.012	0.000	0.000	2.047	0.956	0.958	0.981	0.829	0.607	0.288
	Std	0.060	0.037	0.000	0.000	0.065	0.057	0.048	0.057	0.079	0.102	0.332
	Bias	0.038	-0.012	0.000	0.000	0.047	-0.044	-0.043	-0.019	-0.071	-0.093	-0.312
	RMSE	0.071	0.039	0.000	0.000	0.081	0.072	0.064	0.060	0.106	0.138	0.456
	Coverage	1.000	0.980	1.000	1.000	1.000	1.000	0.990	0.990	0.950	0.930	0.660
GGRM-SI(2)	Mean	1.564	-0.019	-0.001	0.013	2.120	0.943	0.939	0.983	0.730	0.554	0.131
	Std	0.074	0.040	0.090	0.096	0.155	0.104	0.068	0.069	0.208	0.128	0.479
	Bias	0.064	-0.019	-0.001	0.013	0.120	-0.057	-0.061	-0.017	-0.170	-0.146	-0.469
	RMSE	0.097	0.044	0.090	0.097	0.196	0.118	0.092	0.071	0.268	0.194	0.670
	Coverage	1.000	0.980	0.910	0.930	0.950	0.990	0.970	0.970	0.810	0.820	0.370

Notes: This table displays results based on $R = 100$ replicates. The values include the average and standard deviation of the point estimates; the average bias (Bias), the Root Mean Squared Error (RMSE), and the coverage rate (Coverage) across replicates; where Bias $= R^{-1}\sum_{r=1}^{R}(\hat{\alpha}_r - \alpha)$, RMSE $= \sqrt{R^{-1}\sum_{r=1}^{R}(\hat{\alpha}_r - \alpha)^2}$, and Coverage $= R^{-1}\sum_{r=1}^{R}\mathbb{1}\{\alpha \in \widehat{CI}_{0.95,r}\}$.

Table B.2: DGP II: with Spillovers, Normal Distribution

Model(Alg)	Metric	α	$\beta^{(D)}$	$\delta^{(1)}$	$\delta^{(0)}$	$\beta_1^{(1)}$	$\beta_1^{(0)}$	σ_1^2	σ_0^2	ρ_{1D}	ρ_{0D}	ρ_{10}
	True Value	1.500	0.000	1.500	0.500	2.000	1.000	1.000	1.000	0.900	0.700	0.600
GGRM-noSI(1)	Mean	1.512	-0.005	0.000	0.000	2.659	1.219	1.182	1.015	0.826	0.691	0.741
	Std	0.056	0.034	0.000	0.000	0.050	0.042	0.051	0.045	0.025	0.042	0.044
	Bias	0.012	-0.005	-1.500	-0.500	0.659	0.219	0.182	0.015	-0.074	-0.009	0.141
	RMSE	0.057	0.034	1.500	0.500	0.661	0.223	0.189	0.048	0.078	0.043	0.147
	Coverage	0.930	0.950	0.000	0.000	0.000	0.000	0.050	0.930	0.150	0.940	0.950
GGRM-SI(1)	Mean	1.514	-0.009	1.496	0.510	2.005	0.994	1.002	0.997	0.889	0.699	0.707
	Std	0.053	0.034	0.085	0.095	0.054	0.060	0.042	0.045	0.018	0.042	0.050
	Bias	0.014	-0.009	-0.004	0.010	0.005	-0.006	0.002	-0.003	-0.011	-0.001	0.107
	RMSE	0.055	0.035	0.085	0.095	0.054	0.060	0.042	0.045	0.021	0.042	0.118
	Coverage	0.920	0.940	0.940	0.940	0.930	0.960	0.980	0.950	0.960	0.940	0.960
GGRM-noSI(2)	Mean	1.523	-0.009	0.000	0.000	2.714	1.204	1.155	1.026	0.746	0.628	0.280
	Std	0.062	0.036	0.000	0.000	0.127	0.074	0.070	0.189	0.157	0.085	0.404
	Bias	0.023	-0.009	-1.500	-0.500	0.714	0.204	0.155	0.026	-0.154	-0.072	-0.320
	RMSE	0.066	0.037	1.500	0.500	0.725	0.217	0.170	0.191	0.220	0.112	0.515
	Coverage	1.000	0.990	0.000	0.000	0.000	0.560	0.530	1.000	0.230	0.940	0.690
GGRM-SI(2)	Mean	1.546	-0.022	1.509	0.549	2.136	0.945	0.977	10.462	0.666	0.518	-0.098
	Std	0.094	0.040	0.103	0.229	0.172	0.695	0.104	35.229	0.240	0.106	0.364
	Bias	0.046	-0.022	0.009	0.049	0.136	-0.055	-0.024	9.462	-0.234	-0.182	-0.698
	RMSE	0.105	0.046	0.104	0.234	0.219	0.697	0.107	36.478	0.335	0.211	0.787
	Coverage	1.000	1.000	0.970	0.970	0.920	1.000	0.990	1.000	0.770	0.820	0.250

Notes: This table displays results based on $R = 100$ replicates. The values include the average and standard deviation of the point estimates; the average bias (Bias), the Root Mean Squared Error (RMSE), and the coverage rate (Coverage) across replicates; where Bias $= R^{-1}\sum_{r=1}^{R}(\hat{\alpha}_r - \alpha)$, RMSE $= \sqrt{R^{-1}\sum_{r=1}^{R}(\hat{\alpha}_r - \alpha)^2}$, and Coverage $= R^{-1}\sum_{r=1}^{R}\mathbb{1}\{\alpha \in \widehat{CI}_{0.95,r}\}$.

Table B.3: DGP III: without Spillovers, Mixture of Normal Distributions

Model(Alg)	Metric	α	$\beta^{(D)}$	$\delta^{(1)}$	$\delta^{(0)}$	$\beta_1^{(1)}$	$\beta_1^{(0)}$	σ_1^2	σ_0^2	ρ_{1D}	ρ_{0D}	ρ_{10}
	True Value	1.500	0.000	1.500	0.500	2.000	1.000	1.000	1.000	0.900	0.700	0.600
GGRM-noSI(1)	Mean	1.441	-0.002	0.000	0.000	1.959	1.034	1.021	1.006	0.674	0.540	0.671
	Std	0.076	0.033	0.000	0.000	0.039	0.045	0.082	0.081	0.042	0.061	0.082
	Bias	-0.059	-0.002	0.000	0.000	-0.041	0.034	0.021	0.006	0.056	0.060	0.071
	RMSE	0.097	0.033	0.000	0.000	0.056	0.057	0.085	0.081	0.071	0.086	0.108
	Coverage	0.750	0.980	1.000	1.000	0.890	0.890	0.720	0.720	0.820	0.850	1.000
GGRM-SI(1)	Mean	1.441	-0.002	-0.023	0.002	1.969	1.032	1.021	1.006	0.674	0.539	0.681
	Std	0.076	0.033	0.074	0.091	0.049	0.061	0.083	0.080	0.042	0.061	0.079
	Bias	-0.059	-0.002	-0.023	0.002	-0.031	0.032	0.021	0.006	0.056	0.058	0.081
	RMSE	0.096	0.033	0.078	0.091	0.058	0.069	0.085	0.080	0.070	0.085	0.114
	Coverage	0.790	0.980	0.970	0.960	0.970	0.940	0.700	0.710	0.830	0.850	0.990
GGRM-noSI(2)	Mean	1.454	-0.004	0.000	0.000	1.998	1.022	1.005	1.002	0.609	0.498	0.215
	Std	0.081	0.034	0.000	0.000	0.084	0.057	0.083	0.081	0.122	0.086	0.499
	Bias	-0.046	-0.004	0.000	0.000	-0.002	0.022	0.005	0.002	-0.009	0.018	-0.385
	RMSE	0.093	0.034	0.000	0.000	0.084	0.061	0.084	0.081	0.122	0.088	0.630
	Coverage	0.910	0.980	1.000	1.000	0.980	0.980	0.860	0.880	0.960	0.930	0.650
GGRM-SI(2)	Mean	1.468	-0.005	-0.021	0.000	2.098	1.034	1.027	1.016	0.491	0.468	-0.006
	Std	0.094	0.034	0.078	0.098	0.198	0.112	0.161	0.108	0.234	0.106	0.625
	Bias	-0.032	-0.005	-0.021	0.000	0.098	0.034	0.027	0.016	-0.127	-0.013	-0.606
	RMSE	0.099	0.035	0.081	0.098	0.221	0.117	0.163	0.109	0.266	0.107	0.871
	Coverage	0.950	0.990	0.960	0.940	0.970	0.980	0.900	0.880	0.960	0.950	0.470

Notes: This table displays results based on $R = 100$ replicates. The values include the average and standard deviation of the point estimates; the average bias (Bias), the Root Mean Squared Error (RMSE), and the coverage rate (Coverage) across replicates; where Bias $= R^{-1}\sum_{r=1}^{R}(\hat{\alpha}_r - \alpha)$, RMSE $= \sqrt{R^{-1}\sum_{r=1}^{R}(\hat{\alpha}_r - \alpha)^2}$, and Coverage $= R^{-1}\sum_{r=1}^{R}\mathbb{1}\{\alpha \in \widehat{CI}_{0.95,r}\}$.

Table B.4: DGP IV: with Spillovers, Mixture of Normal Distributions

Model(Alg)	Metric	α	$\beta^{(D)}$	$\delta^{(1)}$	$\delta^{(0)}$	$\beta_1^{(1)}$	$\beta_1^{(0)}$	σ_1^2	σ_0^2	ρ_{1D}	ρ_{0D}	ρ_{10}
	True Value	1.500	0.000	1.500	0.500	2.000	1.000	1.000	1.000	0.900	0.700	0.600
GGRM-noSI(1)	Mean	1.462	-0.004	0.000	0.000	2.636	1.249	1.186	1.021	0.601	0.527	0.699
	Std	0.074	0.033	0.000	0.000	0.046	0.045	0.084	0.082	0.052	0.059	0.101
	Bias	-0.038	-0.004	-1.500	-0.500	0.636	0.249	0.186	0.021	-0.016	0.047	0.098
	RMSE	0.084	0.033	1.500	0.500	0.638	0.253	0.204	0.084	0.054	0.076	0.141
	Coverage	0.820	0.980	0.000	0.000	0.000	0.000	0.150	0.700	0.960	0.890	0.950
GGRM-SI(1)	Mean	1.441	-0.002	1.477	0.502	1.970	1.033	1.020	1.006	0.673	0.540	0.659
	Std	0.076	0.033	0.074	0.091	0.050	0.061	0.083	0.080	0.043	0.061	0.092
	Bias	-0.059	-0.002	-0.023	0.002	-0.030	0.033	0.020	0.006	0.056	0.060	0.059
	RMSE	0.096	0.033	0.078	0.091	0.058	0.069	0.085	0.081	0.070	0.085	0.110
	Coverage	0.750	0.980	0.970	0.950	0.970	0.940	0.710	0.710	0.860	0.850	0.990
GGRM-noSI(2)	Mean	1.473	-0.006	0.000	0.000	2.697	1.239	1.180	1.016	0.511	0.479	0.027
	Std	0.070	0.034	0.000	0.000	0.113	0.059	0.099	0.082	0.144	0.072	0.454
	Bias	-0.027	-0.006	-1.500	-0.500	0.697	0.239	0.180	0.016	-0.106	-0.001	-0.573
	RMSE	0.075	0.034	1.500	0.500	0.706	0.246	0.205	0.084	0.179	0.072	0.731
	Coverage	0.960	0.990	0.000	0.000	0.000	0.420	0.360	0.890	0.960	0.980	0.720
GGRM-SI(2)	Mean	1.466	-0.005	1.501	0.530	2.103	1.036	1.036	1.020	0.467	0.468	-0.144
	Std	0.091	0.034	0.093	0.103	0.179	0.112	0.161	0.117	0.230	0.096	0.532
	Bias	-0.034	-0.005	0.001	0.030	0.103	0.036	0.036	0.020	-0.150	-0.012	-0.744
	RMSE	0.098	0.034	0.093	0.107	0.206	0.117	0.166	0.118	0.275	0.097	0.915
	Coverage	0.940	0.990	0.990	0.980	0.990	1.000	0.950	0.920	0.960	0.970	0.460

Notes: This table displays results based on $R = 100$ replicates. The values include the average and standard deviation of the point estimates; the average bias (Bias), the Root Mean Squared Error (RMSE), and the coverage rate (Coverage) across replicates; where Bias $= R^{-1}\sum_{r=1}^{R}(\hat{\alpha}_r - \alpha)$, RMSE $= \sqrt{R^{-1}\sum_{r=1}^{R}(\hat{\alpha}_r - \alpha)^2}$, and Coverage $= R^{-1}\sum_{r=1}^{R}\mathbb{1}\{\alpha \in \widehat{CI}_{0.95,r}\}$.

B.4.4 Causal Effects

Building on estimation results for model parameters, we can implement estimation and inference on causal effects of interest. Two empirical questions of interest arise in the context of SEL-focused after-school programs: First, the presence of direct treatment effects: whether enrolling the After-School Program impacts prosocial behavior of participants? Furthermore, whether there exists heterogeneity on the direct treatment effects? Second, the presence of indirect (spillover) effects: whether the participation of closed friends in the After-School Program impacts prosocial behavior of students. The following section demonstrates how the proposed procedure in previous session help us to answer empirical questions using observational data. In both illustrative scenario, estimates are compatible with true DGP.

