WILLIAM & MARY
CHARTERED 1693

Shuo Qi
Postdoctoral Fellow in Digital Inclusion and Governance
Global Research Institute, William & Mary
400 Landrum Drive, Williamsburg, VA 23186
Phone: (857) 333-8272
Email: shuoqi@wm.edu
Website: sites.google.com/view/shuoqiecon

September 14, 2026

Workshop on Development Economics and Networks for Early-Career Researchers
National Bureau of Economic Research
1050 Massachusetts Avenue
Cambridge, MA 02138

Dear Professors Breza, Viviano, and Chandrasekhar,

I am applying to participate in the Workshop on Development Economics and Networks for Early-Career Researchers on December 5, 2026. I am a Postdoctoral Fellow in Digital Inclusion and Governance and a Faculty Affiliate in the Department of Economics at William & Mary. I received my Ph.D. in Economics from Southern Methodist University in 2025. I propose to present my working paper, "*Recovering Latent Heterogeneity in Network Formation Using Triads*."

The paper develops an identification and estimation strategy for dyadic network formation models with unobserved degree heterogeneity. Comparisons of complementary anchored-triad configurations identify realized agent effects rather than eliminating them as nuisance parameters. Conditional on those effects, dyadic link probabilities identify the systematic component of link formation. The triad-anchored likelihood estimator combines triad anchoring, stabilized dyadic refinement, common-parameter bias correction, and bias-reduced effect profiling. Applications to 43 Indian village networks and the Nyakatoke risk-sharing network show how accounting for latent link propensities changes the interpretation of observed matching patterns.

The workshop directly connects to both this paper and my broader research agenda in econometrics, development economics, and social networks. I would especially value feedback on the empirical interpretation of recovered household effects, the support and network-density conditions behind the estimator, and extensions to partially observed networks. These questions also arise in my ongoing field research on digital-payment adoption and network diffusion in Tanzanian markets. I would contribute experience with network formation, endogenous social interactions, egocentric network data, and field experiments in development settings.

If I am shortlisted, Daniel L. Millimet, Dedman Trustee Professor in the Department of Economics at Southern Methodist University, can provide a recommendation. His email address is millimet@smu.edu. Professor Millimet chaired my dissertation committee and is familiar with the development of the proposed paper.

Thank you for considering my application.

Sincerely,

Shuo Qi

Shuo Qi

Shuo Qi

Last Updated: September 14, 2026

CONTACT INFORMATION

Global Research Institute, William & Mary
400 Landrum Drive
Williamsburg, VA 23186

Phone: (857) 333-8272
Email: shuoqi@wm.edu
Website: sites.google.com/view/shuoqiecon

ACADEMIC APPOINTMENT

Postdoctoral Fellow, Global Research Institute, William & Mary 2025–2027
Faculty Affiliate, Department of Economics, William & Mary 2025–2027

EDUCATION

Ph.D. in Economics, Southern Methodist University, Dallas, TX 2025
Melody Rice Memorial Award for Excellence in Research and Service
Committee: Daniel Millimet (Chair), Hao Dong, and Ömer Özak
M.A. Economics, Boston University, Boston, MA 2018
B.A. Economics & B.S. Psychology, South-Central University for Nationalities, China 2013

PH.D. COURSEWORK

All courses completed; no current course enrollments.

Econometrics and quantitative methods: Intro to Quantitative Economics (ECO 6371); Econometrics I (ECO 6372); Econometrics II (ECO 6374); Econometrics III (ECO 6375); Microeconometrics (ECO 7377); Mathematical Economics (ECO 7305).

Economic theory: Microeconomic Theory I (ECO 6384); Microeconomic Theory II (ECO 6385); Macroeconomic Theory I (ECO 6394); Macroeconomic Theory II (ECO 6395); International Macroeconomic Theory and Policy (ECO 7332); Monetary Theory and Policy (ECO 7362).

Development, labor, and applied fields: Development Economics (ECO 7334); Development of Human Capital (ECO 7322); Labor Economics (ECO 7321); Experimental Economics (ECO 7381).

Research and doctoral seminars: Workshop for Preprospectus (ECO 7004); Research (ECO 8000 and ECO 8105, recurring registrations); Graduate Full-Time Status (ECO 8049, recurring registration).

RESEARCH FIELDS

Fields Econometrics, Development Economics, Labor Economics
Interests Social Networks, Incomplete Data, Model Misspecification, Digital Inclusion

WORKING PAPERS

1. "Recovering Latent Heterogeneity in Network Formation Using Triads"
2. "Identification and Estimation of Social Interactions in Binary Choice Games with Egocentric Network Data"
3. "Inference for Social Interactions in Large Endogenous Networks" with Wan Zhang
   • ***Best Paper Prize in Econometrics, 2026 Asia Meeting of the Econometric Society***
4. "Risk-Optimal Shrinkage for Orthogonal Score Estimators under Local Misspecification" with Daniel Millimet
5. "Why Does Cash Persist? Evidence from Tanzanian Markets" with Philip Roessler, Russell Toth, and Tiffany Tsai
6. "Bridging the AI Divide: An RCT on Voice-Based Generative AI, Social Learning, and Women's Empowerment" with Philip Roessler

WORK IN PROGRESS

1. "Name Your Friends, but How Many? Survey Design under Heterogeneous Network Interference"
2. "Network Diffusion and Spillovers in the Adoption of Digital Payments: Evidence from Tanzanian Markets" with Philip Roessler, Russell Toth, and Tiffany Tsai

BOOK CHAPTERS

"Evidence-Based Policy Making," with Daniel Millimet and Xuchao Gao. Forthcoming in *Handbook of Labor, Human Resources and Population Economics* (Klaus F. Zimmermann, ed.), Springer Nature.

SEMINARS AND CONFERENCE PRESENTATIONS

* scheduled

2026
Southern Economic Association, 96th Annual Meeting (Houston)*; Midwest Econometrics Group, 36th Annual Conference (University of Cincinnati)*; Asia Meeting of the Econometric Society (HKU Business School; coauthor presentation; *Best Paper Prize in Econometrics*); IPA–GPRL Annual Researcher Gathering (Northwestern University); Network Science in Economics Conference (University of Miami, Poster).

2025
World Congress of the Econometric Society (Seoul); Chinese Economists Society (CES) North America Annual Conference, Rising Star Session (University of Michigan); TEXPOP (UT Austin); Texas Camp Econometrics XXVIII (UT Austin); Seminar, City University of New York–Baruch College; American Economic Association Annual Meeting (San Francisco, Lightning Session).

2024
Southern Economic Association, 94th Annual Meeting (Washington, D.C.); Midwest Econometrics Group, 34th Annual Conference (University of Kentucky); North East Universities Development Consortium (Northeastern University); North American Summer Meeting of the Econometric Society (Vanderbilt University); Asian Meeting of the Econometric Society (Zhejiang University); American Society of Health Economists, 13th Annual Conference (San Diego); Stata Texas Empirical Microeconomics Conference (Dallas Fed, Poster); Global GLO-JOPE Conference, Job Market Session (Virtual); TEXPOP (UT Austin); Texas Camp Econometrics XXVII (Montgomery); Seminar, University of Texas at Arlington; American Economic Association Annual Meeting (San Antonio, Poster).

2023
Southern Economic Association, 93rd Annual Meeting (New Orleans); Midwest Econometrics Group, 33rd Annual Conference (Cleveland Fed); American Society of Health Economists, 12th Annual Conference (St. Louis, Poster); Midwest Economics Association, 87th Annual Meeting (Cleveland); CIREQ Interdisciplinary PhD Student Conference on Big Data and AI (McGill University); Causal Data Science Meeting (Virtual); Stata Texas Empirical Microeconomics Conference (Southlake, Poster); Texas Camp Econometrics XXVI (Montgomery); Seminar, Western Illinois University; SMU Economics PhD Alumni Conference (Virtual).

2022
Southern Economic Association, 92nd Annual Meeting (Fort Lauderdale).

TEACHING EXPERIENCE

Instructor

Econometrics (Undergraduate), William & Mary — Spring 2026; Spring 2027 (scheduled)
Intermediate Microeconomics (Undergraduate), SMU — Fall 2022

Guest Lecturer

Text Analysis with Python, DIGLab Research Group, William & Mary — Fall 2026
Topics in Social Network Analysis, DIGLab Research Group, William & Mary — Fall 2025
Introduction to Stata, Econometrics I (Ph.D.), SMU — Fall 2020

Teaching Assistant — 2018–2021

Ph.D. – Econometrics I (Fall 2020); Econometrics II (Spring 2021).
Master – Game Theory (Fall 2018, Fall 2019); Law and Economics (Spring 2020, Spring 2019); Economic Development (Fall 2018); Economic and Business Forecasting (Spring 2021).
Undergraduate – Principles of Macroeconomics (Spring 2018); Intermediate Microeconomics (Spring 2019, Spring 2020); Money and Banking (Fall 2019); Financial Economics (Spring 2018).

MENTORSHIP

Research Assistants

Serena Khan (William & Mary, B.A. Economics, Joint Program with St. Andrews, Class of 2028)
Emma McCarthy (William & Mary, B.A. Economics & International Relations, Class of 2028)
Yen Nhi Pham (William & Mary, B.A. Economics, minor in Computer Science, Class of 2028)
Ridhima Gandhi (William & Mary, B.A. Economics, minor in Mathematics, Class of 2027)
Katriel Amoyaw (William & Mary, B.A. Public Policy & Environmental Science, Class of 2027)
Kara Clapper (William & Mary, B.A. Government & English, Class of 2027)
Ryan Kas (William & Mary, B.A. Economics, minor in History, Class of 2026)

HONORS AND AWARDS

Awards

Best Paper Prize in Econometrics, 2026 Asia Meeting of the Econometric Society — 2026
For *"Inference for Social Interactions in Large Endogenous Networks"* (with Wan Zhang)
Award of Optimization-Conscious Econometrics Summer School, University of Chicago — 2024
Melody Rice Memorial Award for Excellence in Research and Service, SMU — 2024
Graduate Student Award, Southern Economic Association — 2022
Dean's Award for Best Presentation, Research & Innovation Week Graduate Poster Session, SMU — 2022
Cobb Fellowship for Best Third-Year Paper, Department of Economics, SMU — 2021–2022

Fellowships

Digital Inclusion & Governance Fellowship, Global Research Institute, William & Mary — 2025–2027
Research Support Funding, Global Research Institute, William & Mary — 2025–2027
Dean's Dissertation Fellowship, SMU — 2024–2025
Graduate Fellowship in Computational Science and Engineering, SMU — 2022–2024
Doctoral Teaching Fellowship, SMU — 2022
Full Tuition Scholarship and Teaching Assistantship, SMU — 2018–2021

Travel Grants

CES North America Conference 2025 Travel Grant, Chinese Economists Society — 2025
Dedman Graduate Student Assembly Grant (×2), SMU — 2023–2025
Moody Graduate Student Travel Grant (×3), SMU — 2022–2025
TEXPOP Travel Grant (×2), Population Research Center, University of Texas at Austin — 2024–2025
CIREQ Travel Grant, Centre Interuniversitaire de Recherche en Économie Quantitative — 2023

CONFERENCE DISCUSSIONS

Jiusi Xiao, *"The Effect of Own Gender Prosecutor on Charging and Conviction: Evidence from Random Assignment"*, GLO Job Market Session, Virtual Dec. 2024
Torsten Figueiredo Walter and Niclas Moneke, *"Endogenous Sample Selection"*, North East Universities Development Consortium, Northeastern University, Boston, MA Nov. 2024
Jaime Ramirez-Cuellar, *"Testing for Idiosyncratic Treatment Effect Heterogeneity"*, Midwest Economics Association 87th Annual Meetings, Cleveland, OH Apr. 2023

SKILLS

Python, R, Stata; English (fluent), Mandarin Chinese (native), French (basic), Swahili (beginner)

REFERENCES

Daniel L. Millimet (Ph.D. Advisor)
Dedman Trustee Professor
Department of Economics
Southern Methodist University
Phone: +1 (214)768-3269
Email: millimet@smu.edu

Hao Dong
Associate Professor
Department of Economics
Southern Methodist University
Phone: +1 (214)768-3288
Email: haod@smu.edu

Ömer Özak
Associate Professor
Department of Economics
Southern Methodist University
Phone: +1 (214)768-2755
Email: ozak@smu.edu

Tim Salmon
Professor
Department of Economics
Southern Methodist University
Phone: +1 (214)768-3547
Email: tsalmon@smu.edu

Philip Roessler (Postdoctoral Supervisor)
Professor of Government
Director, Digital Inclusion & Governance Lab (DIGLab)
College of William & Mary
Phone: +1 (757)221-3045
Email: proessler@wm.edu

Motivation and Current Research Interests

Shuo Qi

Postdoctoral Fellow, Global Research Institute, William & Mary | shuoqi@wm.edu

I would like to attend the Workshop on Development Economics and Networks because my current work combines econometric methods for network data with empirical research on risk sharing, technology adoption, and market coordination in low-income settings. The workshop would give me a focused setting in which to improve a work-in-progress paper while learning how other researchers connect network measurement and identification to substantive development questions.

I propose to present "*Recovering Latent Heterogeneity in Network Formation Using Triads*." The paper studies a dyadic logit model in which observed homophily and unobserved degree heterogeneity jointly determine links. Standard conditional-likelihood methods identify homophily by removing agent effects. My approach instead uses the log-odds ratio of two complementary anchored-triad configurations to isolate each focal agent's latent link propensity. Once those effects are recovered, dyadic variation identifies the systematic component of link formation. The triad-anchored likelihood estimator combines triad anchoring, stabilized dyadic refinement, common-parameter bias correction, and bias-reduced effect profiling. Large-network asymptotics provide inference when dyads and triads overlap and the nuisance-parameter dimension grows with the number of nodes.

The empirical applications show why the latent effects are economically useful. Across 43 Indian village networks, incorporating household effects changes the sign of several estimated matching coefficients; the recovered effects also show that larger households, households with more housing resources, and households containing a designated leader tend to have higher latent link propensities. In the Nyakatoke risk-sharing network, households with the same observed degree can have substantially different covariate-adjusted link propensities, while kinship, religion, and physical proximity remain important predictors of support links. I would particularly welcome feedback on the support conditions used for anchored-triad comparisons, the dense-network approximation, the interpretation of recovered effects in development networks, and extensions to partial or directed network data.

My broader research asks how to identify social interactions when networks are endogenous, incomplete, or expensive to measure. One project uses expected degree as a control for selection into peer groups in a large endogenous network. Another identifies strategic spillovers from egocentric network samples and studies how survey breadth and depth should be allocated under heterogeneous interference. I also study how approximate economic restrictions can be used through risk-optimal shrinkage rather than imposed exactly.

At William & Mary, I apply these ideas in field research on digital inclusion. One program studies cash persistence, digital-payment adoption, and network diffusion across Tanzanian markets. A second examines whether voice-based generative AI and social learning can reduce barriers to technology use among women. Across these projects, my interest is the same: network formation, measurement, and behavior should be analyzed together when the policy effect depends on who interacts with whom. I hope to contribute this econometric perspective to the workshop and to leave with a sharper empirical strategy for connecting recovered network heterogeneity to development mechanisms.

Recovering Latent Heterogeneity in Network Formation Using Triads*

Shuo Qi†

September 14, 2026

Abstract

This paper studies identification and inference in an undirected dyadic logit model with unobserved degree heterogeneity. A comparison of two complementary anchored-triad configurations eliminates the effects of the comparison agents while retaining the focal agent's effect. Under a diagonal normalization and continuity conditions, this relation identifies each fixed realized agent effect in a single growing network, and dyadic link probabilities then identify the unknown surplus function. I develop the triad-anchored likelihood estimator (TALE), which combines triad anchoring, stabilized dyadic refinement, common-parameter bias correction, and bias-reduced effect profiling. Dense-network asymptotics provide inference for the surplus parameter, individual effects, and parameters governing their conditional mean given observed attributes. Monte Carlo experiments examine precision, coverage, and computational performance. Applications to 43 Indian village networks and the Nyakatoke risk-sharing network illustrate two uses of recovered heterogeneity. Accounting for household effects changes the sign of several estimated matching coefficients in India. In Nyakatoke, households with the same observed degree can have substantially different covariate-adjusted link propensities, while kinship, religion, and physical distance remain important predictors of links.

JEL: C14, C31, D85

Keywords: Dyadic link formation; homophily effects; individual fixed effects; triad configuration; dense networks.

---

*This paper originates from the first chapter of my doctoral dissertation, *Identification and Estimation of Discrete Choice Models with Spillovers Using Partial Network Data*. While the original chapter primarily studied discrete games with partial network data governed by a dyadic link formation model, this paper specifically advances the econometric treatment of the underlying formation process itself. I am deeply indebted to Daniel Millimet for his invaluable guidance. I also thank Hao Dong and Ömer Özak for helpful discussions and support. I thank Eric Auerbach, Xiaohong Chen, Xun Tang, Haiqing Xu, and Wan Zhang for insightful comments. All the errors are mine.

†Global Research Institute and Department of Economics, College of William & Mary. 400 Landrum Drive, Williamsburg, VA 23186, United States. E-mail: shuoqi@wm.edu.

1 Introduction

Empirical models of network formation are increasingly central to understanding how social and economic relationships emerge, persist, and govern the flow of information, credit, and trade. The theoretical and empirical literature has long recognized that the architecture of a network critically dictates economic outcomes, ranging from the diffusion of microfinance (Banerjee et al., 2013) and technology adoption (Oster and Thornton, 2012; Bryan et al., 2014; Ferrali et al., 2020; Islam et al., 2023), to informal risk-sharing and insurance (De Weerdt and Dercon, 2006; De Weerdt and Fafchamps, 2011; Cai et al., 2015), and the manifestation of peer effects in education and labor markets (Manski, 1993; Hoxby, 2000; Calvó-Armengol et al., 2009; Lundborg, 2006; Moretti, 2011; Billings and Hoekstra, 2023). In modeling the formation of these networks, researchers consistently seek to disentangle two fundamental, often confounding forces: assortative matching on observable characteristics (homophily) and persistent, unobserved heterogeneity across agents in their baseline propensity to form links (degree heterogeneity). In standard dyadic link formation models, these two forces enter the latent bilateral surplus simultaneously. Consequently, observed link frequencies alone cannot readily distinguish whether two highly connected individuals are linked because they share similar observable traits, or simply because they both possess high latent sociability. Failing to formally account for this unobserved degree heterogeneity naturally biases estimates of homophily, typically leading researchers to severely overstate the importance of observable similarity in the formation of social ties.

This paper develops identification and inference for both components in an undirected dyadic logit model. For distinct agents i and j, a link forms according to

$$D_{ij} = \mathbf{1}\big(w(X_i, X_j) + A_i + A_j \geq U_{ij}\big),$$

where X_i denotes observed attributes, A_i is a latent agent effect, and w is an unknown symmetric dyadic surplus function. Homophily is an important special case of w, rather than a restriction on its general form. Agent types are independently and identically distributed across agents, while A_i may depend arbitrarily on X_i. Dyadic shocks are independent standard logistic variables, independent of agent types. The researcher observes a single network and agent attributes, but not the realized effects. I show how anchored-triad comparisons identify those effects and, subsequently, the

surplus function without imposing a parametric form on it. I then develop a likelihood estimator under a parametric surplus specification.

The identifying comparison uses three distinct agents, with one designated as the anchor. Fix a focal agent i and comparison agents j and k. In one configuration, i is linked to both comparison agents, who are not linked to each other. In the complementary configuration, the comparison agents are linked to each other, while neither is linked to i. Let $S_{i,jk}$ equal one and minus one for these configurations, respectively, and zero otherwise. For these informative anchored triads, the logistic specification implies

$$\begin{aligned}&\Pr\big(S_{i,jk} = 1 \mid |S_{i,jk}| = 1, X_i, A_i, X_j, X_k\big)\\ &\qquad = \Lambda\big(w(X_i, X_j) + w(X_i, X_k) - w(X_j, X_k) + 2A_i\big).\end{aligned}$$

The effects of the two comparison agents cancel, but the focal agent's effect remains. Thus, conditioning need not remove all the heterogeneity: it can isolate a particular agent's contribution. Triads serve here as econometric comparisons within a dyadic formation model, not as primitives representing preferences over three-agent structures.[1]

The remaining step separates the focal effect from the surplus contrast. In the large-network experiment, comparison pairs vary while the focal agent is held fixed, revealing its conditional configuration probabilities. Under the normalization $w(x, x) = 0$ and continuity at diagonal points, the surplus contrast vanishes as the comparison agents' attributes approach those of the focal agent. The limiting conditional probability is therefore $\Lambda(2A_i)$, which identifies A_i. Discrete attributes permit exact matches, whereas continuously distributed attributes require a diagonal-limit argument. These comparisons involve distinct agents and do not require self-links. The result identifies every fixed realized agent effect in a single growing network. Given these effects, inversion of dyadic link probabilities identifies w almost everywhere, and global continuity extends identification to its full covariate support. The diagonal normalization fixes the decomposition between dyadic surplus and agent heterogeneity, so the recovered effects are interpreted relative to that decomposition.

Identification does not by itself resolve the estimation problem. With n agents,

[1]The anchored-triad relation builds on Qi (2023), which studies discrete-choice spillovers with partial network data. The present paper develops identification and inference for the formation model without requiring a preliminary tetrad-logit step.

the effect vector contains n parameters, and errors in estimating these parameters enter the common-parameter score. Under dense-network asymptotics, the resulting incidental-parameter bias is of the same order as the sampling uncertainty in the common parameter. There is also a distinction between the information used for identification and that used for estimation: the anchored-triad criterion retains selected configurations, whereas the dyadic likelihood uses every observed link and nonlink. Moreover, anchored-triad contributions overlap through shared link indicators even though dyads are independent conditional on agent types. The estimation procedure must therefore combine the anchoring relation with dyadic refinement, bias correction, and appropriate variance estimation.

I propose the triad-anchored likelihood estimator, abbreviated TALE, as a single procedure with four stages. Triad anchoring obtains preliminary estimates of the common parameter and agent effects from informative anchored triads. Stabilized dyadic refinement then uses all dyadic link indicators, initially penalizing departures from the available triad estimates and gradually relaxing the penalty. Common-parameter bias correction removes the leading incidental-parameter bias using a weighted additive projection of the dyadic regressors and a second-order likelihood expansion. Finally, bias-reduced effect profiling holds the corrected common parameter fixed and re-estimates the agent effects using the Jeffreys–Firth adjustment (Firth, 1993). Triad anchoring supplies the starting point, dyadic refinement incorporates the remaining link information, and the final two stages address bias in the two parameter blocks.

The large-sample analysis establishes convergence rates and feasible asymptotic inference for this sequence. Under dense-network regularity conditions, the common dyadic-surplus parameter is estimated at rate n^{-1}, while each fixed agent effect is estimated at rate $n^{-1/2}$. These different rates reflect the order n^2 dyads available for common-parameter estimation and the order n dyads involving any particular agent. Sufficiently small stabilization and numerical errors make the dyadic refinement asymptotically equivalent to joint maximum likelihood. The common-parameter correction removes its leading bias, and the final effect adjustment preserves the fixed-agent first-order distribution. The resulting variance estimators account for estimation of the agent effects rather than treating them as observed.

Recovering the effects also permits analysis of how latent link propensity varies with observed attributes. Subject to integrability, the identification result determines

$\phi(x) = E[A_i \mid X_i = x]$ almost everywhere. For estimation, I specify $\phi(X_i) = h(X_i)'\gamma_0$ and project the final TALE effect estimates on the known regressors $h(X_i)$. Although the effects are estimated from the same network, their covariate-weighted average estimation error is negligible at the $n^{-1/2}$ scale relevant for γ_0. I establish asymptotic normality and a feasible covariance estimator for this conditional-mean parameter. The coefficients describe the relationship between observed attributes and normalized latent link propensities; they do not identify causal effects of those attributes.

Monte Carlo experiments evaluate estimation of the dyadic-surplus parameter, realized agent effects, and conditional-mean coefficients. The designs vary network size, link density, and the dependence between observed attributes and latent effects. The benchmarks are tetrad logit, joint maximum likelihood, and analytically bias-corrected joint maximum likelihood. Intermediate TALE stages are reported to distinguish their contributions. The analysis considers both precision and confidence-interval coverage, together with whether the numerical routines return estimates. The design parameters remain fixed as network size grows, so the lower-density cases are finite-sample stress tests within a dense-network framework, not evidence for a separate sparse-network asymptotic result.