Given an arbitrary generated dataset from the first scenario (DGP1) - with no spillovers and a normal distribution of the error term, we can obtain estimates of Average Partial Indirect Effects, Average Direct Treatment Effect, and Marginal Direct Treatment Effect as depicted in Figures B.5 and B.6: First, *Null Indirect (Spillover) Effects*: Two flat curves in Figure B.5 reveals that the participation of peers from friendship network in the After-School Program has an insignificantly positive impact on prosocial behavior of students, no matter whether they enroll the program or not. Second, *Null Interaction*: The flat curve in Figure B.6(a) shows that increasing neighbourhood treatment doesn't enhance the average direct treatment effect of ASP. Third, *Selection on Gain*: Figure B.6(b) indicates that the direct effects are higher for individuals with values of unobservables that make them more likely to attend the ASP.

Given an arbitrary generated dataset from the second scenario (DGP2) - with the presence of spillovers and a normal distribution of the error term, we can obtain estimates of Average Partial Indirect Effects, Average Direct Treatment Effect, and Marginal Direct Treatment Effect as depicted in Figures B.7 and B.8: First, *Positive Indirect (Spillover) Effects*: Two upward curves in Figure B.7 reveals that the participation of peers from friendship network in the After-School Program significantly improve the prosocial behavior of students, no matter whether they enroll the program or not (which is consistent with Observational-Learning explanation). Second, *Positive Interaction*: The upward curves in Figure B.8(a) implies that increasing neighbourhood treatment magnifies the average direct treatment effect of ASP (which is consistent with Social-Norm explanation). Third, *Selection on Gain*: Figure B.8(b) indicates that the direct effects are higher for individuals with values of unobservables that make them more likely to attend the ASP.

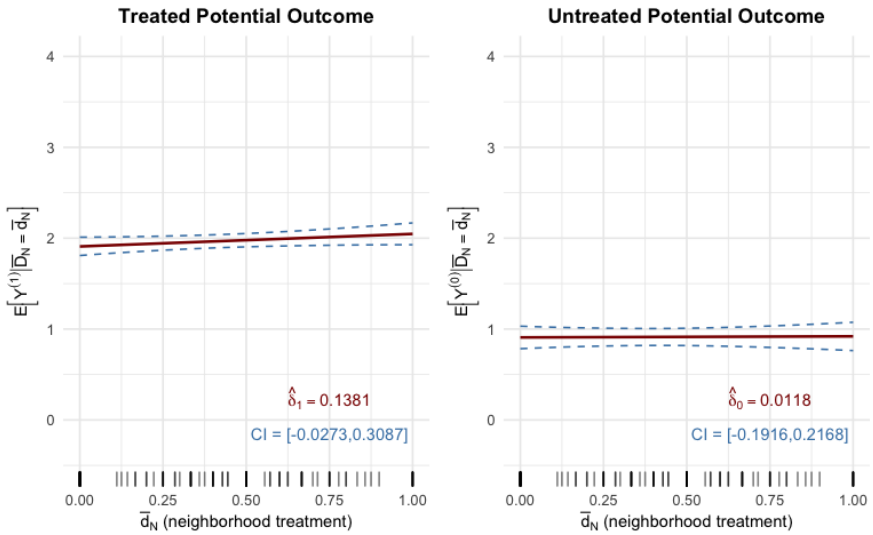

Figure B.5: The average potential outcome when being treated (left-hand side panel) and when being untreated (right-hand side panel), as a function of the neighbourhood treatment $\bar{d}_{\mathcal{N}}$. The dashed curves represent 95% confidence intervals derived from posterior samples. Ticks in the rug plot on the horizontal axis represent empirical distribution of $\bar{d}_{\mathcal{N}}$ based on the real network.

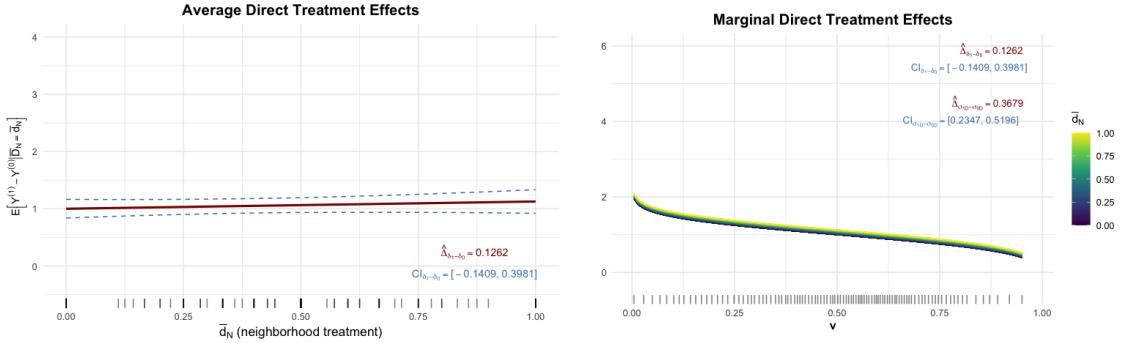

Figure B.6: (a) The average direct treatment effect as a function of the neighbourhood treatment $\bar{d}_{\mathcal{N}}$. The dashed curves represent 95% confidence intervals derived from posterior samples. Ticks in the rug plot on the horizontal axis represent empirical distribution of $\bar{d}_{\mathcal{N}}$ based on the real network. (b) The marginal direct treatment effect as a function of the neighbourhood treatment $\bar{d}_{\mathcal{N}}$ and the unmeasured resistance level v. Ticks in the rug plot on the horizontal axis represent empirical distribution of v, obtained from posterior samples.

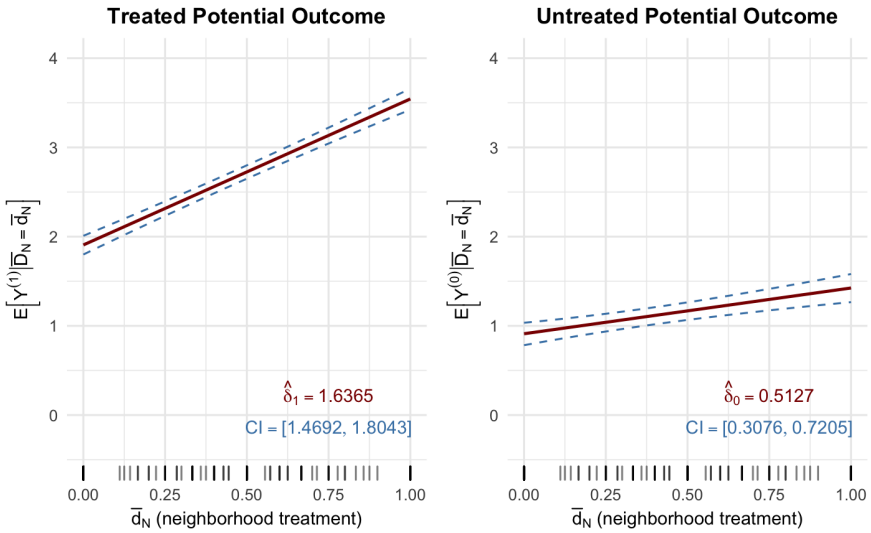

Figure B.7: The average potential outcome when being treated (left-hand side panel) and when being untreated (right-hand side panel), as a function of the neighbourhood treatment $\bar{d}_{\mathcal{N}}$. The dashed curves represent 95% confidence intervals derived from posterior samples. Ticks in the rug plot on the horizontal axis represent empirical distribution of $\bar{d}_{\mathcal{N}}$ based on the real network.

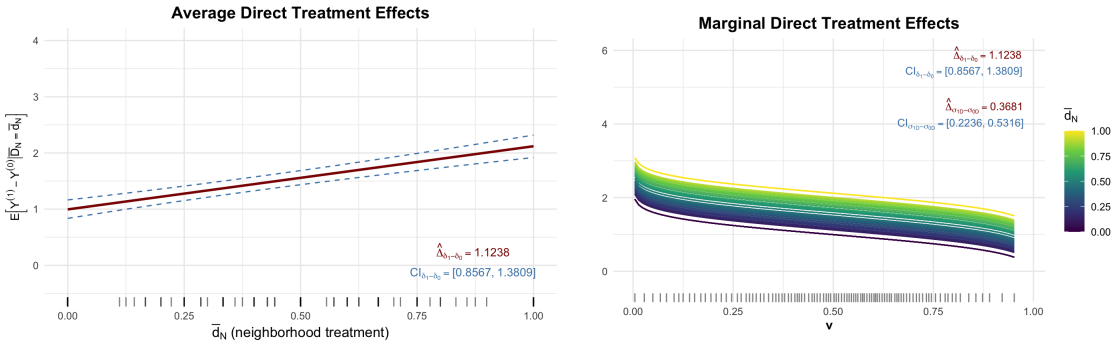

Figure B.8: (a) The average direct treatment effect as a function of the neighbourhood treatment $\bar{d}_{\mathcal{N}}$. The dashed curves represent 95% confidence intervals derived from posterior samples. Ticks in the rug plot on the horizontal axis represent empirical distribution of $\bar{d}_{\mathcal{N}}$ based on the real network. (b) The marginal direct treatment effect as a function of the neighbourhood treatment $\bar{d}_{\mathcal{N}}$ and the unmeasured resistance level v. Ticks in the rug plot on the horizontal axis represent empirical distribution of v, obtained from posterior samples.

B.5 On Empirical Application

Table B.5: Variable Description

Variable	Description
Housing Unit Growth	Growth of total housing units between 2017-2022
QOZ	An indicator equal to one if an eligible tract was selected as an Opportunity Zone, or zero otherwise.
Political Affiliation	An indicator equal to one if a tract's representative to the state's lower house is of the same political party as the state's governor, and zero otherwise.
Poverty Rate	The proportion of residents in a tract whose ratio of income to the poverty threshold is less than or equal to 0.99, scaled by the number of residents.
Median Earnings	Logarithm of the median earnings in a tract.
Employment Rate	The number of individuals in the labor force in a tract that are working, either in civilian or Armed Forces, scaled by the total labor force of the tract.
% White	The proportion of non-Hispanic white residents in a tract.
% Native hc covered	The proportion of native-born individuals covered by health insurance in a tract.
% Higher ed.	The proportion of the population in a tract with at least a high school education.
% Rent	The proportion of rental unit in a tract.
Population	The total population in a tract from the 2010 Census.

Table B.6: Data Sources

Source	URL
IRS (Internal Revenue Service) data: List of Qualified Opportunity Zones	https://www.irs.gov/credits-deductions/businesses/opportunity-zones
Urban Institute's data	https://www.urban.org/policy-centers/metropolitan-housing-and-communities-policy-center/projects/opportunity-zones
American Community Survey (ACS) 5-Year Data (2009-2022)	https://www.census.gov/data/developers/data-sets/acs-5year.html
California State Legislature	https://ballotpedia.org/California_State_Legislature
TIGER Geographic Shapefiles	https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html
SLDU and SLDL Blocks Splits	https://www.census.gov/geographies/mapping-files/2018/dec/rdo/2018-state-legislative-bef.html

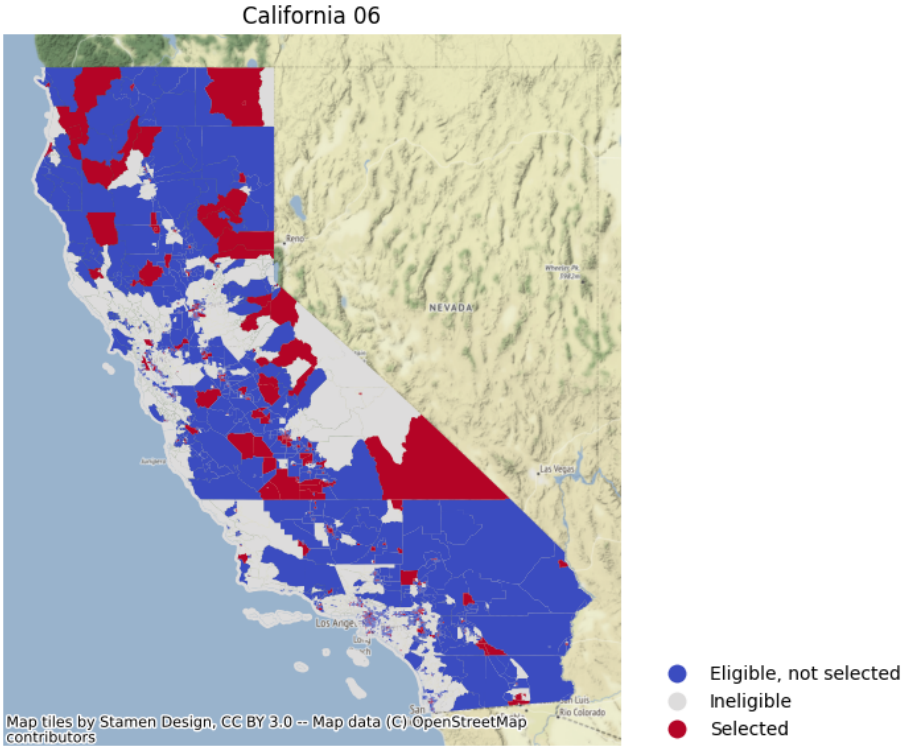

Figure B.9: Map of Opportunity Zone status of census tracts in California.

Appendix C

Appendix for Chapter 4

C.1 Details of Computational Algorithms

C.1.1 Adaptation of Random-Walk Metropolis

Adaptation of random-walk Metropolis algorithm (Andrieu and Thoms, 2008; Atchadé and Rosenthal, 2005; Roberts and Rosenthal, 2009; Vihola, 2022) is designed to improve the efficiency of standard random-walk Metropolis algorithm, by automatically tuning the proposal distribution during the simulation based on historical MCMC samples. The goal is to achieve better convergence and mixing without requiring manual calibration.

Formally, let $\theta \in \Theta \subset \mathbb{R}^d$ be the parameter of interest with target probability density $\pi(\theta)$. Our goal is to generate a dependent random sequence $\{\theta^{[0]}, \ldots, \theta^{[S]}\}$, called a chain, whose stationary density matches π. At iteration s, the standard algorithm uses a normal random walk proposal $\mathcal{N}\left(\theta^{[s]}, \tau^{[s]}\Sigma_\theta^{[s]}\right)$, where $\theta^{[s]}$ is the current state, $\tau^{[s]}$ is a scalar scaling factor (i.e., how far the proposal jumps in the parameter space), and $\Sigma_\theta^{[s]}$ is an empirical estimate of the covariance matrix of the target π. The acceptance probability of a candidate draw $\theta^\star$ is $\alpha = \min\left\{1, \pi\left(\theta^\star\right)/\pi\left(\theta^{[s]}\right)\right\}$. The shape of the proposal distribution has a substantial effect on the algorithm's mixing behaviour[1], while choosing τ and Σ_θ in each iteration by trial and error is both time-consuming and problem-specific. To overcome this tuning challenge, adaptation allows the sampler learn τ and Σ_θ on-the-fly using past draws. The core idea is to adaptively update these parameters to target a prespecified "optimal" acceptance rate ($\bar{\alpha}_{opt}$).

Among several variants, we present in Algorithm C.1 the Adaptive Metropolis algorithm with global adaptive scaling (Atchadé and Rosenthal, 2005). We note that, for the adaptation step, equation (C.1) aims to keep the asymptotic acceptance rate of the algorithm close

[1] Intuitively, if $\tau^{[s]}\Sigma_\theta^{[s]}$ is either too large in some directions or too small in all directions the algorithm has either a very small or a very large acceptance probability, which results in a very poor exploration of the target distribution as the algorithm mix poorly.

to an optimal value $\bar{\alpha}_{opt}$. Commonly, $\bar{\alpha}_{opt} = 0.234$ for a multivariate target ($d > 1$) and $\bar{\alpha}_{opt} = 0.44$ when $d = 1$. The (recommended) adaptation stepsize is a decreasing sequence $\nu^{[s+1]} = (s+1)^{-2/3}$. This approach is backed by Robbins-Monro recursive and more generally, the stochastic approximation framework (Benveniste et al., 2012), and aims to optimize the efficiency of the MCMC sampler by monitoring the acceptance rate toward a theoretically optimal region. The updates are designed to vanish over time (i.e., $\nu^{[s]} \to 0$) to ensure the ergodicity and convergence of the chain.

Algorithm C.1: Adaptive Scaling Metropolis (ASM) algorithm - *General case*

Procedure

Step 0: initialise $s = 0, \theta^{[0]}, \mu_\theta^{[0]} = \theta^{[0]}, \Sigma_\theta^{[0]}, \tau^{[0]}$.

while $s < S$ **do**

Step 1 (Proposal step):

Sample a new candidate $\theta^\star \sim \mathcal{N}\left(\theta^{[s]}, \tau^{[s]}\Sigma_\theta^{[s]}\right)$

Accept $\theta^\star$ with probability $\alpha\left(\theta^\star, \theta^{[s]}\right)$. If accepted, $\theta_{s+1} = \theta^\star$; otherwise, $\theta^{[s+1]} = \theta^{[s]}$.

Step 2 (Adaptation step):

Update the scaling

$$\log\left(\tau^{[s+1]}\right) = \log\left(\tau^{[s]}\right) + \nu^{[s+1]}\left[\alpha\left(\theta^\star, \theta^{[s]}\right) - \bar{\alpha}_{opt}\right], \tag{C.1}$$

Update the empirical covariance

$$\begin{aligned}\mu_\theta^{[s+1]} &= \mu_\theta^{[s]} + \nu^{[s+1]}\left[\theta^{[s+1]} - \mu_\theta^{[s]}\right], \\ \Sigma_\theta^{[s+1]} &= \Sigma_\theta^{[s]} + \nu^{[s+1]}\left[\left(\theta^{[s+1]} - \mu_\theta^{[s]}\right)\left(\theta^{[s+1]} - \mu_\theta^{[s]}\right)^\top - \Sigma_\theta^{[s]}\right].\end{aligned} \tag{C.2}$$

return $\theta^{[s+1]}, \mu_\theta^{[s+1]}, \Sigma_\theta^{[s+1]}, \tau^{[s+1]}$

$s \leftarrow s + 1$

end while

end procedure

For our sampling problem for network interaction effects λ_g ($g = 1, \ldots, G$) in Section 4.3, we apply a univariate version of Algorithm C.1. Specifically, we employ Algorithm C.2 which corresponds to $d = 1$ and the optimal acceptance rate $\alpha_{opt} = 0.44$. In practice, whenever $\alpha\left(\theta^\star, \theta^{[s]}\right)$ falls below $\bar{\alpha}_{opt}$ for most transition attempts, the log-scale update helps increase the scaling parameter $\tau^{[s]}$, and vice versa. This automatic feedback allows the sampler to recover quickly from poor initialisation, enabling an efficient exploration of the posteriors of interest. We embed this Adaptive Scaling Metropolis sampling step for each λ_g into our main

MCMC samplers for the SCHSAR model (Algorithms 4.1 and 4.2).

Algorithm C.2: Adaptive Scaling Metropolis (ASM) algorithm - *Univariate case*

Procedure

Step 0: initialise $s = 0, \theta^{[0]}, \tau^{[0]}$.

while $s < S$ **do**

Step 1 (Proposal step):

Sample a new candidate $\theta^{\star} \sim \mathcal{N}\left(\theta^{[s]}, \tau^{[s]}\right)$

Accept $\theta^{\star}$ with probability $\alpha\left(\theta^{\star}, \theta^{[s]}\right)$. If accepted, $\theta_{s+1} = \theta^{\star}$; otherwise, $\theta^{[s+1]} = \theta^{[s]}$.

Step 2 (Adaptation step): Update the scaling

$$\log\left(\tau^{[s+1]}\right) = \log\left(\tau^{[s]}\right) + \nu^{[s+1]}\left[\alpha\left(\theta^{\star}, \theta^{[s]}\right) - \bar{\alpha}_{opt}\right].$$

return $\theta^{[s+1]}, \tau^{[s+1]}$

$s \leftarrow s + 1$

end while

end procedure

C.1.2 Community Detection Algorithms

When unobserved homophily is present (see Subsection 4.3.4), obtaining a reliable initialisation for the latent individual heterogeneity $\{a_i\}_{i=1}^{N}$ can significantly improve the MCMC convergence. A practical strategy is to apply available community detection algorithms to the observed network, using the $N \times N$ adjacency matrix $\mathbf{W} = [wij]$ to uncover two latent clusters corresponding to the binary types of a_i ("high" vs. "low"). Below we adopt two spectral-clustering variants that exploit the eigenstructure of graph matrices to reveal communities. In both algorithms C.3 and C.4, we extract the top T eigenvectors of a chosen matrix (e.g., the Laplacian or modularity matrix), embed the N nodes into a T-dimensional spectral subspace, and then apply a standard K-means method to partition them into communities (with labels $\{\tilde{a}_i\}_{i=1}^{N}$).

Algorithm C.3: Modularity-Matrix Spectral Clustering algorithm

Data: Adjacency matrix $\mathbf{W} \in \mathbb{R}^{N\times N}$, # of clusters K, # of top eigenvectors T (set $T = K = 2$).

Result: Community assignments $\{\tilde{a}_i\}_{i=1}^N$.

1 Compute degree vector $\mathbf{d} \in \mathbb{R}^N$ where $d_i \coloneqq \sum_j w_{ij}$ and total edges $m = \frac{1}{2}\sum_{i=1}^N d_i$.

2 Form the modularity matrix

$$\mathbf{M} \coloneqq \mathbf{W} - \frac{\mathbf{d}\mathbf{d}^\top}{2m}.$$

3 Compute T eigenvectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_T\}$ of $\mathbf{M}$ with largest positive eigenvalues.

4 Stack these eigenvectors into $\mathbf{V} \coloneqq [\mathbf{v}_1, \ldots, \mathbf{v}_T] \in \mathbb{R}^{N\times T}$.

5 Apply K-means clustering to the rows of $\mathbf{V}$, yielding labels $\{\tilde{a}_i\}_{i=1}^N$.

Algorithm C.4: Normalized-Laplacian Spectral Clustering algorithm

Data: Adjacency matrix $\mathbf{W} \in \mathbb{R}^{N\times N}$, # of clusters K, # of top eigenvectors T (set $T = K = 2$).

Result: Community assignments $\{\tilde{a}_i\}_{i=1}^N$.

1 Compute degree matrix $\mathbf{D} \coloneqq \mathrm{diag}(d_1, \ldots, d_N)$ with $d_i \coloneqq \sum_j w_{ij}$.

2 Form the normalized Laplacian matrix

$$\mathbf{L}_{\text{norm}} \coloneqq \mathbf{I} - \mathbf{D}^{-\frac{1}{2}}\mathbf{W}\mathbf{D}^{-\frac{1}{2}}.$$

3 Compute T eigenvectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_T\}$ of $\mathbf{L}_{\text{norm}}$ with smallest nonzero eigenvalues.

4 Stack these eigenvectors into $\mathbf{V} \coloneqq [\mathbf{v}_1, \ldots, \mathbf{v}_T] \in \mathbb{R}^{N\times T}$.

5 Row-normalize $\mathbf{V}$

$$\widetilde{V}_{i,:} \coloneqq \frac{V_{i,:}}{\|V_{i,:}\|_2}, \quad i = 1, \ldots, N.$$

6 Apply K-means clustering to the rows of $\widetilde{\mathbf{V}}$, yielding labels $\{\tilde{a}_i\}_{i=1}^N$.

By projecting nodes into a low-dimensional spectral subspace ($T \ll N$) instead of clustering directly in the original node space, these approaches effectively capture global connectivity and reveal nodes that are "structurally similar" in the same community. Alternative algorithms are also available in the MATLAB toolbox for community detection by Kehagias (2018). In Monte Carlo experiments, we found that perfect clustering is unnecessary for our proposed MCMC sampler in Algorithm 4.2. Moderate misclassification in the starting values is corrected throughout the MCMC procedure, which refines a_i using both network and outcome information.

C.2 On Simulation Study

C.2.1 Diagnostic Plots

DGP I(a): $N = 1000$, unobserved degree heterogeneity, high SNR

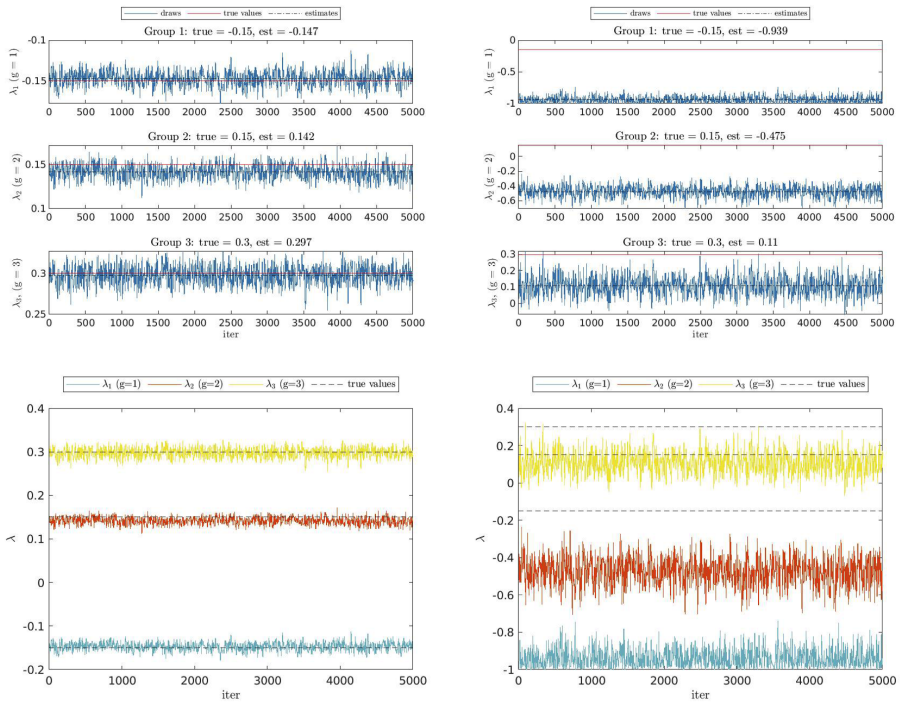

Figure C.1: Draws for λ_g (SCHSAR-left and HSAR-right).

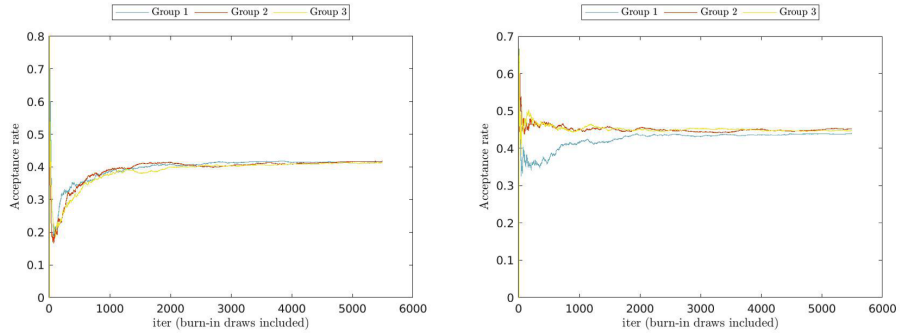

Figure C.2: M-H acceptance rate of λ_g (SCHSAR-left and HSAR-right).