Two applications illustrate the distinction between matching patterns and latent link propensities. The first uses 43 Indian village networks collected by Banerjee et al. (2013), containing 9,126 households in the retained sample. A pooled specification shares dyadic-surplus coefficients across villages while allowing each household to have its own effect. Several estimated matching patterns change after household heterogeneity is incorporated. For example, the coefficient on same leader status changes from -0.289 under pooled logit to 0.159 under TALE, and the coefficient on the rooms gap changes from positive to negative. The recovered effects provide a complementary household-level result: larger households, households with more housing resources, and households containing a designated leader tend to have higher latent link propensities. These associations remain similar after village fixed effects are included in the projection. The application therefore distinguishes the relationship between a household's characteristics and its overall link propensity from the relationship between similarity across households and their propensity to link to one another.

The second application studies reported risk-sharing relationships among 116 households in Nyakatoke, Tanzania (De Weerdt, 2004). Kinship, common religion,

and physical proximity remain important predictors of links after allowing for household heterogeneity. The recovered effects are strongly related to observed degree, but the relationship is not one-to-one. Among the 13 households with five observed links, the effects span approximately 1.71 logit-index units. This difference reflects adjustment for observed dyadic characteristics and the link propensities of potential partners: the same number of connections need not imply the same conditional propensity to form a link. Religion and wealth explain little of the recovered heterogeneity, even though common religion helps predict which households are linked. Together, the applications show why the dyadic and household components should be studied jointly, rather than interpreting either observed matching patterns or degree in isolation.

Related literature.

This paper contributes to the econometrics of network formation with unobserved degree heterogeneity; see de Paula (2020) and Graham (2020) for surveys. Graham (2017) develops both tetrad logit, which conditions out agent effects to estimate common parameters, and joint maximum likelihood, which estimates both parameter blocks and consistently recovers the agent effects under dense-network conditions. Jochmans (2018) develops a related conditional approach for directed networks. The present paper uses a different conditional comparison: anchored triads eliminate the effects of the comparison agents while retaining the focal agent's effect. This relation provides a route to identifying realized heterogeneity and forms the basis of TALE. Its subsequent dyadic refinement uses the full dyadic likelihood, linking the procedure to the joint-likelihood benchmark rather than replacing likelihood-based estimation.

The identification analysis is closely related to Gao (2020), who establishes identification of the surplus function, realized agent effects, and an unknown shock distribution under his maintained conditions, and to Zeleneev (2026), who allows unobserved heterogeneity to enter network interactions nonparametrically. The present paper maintains additive agent effects and logistic shocks. These restrictions yield an explicit anchored-triad relation that, together with diagonal continuity and normalization, identifies each fixed realized agent effect in a single growing network. Dyadic link probabilities then identify the surplus function. The identification argument does not impose a parametric form on this function or on the conditional distribution of agent

effects given observed attributes. Parametric specifications of the surplus function and the conditional mean of the effects are introduced subsequently for estimation.

The estimation analysis builds on likelihood methods for models with growing numbers of incidental parameters. Hahn and Newey (2004) develops analytical and jackknife bias corrections for nonlinear panel models, while Dzemski (2019) and Yan et al. (2019) study likelihood estimation and inference in directed network models with unobserved heterogeneity. Jochmans (2023) develops modified-likelihood methods for fixed-effect models with dyadic data. TALE combines triad anchoring and diminishing dyadic stabilization with an analytical correction for the common parameter and a final effect profile using the Jeffreys–Firth adjustment of Firth (1993). The weighted additive projection underlying the common-parameter correction accounts for the linear nuisance-score component, while a separate second-order correction removes the leading incidental-parameter bias. The analysis also develops inference for individual effects and for parameters governing their conditional mean.

Subnetwork comparisons also arise in models with strategic link interdependence. For example, Sheng (2020) uses subnetwork probabilities to construct identifying restrictions for network formation games with multiple equilibria. A different use of subgraphs appears in Chandrasekhar and Jackson (2025), whose subgraph-generated models form networks as unions of generated subgraphs. The role of triads in the present paper is distinct from both approaches. I maintain a dyadic formation rule with links independent conditional on agent types and use anchored triads as econometric conditioning devices. The resulting comparisons isolate latent heterogeneity without treating triads as primitives of the formation process or introducing strategic link externalities.

Finally, recovering agent heterogeneity is relevant to research on endogenous networks and outcome equations. Johnsson and Moon (2021) develops control-function methods for peer-effect models, including approaches based on estimated agent effects and on degree-based controls. Auerbach (2022) uses similarity in network linking patterns to control for latent heterogeneity in a partially linear outcome model. The present paper studies the formation model itself. It provides identification and inference for realized link propensities and their conditional mean given observed attributes, rather than establishing a correction for endogeneity in an outcome equation. The empirical projections therefore describe how normalized latent link propensities vary with observed characteristics.

Outline. The remainder of the paper is organized as follows. Section 2 introduces the dyadic link formation model, normalizations, and maintained assumptions. Section 3 establishes the nonparametric identification of the realized agent heterogeneity and the homophily function. Section 4 develops the four-step semiparametric estimation procedure and derives the asymptotic properties of the orthogonally corrected and profiled estimators. Section 5 presents Monte Carlo simulation evidence assessing finite-sample performance. Section 6 applies the framework to the Indian village and Nyakatoke applications. Section 7 concludes. Proofs of the main theorems and additional results are collected in the Appendix.

2 A Dyadic Link Formation Model

Consider an undirected network on n agents. For each agent i, $X_i \in \mathcal{X} \subset \mathbb{R}^{d_x}$ denotes a vector of observed attributes and $A_i \in \mathcal{A} \subset \mathbb{R}$ denotes a latent agent effect, where $\mathcal{X}$ and $\mathcal{A}$ are their marginal supports. For each unordered dyad (i, j) with $1 \leq i < j \leq n$, $D_{ij} = D_{ji} \in \{0, 1\}$ indicates whether agents i and j are linked, and $U_{ij} = U_{ji} \in \mathbb{R}$ denotes a dyad-specific shock. The researcher observes $X^n = (X_1', \ldots, X_n')'$ and $\{D_{ij} : 1 \leq i < j \leq n\}$, but does not observe $A^n = (A_1, \ldots, A_n)'$ or $\{U_{ij} : 1 \leq i < j \leq n\}$.

I model link formation through the latent-surplus representation

$$D_{ij} = \mathbb{1}\{w(X_i, X_j) + A_i + A_j - U_{ij} \geq 0\}, \qquad 1 \leq i < j \leq n, \tag{1}$$

where $w : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is an unknown dyadic surplus function. Thus, agents i and j form a link when the sum of the dyadic surplus component $w(X_i, X_j)$ and their latent link propensities $A_i + A_j$ exceeds the dyad-specific shock U_{ij}.

Because the network is undirected, I impose $w(x, \tilde{x}) = w(\tilde{x}, x)$ for every $x, \tilde{x} \in \mathcal{X}$. Homophily is an important special case of the dyadic surplus. For example, under the parametric specification $w(x_i, x_j) = \sum_{k=1}^{d_x} |x_{ik} - x_{jk}|\beta_{0k}$, a negative value of β_{0k} implies that similarity in characteristic k raises link surplus, whereas a positive value implies that dissimilarity raises link surplus. The agent effect A_i captures persistent agent-specific heterogeneity in the propensity to form links that is not represented by w. In the empirical applications, these effects are referred to as economy effects or household effects according to the unit of analysis.

For the large-network analysis, the network observed at sample size n is the subnetwork induced by the first n agents of an infinite population. The maintained data-generating process is as follows.

Assumption 1 (Data generating process) *The following conditions hold.*

(i) The pairs (X_i, A_i) are independent and identically distributed across agents $i \geq 1$, with joint support $\mathcal{S}_{XA} \subseteq \mathcal{X} \times \mathcal{A}$.

(ii) The shocks $\{U_{ij} : 1 \leq i < j < \infty\}$ are independent and identically distributed across unordered dyads, are independent of $\{(X_i, A_i) : i \geq 1\}$, and have the standard logistic distribution, with cumulative distribution function

$$\Lambda(u) = \frac{\exp(u)}{1 + \exp(u)}, \qquad u \in \mathbb{R}.$$

Assumption 1(i) imposes random sampling at the agent level while allowing arbitrary dependence between X_i and A_i. In particular, the joint support $\mathcal{S}_{XA}$ need not equal the product space $\mathcal{X} \times \mathcal{A}$. Although (X_i, A_i) is randomly sampled at the population level, the analysis below takes the realized sequence $(A_1, A_2, \ldots)$ as an object of identification and estimation.

Assumption 1(ii) specifies the dyad-specific shocks to be standard logistic, as in Graham (2017), thereby fixing the scale of latent surplus. Together with (1), it implies

$$\Pr(D_{ij} = 1 \mid X_i, A_i, X_j, A_j) = \Lambda\left(w(X_i, X_j) + A_i + A_j\right), \qquad i < j.$$

Conditional on agent types $\{(X_i, A_i) : i \geq 1\}$, link indicators are independent across unordered dyads. Unconditionally, they are independent for disjoint dyads but generally dependent for dyads sharing an agent due to the common agent type.

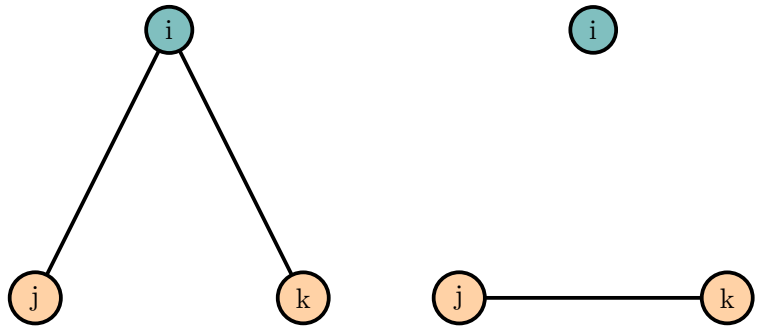
3 Identification

This section establishes identification of the realized agent effects and the dyadic surplus function under the large-network experiment introduced in Section 2. The argument proceeds in two steps. First, anchored triad comparisons identify every fixed realized agent effect from the limiting distribution. Once the realized effects are recovered, dyadic link probabilities identify the dyadic surplus function.

Fix three distinct agents i, j, k. The tuple (i, j, k) is an i-anchored triad, where j and k are referred to as the comparison agents. Define the anchored-triad configuration statistic

$$S_{i,jk} = D_{ij}D_{ik}(1 - D_{jk}) - (1 - D_{ij})(1 - D_{ik})D_{jk}.$$

Because links are undirected, $S_{i,kj} = S_{i,jk} \in \{-1, 0, 1\}$. The event $S_{i,jk} = 1$ corresponds to $D_{ij} = D_{ik} = 1$ and $D_{jk} = 0$, whereas $S_{i,jk} = -1$ corresponds to $D_{ij} = D_{ik} = 0$ and $D_{jk} = 1$. I refer to the cases $|S_{i,jk}| = 1$ as the informative anchored triad configurations. These configurations are depicted in Figure 1.



(a) (b)

Figure 1: Informative i-anchored triad configurations. Panel (a) corresponds to $S_{i,jk} = 1$, with $D_{ij} = D_{ik} = 1$ and $D_{jk} = 0$. Panel (b) corresponds to $S_{i,jk} = -1$, with $D_{ij} = D_{ik} = 0$ and $D_{jk} = 1$.

Let $X_{(i,j,k)} \equiv (X_i, X_j, X_k)$ and $\Delta_w(X_i, X_j, X_k) \equiv w(X_i, X_j) + w(X_i, X_k) - w(X_j, X_k)$. By Assumption 1(ii), conditional on the types of agents i, j, k,

$$\frac{\Pr(S_{i,jk} = 1 \mid X_{(i,j,k)}, A_i, A_j, A_k)}{\Pr(S_{i,jk} = -1 \mid X_{(i,j,k)}, A_i, A_j, A_k)} = \exp\left(\Delta_w(X_i, X_j, X_k) + 2A_i\right).$$

The right-hand side does not depend on A_j or A_k. Integrating over their conditional distribution therefore preserves the odds ratio. Moreover, the logistic specification gives both informative configurations positive conditional probability. Hence,

$$\Pr\left(S_{i,jk} = 1 \mid |S_{i,jk}| = 1, X_{(i,j,k)}, A_i\right) = \Lambda\left(\Delta_w(X_i, X_j, X_k) + 2A_i\right). \tag{2}$$

Thus, the anchored-triad comparison eliminates the two comparison-agent effects,

while retaining the focal-agent effect and a contrast of dyadic surplus terms.

For a fixed focal agent i, varying the comparison pair (j,k) under the large-network experiment identifies the conditional distribution of $(S_{i,jk}, X_j, X_k)$ given the realized type (X_i, A_i). Consequently, the conditional probability

$$\tau_i(t_1,t_2) \equiv \Pr\left(S_{i,jk}=1 \mid |S_{i,jk}|=1, X_j=t_1, X_k=t_2, X_i, A_i\right) \tag{3}$$

is identified for $P_X \otimes P_X$-almost every $(t_1,t_2) \in \mathcal{X}\times\mathcal{X}$.

Identification of A_i evaluates τ_i at the diagonal point (X_i, X_i). If X_i is discretely distributed, then the event $X_j = X_k = X_i$ occurs with positive probability, so $\tau_i(X_i, X_i)$ is directly identified. When X_i contains continuously distributed components, however, (X_i, X_i) may have probability zero, so almost-everywhere identification alone does not determine $\tau_i(X_i, X_i)$. The following continuity and normalization conditions extend identification to these diagonal points.

Assumption 2 (Diagonal continuity and location normalization)

(i) For every $x \in \mathcal{X}$, w is continuous at (x,x).

(ii) For every $x \in \mathcal{X}$, $w(x,x) = 0$.

By Assumption 2(i) and continuity of Λ, $\tau_i(t_1,t_2)$ is continuous at $(X_i,X_i) \in \mathcal{X}\times\mathcal{X}$. Hence, its value at the diagonal is uniquely determined as the limit of identified values $\tau_i(t_{1r},t_{2r})$ along any sequence of identified points (t_{1r},t_{2r}) converging to (X_i,X_i). By Assumption 2(ii), $\tau_i(X_i,X_i) = \Lambda(2A_i)$, which uniquely determines the realized value of A_i.

Theorem 1 (Identification of agent effects) *Suppose Assumptions 1 and 2 hold. Then, for each fixed agent i, the realized agent effect A_i is identified, i.e.,*

$$A_i = \frac{1}{2}\Lambda^{-1}\left(\tau_i(X_i,X_i)\right) = \frac{1}{2}\log\left(\frac{\tau_i(X_i,X_i)}{1-\tau_i(X_i,X_i)}\right).$$

Hence, the realized sequence $(A_1, A_2, \ldots)$ is coordinatewise identified.

Theorem 1 formalizes the anchored-triad argument developed above. Because the identification steps have already been established explicitly, its proof is omitted.

Given the identified agent effects, the population distribution of $(D_{ij}, X_i, A_i, X_j, A_j)$ identifies the conditional link probability

$$p_D(x, a, x', a') \equiv \Pr\left(D_{ij} = 1 \mid (X_i, A_i, X_j, A_j) = (x, a, x', a')\right) = \Lambda\left(w(x, x') + a + a'\right)$$

for $P_{XA} \otimes P_{XA}$-almost every (x, a, x', a'), and hence identifies w $P_X \otimes P_X$-almost everywhere. To extend the result, I impose the condition below.

Assumption 3 (Global continuity) *w is continuous on $\mathcal{X} \times \mathcal{X}$.*

Assumption 3 strengthens the diagonal continuity in Assumption 2(i). The two conditions serve different roles. Assumption 2(i) is sufficient to identify the realized agent effects, whereas Assumption 3 extends the almost-everywhere identified w to its full support. In particular, under Assumption 3 and continuity of Λ, $p_D(x, a, x', a')$ is continuous on $\mathcal{S}_{XA} \times \mathcal{S}_{XA}$, so its almost-everywhere identified values uniquely determine it everywhere. Together, Assumption 2(ii) and Assumption 3 match the normalization and continuity conditions in Gao (2020) and Qi (2023).

Theorem 2 (Identification of dyadic surplus) *Suppose Assumptions 1, 2(ii), and 3 hold. Then w is identified on $\mathcal{X} \times \mathcal{X}$, i.e., for every $(x, a), (x', a') \in \mathcal{S}_{XA}$,*

$$w(x, x') = \Lambda^{-1}\left(p_D(x, a, x', a')\right) - a - a' = \log\left(\frac{p_D(x, a, x', a')}{1 - p_D(x, a, x', a')}\right) - a - a'.$$

Since the result follows directly from the preceding argument, the proof of Theorem 2 is omitted. Because X_i is observed and A_i is identified, the joint distribution of (X_i, A_i) is therefore identified. The following result is immediate.

Corollary 1 (Conditional mean of agent effects) *Suppose Assumptions 1 and 2 hold and $E[|A_i|] < \infty$. Then $\varphi(x) \equiv E[A_i \mid X_i = x]$ is identified P_X-almost surely.*

4 Estimation

This section translates the identification results into a feasible likelihood-based procedure. I impose the parametric specifications

$$w(X_i, X_j) = g(X_i, X_j)'\beta_0, \qquad \varphi(X_i) = h(X_i)'\gamma_0,$$

where $g : \mathcal{X} \times \mathcal{X} \to \mathbb{R}^{d_\beta}$ and $h : \mathcal{X} \to \mathbb{R}^{d_\gamma}$ are known, d_β and d_γ are fixed, $g(x, \tilde{x}) = g(\tilde{x}, x)$, and $g(x, x) = 0$. The latter restriction preserves the location normalization in Assumption 2(ii). Write $W_{ij} = g(X_i, X_j)$ and $Z_i = h(X_i)$. Equivalently,

$$A_i = Z_i'\gamma_0 + V_i, \qquad E[V_i \mid X_i] = 0.$$

Let $\mathcal{B} \subseteq \mathbb{R}^{d_\beta}$, $\mathcal{A}_0 \subseteq \mathbb{R}$, and $\Gamma \subseteq \mathbb{R}^{d_\gamma}$ denote the parameter spaces for the dyadic-surplus parameter, an agent effect, and the conditional-mean parameter, respectively. The parameters of interest are $\beta_0 \in \mathcal{B}$, the realized agent effects $A^n \in \mathcal{A}_0^n$, and $\gamma_0 \in \Gamma$.

The proposed triad-anchored likelihood estimator (TALE) has four stages. The first stage uses informative anchored triads to obtain initial estimates of β_0 and A^n. The second stage refines these estimates using all dyadic link indicators with diminishing stabilization. The third stage corrects the leading incidental-parameter bias in the estimate of β_0, and the fourth stage bias-reduces the agent-effect estimates and uses them to estimate γ_0. The resulting estimate is

$$\widehat{\Theta}^{\text{TALE}} = \left(\widehat{\beta}^{\text{TALE}}, \widehat{A}^{n,\text{TALE}}, \widehat{\gamma}^{\text{TALE}}\right).$$

4.1 The Triad-Anchored Likelihood Estimator

Stage 1: Triad anchoring. For an i-anchored triad (i, j, k), define $\Delta_{i,jk} = W_{ij} + W_{ik} - W_{jk}$. Under the parametric specification above, the argument of Λ in (2) becomes $2A_i + \Delta_{i,jk}'\beta_0$. For $s \in \{-1, 0, 1\}$ and $\eta \in \mathbb{R}$, define

$$\ell_T(s, \eta) = |s| \log \Lambda(s\eta) = |s| \left[s\eta - \log\left(1 + \exp(s\eta)\right)\right].$$

Let $\mathcal{I}_{T,n} \equiv \{i : \{-1, 1\} \subseteq \{S_{i,jk} : 1 \le j < k \le n,\ j, k \ne i\}\}$ contain the agents for which both informative anchored-triad configurations are observed. The anchored-triad conditional log-likelihood is

$$\mathcal{L}_{T,n}\left(\beta, \{a_i : i \in \mathcal{I}_{T,n}\}\right) = \sum_{i \in \mathcal{I}_{T,n}} \sum_{\substack{1 \le j < k \le n \\ j,k \ne i}} \ell_T\left(S_{i,jk}, 2a_i + \Delta_{i,jk}'\beta\right). \tag{4}$$

The first-stage estimates are given by

$$\left(\widehat{\beta}^T, \{\widehat{A}_i^T : i \in \mathcal{I}_{T,n}\}\right) \in \arg \max_{\substack{\beta \in \mathcal{B}, \\ a_i \in \mathcal{A}_0,\, i \in \mathcal{I}_{T,n}}} \mathcal{L}_{T,n}\left(\beta, \{a_i : i \in \mathcal{I}_{T,n}\}\right).$$

For $i \notin \mathcal{I}_{T,n}$, I set $\widehat{A}_i^T = 0$ by convention.

Stage 2: Stabilized dyadic refinement. The second stage uses all dyadic link indicators to refine the first-stage estimates. For $d \in \{0,1\}$ and $\eta \in \mathbb{R}$, define

$$\ell_D(d,\eta) = d\eta - \log\left(1 + \exp(\eta)\right),$$

and let

$$\mathcal{L}_{D,n}(\beta, a^n) = \sum_{1 \le i < j \le n} \ell_D\left(D_{ij}, W_{ij}'\beta + a_i + a_j\right). \tag{5}$$

The full dyadic likelihood uses link information discarded by the anchored-triad criterion and hence has the potential to improve estimation precision. Direct joint estimation of the high-dimensional agent effects, however, can be unstable in finite samples. I therefore refine the first-stage estimates using the full dyadic likelihood while initially penalizing departures of a_i from $\widehat{A}_i^T$ for $i \in \mathcal{I}_{T,n}$, with the penalty gradually relaxed as the refinement proceeds.

Write $\widehat{A}^{n,T} \equiv (\widehat{A}_i^T)_{i=1}^n$ and initialize $(\beta^{(0)}, a^{n,(0)}) = (\widehat{\beta}^T, \widehat{A}^{n,T})$. At iteration t, define

$$p_{ij}^{(t)} = \Lambda\left(W_{ij}'\beta^{(t)} + a_i^{(t)} + a_j^{(t)}\right), \qquad H_{i,t}^D = \sum_{j \neq i} p_{ij}^{(t)}\left(1 - p_{ij}^{(t)}\right).$$

The quantity $H_{i,t}^D$ is the negative second derivative of $\mathcal{L}_{D,n}$ with respect to a_i, evaluated at the current iterate. I use this curvature to scale the agent-specific penalty. Let

$$\kappa_{i,t} = \mathbb{1}(i \in \mathcal{I}_{T,n}) c_t H_{i,t}^D, \qquad c_t = c_0 \rho^t,$$

where $c_0 > 0$ and $\rho \in (0,1)$ are fixed constants. This scales the penalty to the local likelihood curvature for each agent while allowing its influence to decline geometrically as the refinement proceeds.

Given $\kappa_{i,t}$, define

$$\mathcal{Q}_{n,t}(\beta, a^n) = \mathcal{L}_{D,n}(\beta, a^n) - \frac{1}{2}\sum_{i=1}^{n} \kappa_{i,t}\left(a_i - \widehat{A}_i^T\right)^2. \tag{6}$$

Conditional on $\beta^{(t)}$, update the agent effects according to

$$a^{n,(t+1)} \in \arg\max_{a^n \in \mathcal{A}_0^n} \mathcal{Q}_{n,t}\left(\beta^{(t)}, a^n\right),$$

and then update the dyadic-surplus parameter according to

$$\beta^{(t+1)} \in \arg\max_{\beta \in \mathcal{B}} \mathcal{L}_{D,n}\left(\beta, a^{n,(t+1)}\right).$$

Only the agent effects are penalized, whereas β is re-estimated from the full dyadic likelihood at every iteration. Alternating between the two parameter blocks is used for computational convenience, replacing joint optimization over $n + d_\beta$ parameters by successive conditional optimizations.

Let R_n be the first iteration for which $c_{R_n} \leq \delta_n$ and

$$\left\|\beta^{(R_n+1)} - \beta^{(R_n)}\right\| + \max_{1\leq i\leq n}\left|a_i^{(R_n+1)} - a_i^{(R_n)}\right| \leq \varepsilon_n,$$

where δ_n and ε_n are positive sequences converging to zero at rates specified in Subsection 4.2. The second-stage estimates are

$$\left(\widehat{\beta}^S, \widehat{A}^{n,S}\right) = \left(\beta^{(R_n+1)}, a^{n,(R_n+1)}\right).$$

For an interior solution, the β-maximization step implies

$$\sum_{i<j}\left[D_{ij} - \Lambda\left(W_{ij}'\widehat{\beta}^S + \widehat{A}_i^S + \widehat{A}_j^S\right)\right]W_{ij} = 0. \tag{7}$$

Stage 3: Bias correction for the common parameter. The third stage corrects the leading bias in $\widehat{\beta}^S$ induced by estimating the growing vector of agent effects. Holding $\widehat{\mathbf{A}}^S$ fixed, for a candidate b define

$$\widehat{p}_{ij}^S(b) = \Lambda\left(W_{ij}'b + \widehat{A}_i^S + \widehat{A}_j^S\right), \qquad \widehat{\nu}_{ij}^S(b) = \widehat{p}_{ij}^S(b)\left\{1 - \widehat{p}_{ij}^S(b)\right\}.$$

Let $(\widehat{\xi}_{1,S}(b), \ldots, \widehat{\xi}_{n,S}(b))$ solve the weighted additive projection

$$\min_{(\xi_1,\ldots,\xi_n)\in(\mathbb{R}^{d_\beta})^n} \sum_{i<j} \widehat{\nu}_{ij}^S(b) \left\| W_{ij} - \xi_i - \xi_j \right\|^2,$$

and define

$$\widetilde{W}_{ij}^S(b) = W_{ij} - \widehat{\xi}_{i,S}(b) - \widehat{\xi}_{j,S}(b).$$

Thus, $\widetilde{W}_{ij}^S(b)$ is W_{ij} residualized with respect to additive agent-specific terms under the likelihood weights $\widehat{\nu}_{ij}^S(b)$.

Define

$$\widehat{\Gamma}_S(b) = \sum_{i<j} \widehat{\nu}_{ij}^S(b) \widetilde{W}_{ij}^S(b) \widetilde{W}_{ij}^S(b)', \qquad \widehat{H}_{i,S}(b) = \sum_{j\neq i} \widehat{\nu}_{ij}^S(b),$$

and

$$\widehat{R}_{i,S}(b) = \sum_{j\neq i} -\widehat{\nu}_{ij}^S(b) \left\{1 - 2\widehat{p}_{ij}^S(b)\right\} \widetilde{W}_{ij}^S(b).$$

A second-order expansion of (7), accounting for estimation of the agent effects, yields the estimated leading bias at b,

$$\widehat{b}_{\beta,S}(b) = \widehat{\Gamma}_S(b)^{-1} \frac{1}{2} \sum_{i=1}^{n} \frac{\widehat{R}_{i,S}(b)}{\widehat{H}_{i,S}(b)}. \tag{8}$$

Let ζ_n be a positive sequence converging to zero at a rate specified in Subsection 4.2. The third-stage estimate $\widehat{\beta}^{\mathrm{TALE}}$ satisfies

$$\left\| \widehat{\beta}^{\mathrm{TALE}} - \left\{ \widehat{\beta}^S - \widehat{b}_{\beta,S}\left(\widehat{\beta}^{\mathrm{TALE}}\right) \right\} \right\| \leq \zeta_n. \tag{9}$$

Because the bias map depends on the corrected value of b, (9) is solved iteratively. Starting from $b^{(0)} = \widehat{\beta}^S$, iterate $b^{(r+1)} = \widehat{\beta}^S - \widehat{b}_{\beta,S}(b^{(r)})$, holding $\widehat{\mathbf{A}}^S$ fixed, until the tolerance is met.

Stage 4: Bias-reduced effect profiling. The third stage applies a bias correction to the second-stage estimate $\widehat{\beta}^S$ while retaining $\widehat{A}^{n,S}$. The fourth stage therefore fixes β at $\widehat{\beta}^{\mathrm{TALE}}$ and re-estimates the agent effects using a bias-reduced likelihood.

For any (β, a^n), define

$$\nu_{ij}(\beta, a^n) = \Lambda\left(W_{ij}'\beta + a_i + a_j\right)\left[1 - \Lambda\left(W_{ij}'\beta + a_i + a_j\right)\right].$$

Let $J_A(\beta, a^n) = [J_{A,ij}(\beta, a^n)]_{i,j=1}^n$ denote the fixed-β information matrix for the agent effects, with

$$J_{A,ii}(\beta, a^n) = \sum_{j \neq i} \nu_{ij}(\beta, a^n), \qquad J_{A,ij}(\beta, a^n) = \nu_{ij}(\beta, a^n), \quad i \neq j.$$

To reduce finite-sample bias in the agent-effect estimates, I use the Jeffreys–Firth adjustment of Firth (1993). The final agent-effect estimates are

$$\widehat{A}^{n,\mathrm{TALE}} \in \arg\max_{a^n \in \mathcal{A}_0^n} \left\{ \mathcal{L}_{D,n}\left(\widehat{\beta}^{\mathrm{TALE}}, a^n\right) + \frac{1}{2}\log\det J_A\left(\widehat{\beta}^{\mathrm{TALE}}, a^n\right) \right\}. \tag{10}$$

For computation, $\widehat{A}^{n,S}$ serves as the initial value for this fixed-β optimization.

Finally, using the specification $\varphi(X_i) = Z_i'\gamma_0$, estimate the conditional-mean parameter by

$$\widehat{\gamma}^{\mathrm{TALE}} = \left(\sum_{i=1}^n Z_i Z_i'\right)^{-1} \sum_{i=1}^n Z_i \widehat{A}_i^{\mathrm{TALE}}. \tag{11}$$

Together, (9), (10), and (11) complete the construction of TALE.

4.2 Large-sample properties

I now establish the large-sample properties of TALE as $n \to \infty$, with d_β and d_γ fixed. The following conditions are maintained throughout this subsection.

Assumption 4 (Parameter spaces and regressor variation)

(i) *$\mathcal{B}$ is compact and convex, with $\beta_0 \in \mathrm{int}(\mathcal{B})$. The set $\mathcal{A}_0$ is a compact interval containing zero, and $\mathcal{A} \subset \mathrm{int}(\mathcal{A}_0)$.*

(ii) *W_{ij} and Z_i are uniformly bounded.*

(iii) *The matrices $Q_\perp \equiv E[W_{ij}^\perp W_{ij}^{\perp\prime}]$, $Q_Z \equiv E[Z_i Z_i']$, and $\Omega_\gamma \equiv E[Z_i Z_i' V_i^2]$ are positive definite, where $W_{ij}^\perp = W_{ij} - E[W_{ij} \mid X_i] - E[W_{ij} \mid X_j] + E[W_{ij}]$.*

Under Assumption 4(i) and the boundedness of W_{ij} in Assumption 4(ii), the logit index $W_{ij}'\beta + a_i + a_j$ is uniformly bounded over $(\beta, a_i, a_j) \in \mathcal{B} \times \mathcal{A}_0^2$. The resulting link probabilities are therefore uniformly bounded away from zero and one, and expected degrees are of order n. The asymptotic analysis is thus for a dense-network sequence. Assumption 4(iii) supplies the nondegeneracy conditions for estimation of β_0 and γ_0.

Proposition 1 (Availability and convergence rates) *Suppose Assumptions 1 and 4(i)–(ii) hold, $Q_\perp$ is positive definite, $\delta_n = O(n^{-1/2})$, and $\varepsilon_n = O(n^{-1})$. Then* $\Pr(\mathcal{I}_{T,n} = \{1, \dots, n\}) \to 1$. *For $r \in \{T, S, \text{TALE}\}$, $\widehat{\beta}^r - \beta_0 = O_p(n^{-1})$, and $\widehat{A}_i^r - A_i = O_p(n^{-1/2})$ for every fixed agent i. Moreover,*

$$\frac{1}{n}\sum_{i=1}^{n}\left(\widehat{A}_i^r - A_i\right)^2 = O_p(n^{-1}), \qquad \max_{1\le i\le n}\left|\widehat{A}_i^r - A_i\right| = O_p\left(\sqrt{\frac{\log n}{n}}\right).$$

The availability result implies that the fallback $\widehat{A}_i^T = 0$ is asymptotically irrelevant. The proposition also shows that $\widehat{\beta}^r - \beta_0 = O_p(n^{-1})$ and $\widehat{A}_i^r - A_i = O_p(n^{-1/2})$ for each fixed agent.

Let $(\widehat{\beta}^J, \widehat{A}^{n,J})$ denote the joint maximum likelihood (JML) estimator obtained from $\mathcal{L}_{D,n}$. If the rate conditions in Proposition 1 are strengthened to $\delta_n = o(n^{-1/2})$ and $\varepsilon_n = o(n^{-1})$, then $\widehat{\beta}^S - \widehat{\beta}^J = o_p(n^{-1})$ and, for each fixed i, $\widehat{A}_i^S - \widehat{A}_i^J = o_p(n^{-1/2})$. Thus, the stabilization is asymptotically negligible relative to the estimation errors of the JML estimator.

For first-order inference, strengthen the stopping-rate conditions to $\delta_n = o(n^{-1/2})$ and $\varepsilon_n = o(n^{-1})$, and let the Stage-3 fixed-point tolerance satisfy $\zeta_n = o(n^{-1})$. These conditions make the remaining stabilization and numerical errors negligible relative to the n^{-1} stochastic scale of the dyadic-surplus estimator.

For first-stage inference, let ι_i denote the ith coordinate vector and define $x_{i,jk} \equiv (\Delta_{i,jk}', 2\iota_i')'$. At the true parameters, let

$$H_{T,0} \equiv \sum_i \sum_{j<k;\, j,k\neq i} v_{i,jk} x_{i,jk} x_{i,jk}', \qquad \Omega_{T,0} \equiv \operatorname{Var}\left(\sum_i \sum_{j<k;\, j,k\neq i} x_{i,jk} u_{i,jk} \,\middle|\, \mathcal{F}_n\right).$$

Write $V_{T,0} \equiv H_{T,0}^{-1}\Omega_{T,0}H_{T,0}^{-1}$, and let $V_{\beta,T,0}$ denote its β block and $v_{i,T,0}$ its diagonal entry corresponding to A_i.

Assumption 5 (Nondegenerate first-stage variance) *For every fixed nonzero*

$c \in \mathbb{R}^{d_\beta}$ *and every fixed agent* i*, there exist constants* $0 < \underline{c} < \bar{c} < \infty$ *such that, with probability approaching one,* $\underline{c} \le n^2 c' V_{\beta,T,0} c \le \bar{c}$ *and* $\underline{c} \le n v_{i,T,0} \le \bar{c}$*.*

The first-stage variance estimator accounts for the dependence between anchored-triad scores that share link indicators. At the first-stage estimates, define

$$\widehat{q}_{i,jk} \equiv \Lambda(2\widehat{A}_i^T + \Delta'_{i,jk}\widehat{\beta}^T, \qquad \widehat{u}_{i,jk} \equiv \frac{|S_{i,jk}| + S_{i,jk}}{2} - |S_{i,jk}|\widehat{q}_{i,jk},$$
$$\widehat{v}_{i,jk} \equiv |S_{i,jk}|\widehat{q}_{i,jk}(1 - \widehat{q}_{i,jk}), \qquad \widehat{H}_T \equiv \sum_i \sum_{j<k;\, j,k\neq i} \widehat{v}_{i,jk} x_{i,jk} x'_{i,jk}.$$

For an unordered dyad d, let $\widehat{G}_d$ be the sum of $x_{i,jk}\widehat{u}_{i,jk}$ over anchored triads whose three constituent dyads include d. All sums use focal agents in $\mathcal{I}_{T,n}$; when $\mathcal{I}_{T,n}$ is a proper subset, the matrices are formed over the corresponding estimated effect coordinates. Define

$$\widehat{V}_T \equiv \widehat{H}_T^{-1} \left(\sum_d \widehat{G}_d \widehat{G}'_d \right) \widehat{H}_T^{-1}. \tag{12}$$

Write $\widehat{V}_{\beta,T}$ for its β block and $\widehat{v}_{i,T}$ for its diagonal entry corresponding to A_i.

For the dyadic estimates, use the quantities defined in Stage 3 to set

$$\widehat{V}_\beta(b) \equiv \widehat{\Gamma}_S(b)^{-1} \left[\sum_{i<j} \left\{ D_{ij} - \widehat{p}_{ij}^S(b) \right\}^2 \widetilde{W}_{ij}^S(b) \widetilde{W}_{ij}^S(b)' \right] \widehat{\Gamma}_S(b)^{-1}. \tag{13}$$

For $r \in \{S, \text{TALE}\}$, set $\widehat{V}_{\beta,r} \equiv \widehat{V}_\beta(\widehat{\beta}^r)$ and $\widehat{v}_{i,r} \equiv \left[J_A(\widehat{\beta}^r, \widehat{A}^{n,r})^{-1} \right]_{ii}$.

Let $b_{T,n}$ and $b_{D,n}$ denote the leading conditional bias terms in the expansions of $\widehat{\beta}^T$ and the dyadic JML estimator, respectively. Their expressions are derived in Appendix A.2, and both are of order n^{-1}. The dyadic term is the population counterpart of the Stage-3 bias map, whereas the first-stage term additionally reflects the random curvature of the anchored-triad criterion. Set $b_{S,n} \equiv b_{D,n}$ and $b_{\text{TALE},n} \equiv 0$.

Theorem 3 (Asymptotic normality) *Suppose Assumption 1, Assumption 4(i)–(ii), and Assumption 5 hold, and* $Q_\perp$ *is positive definite. Suppose further that* $\delta_n = o(n^{-1/2})$*,* $\varepsilon_n = o(n^{-1})$*, and* $\zeta_n = o(n^{-1})$*. Then, for every fixed nonzero* $c \in \mathbb{R}^{d_\beta}$*, every fixed agent* i*, and* $r \in \{T, S, \text{TALE}\}$*,*

$$\frac{c'(\widehat{\beta}^r - \beta_0 - b_{r,n})}{\sqrt{c'\widehat{V}_{\beta,r}c}} \Rightarrow N(0,1), \qquad \frac{\widehat{A}_i^r - A_i}{\sqrt{\widehat{v}_{i,r}}} \Rightarrow N(0,1).$$

The variance estimators are ratio consistent for the conditional variances of the corresponding leading linear terms. In particular, the dyadic-surplus variances are of order n^{-2} and the fixed-agent variances are of order n^{-1}.

The plug-in first-stage bias estimator defined in Appendix A.2 satisfies

$$\widehat{b}_{T,n} - b_{T,n} = o_p(n^{-1}),$$

and the Stage-3 bias map satisfies

$$\widehat{b}_{\beta,S}(\widehat{\beta}^S) - b_{D,n} = o_p(n^{-1}).$$

Hence the same normal limits hold after replacing $b_{T,n}$ and $b_{D,n}$ by these feasible bias estimates. All limits above also hold conditionally on $\mathcal{F}_n$, in probability.

The first-stage and stabilized dyadic-surplus estimates therefore have leading biases of the same order as their standard errors. Feasible centered inference uses $\widehat{\beta}^T - \widehat{b}_{T,n}$ at the first stage and $\widehat{\beta}^S - \widehat{b}_{\beta,S}(\widehat{\beta}^S)$ after dyadic refinement. Stage 3 incorporates the latter correction into $\widehat{\beta}^{\mathrm{TALE}}$. For a fixed agent, by contrast, the Jeffreys–Firth adjustment changes the fixed-β effect profile by order n^{-1}, which is negligible relative to its $n^{-1/2}$ stochastic scale.

Finally, define $\widehat{V}_i \equiv \widehat{A}_i^{\mathrm{TALE}} - Z_i'\widehat{\gamma}^{\mathrm{TALE}}$, and

$$\widehat{V}_\gamma \equiv \left(\sum_i Z_i Z_i'\right)^{-1} \left(\sum_i Z_i Z_i' \widehat{V}_i^2\right) \left(\sum_i Z_i Z_i'\right)^{-1}. \tag{14}$$

Corollary 2 (Conditional-mean parameter) *Suppose Assumptions 1 and 4 hold and the conditions of Theorem 3 are satisfied. Then $\widehat{\gamma}^{\mathrm{TALE}} \to_p \gamma_0$ and*

$$\sqrt{n}\left(\widehat{\gamma}^{\mathrm{TALE}} - \gamma_0\right) = Q_Z^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i V_i + o_p(1) \Rightarrow N\left(0, Q_Z^{-1}\Omega_\gamma Q_Z^{-1}\right).$$

Moreover, $n\widehat{V}_\gamma \to_p Q_Z^{-1}\Omega_\gamma Q_Z^{-1}$.

Thus, averaging the estimated agent effects makes their estimation error second order for estimation of γ_0. Specifically, $\frac{1}{n}\sum_{i=1}^n Z_i\left(\widehat{A}_i^{\mathrm{TALE}} - A_i\right) = O_p(n^{-1})$.

5 Monte Carlo simulation

This section evaluates the finite-sample performance of the triad-anchored likelihood estimator (TALE) and its four stages in designs that reproduce the main empirical difficulties motivating the paper: unobserved degree heterogeneity, dependence between observables and latent link propensities, and low-density realized networks. The simulations focus on three margins that are central to the paper's contribution. The first is the performance of Stage 1 triad anchoring relative to tetrad logit for estimating the dyadic-surplus parameter. The second is the improvement in the recovered realized agent effects as TALE moves from triad anchoring to stabilized dyadic refinement and then to bias-reduced effect profiling. The third is computational feasibility when informative local conditioning configurations become scarce, a setting in which direct joint fixed-effect methods often return estimates only on a selected subset of replications.

The designs are motivated by Appendix D of Graham (2017), but are tailored to the present paper because the target includes both the dyadic-surplus parameter and the realized vector of agent effects. For each unordered dyad $1 \leq i < j \leq n$, links are generated according to

$$D_{ij} = 1\{W_{ij}\beta_0 + A_i + A_j \geq U_{ij}\},$$

where the shocks U_{ij} are i.i.d. standard logistic. The dyadic regressor is $W_{ij} = |X_i - X_j|$, with $X_i \sim \text{Unif}[-2, 2]$ independently across agents, and I set $\beta_0 = -1$. Thus, in these designs, β_0 is a homophily coefficient: holding the unobserved agent effects fixed, similarity in the observed characteristic raises the probability of a link. The agent effect is generated as

$$A_i = \gamma_0 + \gamma_1 X_i + V_i,$$

with

$$V_i = \text{sign}(X_i)\left[B_i - \frac{\lambda_0}{\lambda_0 + \lambda_1}\right], \qquad B_i \sim \text{Beta}(\lambda_0, \lambda_1),$$

where B_i is independent of X_i. This construction keeps $E[V_i \mid X_i] = 0$, while allowing the shape of latent heterogeneity and its dependence on observables to vary across designs. The parameter γ_0 shifts the overall link propensity and hence network den-

sity, whereas γ_1 controls the systematic dependence of latent link propensity on the observed characteristic.

The main tables use 500 independent replications with $n = 100$ agents. Table 1 reports the six designs and the resulting network statistics. Designs S.1–S.3 set $\gamma_1 = 0$ and $(\lambda_0, \lambda_1) = (1, 1)$, so that latent agent effects are conditionally mean independent of X_i and symmetrically distributed around γ_0. Designs S.4–S.6 set $\gamma_1 = 0.25$ and $(\lambda_0, \lambda_1) = (0.25, 0.75)$, producing agent effects that are correlated with observables and skewed. In both groups, $\gamma_0 \in \{-0.5, -0.75, -1\}$ moves the network from moderately dense to relatively low density.

Table 1: Simulation Designs and Network Summary Statistics

	Symmetric Uncorrelated Agent Effects			Right-Skewed Correlated Agent Effects		
	S.1	S.2	S.3	S.4	S.5	S.6
Panel A: Design parameters						
γ_0	-0.5	-0.75	-1	-0.5	-0.75	-1
γ_1	0	0	0	0.25	0.25	0.25
λ_0	1	1	1	0.25	0.25	0.25
λ_1	1	1	1	0.75	0.75	0.75
Panel B: Network statistics						
Largest Component Fraction	1.000	0.998	0.984	0.998	0.991	0.966
Network Density	0.121	0.079	0.051	0.127	0.085	0.055
Avg. Degree	11.946	7.798	5.013	12.574	8.415	5.456
Std. of Degree	4.521	3.437	2.619	5.941	4.528	3.332
Network Transitivity	0.161	0.108	0.072	0.206	0.149	0.104
Informative Triad Fraction	0.105	0.072	0.048	0.108	0.076	0.051
Tetrad Tetrad Fraction	0.025	0.013	0.006	0.024	0.013	0.007

Notes: Panel A reports the parameter values used in each simulation design. Panel B reports averages across 500 Monte Carlo replications with $n = 100$. Largest Component Fraction is the fraction of agents contained in the largest connected component. Network Density is the fraction of realized links among all unordered dyads. Avg. Degree and Std. of Degree are the mean and standard deviation of agent degree. Network Transitivity is the global clustering coefficient: the number of closed triples divided by the number of connected triples. Informative Triad Fraction is the fraction of anchored triad configurations satisfying $|S_{i,jk}| = 1$. Informative Tetrad Fraction is the fraction of two-link tetrad configurations in which the two links share no agents, as used by the tetrad-logit estimator of Graham (2017).

The designs deliberately avoid a trivial fragmentation problem. Even in the lowest-density correlated design, the largest component contains about 97 percent

of agents. The difficulty instead comes from the declining amount of local conditioning information. As γ_0 falls, density decreases from about 0.12 to about 0.05 and average degree falls from roughly 12 to roughly 5. At the same time, transitivity and the frequency of conditioning configurations fall sharply. The informative anchored-triad fraction declines from about 0.11 to about 0.05, while the tetrad conditioning fraction falls from about 0.025 to about 0.006. Thus the difficult cases are lower-density networks in which the graph remains largely connected but the local configurations used by conditional procedures become increasingly rare.

Table 2 reports the performance of three benchmarks and the four stages of TALE for β_0 and the realized agent effects A^n. Tetrad logit provides a conditional-likelihood benchmark for β_0. JML is direct joint maximum likelihood, and JML + RBC is the residualized bias-correction benchmark applied to JML. The remaining rows correspond to the four stages of the proposed estimator. TALE Stage 1 is triad anchoring. Stage 2 performs stabilized dyadic refinement using the full dyadic likelihood. Stage 3 applies the analytical bias correction to $\widehat{\beta}^S$ while retaining the Stage-2 agent effects. Stage 4, which delivers the final TALE estimate, fixes β at the Stage-3 estimate and bias-reduces the agent-effect profile. Accordingly, the Stage-3 entries for A^n coincide with Stage 2, while the Stage-4 entries for β_0 coincide with Stage 3. This layout shows which stage is responsible for each finite-sample gain without presenting the four stages as separate proposed estimators.

For the dyadic-surplus parameter, the main message is that recovering agent effects does not come at the cost of a weak estimate of β_0. Tetrad logit performs well, as expected: it has small bias and coverage near or above the nominal level in all six designs, but it does not estimate A^n. TALE is competitive on the same parameter. Stage 3, which is also the β component of the final TALE estimate, has smaller RMSE than tetrad logit in every design in Table 2, with RMSE between 0.068 and 0.103 compared with 0.075 to 0.123 for tetrad logit. Its coverage remains close to nominal, ranging from 0.942 to 0.960. Thus the simulations do not indicate a trade-off in which latent heterogeneity can be recovered only by sacrificing precision in the dyadic-surplus parameter.