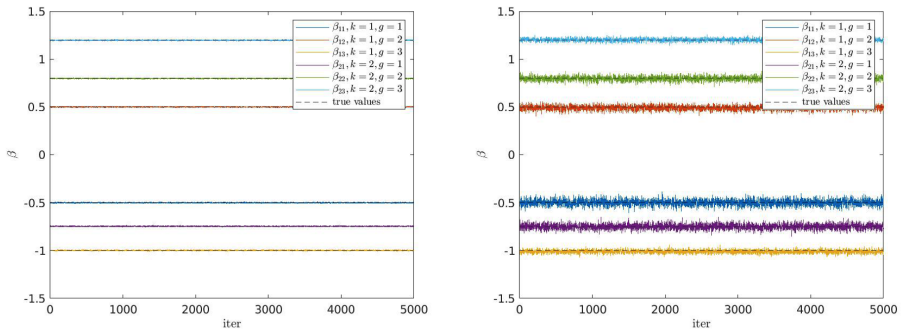

Figure C.3: Draws for β_g (SCHSAR-left and HSAR-right).

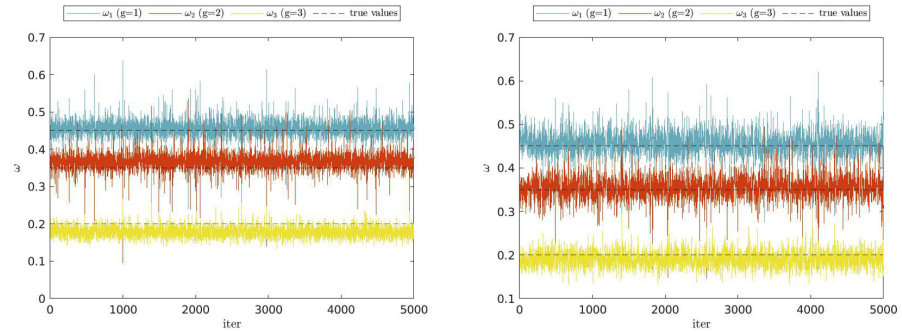

Figure C.4: Draws for ω_g (SCHSAR-left and HSAR-right).

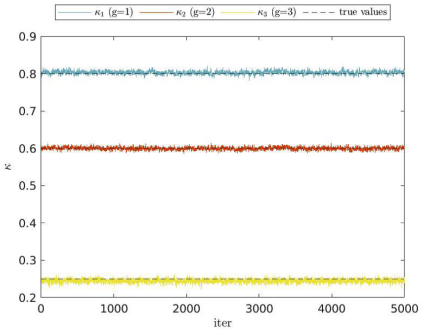

Figure C.5: Draws for κ_g (SCHSAR only).

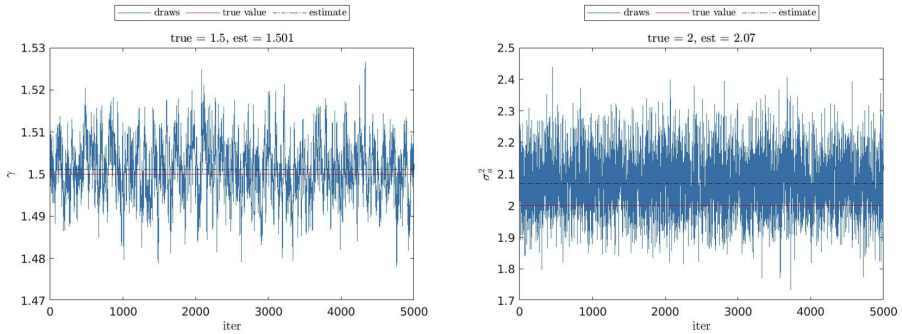

Figure C.6: Draws for γ and σ_a^2 (SCHSAR only).

DGP I(b): $N = 1000$, unobserved degree heterogeneity, medium SNR

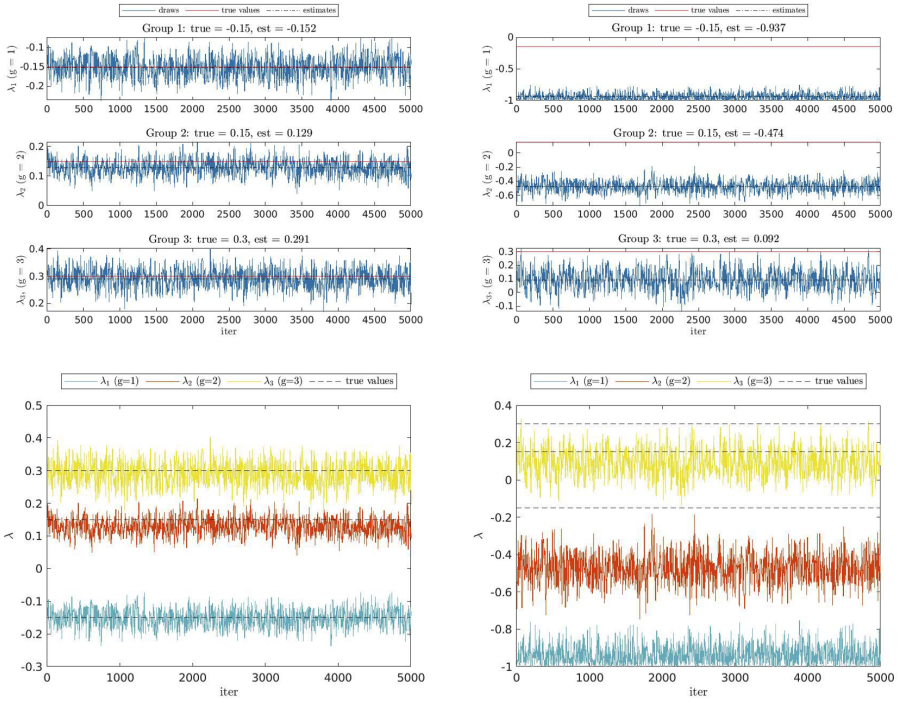

Figure C.7: Draws for λ_g (SCHSAR-left and HSAR-right).

DGP I(c): $N = 1000$, unobserved degree heterogeneity, low SNR

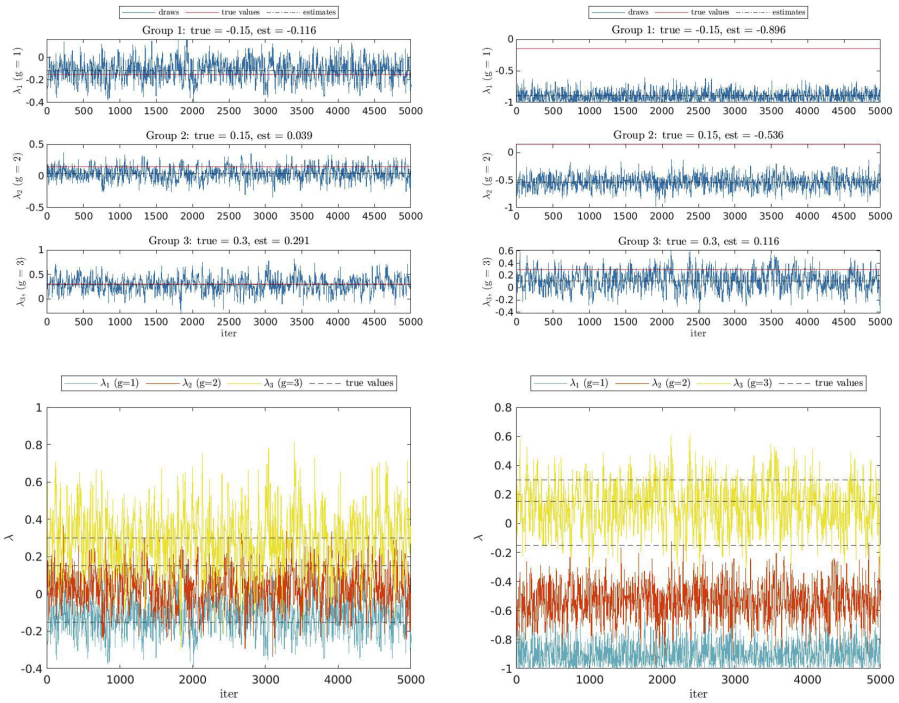

Figure C.8: Draws for λ_g (SCHSAR-left and HSAR-right).

DGP II(a): $N = 1000$, unobserved degree heterogeneity, link misspecification, high SNR

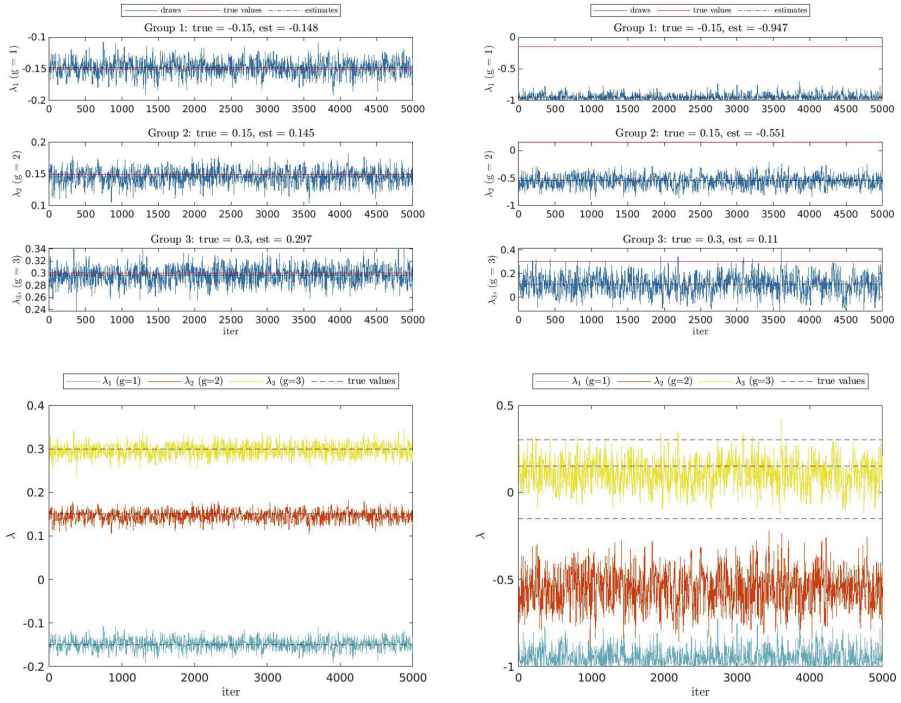

Figure C.9: Draws for λ_g (SCHSAR-left and HSAR-right).

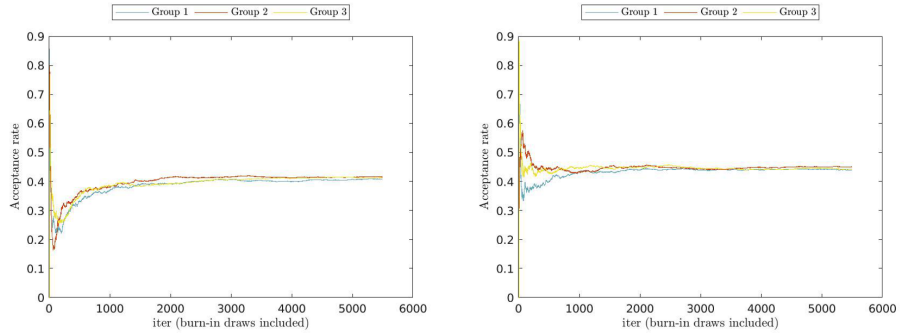

Figure C.10: M-H acceptance rate of λ_g (SCHSAR-left and HSAR-right).

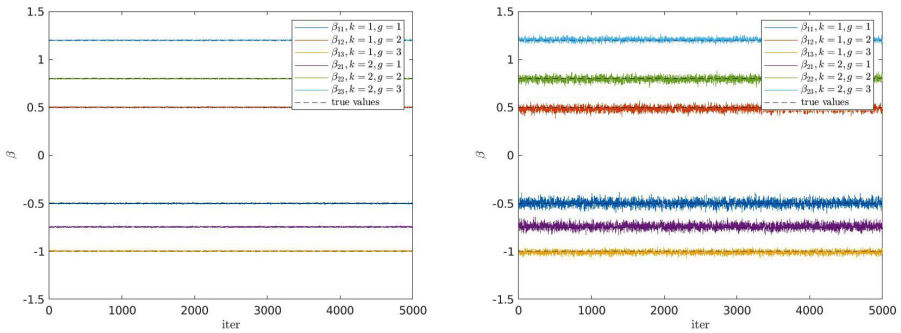

Figure C.11: Draws for β_g (SCHSAR-left and HSAR-right).

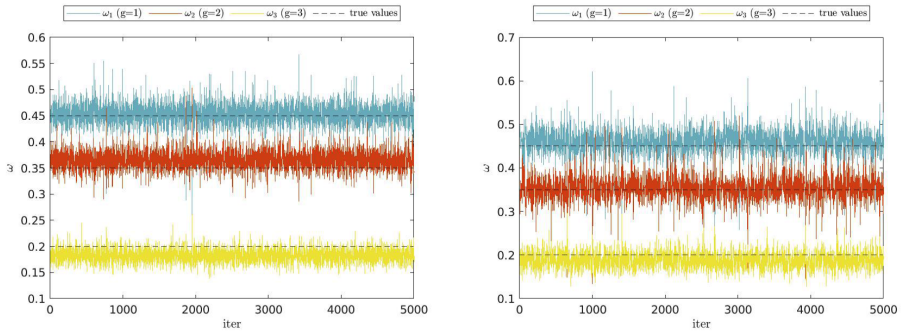

Figure C.12: Draws for ω_g (SCHSAR-left and HSAR-right).