The comparison with JML and JML + RBC highlights the computational role of triad anchoring. The direct JML benchmarks perform reasonably in the denser designs conditional on convergence, but their success counts deteriorate quickly as network density falls. At $n = 100$, JML and JML + RBC succeed in only 116 of 500

Table 2: Monte Carlo Results for (β_0, A^n)

Design	Method	Dyadic-surplus parameter β_0				Realized agent effects A^n				Successful replications
		Bias	SD	RMSE	95% Cov.	Bias	SD	RMSE	95% Cov.	
	Tetrad logit	-0.004	0.075	0.075	0.982					500
	JML	-0.027	0.069	0.074	0.904	-0.085	0.040	0.366	0.967	499
	JML + RBC	-0.003	0.068	0.068	0.942	-0.041	0.038	0.361	0.954	499
S.1	TALE: Stage 1	-0.023	0.071	0.075	0.968	-0.098	0.069	0.540	0.958	500
	TALE: Stage 2	-0.027	0.069	0.074	0.932	-0.032	0.040	0.366	0.953	500
	TALE: Stage 3	-0.003	0.068	0.068	0.956	-0.032	0.040	0.366	0.950	500
	TALE: Stage 4	-0.003	0.068	0.068	0.956	0.000	0.036	0.337	0.955	500
	Tetrad logit	-0.014	0.091	0.092	0.976					500
	JML	-0.036	0.084	0.091	0.899	-0.133	0.049	0.461	0.977	405
	JML + RBC	-0.010	0.081	0.082	0.926	-0.067	0.045	0.450	0.960	405
S.2	TALE: Stage 1	-0.029	0.089	0.094	0.974	-0.323	0.178	1.538	0.953	500
	TALE: Stage 2	-0.033	0.085	0.091	0.948	-0.107	0.116	0.845	0.957	500
	TALE: Stage 3	-0.008	0.082	0.082	0.958	-0.107	0.116	0.845	0.954	500
	TALE: Stage 4	-0.008	0.082	0.082	0.958	-0.002	0.044	0.415	0.955	500
	Tetrad logit	-0.013	0.118	0.119	0.968					500
	JML	-0.033	0.098	0.103	0.879	-0.190	0.060	0.557	0.985	116
	JML + RBC	-0.005	0.095	0.095	0.922	-0.094	0.055	0.537	0.963	116
S.3	TALE: Stage 1	-0.027	0.117	0.120	0.970	-1.089	0.371	3.454	0.901	500
	TALE: Stage 2	-0.029	0.107	0.110	0.940	-0.433	0.292	2.625	0.949	500
	TALE: Stage 3	-0.002	0.103	0.103	0.942	-0.433	0.292	2.625	0.946	500
	TALE: Stage 4	-0.002	0.103	0.103	0.942	-0.004	0.055	0.513	0.952	500
	Tetrad logit	-0.022	0.075	0.078	0.958					500
	JML	-0.026	0.071	0.076	0.906	-0.086	0.044	0.378	0.967	415
	JML + RBC	-0.003	0.070	0.069	0.920	-0.043	0.042	0.374	0.955	415
S.4	TALE: Stage 1	-0.026	0.074	0.078	0.952	-0.206	0.136	1.184	0.950	500
	TALE: Stage 2	-0.027	0.070	0.075	0.934	-0.076	0.108	0.731	0.952	500
	TALE: Stage 3	-0.003	0.068	0.068	0.942	-0.076	0.108	0.731	0.948	500
	TALE: Stage 4	-0.003	0.068	0.068	0.942	-0.001	0.040	0.353	0.954	500
	Tetrad logit	-0.031	0.090	0.095	0.970					500
	JML	-0.024	0.078	0.082	0.919	-0.124	0.052	0.453	0.974	221
	JML + RBC	0.001	0.076	0.076	0.932	-0.061	0.048	0.443	0.957	221
S.5	TALE: Stage 1	-0.026	0.089	0.093	0.964	-0.570	0.257	2.365	0.928	500
	TALE: Stage 2	-0.029	0.084	0.089	0.950	-0.254	0.234	1.821	0.948	500
	TALE: Stage 3	-0.004	0.081	0.081	0.960	-0.254	0.234	1.821	0.944	500
	TALE: Stage 4	-0.004	0.081	0.081	0.960	0.001	0.046	0.427	0.951	500
	Tetrad logit	-0.038	0.117	0.123	0.970					500
	JML	0.010	0.104	0.102	0.885	-0.178	0.070	0.540	0.982	26
	JML + RBC	0.037	0.101	0.105	0.808	-0.085	0.062	0.521	0.962	26
S.6	TALE: Stage 1	-0.031	0.117	0.121	0.970	-1.402	0.425	3.970	0.869	500
	TALE: Stage 2	-0.026	0.104	0.107	0.940	-0.769	0.418	3.900	0.932	500
	TALE: Stage 3	0.001	0.101	0.101	0.946	-0.769	0.418	3.900	0.929	500
	TALE: Stage 4	0.001	0.101	0.101	0.946	0.000	0.058	0.517	0.951	500

Notes: Entries, including 95% coverage rates, are based on successful replications among 500 Monte Carlo replications with $n = 100$, where success means that the computation completed and returned estimates. JML denotes joint maximum likelihood; JML + RBC is the residualized bias-correction benchmark. TALE Stages 1–4 are triad anchoring, stabilized dyadic refinement, bias correction for β, and bias-reduced effect profiling, respectively. Stage 3 retains the Stage-2 A^n, while Stage 4 retains the Stage-3 β. Bias, SD, RMSE, and 95% Cov. denote Monte Carlo mean bias, standard deviation, root mean squared error, and coverage. For A^n, Bias and SD use the average error $n^{-1}\sum_i(\widehat{A}_i - A_i)$, RMSE uses $\{n^{-1}\sum_i(\widehat{A}_i - A_i)^2\}^{1/2}$, and coverage is averaged across agents.

replications in Design S.3 and only 26 of 500 replications in Design S.6. Their reported bias, RMSE, and coverage in those designs are therefore computed on selected successful replications. By contrast, all four TALE stages return estimates in all 500 replications in every design. This distinction is important empirically: the relevant comparison concerns not only accuracy conditional on numerical success, but also whether the procedure remains operational when heterogeneous degree propensities and scarce conditioning events make direct high-dimensional likelihood optimization difficult.

The rows for A^n show why the four-stage construction is useful. Stage 1 already produces estimates of both β_0 and the realized agent effects, reflecting the constructive content of the identification argument, but its role is to provide an anchored starting point. Its RMSE for A^n rises in the lower-density designs because each focal agent is supported by fewer informative anchored triads. Stage 2 improves these estimates by incorporating all dyadic link indicators. The largest improvement in the realized effects comes from Stage 4, which bias-reduces the fixed-β effect profile. In Design S.3, the RMSE for A^n falls from 2.625 at Stage 2 to 0.513 at Stage 4. In Design S.6, it falls from 3.900 to 0.517. The resulting coverage for the agent effects is close to 95 percent across all designs. Hence the simulation evidence supports the paper's central empirical premise: the latent effects can be treated as estimable economic objects, rather than only as nuisance parameters to be conditioned out.

Table 3 examines the next step, namely whether the recovered agent effects can be used to estimate the systematic component $A_i = \gamma_0 + \gamma_1 X_i + V_i$. The table reports this target for the final TALE estimate; the benchmark rows are omitted. The estimates use the Stage-4 agent effects and project them on $(1, X_i)$, exactly as in the final step of the estimation procedure.

The low-dimensional heterogeneity parameters are recovered accurately. For γ_0, the absolute bias is at most 0.006 and the RMSE is at most 0.066 across the six designs. Coverage ranges from 0.956 to 0.976. For γ_1, the absolute bias is at most 0.008, the RMSE is at most 0.049, and coverage ranges from 0.966 to 0.982. These results are important because they evaluate a target that conditional likelihood methods cannot deliver. TALE not only controls for latent degree heterogeneity when estimating the dyadic-surplus parameter; it also recovers enough information about that heterogeneity to study its systematic relationship with observed characteristics.

Appendix Tables B.1–B.3 and B.4–B.6 repeat the exercise with $n = 50$ and $n =$

Table 3: Monte Carlo Results for (γ_0, γ_1)

	γ_0				γ_1			
Design	Bias	SD	RMSE	95% Cov.	Bias	SD	RMSE	95% Cov.
S.1	0.001	0.047	0.047	0.964	-0.002	0.036	0.036	0.982
S.2	-0.004	0.053	0.053	0.974	0.000	0.040	0.040	0.982
S.3	-0.006	0.062	0.062	0.976	0.007	0.047	0.048	0.982
S.4	-0.002	0.050	0.050	0.956	0.001	0.039	0.039	0.972
S.5	-0.003	0.055	0.055	0.968	-0.002	0.044	0.044	0.966
S.6	0.000	0.066	0.066	0.964	-0.008	0.048	0.049	0.972

Notes: Entries, including 95% coverage rates, are based on successful replications among 500 Monte Carlo replications with $n = 100$. The estimates of (γ_0, γ_1) are obtained by projecting the final TALE agent-effect estimates $\widehat{A}^{n,\mathrm{TALE}}$ on $(1, X_i)$. Bias, SD, RMSE, and 95% Cov. denote Monte Carlo mean bias, standard deviation, root mean squared error, and coverage.

150. The same pattern appears at both sample sizes. With $n = 50$, the direct JML benchmarks succeed in only 10 replications in Design S.3 and 5 replications in Design S.6, whereas all four TALE stages remain feasible in all 500 replications. With $n = 150$, the benchmark algorithms improve but still have only 88 successes in the lowest-density correlated design. The final TALE estimates of A^n and (γ_0, γ_1) become more precise as n grows. Overall, the Monte Carlo evidence supports three conclusions: informative anchored triads provide a useful first-stage anchor for β_0 and A^n; stabilized dyadic refinement preserves numerical feasibility while using the full network; and the final bias corrections deliver accurate inference for the dyadic-surplus parameter and economically useful estimates of latent heterogeneity.

6 Empirical Applications

6.1 Microfinance Village Networks

I apply TALE to the village networks collected by Banerjee et al. (2013) in their study of the diffusion of microfinance in Karnataka, India. The data contain household socioeconomic characteristics and nominations across multiple types of social and economic relationships. I define an undirected link between two households if either reports a relationship with the other. For each of the 43 villages, I retain the giant component and treat it as a separate network.

The retained sample contains 9,126 households. Table 4 summarizes the six house-

hold characteristics used in the analysis: private electricity, latrine ownership, numbers of beds and rooms, household size, and an indicator for whether the household contains an individual designated by the microfinance program as a leader. About 63 percent of households have private electricity, 27 percent own a latrine, mean household size is 4.8, and 12.5 percent are leader households.

Table 4: Household characteristics across 43 villages

Variable	Mean	Std. dev.	Min.	Max.
Private electricity	0.628	0.483	0	1
Owned latrine	0.273	0.446	0	1
Beds	0.874	1.424	0	50
Rooms	2.367	1.327	0	20
Household size	4.776	2.218	1	31
Leader	0.125	0.331	0	1

Notes: Summary statistics use all 9,126 retained households. Binary variables are indicators; beds, rooms and household size are counts in levels. Supplied data and existing imputations are retained.

I first estimate a common link-formation specification across the 43 networks. For the binary household characteristics, the dyadic regressors indicate whether the two households have the same status. For beds, rooms, and household size, the regressors are absolute differences. The common dyadic-surplus coefficients are pooled across villages, while TALE allows each of the 9,126 households to have its own latent link propensity. All within-village household pairs enter the likelihood, giving 1,024,042 dyads in total.

Table 5 compares pooled logit with TALE. The pooled specification omits household effects, whereas the TALE column reports the Stage-3 bias-corrected dyadic-surplus estimates. The comparison shows that controlling for latent degree heterogeneity can materially change the estimated pattern of matching on observed characteristics.

For the two binary measures of household infrastructure, TALE finds positive assortative matching. Households with the same electricity status and the same latrine-ownership status are more likely to be linked, with estimated coefficients of 0.323 and 0.324, respectively. For the count variables, similarity in beds and rooms is also associated with link formation: the corresponding gap coefficients are -0.067 and -0.018. The household-size gap is small and statistically insignificant.

The contrast with pooled logit is particularly informative. The pooled coefficient

Table 5: Link Formation Estimates across 43 Indian Villages

Variable	Pooled Logit	TALE
Same private electricity status	0.298***	0.323***
	(0.010)	(0.011)
Same owned latrine status	0.163***	0.324***
	(0.010)	(0.012)
Beds gap	0.004	-0.067***
	(0.003)	(0.006)
Rooms gap	0.028***	-0.018***
	(0.004)	(0.006)
Household size gap	0.039***	-0.005
	(0.002)	(0.004)
Same leader status	-0.289***	0.159***
	(0.011)	(0.017)
Constant	-3.293***	–
	(0.015)	
Household effects	No	Yes
Villages	43	43
Households	9,126	9,126
Dyads	1,024,042	1,024,042
Log likelihood	-181,465.73	-155,580.39
Pseudo R^2	0.0065	0.1483

Notes: All within-village household pairs enter the estimation; cross-village pairs are excluded. The six dyadic covariates have common coefficients across villages. Binary characteristics are represented by same-status indicators, while count characteristics enter as absolute gaps. Pooled Logit includes a common intercept and omits household effects. TALE includes unrestricted household effects and reports the Stage-3 bias-corrected dyadic-surplus estimates $\widehat{\beta}^{\mathrm{TALE}}$. Fit statistics are evaluated at the final TALE estimates. Standard errors are in parentheses; TALE uses the nuisance-adjusted dyadic sandwich estimator. McFadden pseudo-R^2 uses a pooled intercept-only model as the null. $^{*}p < 0.10$, $^{****}p < 0.05$, $^{***}p < 0.01$.

on the beds gap is essentially zero, while the rooms and household-size gaps have positive coefficients. After household effects are introduced, the beds and rooms coefficients become negative and the household-size coefficient moves close to zero. The coefficient on same leader status changes even more sharply, from -0.289 under pooled logit to 0.159 under TALE. Thus, some apparent differences in linking across household types are reversed once persistent household-specific link propensities are accounted for. In this application, treating degree heterogeneity as an omitted component can therefore alter not only the magnitude but also the sign of estimated homophily patterns.

The improvement in in-sample fit is also substantial: the McFadden pseudo-R^2 rises from 0.0065 under pooled logit to 0.1483 under TALE. Because TALE introduces a large collection of household effects, this comparison should not be interpreted as formal model-selection evidence. Its main role is to indicate that household-specific link propensities capture an empirically important dimension of the observed networks.

A distinctive feature of TALE is that this heterogeneity is recovered rather than removed from the analysis. I therefore next examine the final Stage-4 household effects $\widehat{A}_{iv}^{\mathrm{TALE}}$, where i indexes households and v villages. These effects are obtained by bias-reduced profiling with the Stage-3 dyadic-surplus estimate held fixed. To illustrate their network interpretation, I display one of the 43 village networks. The effects are centered within the displayed network for graphical presentation.

Figure 2 shows pronounced heterogeneity in latent link propensity within a single village. Households with large positive effects tend to appear in the densely connected part of the network, whereas households with low effects tend to have fewer observed connections. This relationship is expected because A_{iv} shifts the probability of every link involving household i. The recovered effect, however, is a model-based link propensity conditional on observed household similarity and on the effects of potential partners, rather than simply the household's observed degree.

To examine this relationship directly, let $d_{iv} \equiv \sum_{j \neq i} D_{ij,v}$ denote household i's observed degree within village v.

Figure 3 shows that the two objects are very strongly related in Village 1. The Pearson correlation is approximately 0.935 and the Spearman rank correlation is approximately 0.991. This close relationship is economically natural: households with a high persistent propensity to form links should generally have high observed degree.

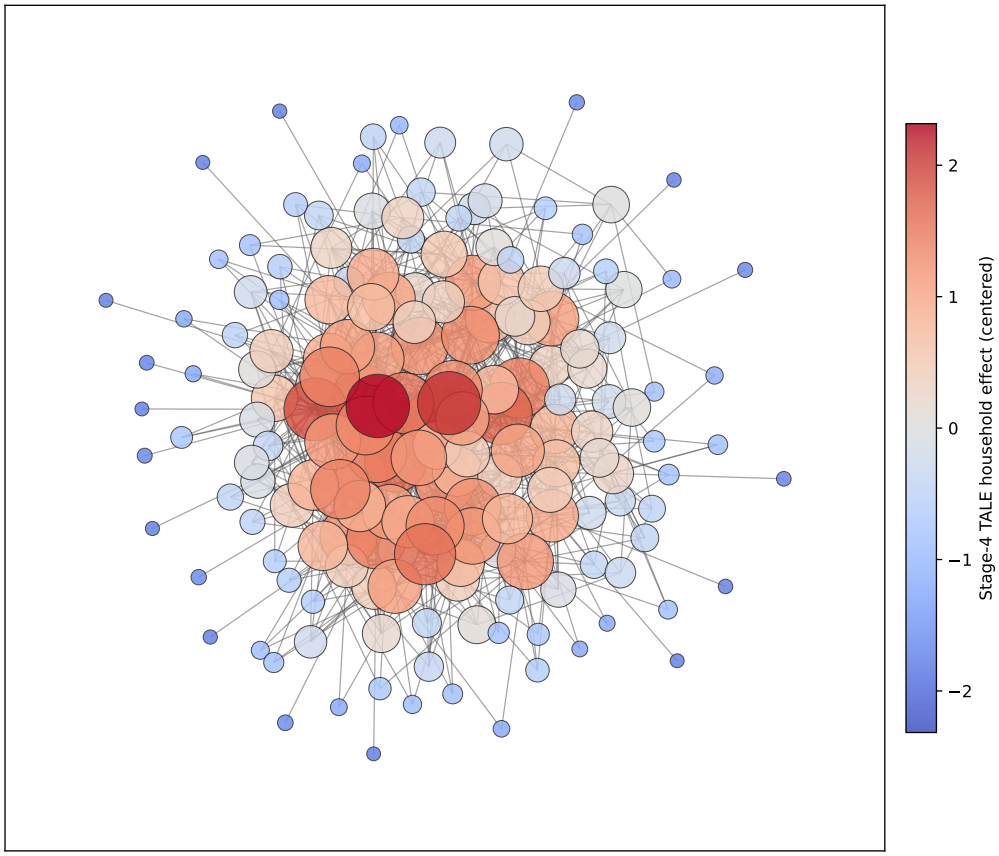


Figure 2: Microfinance Village Network: Recovered Household Effects and Observed Links

Notes: The figure displays the giant component of Village 1. Each node represents a household, and edges indicate observed social or economic relationships. Node size and color reflect the mean-centered Stage-4 TALE household effects; larger nodes and warmer colors indicate higher recovered link propensities.

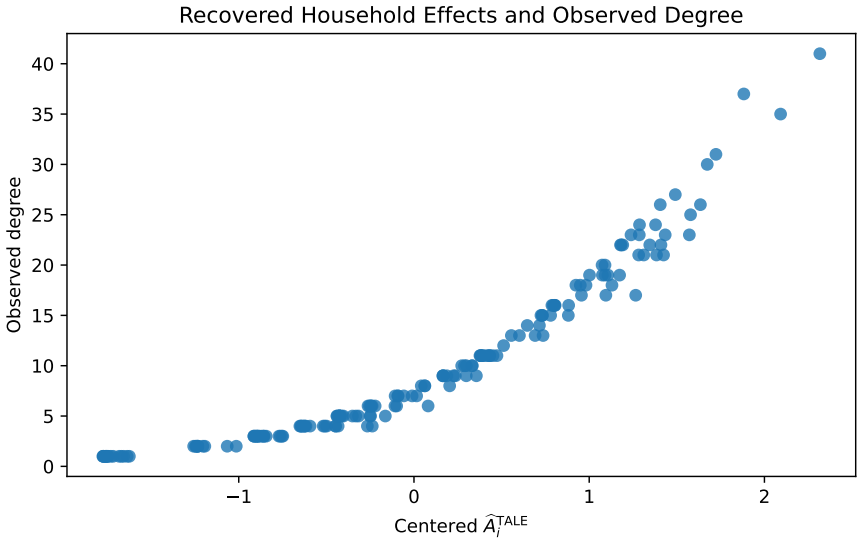


Figure 3: Microfinance Village Network: Recovered Household Effects and Observed Degree

Notes: The figure displays households in the giant component of Village 1. The horizontal axis reports the mean-centered Stage-4 TALE household effect and the vertical axis reports observed degree.

It also clarifies what is gained by estimating the household effects. Degree is a realized network count, whereas $\widehat{A}_{iv}^{\text{TALE}}$ places households on a common covariate-adjusted link-propensity scale. This adjustment is precisely what permits the comparison between pooled logit and TALE above and, more importantly, allows the recovered heterogeneity itself to become an object of economic analysis.

I finally examine the systematic relationship between the recovered household effects and observed household characteristics. This is the empirical counterpart of the conditional-mean specification $A_i = Z_i'\gamma_0 + V_i$ introduced in Section 4.

Table 6 relates the final Stage-4 TALE household effects to observed household characteristics. The first column pools the 43 villages without village fixed effects, while the second adds village fixed effects and therefore identifies the household-characteristic associations from within-village variation.

The estimated relationships are remarkably stable across the two specifications. Private electricity has a small and statistically insignificant coefficient in both columns. By contrast, latrine ownership, the numbers of beds and rooms, household size, and leader status are all positively and precisely associated with the recovered household effects. With village fixed effects, the corresponding coefficients are 0.136 for latrine

Table 6: Recovered Household Effects and Household Characteristics

Variable	No Village FE	Village FE
Constant	-2.663***	-2.598***
	(0.027)	(0.070)
Private electricity	-0.031	-0.028
	(0.019)	(0.020)
Owned latrine	0.114***	0.136***
	(0.022)	(0.022)
Beds	0.037***	0.041***
	(0.008)	(0.008)
Rooms	0.035***	0.029***
	(0.008)	(0.008)
Household size	0.106***	0.108***
	(0.004)	(0.004)
Leader	0.388***	0.386***
	(0.028)	(0.027)
Villages	43	43
Households	9,126	9,126
R^2	0.1412	0.1729

Notes: The dependent variable in both columns is the uncentered final Stage-4 TALE household effect $\widehat{A}_{iv}^{\mathrm{TALE}}$ from the link model pooled across all 43 villages. Both regressions include a constant and all six household characteristics in levels. The first column omits village indicators; the second includes 42 village indicators, with village 1 as the reference category. SEs in parentheses incorporate household-effect estimation uncertainty, covariance across villages from the shared estimated β, and the noise-adjusted residual-variance correction. R^2 includes all regressors in the respective column. These coefficients describe associations. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

ownership, 0.041 for beds, 0.029 for rooms, 0.108 for household size, and 0.386 for leader status. The leader coefficient is the largest: conditional on the other household characteristics and village fixed effects, households containing an individual designated by the microfinance program as a leader have substantially higher recovered link propensities.

The similarity of the coefficients with and without village fixed effects is informative. For example, the leader coefficient changes only from 0.388 to 0.386, while the household-size coefficient changes from 0.106 to 0.108. The positive associations therefore primarily reflect within-village differences across households rather than differences in average latent link propensity across villages. Adding village fixed effects raises the R^2 from 0.141 to 0.173, indicating that village-level heterogeneity contributes additional explanatory power but leaves the main household-level relationships largely unchanged.

These results give economic content to the latent degree heterogeneity that matters for the link-formation estimates. Larger households, households with more housing resources, and especially households containing a designated leader tend to have systematically higher underlying propensities to form social and economic relationships. At the same time, the specification with village fixed effects explains only about 17 percent of the variation in the recovered household effects, leaving most of the household-specific heterogeneity unexplained by the available observables.

Taken together, the 43-village application illustrates two roles of latent degree heterogeneity in network formation. First, allowing for household effects materially changes the estimated pattern of matching on observed characteristics, including several sign reversals relative to pooled logit. Second, because TALE recovers rather than eliminates these effects, their systematic relationship with household characteristics can be studied directly. The application therefore separates observable assortative matching from persistent household-specific propensity to participate in the village network.

6.2 Nyakatoke Village Network

I apply TALE to a household risk-sharing network from Nyakatoke, a village in Tanzania studied by De Weerdt (2004). A link is present if either household names the other as someone it can rely on for help in cash, kind, or labor. The resulting undi-

rected network therefore records reported support relationships among households. The sample contains 116 households and 480 links among 6,670 unordered household pairs, giving a network density of 0.072 and an average degree of 8.276.

Table 7: Nyakatoke Village Network: Summary Statistics

Variable	Obs.	Mean	SD	Min	Med	Max
Panel A: Household characteristics						
Age of household head	116	45.509	15.820	21.000	42.000	89.000
Female head	116	0.267	0.444	0.000	0.000	1.000
Head born in NYK	116	0.526	0.501	0.000	1.000	1.000
Catholic household	116	0.422	0.496	0.000	0.000	1.000
Muslim household	116	0.207	0.407	0.000	0.000	1.000
Log wealth	116	12.531	1.908	0.000	12.678	15.783
Panel B: Dyadic link outcome and link shifters						
Link indicator	6,670	0.072	0.258	0.000	0.000	1.000
Age gap of heads	6,670	17.862	13.474	0.000	15.000	68.000
Same-sex head pair	6,670	0.605	0.489	0.000	1.000	1.000
Same NYK-born status	6,670	0.497	0.500	0.000	0.000	1.000
Same religion	6,670	0.353	0.478	0.000	0.000	1.000
Close kin	6,670	0.016	0.126	0.000	0.000	1.000
Extended kin	6,670	0.015	0.121	0.000	0.000	1.000
Other blood relation	6,670	0.025	0.156	0.000	0.000	1.000
Log wealth gap	6,670	1.437	2.284	0.000	0.882	15.783
Log physical distance	6,670	6.049	0.708	2.667	6.138	7.460

Notes: Obs., Mean, SD, Min, Med, and Max denote the number of observations, sample mean, standard deviation, minimum, median, and maximum, respectively. Panel A reports household-level characteristics. Panel B reports variables defined for unordered household pairs. The link indicator equals one if either household names the other as someone one can rely on for help in cash, kind, or labor. Log wealth is defined as $\log(1 + \text{wealth})$, where wealth is measured in Tanzanian shillings. Age gap of heads is the absolute age difference between the two household heads. Same head sex is an indicator equal to one if the two household heads have the same sex. Log wealth gap is the absolute difference in log wealth. Physical distance is the absolute geographic distance between households, measured in meters before taking logs. The network has 480 links, density 0.072, average degree 8.276, largest-component fraction 1.000, global transitivity 0.188, and usable triad fraction 0.066.

Table 7 summarizes the household and dyadic characteristics used in the analysis, including religion, kinship, wealth, demographic similarity, and physical distance.

I first estimate the link-formation model. The baseline specification includes same religion, indicators for close kin, extended kin, and other blood relations, the log

wealth gap, and log physical distance. The extended specification also includes the age gap between household heads, same head sex, and same Nyakatoke-born status.

Table 8: Nyakatoke Village Network: Link Formation Model Estimates

	Pooled Logit		TALE	
	Baseline	Extended	Baseline	Extended
Constant	2.586***	2.582***		
	(0.378)	(0.396)		
Same religion	0.350***	0.357***	0.430***	0.429***
	(0.105)	(0.105)	(0.106)	(0.106)
Close kin	2.615***	2.562***	2.666***	2.612***
	(0.244)	(0.248)	(0.291)	(0.289)
Extended kin	1.757***	1.811***	1.719***	1.781***
	(0.245)	(0.248)	(0.250)	(0.254)
Other blood relation	1.390***	1.403***	1.295***	1.297***
	(0.205)	(0.205)	(0.234)	(0.236)
Log wealth gap	-0.053*	-0.048*	-0.095*	-0.076
	(0.027)	(0.027)	(0.056)	(0.057)
Log physical distance	-0.936***	-0.952***	-0.948***	-0.947***
	(0.065)	(0.065)	(0.071)	(0.071)
Age gap of heads		-0.007*		-0.013***
		(0.004)		(0.005)
Same head sex		0.336***		0.315**
		(0.110)		(0.126)
Same NYK-born status		-0.003		0.021
		(0.105)		(0.099)
Obs.	6,670	6,670	6,670	6,670
Log likelihood	-1434.432	-1427.352	-1269.045	-1262.199
Pseudo R^2	0.169	0.173	0.265	0.268

Notes: The dependent variable is the link indicator. The pooled logit columns omit household effects and include a constant. The TALE columns report the bias-corrected dyadic-surplus estimates $\widehat{\beta}^{\mathrm{TALE}}$ from Stage 3. Quantities that depend on household effects are evaluated using the final Stage-4 estimates $\widehat{A}^{n,\mathrm{TALE}}$. Standard errors are in parentheses; TALE uses the nuisance-adjusted dyadic sandwich evaluated at the final estimates. Log likelihood and McFadden pseudo-R^2 use the unpenalized dyadic likelihood at the reported estimates. Significance levels: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 8 compares pooled logit with TALE. The pooled specifications impose a common link propensity across households, whereas TALE allows each household to have its own latent propensity to form links. The TALE columns report the Stage-3

bias-corrected dyadic-surplus estimates.

The main patterns in link formation remain after allowing for household heterogeneity. In the extended TALE specification, the coefficient on same religion is 0.429. The coefficients on close kin, extended kin, and other blood relations are 2.612, 1.781, and 1.297, respectively, while the coefficient on log physical distance is -0.947. Kinship, common religion, and geographic proximity are therefore strongly associated with whom households rely on even after household-specific differences in overall link propensity are taken into account.

The extended specification also identifies additional matching patterns. The age-gap coefficient is -0.013, indicating that households whose heads are farther apart in age are less likely to be linked. The same-head-sex coefficient is 0.315, while same Nyakatoke-born status has little association with link formation. I therefore use the extended specification for the analysis of household heterogeneity below.

Allowing for household effects also changes the estimated association with wealth differences. In the extended specification, the coefficient on the log wealth gap changes from -0.048 under pooled logit to -0.076 under TALE, while its standard error rises from 0.027 to 0.057. The point estimate becomes somewhat more negative, but the evidence for wealth homophily becomes less precise once latent household heterogeneity is allowed for. The log likelihood and pseudo-R^2 also improve relative to the pooled specification, although these comparisons are descriptive because TALE introduces household-specific parameters.

A central advantage of TALE is that the household-specific heterogeneity is recovered rather than eliminated as a nuisance component. I therefore turn to the final Stage-4 household effects $\widehat{A}_i^{\mathrm{TALE}}$ from the extended specification. These effects are obtained by bias-reduced profiling with the Stage-3 dyadic-surplus estimate held fixed. For graphical presentation, I center the recovered effects at their sample mean.

Figure 4 shows substantial variation in the recovered household effects. Households with larger effects tend to have more observed connections, as expected. Holding the partner's effect and the observed dyadic characteristics fixed, a larger A_i implies a higher propensity for household i to form a support relationship. The recovered effect is therefore a conditional link propensity rather than simply the household's observed number of connections.

To clarify this distinction, let $d_i \equiv \sum_{j \neq i} D_{ij}$ denote observed degree. Under the link-formation model, observed degree depends on A_i, but also on the observed

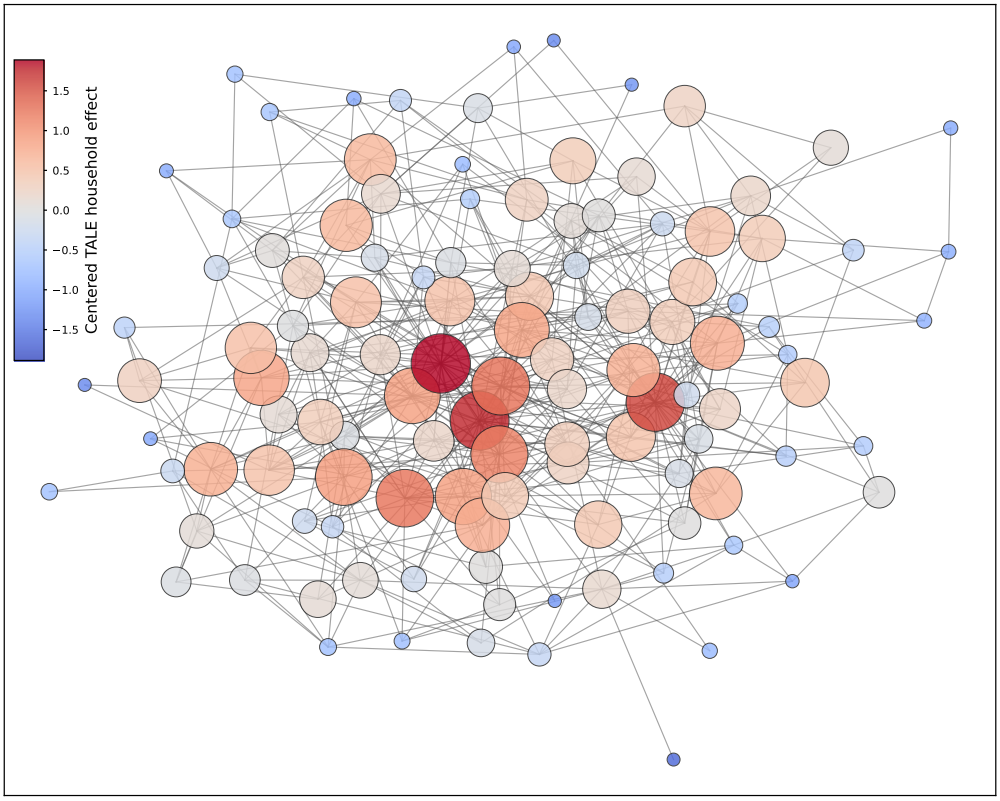


Figure 4: Nyakatoke Village Network: Recovered Household Effects and Observed Links

Notes: Each node represents a household, and edges indicate reported support relationships. Node size and color reflect the mean-centered Stage-4 TALE household effects from the extended specification; larger nodes and warmer colors indicate higher recovered link propensities.

characteristics of household i's potential dyads and on the latent effects of its potential partners. Recovering A_i therefore adjusts observed connectivity for these differences in link opportunities.

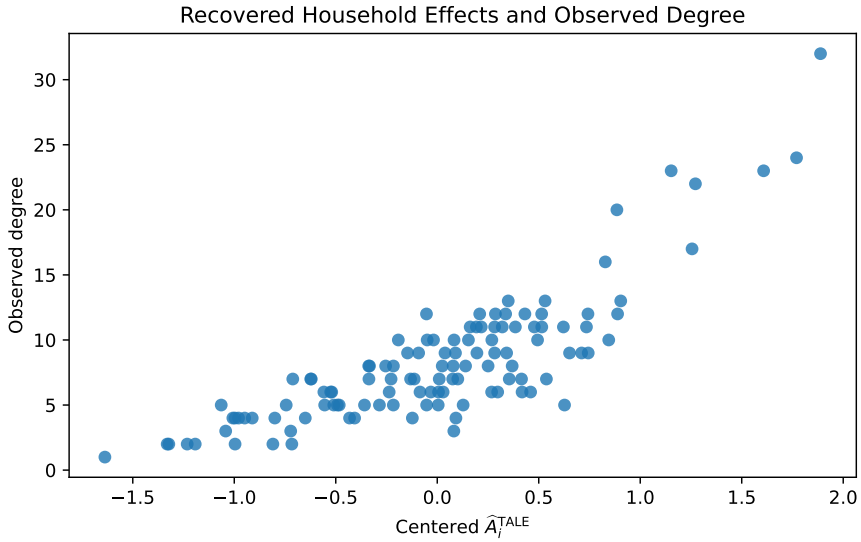


Figure 5: Nyakatoke Village Network: Recovered Household Effects and Observed Degree

Notes: Each point represents a household. The horizontal axis reports the mean-centered Stage-4 TALE household effect from the extended specification, and the vertical axis reports observed degree. Mean-centering does not affect cross-household differences or the correlation with degree.

Figure 5 compares the recovered effects with observed degree directly. The two are strongly related: the Pearson correlation is approximately 0.826 and the Spearman rank correlation is 0.794. The relationship is nevertheless not one-to-one. Among the 13 households with five observed links, for example, the centered household effects span approximately 1.71 logit-index units. Thus, households with the same observed degree can have substantially different recovered link propensities once their dyadic characteristics and potential partners are taken into account.

Having recovered the household-specific component of link propensity, I finally examine its systematic relationship with observed household attributes. This exercise corresponds to the conditional-mean specification $A_i = Z_i'\gamma_0 + V_i$ introduced in Section 4.

Table 9 projects the final Stage-4 TALE household effects on observed household characteristics. The baseline projection includes religion and log wealth, while the extended projection additionally includes female head, Nyakatoke-born status, and

Table 9: Recovered Household Effects and Household Characteristics

	Baseline	Extended
Constant	0.530	1.074
	(0.881)	(0.906)
Catholic household	-0.046	-0.092
	(0.141)	(0.146)
Muslim household	-0.021	-0.027
	(0.176)	(0.180)
Log wealth	0.059	0.071
	(0.067)	(0.064)
Female head		-0.115
		(0.158)
Head born in NYK		-0.166
		(0.143)
Head born in 1930s		-0.375
		(0.310)
Head born in 1940s		-0.628**
		(0.290)
Head born in 1950s		-0.546**
		(0.249)
Head born in 1960s		-0.643***
		(0.240)
Head born in 1970s+		-0.677***
		(0.249)
Households	116	116
R^2	0.030	0.127

Notes: The dependent variable is the final Stage-4 TALE household effect $\widehat{A}_i^{\text{TALE}}$ from the extended link-formation specification. Effects enter the projections without mean-centering. The regressions are descriptive projections of recovered household effects and should not be interpreted as causal effects. The baseline column includes religion and log wealth. The second column additionally includes female head, head born in Nyakatoke, and household-head birth-cohort indicators. Birth cohorts are decade indicators constructed from the 2000 survey year and the household head's reported age; the omitted category is pre-1930. The cohort indicators capture cross-sectional cohort differences and do not imply that the time-invariant household effect changes as the head ages. Standard errors account for uncertainty in the bias-reduced household effects and include the beta plug-in correction. Significance levels: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

household-head birth-cohort indicators.

Religion and wealth explain little of the recovered heterogeneity. The coefficients on Catholic household, Muslim household, and log wealth are small and statistically insignificant in both specifications. Female head and Nyakatoke-born status are also statistically insignificant in the extended specification. The additional household characteristics raise the R^2 from 0.030 to 0.127, but most of the variation in the recovered effects remains unexplained by the available observables.

The clearest systematic pattern is across birth cohorts. Relative to households whose heads were born before 1930, the estimated effects are lower for the later cohorts. The coefficients for heads born in the 1940s, 1950s, 1960s, and 1970s or later are statistically significant, whereas the coefficient for the 1930s is not. These estimates describe cross-sectional cohort differences and should not be interpreted as implying that the time-invariant household effect changes as a household head ages.

Overall, the Nyakatoke application separates two dimensions of network formation. Observable dyadic characteristics—most prominently kinship, religion, demographic similarity, and physical proximity—help determine which households form support relationships. At the same time, households differ substantially in their underlying propensity to form links even after these characteristics are held fixed. Recovering these household effects makes it possible to distinguish latent link propensity from observed degree and then to study its systematic relationship with household characteristics.

7 Concluding Remarks

This paper develops a framework for identifying and estimating dyadic link formation models with latent agent heterogeneity. Informative anchored triads identify realized agent effects rather than merely difference them out, after which dyadic variation identifies the systematic component of link formation. Building on this identification strategy, the triad-anchored likelihood estimator (TALE) combines triad anchoring, stabilized dyadic refinement, bias correction for the dyadic-surplus parameter, and bias-reduced profiling of the agent effects. The large-sample analysis establishes convergence rates and asymptotic inference for both the dyadic-surplus parameter and fixed agents' latent effects.

Monte Carlo experiments show that the different stages of TALE improve finite-

sample accuracy and stability relative to simpler alternatives, while the empirical applications illustrate the economic value of recovering agent heterogeneity. In the Nyakatoke risk-sharing network and across 43 Indian village networks, latent degree heterogeneity both affects estimates of observable matching patterns and contains systematic information about household link propensities. More broadly, the framework makes latent agent heterogeneity an estimable object rather than a nuisance to be eliminated. Extensions to directed networks, downstream outcome models, and strategic link formation are directions for future research.

References

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

Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2013): "The diffusion of microfinance," *Science*, 341, 1236498.

Billings, S. B. and M. Hoekstra (2023): "The Effect of School and Neighborhood Peers on Achievement, Misbehavior, and Adult Crime," *Journal of Labor Economics*.

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.

Cai, J., A. D. Janvry, and E. Sadoulet (2015): "Social networks and the decision to insure," *American Economic Journal: Applied Economics*, 7, 81–108.

Calvó-Armengol, A., E. Patacchini, and Y. Zenou (2009): "Peer effects and social networks in education," *The Review of Economic Studies*, 76, 1239–1267.

Chandrasekhar, A. G. and M. O. Jackson (2025): "A Network Formation Model Based on Subgraphs," *Review of Economic Studies*, advance Access publication.

de Paula, Á. (2020): "Econometric models of network formation," *Annual Review of Economics*, 12, 775–799.

De Weerdt, J. (2004): "Risk-Sharing and Endogenous Network Formation," in *Insurance Against Poverty*, ed. by S. Dercon, Oxford: Oxford University Press, 211–246, Nyakatoke Network Dataset available at https://doi.org/10.7910/DVN/U2OJE3.

De Weerdt, J. and S. Dercon (2006): "Risk-sharing networks and insurance against illness," *Journal of Development Economics*, 81, 337–356.

De Weerdt, J. and M. Fafchamps (2011): "Social identity and the formation of health insurance networks," *Journal of Development Studies*, 47, 1152–1177.

Dzemski, A. (2019): "An empirical model of dyadic link formation in a network with unobserved heterogeneity," *Review of Economics and Statistics*, 101, 763–776.

Ferrali, R., G. Grossman, M. R. Platas, and J. Rodden (2020): "It takes a village: Peer effects and externalities in technology adoption," *American Journal of Political Science*, 64, 536–553.

Firth, D. (1993): "Bias Reduction of Maximum Likelihood Estimates," *Biometrika*, 80, 27–38.

Gao, W. Y. (2020): "Nonparametric identification in index models of link formation," *Journal of Econometrics*, 215, 399–413.

Graham, B. S. (2017): "An econometric model of network formation with degree heterogeneity," *Econometrica*, 85, 1033–1063.

——— (2020): "Network data," in *Handbook of Econometrics*, Elsevier, vol. 7, 111–218.

Hahn, J. and W. K. Newey (2004): "Jackknife and Analytical Bias Reduction for Nonlinear Panel Models," *Econometrica*, 72, 1295–1319.

Hoxby, C. M. (2000): "Peer effects in the classroom: Learning from gender and race variation," *NBER Working Paper*.

Islam, A., G. Kusnadi, J. Rezki, A. Sim, G. van Empel, M. Vlassopoulos, and Y. Zenou (2023): "Addressing Vaccine Hesitancy Using Local Ambassadors: A Randomized Controlled Trial in Indonesia," Tech. rep., Institute of Labor Economics (IZA).

Jochmans, K. (2018): "Semiparametric analysis of network formation," *Journal of Business & Economic Statistics*, 36, 705–713.

——— (2023): "Modified-likelihood estimation of fixed-effect models for dyadic data," *SERIEs*, 14, 417–433.

Johnsson, I. and H. R. Moon (2021): "Estimation of peer effects in endogenous social networks: Control function approach," *Review of Economics and Statistics*, 103, 328–345.

Lundborg, P. (2006): "Having the wrong friends? Peer effects in adolescent substance use," *Journal of Health Economics*, 25, 214–233.

Manski, C. F. (1993): "Identification of endogenous social effects: The reflection problem," *The Review of Economic Studies*, 60, 531–542.

Moretti, E. (2011): "Social learning and peer effects in consumption: Evidence from movie sales," *The Review of Economic Studies*, 78, 356–393.

Oster, E. and R. Thornton (2012): "Determinants of technology adoption: Peer effects in menstrual cup take-up," *Journal of the European Economic Association*, 10, 1263–1293.

Qi, S. (2023): "Identification and Estimation of Discrete Choice Models with Spillovers Using Partial Network Data," *Unpublished manuscript*.

Sheng, S. (2020): "A structural econometric analysis of network formation games through subnetworks," *Econometrica*, 88, 1829–1858.

Yan, T., B. Jiang, S. E. Fienberg, and C. Leng (2019): "Asymptotics for directed networks with reciprocated links," *Statistica Sinica*, 29, 199–226.

Zeleneev, A. (2026): "Identification and estimation of network models with nonparametric unobserved heterogeneity," *arXiv preprint arXiv:2602.06885*.

A Proofs

A.1 Proof of Proposition 1

Throughout the appendix, let $\mathcal{F}_n \equiv \sigma(X^n, A^n)$. For any $(\beta, a^n) \in \mathcal{B} \times \mathcal{A}_0^n$, define $p_{ij}(\beta, a^n) \equiv \Lambda(W_{ij}'\beta + a_i + a_j)$, and write $p_{ij,0} \equiv p_{ij}(\beta_0, A^n)$, $\nu_{ij,0} \equiv p_{ij,0}(1 - p_{ij,0})$, and $e_{ij} \equiv D_{ij} - p_{ij,0}$. For an anchored triad (i, j, k), define $q_{i,jk,0} \equiv \Lambda(2A_i + \Delta_{i,jk}'\beta_0)$ and $\mu_{i,jk,0} \equiv \Pr(|S_{i,jk}| = 1 \mid \mathcal{F}_n)$.

By Assumption 4(i)–(ii), there exists $\underline{p} > 0$, independent of n, such that $\underline{p} \leq p_{ij}(\beta, a^n) \leq 1 - \underline{p}$ uniformly over $(\beta, a^n) \in \mathcal{B} \times \mathcal{A}_0^n$ and $i \neq j$. Consequently, $q_{i,jk,0}$ is uniformly bounded away from zero and one, and $\mu_{i,jk,0}$ is uniformly bounded away from zero.

Lemma A.1 (Availability of informative configurations) *Suppose Assumptions 1 and 4(i)–(ii) hold. Then* $\Pr(\mathcal{I}_{T,n} = \{1, \ldots, n\}) \to 1$.

Proof. Let $\mathcal{F}_n = \sigma(X^n, A^n)$. By compactness of the parameter spaces and boundedness of W_{ij}, there exists $\underline{p} > 0$ such that $\underline{p} \leq \Pr(D_{ij} = 1 \mid \mathcal{F}_n) \leq 1 - \underline{p}$ uniformly over $i \neq j$.

Fix i and choose $m_n = \lfloor (n-1)/2 \rfloor$ disjoint comparison pairs (j_r, k_r) among the remaining agents. Conditional on $\mathcal{F}_n$, the corresponding anchored triads depend on disjoint sets of dyadic shocks and are therefore independent. For each pair, $\Pr(S_{i,j_rk_r} = 1 \mid \mathcal{F}_n) \geq \underline{p}^3$ and $\Pr(S_{i,j_rk_r} = -1 \mid \mathcal{F}_n) \geq \underline{p}^3$. Hence,

$$\begin{aligned}\Pr(i \notin \mathcal{I}_{T,n} \mid \mathcal{F}_n) &\leq \Pr(S_{i,j_rk_r} \neq 1, \forall\, r = 1, \ldots, m_n \mid \mathcal{F}_n) \\ &\quad + \Pr(S_{i,j_rk_r} \neq -1, \forall\, r = 1, \ldots, m_n \mid \mathcal{F}_n) \leq 2(1 - \underline{p}^3)^{m_n}.\end{aligned}$$

Taking expectations and applying a union bound over i gives

$$\Pr(\mathcal{I}_{T,n} \neq \{1, \ldots, n\}) \leq 2n(1 - \underline{p}^3)^{\lfloor (n-1)/2 \rfloor} \to 0.$$

■

For the remaining arguments, work on the event $\mathcal{I}_{T,n} = \{1, \ldots, n\}$, whose probability approaches one by Lemma A.1. For a triad with focal agent i, define $u_{i,jk} \equiv \mathbb{1}(S_{i,jk} = 1) - |S_{i,jk}| q_{i,jk,0}$ and $v_{i,jk} \equiv \mu_{i,jk,0} q_{i,jk,0}(1 - q_{i,jk,0})$. The first quantity is the derivative of $\ell_T(S_{i,jk}, \eta)$ with respect to η, evaluated at $\eta = 2A_i + \Delta_{i,jk}'\beta_0$, and satisfies

$E[u_{i,jk} \mid \mathcal{F}_n] = 0$. The second is the conditional expected negative second derivative of the same contribution.