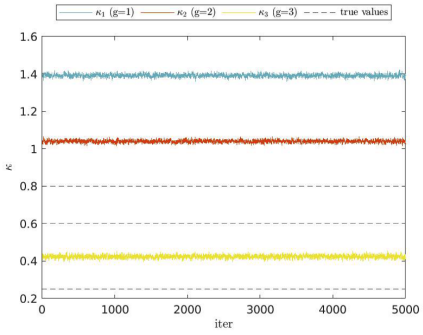

Figure C.13: Draws for κ_g (SCHSAR only).

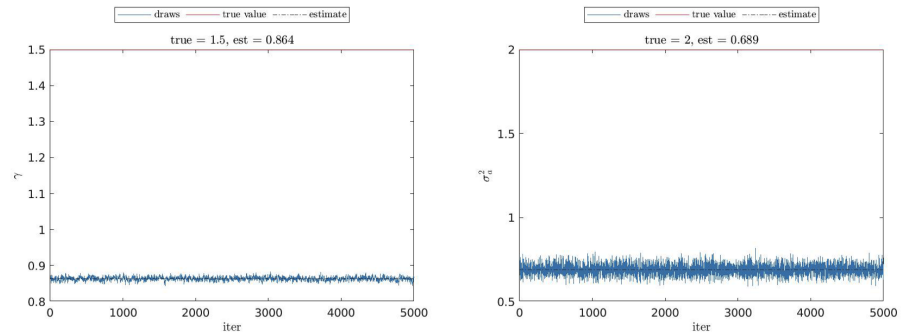

Figure C.14: Draws for γ and σ_a^2 (SCHSAR only).

DGP II(b). $N = 1000$, unobserved degree heterogeneity, link misspecification, medium SNR

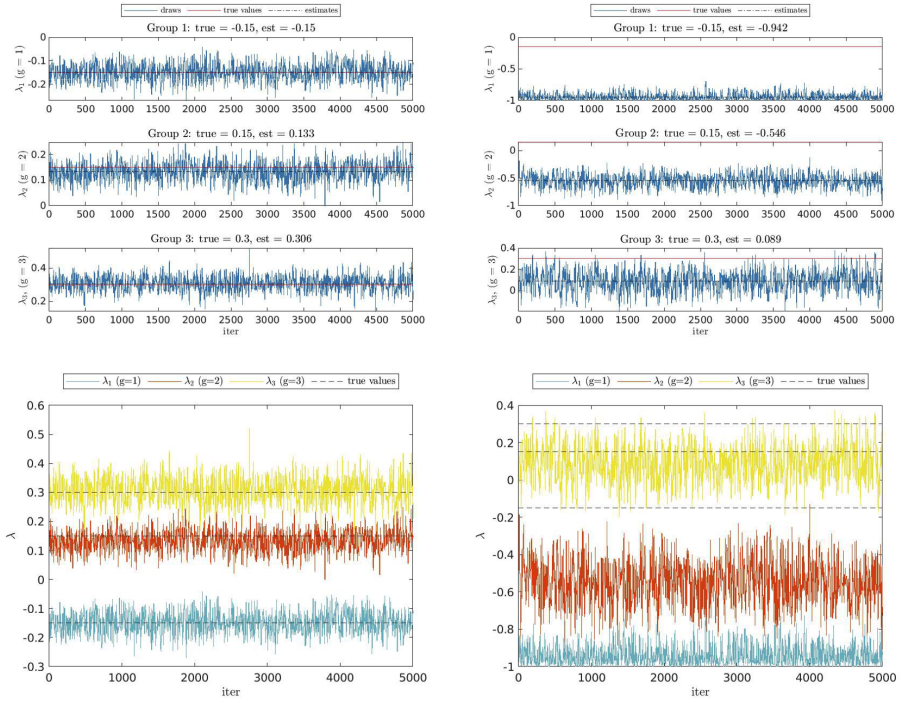

Figure C.15: Draws for λ_g (SCHSAR-left and HSAR-right).

DGP II(c). $N = 1000$, unobserved degree heterogeneity, link misspecification, low SNR

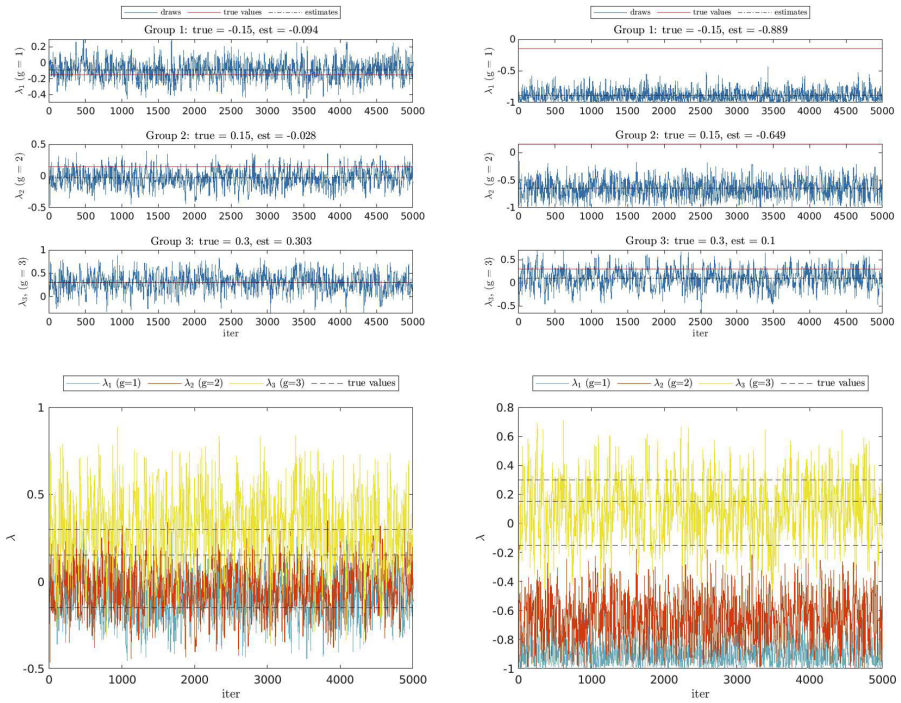

Figure C.16: Draws for λ_g (SCHSAR-left and HSAR-right).

DGP III(a). $N = 1000$, unobserved homophily, high SNR

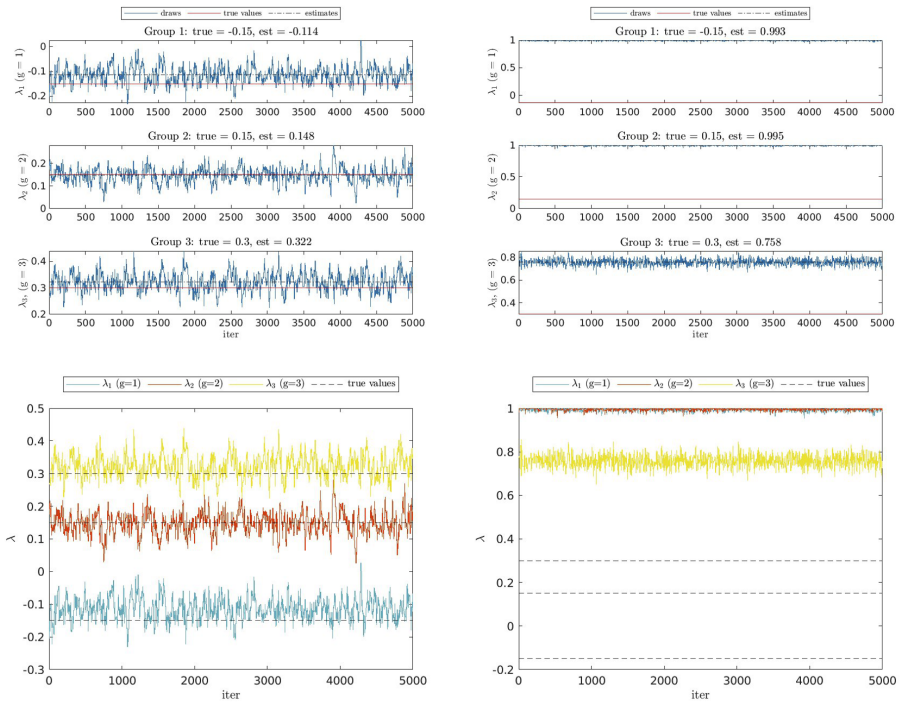

Figure C.17: Draws for λ_g (SCHSAR-left and HSAR-right).

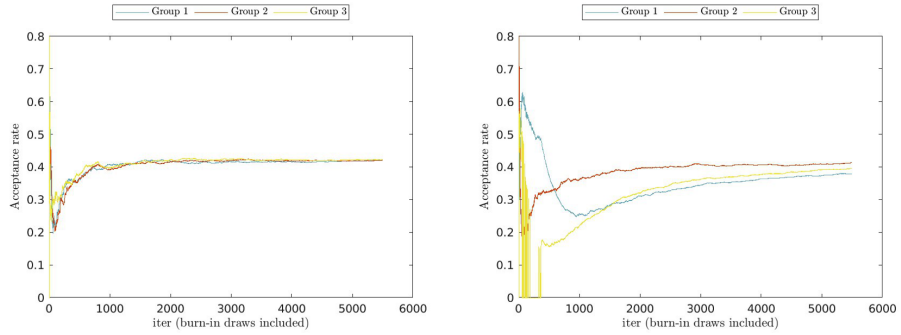

Figure C.18: M-H acceptance rate of λ_g (SCHSAR-left and HSAR-right).

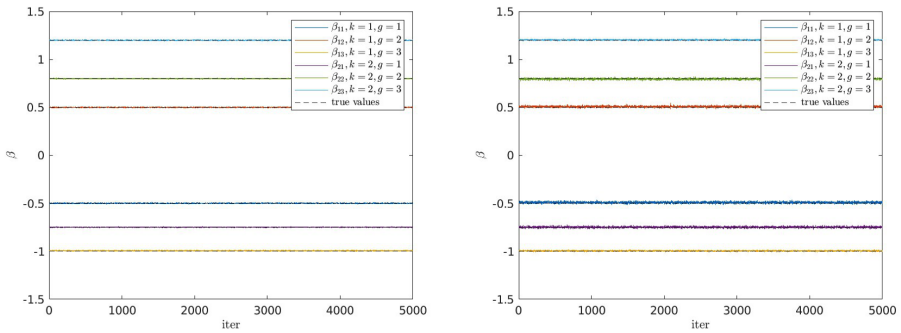

Figure C.19: Draws for β_g (SCHSAR-left and HSAR-right).

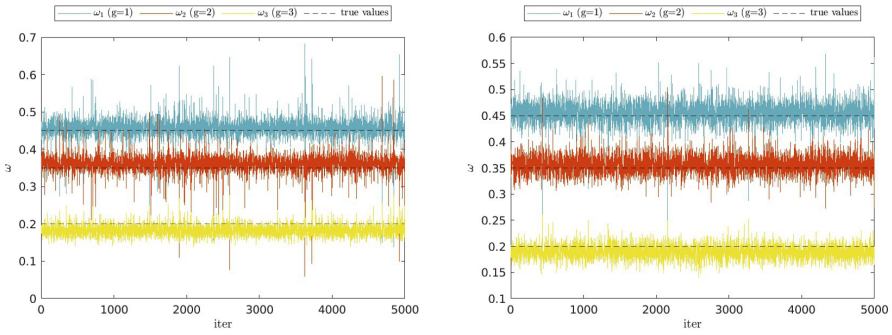

Figure C.20: Draws for ω_g (SCHSAR-left and HSAR-right).

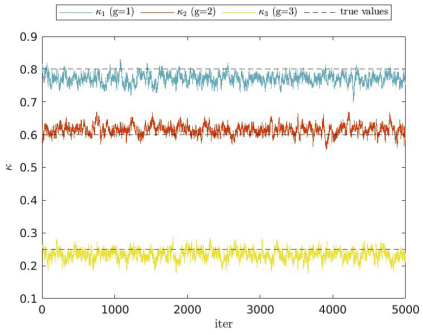

Figure C.21: Draws for κ_g (SCHSAR only).

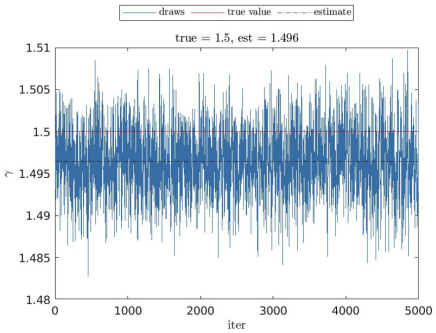

Figure C.22: Draws for γ (SCHSAR only).