For each focal agent i, let $\mathcal{L}_{T,n,i}(\beta, a_i) \equiv \sum_{j<k;\, j,k\neq i} \ell_T(S_{i,jk}, 2a_i + \Delta'_{i,jk}\beta)$, so that $\mathcal{L}_{T,n}(\beta, a^n) = \sum_i \mathcal{L}_{T,n,i}(\beta, a_i)$. At (β_0, A_i), define the score contributions

$$U^T_{\beta,i} \equiv \partial_\beta \mathcal{L}_{T,n}(\beta_0, A^n) = \sum_{j<k;\, j,k\neq i} \Delta_{i,jk} u_{i,jk}, \quad U^T_{A,i} \equiv \partial_{a_i} \mathcal{L}_{T,n}(\beta_0, A^n) = 2 \sum_{j<k;\, j,k\neq i} u_{i,jk}.$$

The conditional expected negative Hessian blocks of $\mathcal{L}_{T,n,i}$ at the true parameters are

$$\begin{aligned}
\mathcal{H}^T_{\beta\beta,i} &\equiv -E\left[\left. \frac{\partial^2 \mathcal{L}_{T,n,i}(\beta, a_i)}{\partial\beta\partial\beta'} \right|_{(\beta,a_i)=(\beta_0,A_i)} \,\middle|\, \mathcal{F}_n \right] = \sum_{j<k;\, j,k\neq i} v_{i,jk} \Delta_{i,jk} \Delta'_{i,jk}, \\
\mathcal{H}^T_{\beta A,i} &\equiv -E\left[\left. \frac{\partial^2 \mathcal{L}_{T,n,i}(\beta, a_i)}{\partial\beta\partial a_i} \right|_{(\beta,a_i)=(\beta_0,A_i)} \,\middle|\, \mathcal{F}_n \right] = 2 \sum_{j<k;\, j,k\neq i} v_{i,jk} \Delta_{i,jk}, \\
\mathcal{H}^T_{AA,i} &\equiv -E\left[\left. \frac{\partial^2 \mathcal{L}_{T,n,i}(\beta, a_i)}{\partial a_i^2} \right|_{(\beta,a_i)=(\beta_0,A_i)} \,\middle|\, \mathcal{F}_n \right] = 4 \sum_{j<k;\, j,k\neq i} v_{i,jk} \equiv 4K_{i,T}.
\end{aligned}$$

The residualization used below follows directly from the linearized score equations. To see this, let $d_\beta \equiv \beta - \beta_0$ and $d_i \equiv a_i - A_i$ be deviations from their true values. A Taylor expansion of the global β first-order condition and the ith agent-effect first-order condition takes the form

$$\begin{aligned}
0 &= \sum_i U^T_{\beta,i} - \sum_i \mathcal{H}^T_{\beta\beta,i} d_\beta - \sum_i \mathcal{H}^T_{\beta A,i} d_i + r_\beta, \\
0 &= U^T_{A,i} - \mathcal{H}^{T\prime}_{\beta A,i} d_\beta - \mathcal{H}^T_{AA,i} d_i + r_{A,i}, \qquad i = 1, \ldots, n,
\end{aligned}$$

where r_β and $r_{A,i}$ collect the centered-Hessian and higher-order Taylor terms. Solving the second equation for d_i and substituting into the first gives

$$0 = \sum_i \left[U^T_{\beta,i} - \mathcal{H}^T_{\beta A,i} (\mathcal{H}^T_{AA,i})^{-1} U^T_{A,i} \right] - \sum_i \left[\mathcal{H}^T_{\beta\beta,i} - \mathcal{H}^T_{\beta A,i} (\mathcal{H}^T_{AA,i})^{-1} \mathcal{H}^{T\prime}_{\beta A,i} \right] d_\beta + r_T,$$

where $r_T \equiv r_\beta - \sum_i \mathcal{H}^T_{\beta A,i} (\mathcal{H}^T_{AA,i})^{-1} r_{A,i}$.

Now define $\overline{\Delta}_i \equiv K^{-1}_{i,T} \sum_{j<k;\, j,k\neq i} v_{i,jk} \Delta_{i,jk}$ and $\widetilde{\Delta}_{i,jk} \equiv \Delta_{i,jk} - \overline{\Delta}_i$. Since $\mathcal{H}^T_{\beta A,i} =$

$2K_{i,T}\overline{\Delta}_i$ and $\mathcal{H}^T_{AA,i} = 4K_{i,T}$, the first bracket above satisfies

$$U^T_{\beta,i} - \mathcal{H}^T_{\beta A,i}(\mathcal{H}^T_{AA,i})^{-1}U^T_{A,i} = \sum_{j<k;\, j,k\neq i} \widetilde{\Delta}_{i,jk}u_{i,jk}.$$

Similarly,

$$\mathcal{H}^T_{\beta\beta,i} - \mathcal{H}^T_{\beta A,i}(\mathcal{H}^T_{AA,i})^{-1}\mathcal{H}^{T\prime}_{\beta A,i} = \sum_{j<k;\, j,k\neq i} v_{i,jk}\widetilde{\Delta}_{i,jk}\widetilde{\Delta}'_{i,jk}.$$

Accordingly, define $\Gamma_{T,n} \equiv \sum_i \sum_{j<k;\, j,k\neq i} v_{i,jk}\widetilde{\Delta}_{i,jk}\widetilde{\Delta}'_{i,jk}$. The leading part of the profiled β score is therefore $\sum_i \sum_{j<k;\, j,k\neq i} \widetilde{\Delta}_{i,jk}u_{i,jk}$, while the remaining terms are collected in r_T.

The dyadic likelihood admits the same type of residualization. Recall $e_{ij} \equiv D_{ij} - p_{ij,0}$ and $\nu_{ij,0} \equiv p_{ij,0}(1-p_{ij,0})$. At the true parameters, dyad (i,j) contributes $W_{ij}e_{ij}$ to the β score and e_{ij} to each of the A_i and A_j scores. Eliminating the additive agent-effect directions from the linearized β score is equivalent to the weighted projection

$$\{\alpha_{i,D}\}_{i=1}^n \in \arg\min_{\alpha_1,\ldots,\alpha_n\in\mathbb{R}^{d_\beta}} \sum_{i<j} \nu_{ij,0}\left\|W_{ij} - \alpha_i - \alpha_j\right\|^2.$$

Define $\widetilde{W}_{ij,D} \equiv W_{ij} - \alpha_{i,D} - \alpha_{j,D}$. The profiled score and conditional expected negative Hessian for β are $\sum_{i<j}\widetilde{W}_{ij,D}e_{ij}$ and $\Gamma_{D,n} \equiv \sum_{i<j}\nu_{ij,0}\widetilde{W}_{ij,D}\widetilde{W}'_{ij,D}$, respectively.

Lemma A.2 (Profiled curvature and score orders) *Suppose Assumption 1, Assumption 4(i)–(ii) hold, and $Q_\perp$ is positive definite. Then there exists $c_\Gamma > 0$ such that, with probability approaching one, $\lambda_{\min}(\Gamma_{T,n}) \geq c_\Gamma n^3$ and $\lambda_{\min}(\Gamma_{D,n}) \geq c_\Gamma n^2$. Moreover, uniformly in i, $K_{i,T} \asymp n^2$ and $\sum_{j\neq i}\nu_{ij,0} \asymp n$. The corresponding score orders are*

$$\left\|\sum_i \sum_{j<k;\, j,k\neq i} \widetilde{\Delta}_{i,jk}u_{i,jk}\right\| = O_p(n^2), \qquad \max_i \left|\sum_{j<k;\, j,k\neq i} u_{i,jk}\right| = O_p\left(n^{3/2}\sqrt{\log n}\right),$$
$$\left\|\sum_{i<j}\widetilde{W}_{ij,D}e_{ij}\right\| = O_p(n), \qquad \max_i \left|\sum_{j\neq i} e_{ij}\right| = O_p(\sqrt{n\log n}).$$

Proof. By Assumption 4(i)–(ii), the logit indices are uniformly bounded. Hence there exist constants $0 < \underline{v} < \overline{v} < \infty$ and $0 < \underline{\nu} < \overline{\nu} < \infty$ such that, uniformly over

all relevant indices,

$$\underline{v} \leq v_{i,jk} \leq \overline{v}, \qquad \underline{\nu} \leq \nu_{ij,0} \leq \overline{\nu}.$$

Since each focal agent has $\binom{n-1}{2}$ comparison pairs, it follows immediately that $K_{i,T} \asymp n^2$ uniformly in i. Similarly, $\sum_{j \neq i} \nu_{ij,0} \asymp n$ uniformly in i.

We next establish the curvature bounds. Define $m(x) \equiv E[W_{12} \mid X_1 = x]$, $\overline{W} \equiv E[W_{12}]$, and $f(x) \equiv m(x) - \overline{W}$. Then

$$W_{ij} = \overline{W} + f(X_i) + f(X_j) + W_{ij}^{\perp},$$

where $E[W_{ij}^{\perp} \mid X_i] = E[W_{ij}^{\perp} \mid X_j] = 0$.

For any unit vector $\vartheta \in \mathbb{R}^{d_\beta}$, the definition of the weighted additive projection implies

$$\vartheta' \Gamma_{D,n} \vartheta = \min_{\{\xi_i\}} \sum_{i<j} \nu_{ij,0} \left(\vartheta' W_{ij} - \xi_i - \xi_j\right)^2 \geq \underline{\nu} \min_{\{\xi_i\}} \sum_{i<j} \left(\vartheta' W_{ij} - \xi_i - \xi_j\right)^2.$$

Because $\overline{W} + f(X_i) + f(X_j)$ is additive in i and j, the last minimum equals

$$\min_{\{\zeta_i\}} \sum_{i<j} \left(\vartheta' W_{ij}^{\perp} - \zeta_i - \zeta_j\right)^2.$$

Let $R_{i,n}^D(\vartheta) \equiv \sum_{j \neq i} \vartheta' W_{ij}^{\perp}$. The normal equations for the additive projection imply that its fitted sum of squares is bounded by a constant multiple of $n^{-1} \sum_i \{R_{i,n}^D(\vartheta)\}^2$. Conditional degeneracy gives $E[\{R_{i,n}^D(\vartheta)\}^2] = O(n)$ uniformly over $\|\vartheta\| = 1$, and

$$\frac{1}{n^2} \left[\sum_{i<j} (\vartheta' W_{ij}^{\perp})^2 - \min_{\{\zeta_i\}} \sum_{i<j} (\vartheta' W_{ij}^{\perp} - \zeta_i - \zeta_j)^2 \right] = o_p(1).$$

On the other hand, the dyadic law of large numbers gives

$$\frac{1}{n^2} \sum_{i<j} (\vartheta' W_{ij}^{\perp})^2 \to_p \frac{1}{2} \vartheta' Q_{\perp} \vartheta.$$

Consequently,

$$\frac{1}{n^2} \min_{\{\xi_i\}} \sum_{i<j} (\vartheta' W_{ij} - \xi_i - \xi_j)^2 \to_p \frac{1}{2} \vartheta' Q_{\perp} \vartheta.$$

Since d_β is fixed, the convergence holds uniformly over the unit sphere. Because $Q_{\perp}$

is positive definite, this yields $\lambda_{\min}(\Gamma_{D,n}) \geq c_\Gamma n^2$ with probability approaching one.

For the triad criterion,

$$\Delta_{i,jk} = \overline{W} + 2f(X_i) + R_{i,jk}, \qquad R_{i,jk} \equiv W_{ij}^{\perp} + W_{ik}^{\perp} - W_{jk}^{\perp}.$$

Conditional degeneracy implies $E[R_{i,jk} \mid X_i] = 0$. Moreover, the cross-products among $W_{ij}^{\perp}, W_{ik}^{\perp}$, and $W_{jk}^{\perp}$ have zero expectation, so

$$E[R_{1,23}R_{1,23}'] = 3Q_\perp.$$

For any unit vector ϑ, the definition of $\overline{\Delta}_i$ gives the exact representation

$$\vartheta'\Gamma_{T,n}\vartheta = \sum_i \min_{\xi\in\mathbb{R}} \sum_{j<k;\, j,k\neq i} v_{i,jk} \left(\vartheta'\Delta_{i,jk} - \xi\right)^2 \geq \underline{v} \sum_i \min_{\xi\in\mathbb{R}} \sum_{j<k;\, j,k\neq i} \left(\vartheta'\Delta_{i,jk} - \xi\right)^2.$$

Since $\overline{W} + 2f(X_i)$ is constant across the comparison pairs for a given focal agent, the last expression equals

$$\underline{v} \sum_i \min_{\xi\in\mathbb{R}} \sum_{j<k;\, j,k\neq i} \left(\vartheta' R_{i,jk} - \xi\right)^2.$$

The inner minimum is $\sum_{j<k;\, j,k\neq i} \left\{\vartheta'(R_{i,jk} - \overline{R}_{i,n})\right\}^2$, where $\overline{R}_{i,n}$ is the within-i average of $R_{i,jk}$. Hence

$$\frac{1}{n^3}\sum_i \min_{\xi} \sum_{j<k;\, j,k\neq i} (\vartheta' R_{i,jk} - \xi)^2 = \frac{1}{n^3}\sum_i \sum_{j<k;\, j,k\neq i} (\vartheta' R_{i,jk})^2 - \frac{1}{n^3}\sum_i \binom{n-1}{2} (\vartheta'\overline{R}_{i,n})^2.$$

The first term converges by an order-three law of large numbers to

$$\frac{1}{2}E[(\vartheta' R_{1,23})^2] = \frac{3}{2}\vartheta' Q_\perp \vartheta.$$

The second term is $o_p(1)$ because conditional degeneracy implies $\overline{R}_{i,n} = O_p(n^{-1/2})$ in mean square, uniformly after averaging over i. Therefore

$$\frac{1}{n^3}\sum_i \min_{\xi} \sum_{j<k;\, j,k\neq i} (\vartheta'\Delta_{i,jk} - \xi)^2 \to_p \frac{3}{2}\vartheta' Q_\perp \vartheta.$$

Again, fixed d_β makes the convergence uniform over the unit sphere, and positive definiteness of $Q_\perp$ yields $\lambda_{\min}(\Gamma_{T,n}) \geq c_\Gamma n^3$ with probability approaching one.

It remains to establish the score orders. Conditional on $\mathcal{F}_n$, the dyadic shocks are independent, while each $u_{i,jk}$ is centered and uniformly bounded. For a fixed focal agent i, changing one dyadic shock associated with (i,j) can affect at most $n-2$ terms in $\sum_{j<k;\, j,k\neq i} u_{i,jk}$, whereas changing a shock associated with a dyad between two comparison agents affects only one term. Thus the sum of squared bounded-difference constants is $O(n^3)$. McDiarmid's inequality gives

$$\Pr\left(\left|\sum_{j<k;\, j,k\neq i} u_{i,jk}\right| > x \;\middle|\; \mathcal{F}_n\right) \leq 2\exp\left(-\frac{cx^2}{n^3}\right).$$

A union bound over i then yields

$$\max_i \left|\sum_{j<k;\, j,k\neq i} u_{i,jk}\right| = O_p\left(n^{3/2}\sqrt{\log n}\right).$$

Because W_{ij} is uniformly bounded, so are $\Delta_{i,jk}$, $\overline{\Delta}_i$, and $\widetilde{\Delta}_{i,jk}$. A change in a single dyadic shock affects only $O(n)$ summands in

$$\sum_i \sum_{j<k;\, j,k\neq i} \widetilde{\Delta}_{i,jk} u_{i,jk}.$$

There are $O(n^2)$ independent dyadic shocks, so the sum of squared bounded-difference constants is $O(n^4)$. Applying McDiarmid's inequality coordinatewise and using fixed d_β gives

$$\left\|\sum_i \sum_{j<k;\, j,k\neq i} \widetilde{\Delta}_{i,jk} u_{i,jk}\right\| = O_p(n^2).$$

Finally, conditional on $\mathcal{F}_n$, the dyadic residuals e_{ij} are independent, centered, and satisfy $E[e_{ij}^2 \mid \mathcal{F}_n] = \nu_{ij,0}$. Since $\widetilde{W}_{ij,D}$ is $\mathcal{F}_n$-measurable,

$$E\left[\left(\sum_{i<j} \widetilde{W}_{ij,D} e_{ij}\right)\left(\sum_{i<j} \widetilde{W}_{ij,D} e_{ij}\right)' \;\middle|\; \mathcal{F}_n\right] = \sum_{i<j} \nu_{ij,0} \widetilde{W}_{ij,D} \widetilde{W}_{ij,D}' = \Gamma_{D,n}.$$

Moreover, $\operatorname{tr}(\Gamma_{D,n}) = O_p(n^2)$, so the profiled dyadic score is $O_p(n)$. For each i, $\sum_{j\neq i} e_{ij}$ is a sum of $n-1$ conditionally independent, centered, uniformly bounded variables with conditional variance of order n. Bernstein's inequality followed by a

union bound therefore gives $\max_i \left|\sum_{j\neq i} e_{ij}\right| = O_p(\sqrt{n\log n})$. ■

The preceding lemma provides the two ingredients needed for localization: the curvature of the profiled criteria and the stochastic orders of their scores. The next lemma converts these bounds into convergence rates for the first-stage estimates and for the unpenalized dyadic joint maximum likelihood estimator. The latter is introduced only as a benchmark for the stabilized dyadic refinement.

Lemma A.3 (First-stage and JML rates) *Suppose Assumptions 1, 4(i)–(ii) hold, and $Q_\perp$ is positive definite. Then, for $r \in \{T, J\}$, $\widehat{\beta}^r - \beta_0 = O_p(n^{-1})$, and, for every fixed i, $\widehat{A}_i^r - A_i = O_p(n^{-1/2})$. Moreover,*

$$\frac{1}{n}\sum_{i=1}^{n}(\widehat{A}_i^r - A_i)^2 = O_p(n^{-1}), \qquad \max_{1\leq i\leq n}|\widehat{A}_i^r - A_i| = O_p\left(\sqrt{\frac{\log n}{n}}\right).$$

Proof. For $d = (b', a_1, \ldots, a_n)'$, define the scaled norm

$$\|d\|_n^2 \equiv \|b\|^2 + \frac{1}{n}\sum_{i=1}^{n} a_i^2.$$

Equivalently, if $z(\beta, a^n) \equiv (\beta', a^{n\prime}/\sqrt{n})'$, then $\|z(\beta, a^n) - z(\beta_0, A^n)\|$ is the corresponding Euclidean distance.

We first obtain preliminary rates. By Lemma A.2, the profiled curvature for β is of order n^3 for the triad criterion and n^2 for the dyadic criterion. The curvature for each agent effect is of order n^2 and n, respectively. After the rescaling $a^n \mapsto a^n/\sqrt{n}$, the smallest eigenvalues of the full negative Hessians are therefore of order n^3 for $\mathcal{L}_{T,n}$ and n^2 for $\mathcal{L}_{D,n}$.

The score vectors have the corresponding orders. For the triad criterion,

$$E\left[\frac{1}{n}\sum_{i=1}^{n}(U_{A,i}^T)^2 \;\middle|\; \mathcal{F}_n\right] = O(n^3),$$

so that $\|(U_{A,1}^T, \ldots, U_{A,n}^T)'\| = O_p(n^2)$. The β score is $O_p(n^2)$ by the same bounded-difference argument used in Lemma A.2. Hence the score with respect to $z = (\beta', a^{n\prime}/\sqrt{n})'$ is $O_p(n^{5/2})$.

For the dyadic likelihood, let $U_{A,i}^D \equiv \sum_{j\neq i} e_{ij}$. Conditional independence of the

dyadic residuals gives

$$E\left[\frac{1}{n}\sum_{i=1}^{n}(U_{A,i}^{D})^2 \;\middle|\; \mathcal{F}_n\right] = O(n),$$

while $\sum_{i<j} W_{ij}e_{ij} = O_p(n)$. The score with respect to the rescaled parameter z is therefore $O_p(n^{3/2})$.

Bounded regressors, compact parameter spaces, and bounded derivatives of the logit likelihood imply that the realized Hessians inherit these curvature orders uniformly on a sufficiently small neighborhood of the truth. A standard concavity argument therefore yields

$$\|\widehat{\beta}^T - \beta_0\|^2 + \frac{1}{n}\sum_{i=1}^{n}(\widehat{A}_i^T - A_i)^2 = O_p(n^{-1}),$$
$$\|\widehat{\beta}^J - \beta_0\|^2 + \frac{1}{n}\sum_{i=1}^{n}(\widehat{A}_i^J - A_i)^2 = O_p(n^{-1}).$$

We next sharpen the rate for the common parameter. For the first stage, eliminating the agent-effect components from the Taylor expansion of the score equations, as in the construction preceding Lemma A.2, gives

$$0 = \sum_{i}\sum_{j<k;\, j,k\neq i}\widetilde{\Delta}_{i,jk}u_{i,jk} - \Gamma_{T,n}(\widehat{\beta}^T - \beta_0) + \mathcal{R}_{T,n}.$$

The bounded third derivatives of the triad likelihood and the preliminary rates imply

$$\|\mathcal{R}_{T,n}\| \le O_p(n^2) + C\left\{n^3\|\widehat{\beta}^T - \beta_0\|^2 + n^2\sum_{i=1}^{n}(\widehat{A}_i^T - A_i)^2\right\} = O_p(n^2).$$

Here the first term collects the centered-Hessian contribution, whose order follows from the same bounded-difference argument as in Lemma A.2. Since

$$\left\|\sum_{i}\sum_{j<k;\, j,k\neq i}\widetilde{\Delta}_{i,jk}u_{i,jk}\right\| = O_p(n^2)$$

and $\lambda_{\min}(\Gamma_{T,n}) \ge cn^3$ with probability approaching one,

$$\widehat{\beta}^T - \beta_0 = O_p(n^{-1}).$$

The same argument applies to the dyadic likelihood. Profiling out the agent effects gives

$$0 = \sum_{i<j} \widetilde{W}_{ij,D} e_{ij} - \Gamma_{D,n}(\widehat{\beta}^J - \beta_0) + \mathcal{R}_{D,n},$$

where

$$\|\mathcal{R}_{D,n}\| \leq O_p(n) + C\left\{ n^2 \|\widehat{\beta}^J - \beta_0\|^2 + n \sum_{i=1}^{n} (\widehat{A}_i^J - A_i)^2 \right\} = O_p(n).$$

By Lemma A.2, $\| \sum_{i<j} \widetilde{W}_{ij,D} e_{ij} \| = O_p(n)$ and $\lambda_{\min}(\Gamma_{D,n}) \geq cn^2$ with probability approaching one. Hence

$$\widehat{\beta}^J - \beta_0 = O_p(n^{-1}).$$

It remains to sharpen the coordinatewise rates for the agent effects. For the triad criterion, the A_i score is monotone in a_i, with curvature of order $K_{i,T} \asymp n^2$. A mean-value expansion around (β_0, A_i) therefore gives

$$|\widehat{A}_i^T - A_i| \leq C\left\{ n^{-2} |U_{A,i}^T| + \|\widehat{\beta}^T - \beta_0\| \right\}.$$

For every fixed i, $U_{A,i}^T = O_p(n^{3/2})$, whereas Lemma A.2 gives

$$\max_i |U_{A,i}^T| = O_p(n^{3/2}\sqrt{\log n}).$$

Together with $\widehat{\beta}^T - \beta_0 = O_p(n^{-1})$, these bounds imply

$$\widehat{A}_i^T - A_i = O_p(n^{-1/2})$$

for every fixed i, and

$$\max_i |\widehat{A}_i^T - A_i| = O_p\left(\sqrt{\frac{\log n}{n}} \right).$$

The average squared-error rate was already obtained from the preliminary localization.

For the dyadic likelihood, the A_i first-order condition and a mean-value expansion give

$$|\widehat{A}_i^J - A_i| \leq C\left\{ n^{-1} |U_{A,i}^D| + \|\widehat{\beta}^J - \beta_0\| + \frac{1}{n} \sum_{j=1}^{n} |\widehat{A}_j^J - A_j| \right\}.$$

By the preliminary mean-square rate and Cauchy–Schwarz, $\frac{1}{n}\sum_{j=1}^{n}|\widehat{A}_j^J - A_j| = O_p(n^{-1/2})$. For every fixed i, $U_{A,i}^D = O_p(\sqrt{n})$, while Lemma A.2 gives $\max_i |U_{A,i}^D| = O_p(\sqrt{n\log n})$. The stated fixed-agent and maximal rates for $\widehat{A}^{n,J}$ follow. ■

Lemma A.3 provides two reference points for the remaining stages. The first-stage estimates are sufficiently localized to serve as stabilization centers, while the JML estimator provides the unpenalized dyadic benchmark. The next lemma shows that the diminishing stabilization does not alter these rates and that the subsequent bias correction and effect profiling preserve them.

Lemma A.4 (Propagation through Stages 2–4) *Suppose Assumptions 1, 4(i)–(ii) hold, $Q_\perp$ is positive definite, $\delta_n = O(n^{-1/2})$ and $\varepsilon_n = O(n^{-1})$. Then*

$$\|\widehat{\beta}^S - \widehat{\beta}^J\| + \left\{\frac{1}{n}\sum_{i=1}^{n}(\widehat{A}_i^S - \widehat{A}_i^J)^2\right\}^{1/2} = O_p(n^{-1}).$$

Moreover, $\widehat{\beta}^{\mathrm{TALE}} - \widehat{\beta}^S = O_p(n^{-1})$, and, for $r \in \{S, \mathrm{TALE}\}$, $\widehat{A}_i^r - A_i = O_p(n^{-1/2})$ for every fixed i, $\frac{1}{n}\sum_{i=1}^{n}(\widehat{A}_i^r - A_i)^2 = O_p(n^{-1})$ and $\max_i |\widehat{A}_i^r - A_i| = O_p\left(\sqrt{\frac{\log n}{n}}\right)$.

If, in addition, $\delta_n = o(n^{-1/2})$ and $\varepsilon_n = o(n^{-1})$, then $\widehat{\beta}^S - \widehat{\beta}^J = o_p(n^{-1})$, and, for every fixed i, $\widehat{A}_i^S - \widehat{A}_i^J = o_p(n^{-1/2})$.

Proof. We first compare the stabilized dyadic refinement with JML. Define

$$z(\beta, a^n) \equiv (\beta', a^{n\prime}/\sqrt{n})',$$

and write z^T, z^S, and z^J for the corresponding parameter vectors.

For the comparison it is convenient to divide the dyadic criterion by n^2. This rescaling does not change its maximizer. By Lemma A.2 and the boundedness of the logit derivatives, the negative Hessian of $n^{-2}\mathcal{L}_{D,n}$ with respect to z has eigenvalues bounded above and away from zero on the local neighborhood containing z^T and z^J, with probability approaching one.

At iteration R_n, the stabilization weight for agent i is

$$\kappa_{i,R_n} = c_{R_n} H_{i,R_n}^D, \qquad c_{R_n} \le \delta_n.$$

Since $H_{i,R_n}^D \asymp n$ uniformly on the local neighborhood, the gradient of the normalized

penalty with respect to z satisfies

$$\left\| \nabla_z \left[\frac{1}{2n^2} \sum_i \kappa_{i,R_n} (a_i - \widehat{A}_i^T)^2 \right] \right\| \leq C c_{R_n} \|z - z^T\|.$$

The final agent-effect update solves its conditional penalized first-order conditions, and the final β update solves the corresponding unpenalized β first-order condition. Lipschitz continuity of the score, together with the stopping rule, implies

$$\left\| \nabla_z \{n^{-2} \mathcal{L}_{D,n}\}(z^S) \right\| \leq C \left\{ c_{R_n} \|z^S - z^T\| + \varepsilon_n \right\}.$$

Since $\nabla_z \mathcal{L}_{D,n}(z^J) = 0$ and the normalized dyadic criterion is locally strongly concave, the mean-value theorem gives

$$\|z^S - z^J\| \leq C \left\{ c_{R_n} \|z^S - z^T\| + \varepsilon_n \right\}.$$

Using $\|z^S - z^T\| \leq \|z^S - z^J\| + \|z^J - z^T\|$ and Lemma A.3,

$$\|z^J - z^T\| = O_p(n^{-1/2}).$$

Since $c_{R_n} \leq \delta_n \to 0$, the term involving $\|z^S - z^J\|$ can be absorbed into the left-hand side. Hence

$$\|z^S - z^J\| = O_p\left(\delta_n n^{-1/2} + \varepsilon_n\right) = O_p(n^{-1}).$$

By the definition of z, this is exactly

$$\|\widehat{\beta}^S - \widehat{\beta}^J\| + \left\{ \frac{1}{n} \sum_i (\widehat{A}_i^S - \widehat{A}_i^J)^2 \right\}^{1/2} = O_p(n^{-1}).$$

Combining this result with Lemma A.3 immediately gives the average mean-square rate for $\widehat{A}^{n,S}$. Moreover,

$$\max_i |\widehat{A}_i^S - \widehat{A}_i^J| \leq \left\{ \sum_i (\widehat{A}_i^S - \widehat{A}_i^J)^2 \right\}^{1/2} = O_p(n^{-1/2}),$$

so the fixed-agent and maximal rates of the JML benchmark carry over to Stage 2.

Under the stronger conditions $\delta_n = o(n^{-1/2})$ and $\varepsilon_n = o(n^{-1})$, the preceding

argument instead gives $\|z^S - z^J\| = o_p(n^{-1})$. It follows that $\widehat{\beta}^S - \widehat{\beta}^J = o_p(n^{-1})$ and, for every fixed i, $\widehat{A}_i^S - \widehat{A}_i^J = o_p(n^{-1/2})$.

We next consider the Stage-3 bias correction. On an $O_p(n^{-1})$ neighborhood of β_0, the profiled information matrix satisfies $\widehat{\Gamma}_S(b)^{-1} = O_p(n^{-2})$ uniformly. Moreover, $\widehat{H}_{i,S}(b) \asymp n$ and $\widehat{R}_{i,S}(b) = O_p(n)$ uniformly in i. Consequently,

$$\sup_b \|\widehat{b}_{\beta,S}(b)\| = O_p(n^{-1})$$

over this neighborhood.

The fixed-point iteration is initialized at $b^{(0)} = \widehat{\beta}^S$ and satisfies $b^{(r+1)} = \widehat{\beta}^S - \widehat{b}_{\beta,S}(b^{(r)})$. Since $\widehat{\beta}^S - \beta_0 = O_p(n^{-1})$, the preceding bound implies inductively that every iterate remains in an $O_p(n^{-1})$ neighborhood of both $\widehat{\beta}^S$ and β_0. Because the returned $\widehat{\beta}^{\mathrm{TALE}}$ is one of these iterates,

$$\widehat{\beta}^{\mathrm{TALE}} - \widehat{\beta}^S = O_p(n^{-1}).$$

The tolerance used to stop the fixed-point iteration is needed for the later first-order expansion of the bias correction, but not for this rate statement.

It remains to consider the final bias-reduced effect profile. For $c_{ij} \equiv e_i + e_j$, where e_i is the ith canonical basis vector,

$$J_A(\beta, a^n) = \sum_{i<j} \nu_{ij}(\beta, a^n) c_{ij} c_{ij}'.$$

Uniformly on the local neighborhood, the logit weights are bounded above and away from zero. Since

$$\sum_{i<j} c_{ij} c_{ij}' = (n-2) I_n + \mathbf{1}_n \mathbf{1}_n',$$

the eigenvalues of $J_A(\beta, a^n)$ are of order n and $\|J_A(\beta, a^n)^{-1}\| = O(n^{-1})$.

For the ith coordinate, the Jeffreys–Firth contribution to the score is

$$\begin{aligned} \frac{1}{2} \frac{\partial}{\partial a_i} \log\det J_A(\beta, a^n) &= \frac{1}{2} \operatorname{tr}\left\{ J_A(\beta, a^n)^{-1} \frac{\partial J_A(\beta, a^n)}{\partial a_i} \right\} \\ &= \frac{1}{2} \sum_{j \neq i} \frac{\partial \nu_{ij}(\beta, a^n)}{\partial a_i} c_{ij}' J_A(\beta, a^n)^{-1} c_{ij}. \end{aligned}$$

The derivative of the logit weight is uniformly bounded, while

$$c_{ij}' J_A(\beta, a^n)^{-1} c_{ij} = O(n^{-1}).$$

There are $n-1$ terms in the sum, so the Jeffreys–Firth score adjustment is $O(1)$ uniformly in i.

Let $\overline{A}^n$ denote the unadjusted fixed-β dyadic profile evaluated at $\widehat{\beta}^{\mathrm{TALE}}$. Since $\widehat{\beta}^{\mathrm{TALE}} - \beta_0 = O_p(n^{-1})$, the same score and curvature argument used for the JML agent effects gives

$$\frac{1}{n}\sum_i (\overline{A}_i - A_i)^2 = O_p(n^{-1}), \qquad \overline{A}_i - A_i = O_p(n^{-1/2})$$

for every fixed i, and

$$\max_i |\overline{A}_i - A_i| = O_p\left(\sqrt{\frac{\log n}{n}}\right).$$

The vector of Jeffreys–Firth score adjustments has Euclidean norm $O_p(\sqrt{n})$, whereas the smallest eigenvalue of the fixed-β agent-effect information matrix is of order n. A mean-value expansion of the adjusted first-order conditions around $\overline{A}^n$ therefore gives

$$\|\widehat{A}^{n,\mathrm{TALE}} - \overline{A}^n\| = O_p(n^{-1/2}).$$

Hence

$$\left\{\frac{1}{n}\sum_i (\widehat{A}_i^{\mathrm{TALE}} - \overline{A}_i)^2\right\}^{1/2} = O_p(n^{-1}),$$

and

$$\max_i |\widehat{A}_i^{\mathrm{TALE}} - \overline{A}_i| = O_p(n^{-1/2}).$$

Combining these bounds with the rates for $\overline{A}^n$ proves the claimed fixed-agent, average, and maximal rates for the final agent-effect estimates. ■

The proof of Proposition 1 is now immediate from the preceding lemmas.

Proof of Proposition 1. By Lemma A.1, $\Pr(\mathcal{I}_{T,n} = \{1, \ldots, n\}) \to 1$, so the fallback definition for agents outside $\mathcal{I}_{T,n}$ is asymptotically irrelevant. On this event, Lemma A.3 gives the stated rates for the first-stage estimates and for the JML benchmark. Lemma A.4 transfers these rates to the stabilized dyadic estimates and then to the final TALE estimates. This proves all claims. ■

A.2 Second-order bias expansions

This subsection derives the second-order bias terms used in Theorem 3. Throughout, work on the event $\mathcal{I}_{T,n} = \{1, \dots, n\}$, whose probability approaches one by Lemma A.1. For the first-stage criterion, let $\theta_0^T \equiv (\beta_0', A^{n\prime})'$ and, for each anchored triad, define $x_{i,jk} \equiv (\Delta_{i,jk}', 2\iota_i')'$, where ι_i is the ith coordinate vector in $\mathbb{R}^n$. Recall $u_{i,jk} = \mathbb{1}(S_{i,jk} = 1) - |S_{i,jk}| q_{i,jk,0}$. Define the first-stage score at the truth by $s_{T,n} \equiv \sum_i \sum_{j<k;\, j,k\neq i} x_{i,jk} u_{i,jk}$.

Let $h_{i,jk,0} \equiv |S_{i,jk}| q_{i,jk,0}(1-q_{i,jk,0})$, and define the realized and conditional expected negative Hessians

$$\mathscr{H}_{T,n} \equiv \sum_i \sum_{j<k;\, j,k\neq i} h_{i,jk,0} x_{i,jk} x_{i,jk}', \qquad H_{T,0} \equiv E[\mathscr{H}_{T,n} \mid \mathcal{F}_n].$$

Since $E[h_{i,jk,0} \mid \mathcal{F}_n] = v_{i,jk}$, this agrees with

$$H_{T,0} = \sum_i \sum_{j<k;\, j,k\neq i} v_{i,jk} x_{i,jk} x_{i,jk}'.$$

Write $C_{T,n} \equiv \mathscr{H}_{T,n} - H_{T,0}$.

For vectors a and b of the same dimension as θ_0^T, define the bilinear action of the third derivative of the first-stage criterion by

$$\mathscr{T}_{T,n}[a,b] \equiv -\sum_i \sum_{j<k;\, j,k\neq i} |S_{i,jk}| q_{i,jk,0}(1-q_{i,jk,0})(1-2q_{i,jk,0}) x_{i,jk}(x_{i,jk}'a)(x_{i,jk}'b).$$

Finally, let $\psi_{T,n} \equiv H_{T,0}^{-1} s_{T,n}$.

Let P_β select the β coordinates from $(\beta', a^{n\prime})'$, and let P_i select the coordinate corresponding to A_i. Define

$$b_{T,n} \equiv P_\beta H_{T,0}^{-1} E\left[-C_{T,n}\psi_{T,n} + \frac{1}{2}\mathscr{T}_{T,n}[\psi_{T,n}, \psi_{T,n}] \;\middle|\; \mathcal{F}_n \right]. \tag{A.1}$$

The first term in (A.1) accounts for the random curvature of the anchored-triad criterion, while the second is the usual nonlinear likelihood term.

For the dyadic likelihood, define

$$H_{i,D,0} \equiv \sum_{j\neq i} \nu_{ij,0}$$

and

$$R_{i,D,0} \equiv -\sum_{j \neq i} \nu_{ij,0}(1 - 2p_{ij,0})\widetilde{W}_{ij,D}.$$

The leading incidental-parameter bias is

$$b_{D,n} \equiv \Gamma_{D,n}^{-1} \frac{1}{2} \sum_{i=1}^{n} \frac{R_{i,D,0}}{H_{i,D,0}}. \tag{A.2}$$

By Lemma A.2, $\Gamma_{D,n}^{-1} = O_p(n^{-2})$, $H_{i,D,0} \asymp n$, and $R_{i,D,0} = O(n)$, so $b_{D,n} = O_p(n^{-1})$. The same order for $b_{T,n}$ follows from the curvature and score bounds in Lemma A.2.

Lemma A.5 (Second-order bias expansions) *Suppose Assumption 1, Assumption 4(i)–(ii) hold and $Q_\perp$ is positive definite. Then*

$$\begin{aligned} \widehat{\beta}^T - \beta_0 - b_{T,n} &= P_\beta H_{T,0}^{-1} s_{T,n} + o_p(n^{-1}), \\ \widehat{A}_i^T - A_i &= P_i H_{T,0}^{-1} s_{T,n} + o_p(n^{-1/2}) \end{aligned}$$

for every fixed i.

Let $s_{D,n} \equiv \sum_{i<j} \widetilde{W}_{ij,D} e_{ij}$. Then $\widehat{\beta}^J - \beta_0 - b_{D,n} = \Gamma_{D,n}^{-1} s_{D,n} + o_p(n^{-1})$. Moreover, if

$$s_{A,n} \equiv \left(\sum_{j \neq 1} e_{1j}, \ldots, \sum_{j \neq n} e_{nj} \right)'$$

and $J_{A,0} \equiv J_A(\beta_0, A^n)$, then, for every fixed i,

$$\widehat{A}_i^J - A_i = \iota_i' J_{A,0}^{-1} s_{A,n} + o_p(n^{-1/2}).$$

Proof. Let $d_T \equiv (\widehat{\beta}^{T\prime} - \beta_0', \widehat{A}^{n,T\prime} - A^{n\prime})'$. A third-order Taylor expansion of the first-stage score around θ_0^T gives

$$0 = s_{T,n} - \mathscr{H}_{T,n} d_T + \frac{1}{2} \mathscr{T}_{T,n}[d_T, d_T] + r_{T,n}^{(3)},$$

where bounded fourth derivatives of the logit likelihood imply that the contribution of $r_{T,n}^{(3)}$ to any fixed β coordinate after premultiplication by $H_{T,0}^{-1}$ is $o_p(n^{-1})$, and its contribution to any fixed agent-effect coordinate is $o_p(n^{-1/2})$.

Since $\mathscr{H}_{T,n} = H_{T,0} + C_{T,n}$, the preceding equation can be written as

$$H_{T,0} d_T = s_{T,n} - C_{T,n} d_T + \frac{1}{2} \mathscr{T}_{T,n}[d_T, d_T] + r^{(3)}_{T,n}.$$

The preliminary rates in Lemma A.3 imply that, inside the two second-order terms, d_T may be replaced by $\psi_{T,n} = H^{-1}_{T,0} s_{T,n}$ at a cost that is $o_p(n^{-1})$ for the β coordinates and $o_p(n^{-1/2})$ for every fixed effect coordinate. Hence

$$d_T = \psi_{T,n} + H^{-1}_{T,0} \left\{ -C_{T,n}\psi_{T,n} + \frac{1}{2} \mathscr{T}_{T,n}[\psi_{T,n}, \psi_{T,n}] \right\} + r_{T,n},$$

where

$$P_\beta r_{T,n} = o_p(n^{-1}), \qquad P_i r_{T,n} = o_p(n^{-1/2})$$

for every fixed i.

The dependency structure of the triad scores and the bounds in Lemma A.2 further imply that the centered part of the second-order term is of smaller order:

$$P_\beta H^{-1}_{T,0} \Bigg[- C_{T,n}\psi_{T,n} + \frac{1}{2} \mathscr{T}_{T,n}[\psi_{T,n}, \psi_{T,n}] \\ - E \left[-C_{T,n}\psi_{T,n} + \frac{1}{2} \mathscr{T}_{T,n}[\psi_{T,n}, \psi_{T,n}] \;\middle|\; \mathcal{F}_n \right] \Bigg] = o_p(n^{-1}).$$

Indeed, each component is a finite sum of products of bounded triad-score contributions. Terms based on disjoint sets of dyadic shocks are conditionally independent, while the number of overlapping terms is lower by one order in n than the number of unrestricted pairs. The curvature bounds in Lemma A.2 then give the stated order after premultiplication by $H^{-1}_{T,0}$. This proves

$$\widehat{\beta}^T - \beta_0 - b_{T,n} = P_\beta H^{-1}_{T,0} s_{T,n} + o_p(n^{-1}).$$

The second-order term in a fixed effect coordinate is $O_p(n^{-1})$, and hence negligible relative to its $n^{-1/2}$ first-order scale, which proves the expansion for $\widehat{A}^T_i$.

We next consider the dyadic likelihood. Let $d^J_\beta \equiv \widehat{\beta}^J - \beta_0$ and $d^J_i \equiv \widehat{A}^J_i - A_i$. The

linearized agent-effect score equations and Lemma A.3 yield

$$d_i^J = H_{i,D,0}^{-1} \sum_{j \neq i} e_{ij} + O_p(n^{-1})$$

for the purpose of evaluating the leading quadratic contribution to the β score. Consequently,

$$E[(d_i^J)^2 \mid \mathcal{F}_n] = H_{i,D,0}^{-1} + O_p(n^{-2}),$$

whereas, for $i \neq j$,

$$E[d_i^J d_j^J \mid \mathcal{F}_n] = O_p(n^{-2}).$$

Expanding the profiled β score to second order gives

$$0 = s_{D,n} - \Gamma_{D,n} d_\beta^J + \frac{1}{2} \sum_{i<j} \left\{ -\nu_{ij,0}(1 - 2p_{ij,0}) \widetilde{W}_{ij,D} \right\} (d_i^J + d_j^J)^2 + r_{D,n}.$$

Terms containing d_β^J in the second-order part, together with the third-order remainder, are $o_p(n)$ before premultiplication by $\Gamma_{D,n}^{-1}$. Since $E[d_i^J d_j^J \mid \mathcal{F}_n] = O_p(n^{-2})$ for $i \neq j$, the aggregate contribution of the cross-products $d_i^J d_j^J$ is also $o_p(n)$.

The conditional expectation of the remaining quadratic term is

$$\frac{1}{2} \sum_{i<j} \left\{ -\nu_{ij,0}(1 - 2p_{ij,0}) \widetilde{W}_{ij,D} \right\} \left(H_{i,D,0}^{-1} + H_{j,D,0}^{-1} \right) = \frac{1}{2} \sum_{i=1}^{n} \frac{R_{i,D,0}}{H_{i,D,0}}.$$

Its centered fluctuation is $o_p(n)$. Therefore

$$0 = s_{D,n} - \Gamma_{D,n}(\widehat{\beta}^J - \beta_0) + \frac{1}{2} \sum_i \frac{R_{i,D,0}}{H_{i,D,0}} + o_p(n).$$

Premultiplying by $\Gamma_{D,n}^{-1}$ and using (A.2) gives

$$\widehat{\beta}^J - \beta_0 - b_{D,n} = \Gamma_{D,n}^{-1} s_{D,n} + o_p(n^{-1}).$$

For the fixed effects, the full vector score expansion gives

$$\widehat{A}^{n,J} - A^n = J_{A,0}^{-1} s_{A,n} + r_{A,n},$$

where every fixed coordinate of $r_{A,n}$ is $o_p(n^{-1/2})$. The result follows. ■

We next define the feasible first-stage bias correction. Let $\widehat{E}_T$ denote expectation under the fitted first-stage model in which, conditional on $(X^n, \widehat{A}^{n,T})$, the dyads D^*_{ij} are independent Bernoulli variables with probabilities

$$\widehat{p}^T_{ij} \equiv \Lambda(W'_{ij}\widehat{\beta}^T + \widehat{A}^T_i + \widehat{A}^T_j).$$

For a generic draw $D^{n,*}$, construct $S^*_{i,jk}$ and define

$$\widehat{q}^T_{i,jk} \equiv \Lambda(2\widehat{A}^T_i + \Delta'_{i,jk}\widehat{\beta}^T), \;\; u^*_{i,jk} \equiv \mathbb{1}(S^*_{i,jk} = 1) - |S^*_{i,jk}|\widehat{q}^T_{i,jk}, \;\; h^*_{i,jk} \equiv |S^*_{i,jk}|\widehat{q}^T_{i,jk}(1 - \widehat{q}^T_{i,jk}).$$

Set

$$s^*_{T,n} \equiv \sum_i \sum_{j<k:\, j,k\neq i} x_{i,jk}u^*_{i,jk}, \quad \mathscr{H}^*_{T,n} \equiv \sum_i \sum_{j<k:\, j,k\neq i} h^*_{i,jk}x_{i,jk}x'_{i,jk}, \quad \overline{H}^*_{T,n} \equiv \widehat{E}_T[\mathscr{H}^*_{T,n}],$$

and $C^*_{T,n} \equiv \mathscr{H}^*_{T,n} - \overline{H}^*_{T,n}$. Also define $\psi^*_{T,n} \equiv (\overline{H}^*_{T,n})^{-1}s^*_{T,n}$. For vectors a and b, let

$$\mathscr{T}^*_{T,n}[a,b] \equiv -\sum_i \sum_{j<k;\, j,k\neq i} |S^*_{i,jk}|\widehat{q}^T_{i,jk}(1 - \widehat{q}^T_{i,jk})(1 - 2\widehat{q}^T_{i,jk})x_{i,jk}(x'_{i,jk}a)(x'_{i,jk}b).$$

The feasible first-stage bias estimate is

$$\widehat{b}_{T,n} \equiv P_\beta(\overline{H}^*_{T,n})^{-1}\widehat{E}_T\left[-C^*_{T,n}\psi^*_{T,n} + \frac{1}{2}\mathscr{T}^*_{T,n}[\psi^*_{T,n}, \psi^*_{T,n}]\right]. \tag{A.3}$$

Because the fitted probabilities converge uniformly on the local neighborhood relevant for the bias functional and all derivatives of the logit likelihood are bounded, the rates in Lemma A.3 imply

$$\widehat{b}_{T,n} - b_{T,n} = o_p(n^{-1}).$$

Likewise, uniformly over $\|b - \beta_0\| = O_p(n^{-1})$, the Stage-3 bias map satisfies

$$\widehat{b}_{\beta,S}(b) = b_{D,n} + o_p(n^{-1}).$$

In particular,

$$\widehat{b}_{\beta,S}(\widehat{\beta}^S) - b_{D,n} = o_p(n^{-1}),$$

and the same relation holds when the map is evaluated at $\widehat{\beta}^{\mathrm{TALE}}$.

A.3 Proof of Theorem 3

The rate results and second-order expansions are now combined with conditional central limit arguments.