DGP III(b). $N = 1000$, unobserved homophily, medium SNR

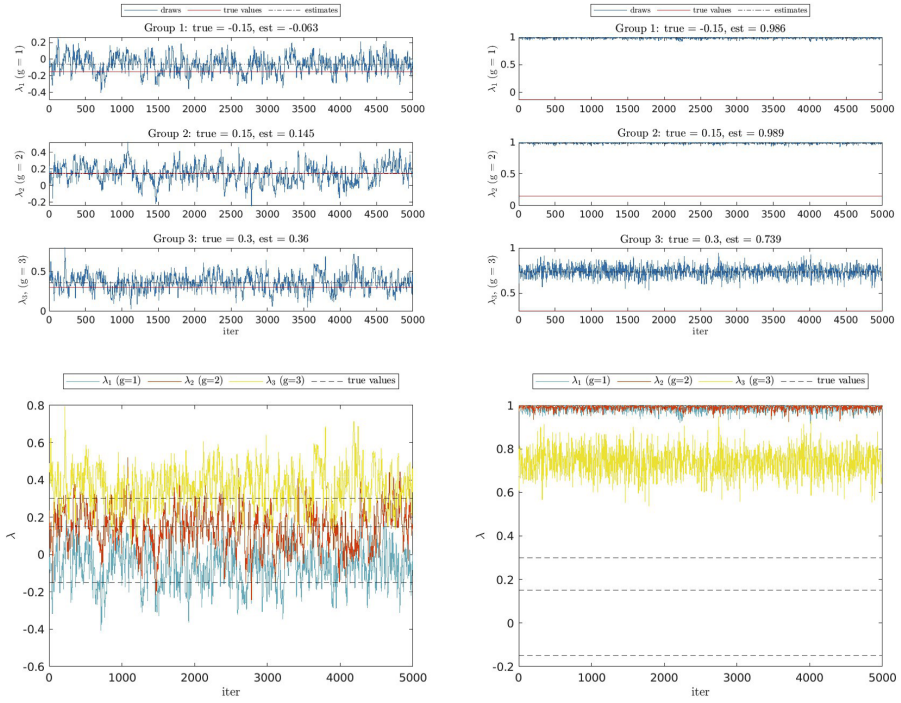

Figure C.23: Draws for λ_g (SCHSAR-left and HSAR-right).

DGP III(c). $N = 1000$, unobserved homophily, low SNR

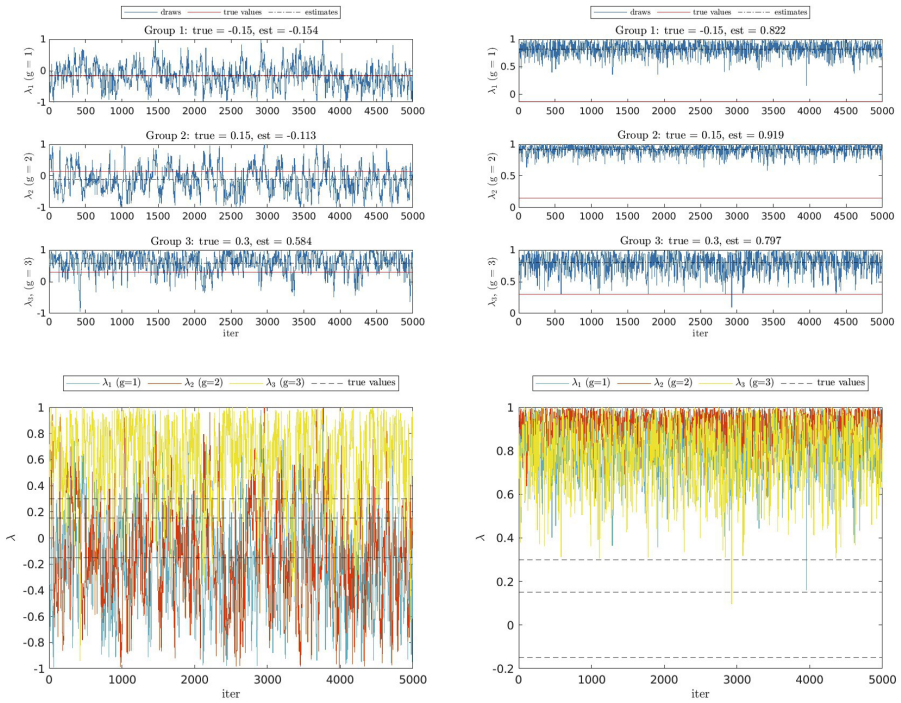

Figure C.24: Draws for λ_g (SCHSAR-left and HSAR-right).

DGP IV(a). $N = 1000$, unobserved homophily, link misspecification, high SNR

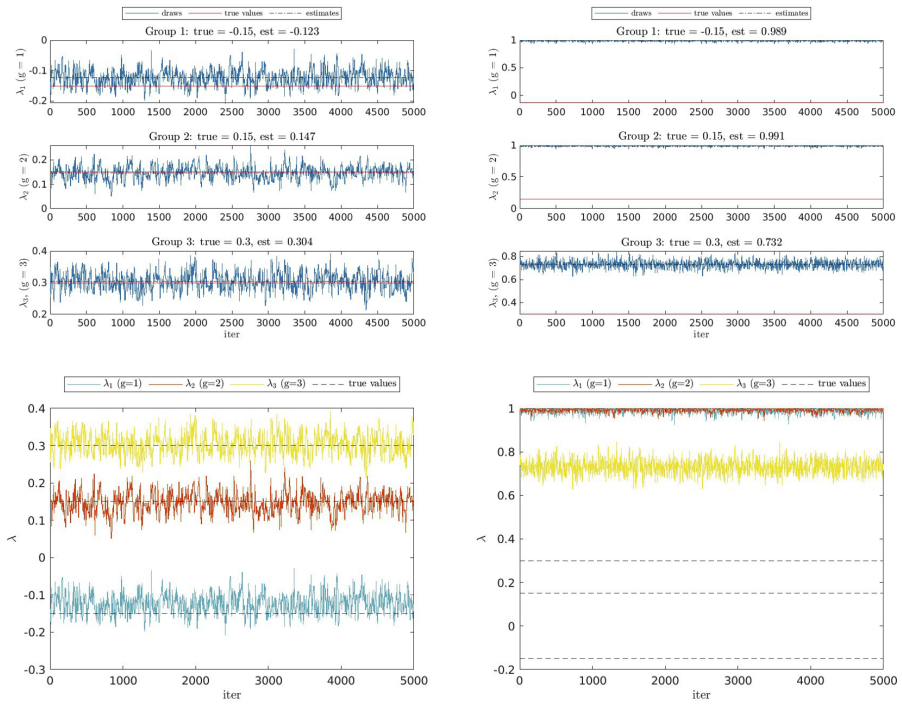

Figure C.25: Draws for λ_g (SCHSAR-left and HSAR-right).

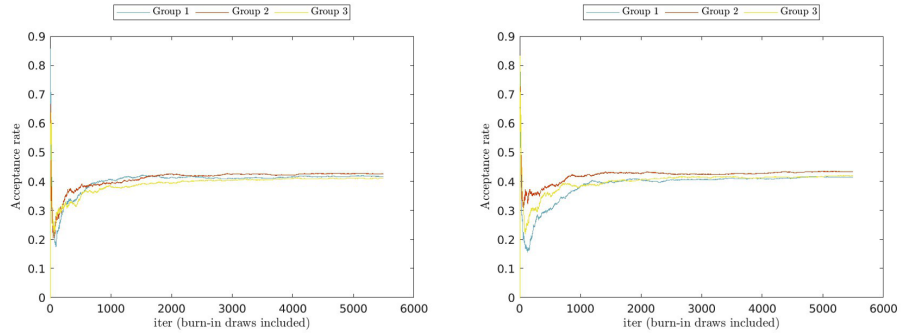

Figure C.26: M-H acceptance rate of λ_g (SCHSAR-left and HSAR-right).

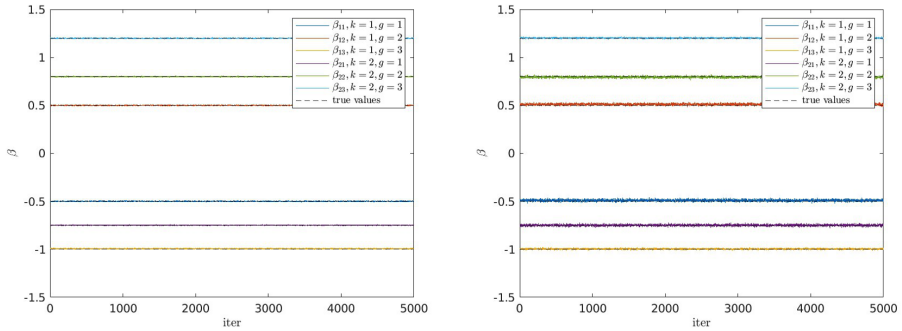

Figure C.27: Draws for β_g (SCHSAR-left and HSAR-right).

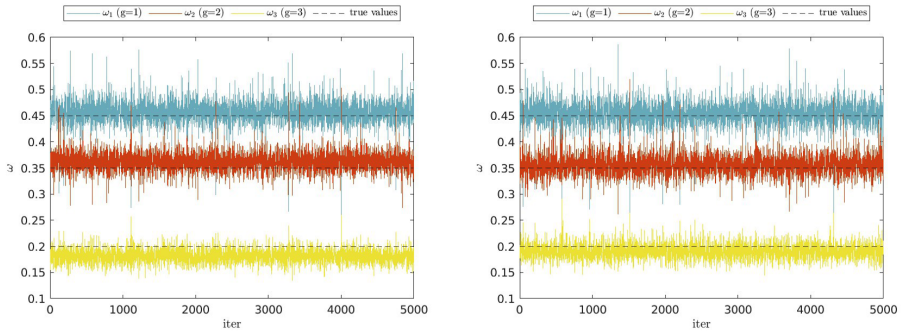

Figure C.28: Draws for ω_g (SCHSAR-left and HSAR-right).

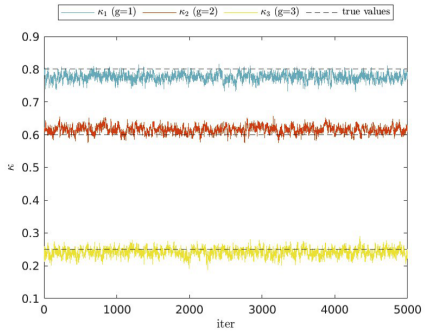

Figure C.29: Draws for κ_g (SCHSAR only).

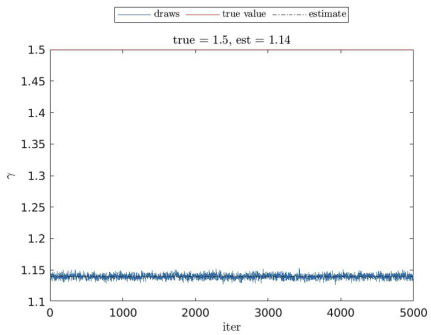

Figure C.30: Draws for γ (SCHSAR only).

DGP IV(b). $N = 1000$, unobserved homophily, link misspecification, medium SNR

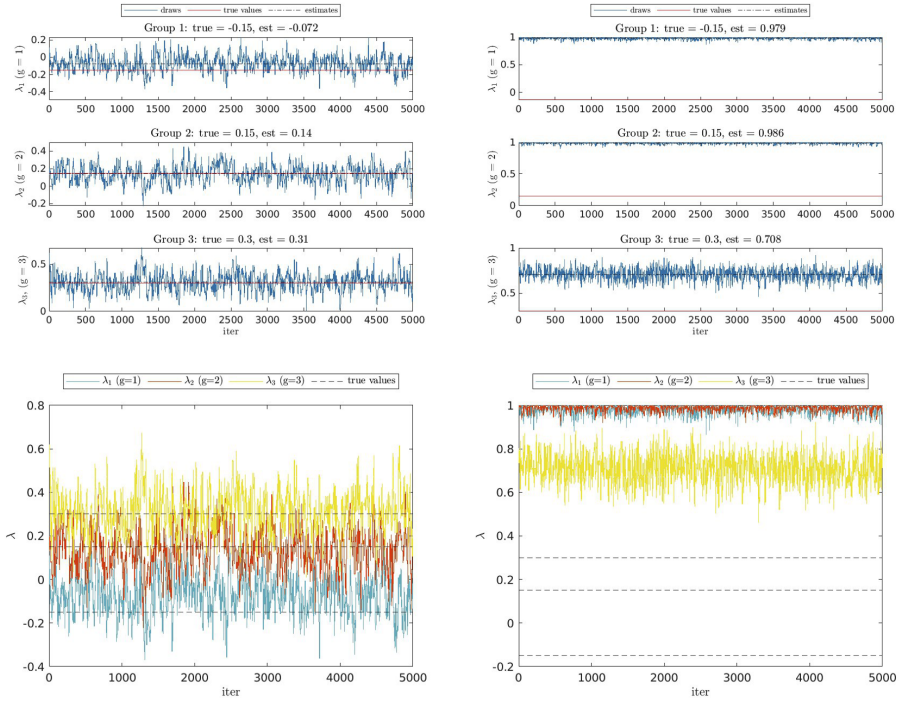

Figure C.31: Draws for λ_g (SCHSAR-left and HSAR-right).

DGP IV(c). $N = 1000$, unobserved homophily, link misspecification, low SNR

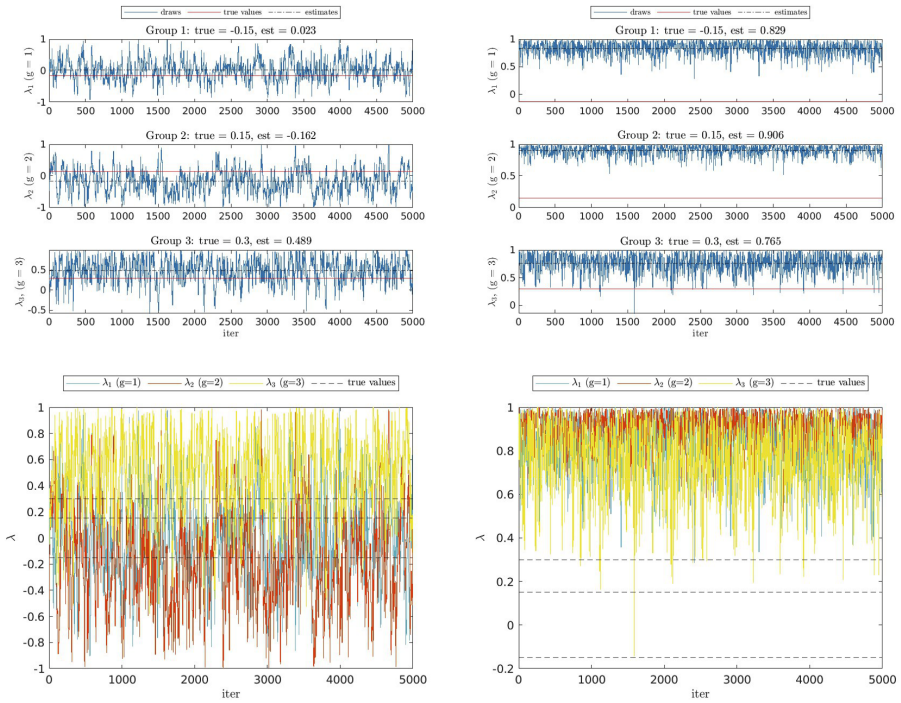

Figure C.32: Draws for λ_g (SCHSAR-left and HSAR-right).

C.2.2 Additional Monte Carlo Experiments

C.2.2.1 Results for small sample size

Table C.1: DGP V: N= 200, Unobserved Degree Heterogeneity

			SCHSAR					HSAR				
SNR	Parameter	True Value	Mean	Std	Bias	RMSE	Coverage	Mean	Std	Bias	RMSE	Coverage
	λ_1	-0.15	-0.146	0.074	0.004	0.074	0.88	-0.710	0.205	-0.560	0.596	0.14
	λ_2	0.15	0.136	0.067	-0.014	0.068	0.92	-0.446	0.265	-0.596	0.652	0.20
	λ_3	0.30	0.301	0.034	0.001	0.034	0.97	0.022	0.206	-0.278	0.346	0.42
	ω_1	0.45	0.454	0.033	0.004	0.033	0.99	0.452	0.037	0.002	0.037	1.00
	ω_2	0.35	0.345	0.032	-0.005	0.032	0.98	0.345	0.037	-0.005	0.037	0.99
	ω_3	0.20	0.200	0.026	0.000	0.026	0.97	0.203	0.031	0.003	0.031	0.98
	β_{11}	-0.50	-0.460	0.195	0.040	0.199	0.89	-0.419	0.288	0.081	0.299	0.84
	β_{12}	0.50	0.443	0.243	-0.057	0.250	0.89	0.404	0.298	-0.096	0.313	0.88
High	β_{13}	-1.00	-0.986	0.150	0.014	0.151	0.95	-1.001	0.034	-0.001	0.034	0.97
	β_{21}	-0.75	-0.688	0.304	0.062	0.311	0.90	-0.617	0.443	0.133	0.463	0.88
	β_{22}	0.80	0.741	0.306	-0.059	0.312	0.91	0.654	0.443	-0.146	0.467	0.86
	β_{23}	1.20	1.195	0.042	-0.005	0.042	0.95	1.198	0.034	-0.002	0.034	0.93
	κ_1	0.80	0.777	0.040	-0.023	0.046	0.81	–	–	–	–	–
	κ_2	0.60	0.595	0.055	-0.005	0.055	0.86	–	–	–	–	–
	κ_3	0.25	0.252	0.035	0.002	0.035	0.97	–	–	–	–	–
	γ	1.50	1.514	0.042	0.014	0.044	0.86	–	–	–	–	–
	σ_a^2	2.00	2.038	0.239	0.038	0.242	0.91	–	–	–	–	–
	λ_1	-0.15	-0.153	0.229	-0.003	0.229	0.97	-0.587	0.236	-0.437	0.497	0.49
	λ_2	0.15	0.054	0.251	-0.096	0.269	0.94	-0.362	0.297	-0.512	0.592	0.62
	λ_3	0.30	0.212	0.272	-0.088	0.286	0.97	0.000	0.295	-0.300	0.420	0.93
	ω_1	0.45	0.452	0.040	0.002	0.040	0.99	0.453	0.043	0.003	0.043	0.99
	ω_2	0.35	0.341	0.039	-0.009	0.040	0.99	0.337	0.040	-0.013	0.042	1.00
	ω_3	0.20	0.207	0.037	0.007	0.037	0.99	0.210	0.039	0.010	0.040	0.99
	β_{11}	-0.50	-0.394	0.322	0.106	0.339	0.82	-0.330	0.396	0.170	0.431	0.78
	β_{12}	0.50	0.321	0.408	-0.179	0.446	0.82	0.207	0.477	-0.293	0.560	0.78
Low	β_{13}	-1.00	-0.934	0.306	0.066	0.313	0.93	-0.881	0.359	0.119	0.379	0.87
	β_{21}	-0.75	-0.568	0.486	0.182	0.519	0.83	-0.471	0.576	0.279	0.640	0.80
	β_{22}	0.80	0.643	0.481	-0.157	0.506	0.84	0.541	0.598	-0.259	0.651	0.75
	β_{23}	1.20	1.127	0.290	-0.073	0.299	0.93	1.118	0.274	-0.082	0.286	0.91
	κ_1	0.80	0.747	0.125	-0.053	0.136	0.92	–	–	–	–	–
	κ_2	0.60	0.581	0.145	-0.019	0.146	0.91	–	–	–	–	–
	κ_3	0.25	0.254	0.274	0.004	0.274	0.96	–	–	–	–	–
	γ	1.50	1.514	0.044	0.014	0.046	0.84	–	–	–	–	–
	σ_a^2	2.00	2.045	0.242	0.045	0.246	0.90	–	–	–	–	–

Notes: This table displays results based on $R = 100$ replicates. The values include the average and standard deviation of the point estimates; the average bias (Bias), the Root Mean Squared Error (RMSE), and the coverage rate (Coverage) across replicates; where Bias $= R^{-1}\sum_{r=1}^{R}(\hat{\alpha}_r - \alpha)$, RMSE $= \sqrt{R^{-1}\sum_{r=1}^{R}(\hat{\alpha}_r - \alpha)^2}$, and Coverage $= R^{-1}\sum_{r=1}^{R}\mathbb{1}\{\alpha \in \widehat{CI}_{0.95,r}\}$.

Table C.2: DGP VI: N= 200, Unobserved Degree Heterogeneity, Link Misspecification

			SCHSAR					HSAR				
SNR	Parameter	True Value	Mean	Std	Bias	RMSE	Coverage	Mean	Std	Bias	RMSE	Coverage
	λ_1	-0.15	-0.138	0.082	0.012	0.083	0.93	-0.673	0.217	-0.523	0.566	0.27
	λ_2	0.15	0.131	0.080	-0.019	0.082	0.90	-0.448	0.286	-0.598	0.663	0.34
	λ_3	0.30	0.301	0.044	0.001	0.044	0.98	0.002	0.257	-0.298	0.393	0.59
	ω_1	0.45	0.454	0.033	0.004	0.033	0.98	0.453	0.038	0.003	0.038	0.99
	ω_2	0.35	0.344	0.031	-0.006	0.032	0.99	0.344	0.037	-0.006	0.038	0.99
	ω_3	0.20	0.201	0.027	0.001	0.027	0.96	0.203	0.031	0.003	0.031	0.98
	β_{11}	-0.50	-0.451	0.217	0.049	0.223	0.91	-0.428	0.276	0.072	0.285	0.85
	β_{12}	0.50	0.449	0.217	-0.051	0.223	0.94	0.413	0.283	-0.087	0.296	0.89
High	β_{13}	-1.00	-1.001	0.008	-0.001	0.009	0.94	-1.002	0.036	-0.002	0.036	0.97
	β_{21}	-0.75	-0.672	0.338	0.078	0.347	0.88	-0.633	0.417	0.117	0.433	0.91
	β_{22}	0.80	0.721	0.337	-0.079	0.346	0.89	0.667	0.421	-0.133	0.442	0.86
	β_{23}	1.20	1.199	0.007	-0.001	0.007	0.96	1.199	0.036	-0.001	0.036	0.93
	κ_1	0.80	1.345	0.076	0.545	0.550	0.00	–	–	–	–	–
	κ_2	0.60	1.039	0.076	0.439	0.446	0.00	–	–	–	–	–
	κ_3	0.25	0.431	0.024	0.181	0.182	0.00	–	–	–	–	–
	γ	1.50	0.872	0.034	-0.628	0.629	0.00	–	–	–	–	–
	σ_a^2	2.00	0.675	0.078	-1.325	1.327	0.00	–	–	–	–	–
	λ_1	-0.15	-0.153	0.252	-0.003	0.252	0.97	-0.526	0.275	-0.376	0.466	0.68
	λ_2	0.15	0.065	0.257	-0.085	0.271	0.97	-0.329	0.303	-0.479	0.567	0.75
	λ_3	0.30	0.179	0.286	-0.121	0.310	0.97	-0.018	0.309	-0.318	0.444	0.93
	ω_1	0.45	0.453	0.040	0.003	0.041	1.00	0.452	0.044	0.002	0.044	1.00
	ω_2	0.35	0.341	0.039	-0.009	0.040	0.99	0.337	0.040	-0.013	0.042	1.00
	ω_3	0.20	0.206	0.037	0.006	0.038	0.97	0.211	0.043	0.011	0.044	0.99
	β_{11}	-0.50	-0.388	0.330	0.112	0.349	0.82	-0.310	0.419	0.190	0.460	0.74
	β_{12}	0.50	0.328	0.390	-0.172	0.426	0.83	0.193	0.474	-0.307	0.565	0.78
Low	β_{13}	-1.00	-0.942	0.284	0.058	0.290	0.92	-0.886	0.333	0.114	0.352	0.87
	β_{21}	-0.75	-0.573	0.484	0.177	0.515	0.84	-0.432	0.616	0.318	0.693	0.77
	β_{22}	0.80	0.630	0.495	-0.170	0.524	0.83	0.511	0.611	-0.289	0.676	0.74
	β_{23}	1.20	1.165	0.126	-0.035	0.131	0.94	1.103	0.303	-0.097	0.318	0.89
	κ_1	0.80	1.299	0.204	0.499	0.539	0.28	–	–	–	–	–
	κ_2	0.60	1.026	0.233	0.426	0.486	0.44	–	–	–	–	–
	κ_3	0.25	0.478	0.310	0.228	0.385	0.90	–	–	–	–	–
	γ	1.50	0.872	0.039	-0.628	0.629	0.00	–	–	–	–	–
	σ_a^2	2.00	0.674	0.081	-1.325	1.328	0.00	–	–	–	–	–

Notes: This table displays results based on $R = 100$ replicates. The values include the average and standard deviation of the point estimates; the average bias (Bias), the Root Mean Squared Error (RMSE), and the coverage rate (Coverage) across replicates; where Bias $= R^{-1}\sum_{r=1}^{R}(\hat{\alpha}_r - \alpha)$, RMSE $= \sqrt{R^{-1}\sum_{r=1}^{R}(\hat{\alpha}_r - \alpha)^2}$, and Coverage $= R^{-1}\sum_{r=1}^{R}\mathbb{1}\{\alpha \in \widehat{CI}_{0.95,r}\}$.

C.2.2.2 Results for exogenous network formation

Table C.3: DGP VII: N = 1000, Unobserved Degree Heterogeneity, Exogenous Network Formation

		SCHSAR					**HSAR**				
Parameter	True Value	Mean	Std	Bias	RMSE	Coverage	Mean	Std	Bias	RMSE	Coverage
λ_1	-0.15	-0.157	0.047	-0.007	0.048	0.93	-0.158	0.044	-0.008	0.045	0.94
λ_2	0.15	0.153	0.044	0.003	0.044	0.94	0.153	0.042	0.003	0.042	0.95
λ_3	0.30	0.308	0.046	0.008	0.047	0.99	0.310	0.046	0.010	0.047	0.98
ω_1	0.45	0.452	0.015	0.002	0.015	0.95	0.453	0.015	0.003	0.015	0.99
ω_2	0.35	0.348	0.015	-0.002	0.015	0.97	0.348	0.015	-0.002	0.015	0.98
ω_3	0.20	0.199	0.013	-0.001	0.013	0.93	0.199	0.013	-0.001	0.013	0.96
β_{11}	-0.50	-0.500	0.003	0.000	0.003	0.93	-0.500	0.003	0.000	0.003	0.92
β_{12}	0.50	0.499	0.011	-0.001	0.011	0.85	0.500	0.003	0.000	0.003	0.85
β_{13}	-1.00	-0.999	0.011	0.002	0.011	0.96	-1.000	0.002	0.000	0.002	0.96
β_{21}	-0.75	-0.750	0.003	0.000	0.003	0.94	-0.750	0.003	0.000	0.003	0.95
β_{22}	0.80	0.798	0.017	-0.002	0.017	0.93	0.800	0.002	0.000	0.002	0.94
β_{23}	1.20	1.200	0.004	0.000	0.004	0.96	1.200	0.003	0.000	0.003	0.96
κ_1	0.00	0.001	0.004	0.001	0.004	0.94	–	–	–	–	–
κ_2	0.00	0.000	0.003	0.000	0.003	0.96	–	–	–	–	–
κ_3	0.00	0.000	0.004	0.000	0.004	0.96	–	–	–	–	–
γ	1.50	1.501	0.010	0.001	0.010	0.79	–	–	–	–	–
σ_a^2	2.00	2.002	0.009	0.002	0.010	0.99	–	–	–	–	–

Notes: This table displays results based on $R = 100$ replicates. The values include the average and standard deviation of the point estimates; the average bias (Bias), the Root Mean Squared Error (RMSE), and the coverage rate (Coverage) across replicates. We consider a model misspecification scenario where the network structure is exogenous rather than endogenous. Estimating the SCHSAR model is able to identify zero correlation and still recover true values of the parameters, similar to a correctly-specified HSAR model.

C.2.2.3 Results for homogeneous network effects

Table C.4: DGP VIII: N = 1000, Unobserved Degree Heterogeneity, Homogeneous Network Effects

		SCHSAR					**HSAR**				
Parameter	True Value	Mean	Std	Bias	RMSE	Coverage	Mean	Std	Bias	RMSE	Coverage
λ_1	0.15	0.149	0.010	-0.001	0.010	0.92	-0.585	0.192	-0.735	0.759	0.00
λ_2	0.15	0.149	0.008	-0.001	0.008	0.96	-0.424	0.154	-0.574	0.594	0.00
λ_3	0.15	0.150	0.008	0.000	0.008	0.98	-0.090	0.070	-0.240	0.250	0.00
ω_1	0.45	0.452	0.014	0.002	0.015	0.99	0.451	0.020	0.001	0.020	0.98
ω_2	0.35	0.348	0.015	-0.002	0.015	0.97	0.350	0.021	0.000	0.021	0.99
ω_3	0.20	0.200	0.013	0.000	0.013	0.92	0.199	0.016	-0.001	0.016	0.92
β_{11}	-0.50	-0.500	0.003	0.000	0.003	0.92	-0.493	0.090	0.007	0.091	0.92
β_{12}	0.50	0.500	0.003	0.000	0.003	0.86	0.487	0.098	-0.013	0.098	0.90
β_{13}	-1.00	-1.000	0.002	0.000	0.002	0.99	-1.000	0.015	0.000	0.015	0.96
β_{21}	-0.75	-0.750	0.003	0.000	0.003	0.95	-0.732	0.158	0.018	0.159	0.91
β_{22}	0.80	0.800	0.003	0.000	0.003	0.93	0.786	0.160	-0.014	0.160	0.96
β_{23}	1.20	1.200	0.003	0.000	0.003	0.96	1.198	0.042	-0.002	0.042	0.97
κ_1	0.80	0.798	0.004	-0.002	0.005	0.90	–	–	–	–	–
κ_2	0.60	0.598	0.004	-0.002	0.004	0.97	–	–	–	–	–
κ_3	0.25	0.249	0.004	-0.001	0.004	0.96	–	–	–	–	–
γ	1.50	1.501	0.010	0.001	0.010	0.77	–	–	–	–	–
σ_a^2	2.00	2.000	0.010	0.000	0.010	0.99	–	–	–	–	–

Notes: This table displays results based on $R = 100$ replicates. The values include the average and standard deviation of the point estimates; the average bias (Bias), the Root Mean Squared Error (RMSE), and the coverage rate (Coverage) across replicates. We consider a model misspecification scenario where network interaction effect is homogeneous rather than heterogeneous. Estimating the SCHSAR model is able to recover true values of the parameters without generating spurious heterogeneous effects if these parterns are indeed non-existent in the real data.

C.3 On Empirical Application

C.3.1 Unobserved Heterogeneity

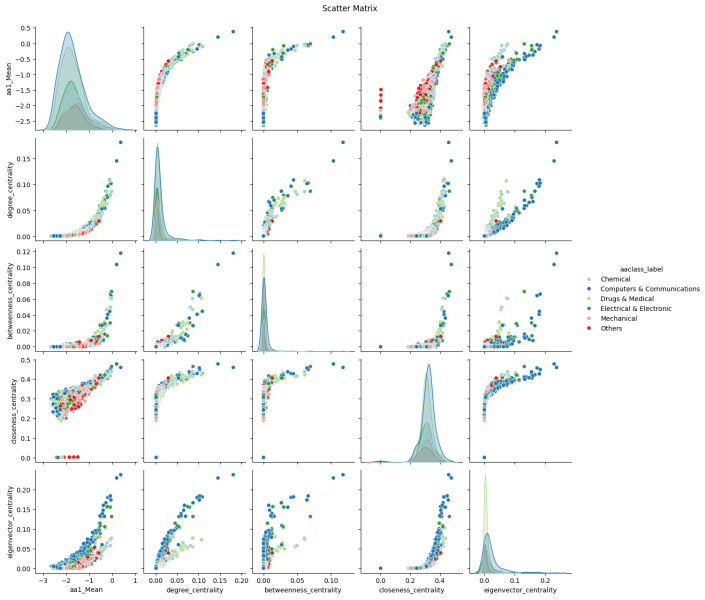

Figure C.33: Positive correlation between unobserved heterogeneity and centrality measures.

Table C.5: The Top 20 Firms with the Highest Unobserved Heterogeneity

Rank	Firm	SIC	Field	Size	Degree	Between	Close	Eigen	UDH
1	Ibm	7370	Computers & Communications	99	0.181	0.118	0.46	0.238	0.38
2	Hp	3570	Computers & Communications	97	0.145	0.104	0.477	0.229	0.203
3	Ge	3724	Electrical & Electronic	100	0.087	0.07	0.464	0.131	-0.016
4	Du Pont	2820	Chemical	97	0.086	0.061	0.438	0.077	-0.027
5	Siemens	9997	Computers & Communications	99	0.102	0.064	0.457	0.173	-0.045
6	Dupont De Nemours Inc	2820	Chemical	96	0.081	0.05	0.416	0.057	-0.066
7	Motorola	3663	Computers & Communications	93	0.103	0.066	0.449	0.184	-0.07
8	Toshiba	3600	Electrical & Electronic	97	0.079	0.03	0.426	0.156	-0.078
9	Pfizer	2834	Drugs & Medical	98	0.107	0.061	0.422	0.072	-0.079
10	Intel	3674	Computers & Communications	96	0.109	0.045	0.427	0.181	-0.096
11	Microsoft	7372	Computers & Communications	98	0.097	0.041	0.429	0.18	-0.179
12	Ti	3674	Electrical & Electronic	89	0.075	0.014	0.405	0.157	-0.202
13	Bayer	2800	Chemical	98	0.064	0.028	0.422	0.065	-0.212
14	Hitachi	9997	Computers & Communications	99	0.078	0.036	0.437	0.155	-0.218
15	J&J	2834	Drugs & Medical	97	0.086	0.047	0.413	0.055	-0.226
16	Basf	2800	Chemical	98	0.062	0.026	0.422	0.056	-0.262
17	Oracle	7372	Computers & Communications	94	0.087	0.027	0.422	0.166	-0.263
18	Eastman	3861	Chemical	91	0.057	0.029	0.432	0.103	-0.28
19	Danaher	3826	Electrical & Electronic	87	0.054	0.029	0.421	0.106	-0.315
20	Alcatel	3661	Computers & Communications	97	0.07	0.016	0.403	0.132	-0.327

Notes: A node's centrality measures include degree centrality, betweenness centrality, closeness centrality, and eigenvector centrality; Size is the percentile ranking of a firm's total asset; SIC is the primary four-digit SIC code according to Compustat U.S. fundamentals database.

C.3.2 Choice of the Number of Latent Types (G)

Table C.6: Parameter Estimates for the HSAR and SCHSAR Models (when G=3)

		HSAR			SCHSAR		
		1st Type	2nd Type	3rd Type	1st Type	2nd Type	3rd Type
Network Interaction							
Assignment Probability	π	0.416 (0.272) [0.00, 0.68]	0.412 (0.043) [0.33, 0.50]	0.185 (0.274) [0.00, 0.66]	0.541 (0.100) [0.36, 0.72]	0.321 (0.046) [0.23, 0.41]	0.138 (0.098) [0.00, 0.30]
Interaction Effect	λ	0.174 (0.371) [-0.58, 0.97]	0.226 (0.035) [0.16, 0.29]	-0.012 (0.461) [-0.77, 0.92]	0.146 (0.053) [0.07, 0.27]	0.189 (0.055) [0.06, 0.28]	0.217 (0.220) [-0.17, 0.74]
Intercept	β_1	-0.022 (4.452) [-11.07, 5.35]	-3.171 (0.464) [-4.08, -2.23]	-2.026 (8.383) [-18.89, 15.08]	2.735 (1.161) [-0.35, 4.09]	-1.462 (2.395) [-4.08, 3.29]	1.740 (5.928) [-13.41, 14.74]
logTaxPrice	β_2	-8.289 (6.356) [-12.88, 7.83]	-2.428 (0.853) [-4.08, -0.77]	-3.012 (9.137) [-18.84, 14.52]	-9.322 (1.855) [-12.30, -4.62]	-3.383 (3.066) [-12.25, 1.49]	-2.268 (6.727) [-14.18, 12.51]
logCapitalExpense	β_3	-0.727 (4.625) [-13.56, 10.60]	0.292 (0.069) [0.16, 0.43]	1.340 (7.158) [-13.78, 17.34]	0.540 (0.134) [0.19, 0.73]	0.354 (0.231) [0.06, 0.86]	0.186 (4.410) [-11.53, 10.16]
EBIT	β_4	-0.792 (3.744) [-9.53, 7.47]	0.052 (0.027) [-0.00, 0.11]	2.006 (8.052) [-16.05, 17.77]	0.037 (0.044) [-0.04, 0.13]	0.102 (0.151) [-0.04, 0.56]	0.670 (4.645) [-9.87, 14.41]
logEmployment	β_5	-2.932 (8.046) [-25.44, 3.95]	-0.617 (0.084) [-0.79, -0.45]	-0.525 (7.022) [-16.96, 16.20]	-0.112 (0.145) [-0.46, 0.13]	-0.514 (0.234) [-0.82, 0.10]	-0.460 (4.594) [-12.37, 11.30]
logRevenue	β_6	-1.459 (6.745) [-13.64, 15.18]	0.109 (0.094) [-0.08, 0.29]	0.059 (8.619) [-13.60, 19.21]	-0.966 (0.283) [-1.24, -0.22]	-0.178 (0.529) [-1.25, 0.34]	-0.457 (3.974) [-8.10, 12.77]
Correlation	κ	–	–	–	0.977 (0.206) [0.53, 1.36]	0.261 (0.259) [-0.15, 0.85]	0.977 (0.206) [0.53, 1.36]
Unobserved Heterogeneity	σ_a^2	–	–	–		3.102 (0.132) [2.85, 3.37]	
Network Formation							
SIC homophily	γ_1	–	–	–		0.725 (0.017) [0.69, 0.76]	
Tech homophily	γ_2	–	–	–		0.692 (0.010) [0.67, 0.71]	
Criteria							
Log likelihood			-1958.68 (20.94)			-44949.63 (9729.57)	
AICM			4794.66 (20.68)			189418920.84 (4440159.24)	
Observations			1150			1150	

Notes: This table presents the estimation results for the HSAR and SCHSAR models with $G = 3$. MCMC sampling runs a total of $50,491$ iterations, where the first 500 iterations discarded as burn-in and every 10th draw is retained, yielding $5,000$ effective draws. Posterior means, standard deviations, and 95% equal-tailed intervals (ETI) are computed using these MCMC draws.

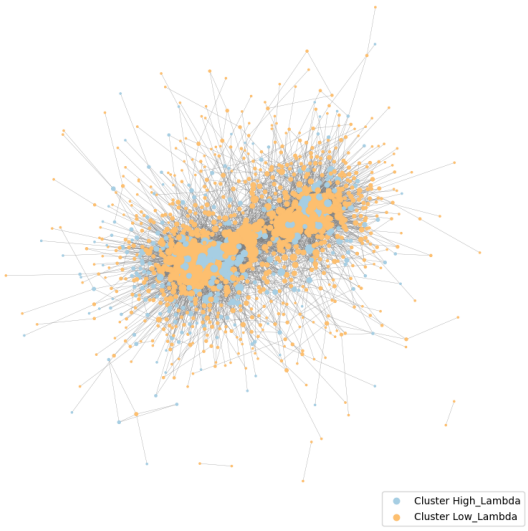

Figure C.34: Network of interactions by latent types defined from SCHSAR (G=2) estimation results.

C.3.3 Direct, Indirect, and Total Effects

Table C.7: Regressions of Effects of Interest on Firm Characteristics

	Direct Effects (1)	Indirect Spillin (2)	Indirect Spillout (3)	Total Spillin (4)	Total Spillout (5)
logTotalAsset	-0.092^{***}	0.010^{**}	0.017^{**}	-0.082^{***}	-0.075^{**}
	(0.025)	(0.004)	(0.007)	(0.024)	(0.030)
Computers & Communications	-1.511^{***}	-0.027	-0.105^{**}	-1.539^{***}	-1.616^{***}
	(0.158)	(0.025)	(0.044)	(0.148)	(0.185)
Drugs & Medical	-0.438^{***}	0.041	-0.021	-0.397^{**}	-0.458^{**}
	(0.167)	(0.026)	(0.047)	(0.156)	(0.195)
Electrical & Electronic	-1.388^{***}	-0.023	-0.089^{*}	-1.411^{***}	-1.477^{***}
	(0.183)	(0.029)	(0.052)	(0.171)	(0.215)
Mechanical	-1.026^{***}	-0.011	-0.041	-1.037^{***}	-1.067^{***}
	(0.214)	(0.034)	(0.060)	(0.200)	(0.251)
Others	-0.292	0.002	-0.003	-0.290	-0.295
	(0.222)	(0.035)	(0.063)	(0.208)	(0.260)
Degree Centrality	-6.974^{*}	1.084^{*}	25.125^{***}	-5.890^{*}	18.151^{***}
	(3.722)	(0.591)	(1.048)	(3.478)	(4.352)
Intercept	8.685^{***}	1.090^{***}	0.896^{***}	9.775^{***}	9.581^{***}
	(0.209)	(0.033)	(0.059)	(0.195)	(0.244)
Observations	1150	1150	1150	1150	1150
Adjusted R^2	0.137	0.021	0.447	0.157	0.105

Notes: Standard errors are reported in parentheses. Coefficients marked with *, **, and *** are significant at the 10%, 5%, and 1% levels, respectively.