Proof of Theorem 3. We begin with the first-stage estimator. By Lemma A.5,

$$\widehat{\beta}^T - \beta_0 - b_{T,n} = P_\beta H_{T,0}^{-1} s_{T,n} + o_p(n^{-1}),$$
$$\widehat{A}_i^T - A_i = P_i H_{T,0}^{-1} s_{T,n} + o_p(n^{-1/2})$$

for every fixed i.

Conditional on $\mathcal{F}_n$, each contribution $x_{i,jk}u_{i,jk}$ is a bounded function of the three dyadic shocks associated with the triad (i,j,k). Two contributions are conditionally independent when their sets of dyadic shocks are disjoint. The resulting dependency graph has $O(n^3)$ vertices and maximal degree $O(n)$.

For a fixed nonzero $c \in \mathbb{R}^{d_\beta}$, write $a_{i,jk,n} \equiv c' P_\beta H_{T,0}^{-1} x_{i,jk}$. The curvature bounds in Lemma A.2 imply that the largest coefficient is negligible relative to the conditional standard deviation of the sum. Together with Assumption 5, the dependency-graph central limit theorem therefore gives

$$\frac{c' P_\beta H_{T,0}^{-1} s_{T,n}}{\sqrt{\mathrm{Var}\left(c' P_\beta H_{T,0}^{-1} s_{T,n} \mid \mathcal{F}_n\right)}} \Rightarrow N(0,1)$$

conditionally on $\mathcal{F}_n$, in probability.

The same argument applies to every fixed effect coordinate. In particular,

$$\frac{P_i H_{T,0}^{-1} s_{T,n}}{\sqrt{\mathrm{Var}\left(P_i H_{T,0}^{-1} s_{T,n} \mid \mathcal{F}_n\right)}} \Rightarrow N(0,1)$$

conditionally on $\mathcal{F}_n$, in probability.

We next verify the first-stage covariance estimator. For an unordered dyad d, define at the truth

$$G_{d,0} \equiv \sum_{(i,j,k):\, d \subset \{ij,ik,jk\}} x_{i,jk} u_{i,jk}.$$

Distinct triad-score contributions can have nonzero conditional covariance only when they share at least one dyadic shock. Hence $\sum_d G_{d,0} G_{d,0}'$ contains every leading covariance contribution from overlapping distinct triads.

The same-triad terms are counted three times rather than once, and distinct triads based on the same three dyads can also be counted more than once. There are only $O(n^3)$ such contributions, whereas the leading order of the β-score covariance is n^4. For any fixed effect coordinate, the corresponding excess terms are of order n^2, whereas the leading score variance is of order n^3. These overcounting terms are therefore asymptotically negligible.

The same dependency-graph counting argument gives concentration of the overlap covariance estimator around its conditional expectation. Together with $\widehat{H}_T - H_{T,0}$ being negligible at the relevant blockwise orders, it yields, for every fixed nonzero c,

$$\frac{c'\widehat{V}_{\beta,T}c}{\mathrm{Var}\left(c'P_\beta H_{T,0}^{-1}s_{T,n} \mid \mathcal{F}_n\right)} \to_p 1,$$

and, for every fixed i,

$$\frac{\widehat{v}_{i,T}}{\mathrm{Var}\left(P_i H_{T,0}^{-1}s_{T,n} \mid \mathcal{F}_n\right)} \to_p 1.$$

Slutsky's theorem proves the stated first-stage normal limits.

We next turn to the dyadic estimator. By Lemma A.5,

$$\widehat{\beta}^J - \beta_0 - b_{D,n} = \Gamma_{D,n}^{-1}s_{D,n} + o_p(n^{-1}).$$

Conditional on $\mathcal{F}_n$, the residuals e_{ij} are independent, centered, and uniformly bounded. Moreover,

$$\mathrm{Var}(s_{D,n} \mid \mathcal{F}_n) = \sum_{i<j} \nu_{ij,0}\widetilde{W}_{ij,D}\widetilde{W}'_{ij,D} = \Gamma_{D,n}.$$

Since $\lambda_{\min}(\Gamma_{D,n}) \geq cn^2$ with probability approaching one, the conditional Lindeberg condition holds. Thus, for every fixed nonzero c,

$$\frac{c'\Gamma_{D,n}^{-1}s_{D,n}}{\sqrt{c'\Gamma_{D,n}^{-1}c}} \Rightarrow N(0,1)$$

conditionally on $\mathcal{F}_n$, in probability.

Under $\delta_n = o(n^{-1/2})$ and $\varepsilon_n = o(n^{-1})$, Lemma A.4 gives

$$\widehat{\beta}^S - \widehat{\beta}^J = o_p(n^{-1}).$$

Therefore

$$\widehat{\beta}^S - \beta_0 - b_{D,n} = \Gamma_{D,n}^{-1} s_{D,n} + o_p(n^{-1}).$$

For Stage 3, the approximate fixed-point condition gives

$$\widehat{\beta}^{\mathrm{TALE}} - \widehat{\beta}^S + \widehat{b}_{\beta,S}(\widehat{\beta}^{\mathrm{TALE}}) = o_p(n^{-1}),$$

because $\zeta_n = o(n^{-1})$. By the uniform bias-map consistency established in Appendix A.2,

$$\widehat{b}_{\beta,S}(\widehat{\beta}^{\mathrm{TALE}}) = b_{D,n} + o_p(n^{-1}).$$

Consequently,

$$\widehat{\beta}^{\mathrm{TALE}} - \beta_0 = \Gamma_{D,n}^{-1} s_{D,n} + o_p(n^{-1}),$$

so the leading incidental-parameter bias has been removed.

We next establish consistency of the dyadic variance estimator. Uniformly over $\|b - \beta_0\| = O_p(n^{-1})$, the convergence rates already established and bounded logit derivatives give $\widehat{\Gamma}_S(b) = \Gamma_{D,n} + o_p(n^2)$. Similarly,

$$\sum_{i<j} \left\{ D_{ij} - \widehat{p}_{ij}^S(b) \right\}^2 \widetilde{W}_{ij}^S(b) \widetilde{W}_{ij}^S(b)' = \Gamma_{D,n} + o_p(n^2).$$

It follows that, for $r \in \{S, \mathrm{TALE}\}$ and every fixed nonzero c,

$$\frac{c' \widehat{V}_{\beta,r} c}{c' \Gamma_{D,n}^{-1} c} \to_p 1.$$

Together with the preceding linear representations, this proves the dyadic-surplus limits for Stages 2 and 3.

It remains to establish the fixed-agent limits for the dyadic stages. Recall

$$s_{A,n} = \left(\sum_{j \neq 1} e_{1j}, \ldots, \sum_{j \neq n} e_{nj} \right)'$$

and $J_{A,0} = J_A(\beta_0, A^n)$. Since $\mathrm{Var}(s_{A,n} \mid \mathcal{F}_n) = J_{A,0}$, for every fixed i, Lemma A.5 gives

$$\widehat{A}_i^J - A_i = \iota_i' J_{A,0}^{-1} s_{A,n} + o_p(n^{-1/2}).$$

Conditional on $\mathcal{F}_n$, the leading term is a linear combination of independent dyadic

residuals. Uniform boundedness of the logit probabilities and $\|J_{A,0}^{-1}\| = O(n^{-1})$ imply the conditional Lindeberg condition. Hence

$$\frac{\iota_i' J_{A,0}^{-1} s_{A,n}}{\sqrt{[J_{A,0}^{-1}]_{ii}}} \Rightarrow N(0,1)$$

conditionally on $\mathcal{F}_n$, in probability.

The stronger Stage-2 rate conditions imply $\widehat{A}_i^S - \widehat{A}_i^J = o_p(n^{-1/2})$ for every fixed i. For the final effect profile, $\widehat{\beta}^{\mathrm{TALE}} - \widehat{\beta}^J = O_p(n^{-1})$, whose effect on a fixed agent-effect estimate is $O_p(n^{-1})$. Moreover, by Lemma A.4, the Jeffreys–Firth score adjustment is $O_p(1)$ while the fixed-effect likelihood curvature is of order n. Therefore, $\widehat{A}_i^{\mathrm{TALE}} - \widehat{A}_i^J = o_p(n^{-1/2})$.

Continuity of J_A, together with the established convergence rates, gives

$$\frac{\widehat{v}_{i,r}}{[J_{A,0}^{-1}]_{ii}} \to_p 1, \qquad r \in \{S, \mathrm{TALE}\}.$$

The fixed-agent limits follow by Slutsky's theorem.

Finally, (A.3) and the bias-map consistency established in Appendix A.2 give

$$\widehat{b}_{T,n} - b_{T,n} = o_p(n^{-1}), \qquad \widehat{b}_{\beta,S}(\widehat{\beta}^S) - b_{D,n} = o_p(n^{-1}).$$

Replacing the theoretical bias terms by their feasible estimates is therefore asymptotically negligible.

All preceding central limit arguments hold conditionally on $\mathcal{F}_n$, in probability. The corresponding unconditional convergences follow by iterated expectations. This completes the proof. ■

A.4 Proof of Corollary 2

The asymptotic distribution of the conditional-mean estimator follows by averaging the final agent-effect estimates.

Proof of Corollary 2. Define $Q_{Z,n} \equiv \frac{1}{n}\sum_{i=1}^n Z_i Z_i'$. Since $A_i = Z_i'\gamma_0 + V_i$ and

$$\widehat{\gamma}^{\mathrm{TALE}} = Q_{Z,n}^{-1} \frac{1}{n} \sum_{i=1}^n Z_i \widehat{A}_i^{\mathrm{TALE}},$$

we have the exact decomposition

$$\widehat{\gamma}^{\mathrm{TALE}} - \gamma_0 = Q_{Z,n}^{-1}\left\{ \frac{1}{n}\sum_{i=1}^{n} Z_i V_i + \frac{1}{n}\sum_{i=1}^{n} Z_i \left(\widehat{A}_i^{\mathrm{TALE}} - A_i \right) \right\}.$$

We first show that the second term is of order n^{-1}. The fixed-effect expansion in the proof of Theorem 3 gives

$$\widehat{A}^{n,\mathrm{TALE}} - A^n = J_{A,0}^{-1} s_{A,n} + \rho_{A,n},$$

where the effects of estimating β, the Jeffreys–Firth adjustment, and the nonlinear likelihood remainder satisfy

$$\frac{1}{n}\sum_{i=1}^{n} Z_i \rho_{A,i,n} = O_p(n^{-1}).$$

To see the order of the leading term, let Z denote the $n \times d_\gamma$ matrix with ith row Z_i'. Conditional on $\mathcal{F}_n$, $\mathrm{Var}(s_{A,n} \mid \mathcal{F}_n) = J_{A,0}$. Therefore

$$\mathrm{Var}\left(\frac{1}{n} Z' J_{A,0}^{-1} s_{A,n} \,\middle|\, \mathcal{F}_n \right) = \frac{1}{n^2} Z' J_{A,0}^{-1} J_{A,0} J_{A,0}^{-1} Z = \frac{1}{n^2} Z' J_{A,0}^{-1} Z.$$

The regressors Z_i are uniformly bounded and $\|J_{A,0}^{-1}\| = O(n^{-1})$. Since $\|Z\|^2 = O_p(n)$, it follows that $Z' J_{A,0}^{-1} Z = O_p(1)$. Hence

$$\frac{1}{n}\sum_{i=1}^{n} Z_i \left(\widehat{A}_i^{\mathrm{TALE}} - A_i \right) = O_p(n^{-1}).$$

Consequently,

$$\sqrt{n}\left(\widehat{\gamma}^{\mathrm{TALE}} - \gamma_0 \right) = Q_{Z,n}^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_i V_i + o_p(1).$$

By the law of large numbers, $Q_{Z,n} \to_p Q_Z$. Because Z_i is a function of X_i and $E[V_i \mid X_i] = 0$, the random vectors $Z_i V_i$ are iid with mean zero and covariance Ω_γ.

The multivariate central limit theorem gives

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_i V_i \Rightarrow N(0, \Omega_\gamma).$$

Slutsky's theorem therefore yields

$$\sqrt{n}\left(\widehat{\gamma}^{\mathrm{TALE}} - \gamma_0\right) \Rightarrow N\left(0, Q_Z^{-1} \Omega_\gamma Q_Z^{-1}\right).$$

Consistency of $\widehat{\gamma}^{\mathrm{TALE}}$ follows immediately.

It remains to establish consistency of the covariance estimator. By definition,

$$\widehat{V}_i = V_i + \left(\widehat{A}_i^{\mathrm{TALE}} - A_i\right) - Z_i' \left(\widehat{\gamma}^{\mathrm{TALE}} - \gamma_0\right).$$

The boundedness of Z_i, the mean-square rate in Proposition 1, and $\widehat{\gamma}^{\mathrm{TALE}} - \gamma_0 = O_p(n^{-1/2})$ imply

$$\frac{1}{n} \sum_{i=1}^{n} Z_i Z_i' \left(\widehat{V}_i - V_i\right)^2 = O_p(n^{-1}).$$

By Cauchy–Schwarz,

$$\frac{1}{n} \sum_{i=1}^{n} Z_i Z_i' V_i \left(\widehat{V}_i - V_i\right) = o_p(1).$$

Therefore

$$\frac{1}{n} \sum_{i=1}^{n} Z_i Z_i' \left(\widehat{V}_i^2 - V_i^2\right) = o_p(1).$$

The law of large numbers then gives

$$\frac{1}{n} \sum_{i=1}^{n} Z_i Z_i' \widehat{V}_i^2 \to_p \Omega_\gamma.$$

Finally,

$$n\widehat{V}_\gamma = Q_{Z,n}^{-1} \left(\frac{1}{n} \sum_{i=1}^{n} Z_i Z_i' \widehat{V}_i^2 \right) Q_{Z,n}^{-1},$$

so $n\widehat{V}_\gamma \to_p Q_Z^{-1} \Omega_\gamma Q_Z^{-1}$. This completes the proof. ■

B Monte Carlo Results

Table B.1: Simulation Designs and Network Summary Statistics

	Symmetric Uncorrelated Agent Effects			Right-Skewed Correlated Agent Effects		
	S.1	S.2	S.3	S.4	S.5	S.6
Panel A: Design parameters						
γ_0	-0.5	-0.75	-1	-0.5	-0.75	-1
γ_1	0	0	0	0.25	0.25	0.25
λ_0	1	1	1	0.25	0.25	0.25
λ_1	1	1	1	0.75	0.75	0.75
Panel B: Network statistics						
Largest Component Fraction	0.994	0.963	0.866	0.982	0.942	0.843
Network Density	0.121	0.078	0.051	0.128	0.085	0.055
Avg. Degree	5.930	3.826	2.484	6.276	4.180	2.691
Std. of Degree	2.710	2.089	1.656	3.292	2.546	1.941
Network Transitivity	0.161	0.103	0.068	0.205	0.142	0.098
Informative Triad Fraction	0.105	0.072	0.048	0.109	0.077	0.051
Tetrad Tetrad Fraction	0.025	0.013	0.006	0.024	0.013	0.006

Notes: Panel A reports the parameter values used in each simulation design. Panel B reports averages across 500 Monte Carlo replications with $n = 50$. Largest Component Fraction is the fraction of agents contained in the largest connected component. Network Density is the fraction of realized links among all unordered dyads. Avg. Degree and Std. of Degree are the mean and standard deviation of agent degree. Network Transitivity is the global clustering coefficient: the number of closed triples divided by the number of connected triples. Informative Triad Fraction is the fraction of anchored triad configurations satisfying $|S_{i,jk}| = 1$. Informative Tetrad Fraction is the fraction of two-link tetrad configurations in which the two links share no agents, as used by the tetrad-logit estimator of Graham (2017).

Table B.2: Monte Carlo Results for (β_0, A^n)

Design	Method	Dyadic-surplus parameter β_0				Realized agent effects A^n				Successful replications
		Bias	SD	RMSE	95% Cov.	Bias	SD	RMSE	95% Cov.	
	Tetrad logit	-0.016	0.162	0.163	0.978					500
	JML	-0.058	0.151	0.161	0.872	-0.158	0.092	0.544	0.979	383
	JML + RBC	-0.004	0.142	0.142	0.909	-0.080	0.084	0.535	0.960	383
S.1	TALE: Stage 1	-0.058	0.164	0.173	0.952	-0.657	0.389	2.515	0.930	500
	TALE: Stage 2	-0.058	0.153	0.163	0.932	-0.188	0.269	1.270	0.952	500
	TALE: Stage 3	-0.007	0.144	0.144	0.940	-0.188	0.269	1.270	0.946	500
	TALE: Stage 4	-0.007	0.144	0.144	0.940	-0.002	0.080	0.492	0.958	500
	Tetrad logit	-0.032	0.205	0.207	0.984					500
	JML	-0.076	0.195	0.208	0.795	-0.199	0.099	0.605	0.988	112
	JML + RBC	-0.018	0.182	0.182	0.857	-0.087	0.089	0.588	0.969	112
S.2	TALE: Stage 1	-0.076	0.213	0.226	0.976	-1.916	0.710	4.717	0.832	500
	TALE: Stage 2	-0.077	0.190	0.205	0.946	-0.825	0.620	3.788	0.934	500
	TALE: Stage 3	-0.020	0.179	0.180	0.936	-0.825	0.620	3.788	0.928	500
	TALE: Stage 4	-0.020	0.179	0.180	0.936	-0.004	0.090	0.587	0.960	500
	Tetrad logit	-0.047	0.256	0.260	0.988					500
	JML	0.016	0.186	0.177	0.800	-0.249	0.128	0.629	0.992	10
	JML + RBC	0.070	0.174	0.180	0.800	-0.097	0.104	0.594	0.972	10
S.3	TALE: Stage 1	-0.088	0.270	0.284	0.972	-3.988	1.018	6.908	0.647	500
	TALE: Stage 2	-0.072	0.226	0.237	0.938	-2.376	1.135	7.208	0.862	500
	TALE: Stage 3	-0.008	0.211	0.211	0.918	-2.376	1.135	7.208	0.856	500
	TALE: Stage 4	-0.008	0.211	0.211	0.918	0.002	0.108	0.684	0.950	500
	Tetrad logit	-0.048	0.153	0.160	0.992					500
	JML	-0.059	0.130	0.142	0.892	-0.138	0.084	0.525	0.979	223
	JML + RBC	-0.004	0.122	0.122	0.951	-0.062	0.077	0.518	0.959	223
S.4	TALE: Stage 1	-0.081	0.161	0.180	0.956	-0.956	0.498	3.214	0.900	500
	TALE: Stage 2	-0.078	0.145	0.164	0.940	-0.422	0.453	2.445	0.938	500
	TALE: Stage 3	-0.025	0.136	0.138	0.948	-0.422	0.453	2.445	0.933	500
	TALE: Stage 4	-0.025	0.136	0.138	0.948	0.009	0.080	0.503	0.956	500
	Tetrad logit	-0.034	0.186	0.189	0.988					500
	JML	-0.060	0.157	0.167	0.887	-0.167	0.099	0.590	0.985	53
	JML + RBC	-0.003	0.147	0.146	0.943	-0.061	0.089	0.579	0.958	53
S.5	TALE: Stage 1	-0.055	0.198	0.205	0.980	-2.113	0.738	4.941	0.809	500
	TALE: Stage 2	-0.053	0.177	0.185	0.948	-1.189	0.795	4.802	0.912	500
	TALE: Stage 3	0.003	0.167	0.166	0.944	-1.189	0.795	4.802	0.907	500
	TALE: Stage 4	0.003	0.167	0.166	0.944	0.002	0.095	0.589	0.956	500
	Tetrad logit	-0.069	0.242	0.251	0.990					500
	JML	0.014	0.180	0.162	0.800	-0.225	0.129	0.619	0.992	5
	JML + RBC	0.067	0.169	0.165	0.800	-0.076	0.116	0.590	0.976	5
S.6	TALE: Stage 1	-0.096	0.246	0.264	0.982	-3.991	0.974	6.860	0.640	500
	TALE: Stage 2	-0.078	0.219	0.232	0.938	-2.680	1.201	7.673	0.846	500
	TALE: Stage 3	-0.013	0.203	0.203	0.952	-2.680	1.201	7.673	0.839	500
	TALE: Stage 4	-0.013	0.203	0.203	0.952	0.026	0.104	0.668	0.949	500

Notes: Entries, including 95% coverage rates, are based on successful replications among 500 Monte Carlo replications with $n = 50$, where success means that the computation completed and returned estimates. JML denotes joint maximum likelihood; JML + RBC is the residualized bias-correction benchmark. TALE Stages 1–4 are triad anchoring, stabilized dyadic refinement, bias correction for β, and bias-reduced effect profiling, respectively. Stage 3 retains the Stage-2 A^n, while Stage 4 retains the Stage-3 β. Bias, SD, RMSE, and 95% Cov. denote Monte Carlo mean bias, standard deviation, root mean squared error, and coverage. For A^n, Bias and SD use the average error $n^{-1}\sum_i(\widehat{A}_i - A_i)$, RMSE uses $\{n^{-1}\sum_i(\widehat{A}_i - A_i)^2\}^{1/2}$, and coverage is averaged across agents.

Table B.3: Monte Carlo Results for (γ_0, γ_1)

	γ_0				γ_1			
Design	Bias	SD	RMSE	95% Cov.	Bias	SD	RMSE	95% Cov.
S.1	-0.001	0.092	0.092	0.966	-0.002	0.069	0.069	0.980
S.2	-0.007	0.100	0.101	0.984	0.004	0.074	0.074	0.984
S.3	0.005	0.118	0.118	0.962	0.004	0.077	0.077	0.992
S.4	0.010	0.093	0.093	0.972	0.000	0.068	0.068	0.986
S.5	0.001	0.108	0.108	0.968	-0.017	0.071	0.073	0.982
S.6	0.026	0.113	0.116	0.964	-0.034	0.076	0.083	0.978

Notes: Entries, including 95% coverage rates, are based on successful replications among 500 Monte Carlo replications with $n = 50$. The estimates of (γ_0, γ_1) are obtained by projecting the final TALE agent-effect estimates $\widehat{A}^{n,\text{TALE}}$ on $(1, X_i)$. Bias, SD, RMSE, and 95% Cov. denote Monte Carlo mean bias, standard deviation, root mean squared error, and coverage.

Table B.4: Simulation Designs and Network Summary Statistics

	Symmetric Uncorrelated Agent Effects			Right-Skewed Correlated Agent Effects		
	S.1	S.2	S.3	S.4	S.5	S.6
Panel A: Design parameters						
γ_0	-0.5	-0.75	-1	-0.5	-0.75	-1
γ_1	0	0	0	0.25	0.25	0.25
λ_0	1	1	1	0.25	0.25	0.25
λ_1	1	1	1	0.75	0.75	0.75
Panel B: Network statistics						
Largest Component Fraction	1.000	1.000	0.997	1.000	0.998	0.988
Network Density	0.120	0.079	0.051	0.127	0.086	0.055
Avg. Degree	17.934	11.769	7.537	18.932	12.768	8.250
Std. of Degree	6.220	4.651	3.466	8.463	6.519	4.733
Network Transitivity	0.161	0.109	0.072	0.205	0.152	0.106
Informative Triad Fraction	0.105	0.072	0.048	0.108	0.077	0.052
Tetrad Tetrad Fraction	0.024	0.013	0.006	0.024	0.013	0.007

Notes: Panel A reports the parameter values used in each simulation design. Panel B reports averages across 500 Monte Carlo replications with $n = 150$. Largest Component Fraction is the fraction of agents contained in the largest connected component. Network Density is the fraction of realized links among all unordered dyads. Avg. Degree and Std. of Degree are the mean and standard deviation of agent degree. Network Transitivity is the global clustering coefficient: the number of closed triples divided by the number of connected triples. Informative Triad Fraction is the fraction of anchored triad configurations satisfying $|S_{i,jk}| = 1$. Informative Tetrad Fraction is the fraction of two-link tetrad configurations in which the two links share no agents, as used by the tetrad-logit estimator of Graham (2017).

Table B.5: Monte Carlo Results for (β_0, A^n)

Design	Method	Dyadic-surplus parameter β_0				Realized agent effects A^n				Successful replications
		Bias	SD	RMSE	95% Cov.	Bias	SD	RMSE	95% Cov.	
	Tetrad logit	-0.002	0.050	0.050	0.958					500
	JML	-0.015	0.046	0.048	0.914	-0.057	0.026	0.288	0.962	500
	JML + RBC	0.000	0.045	0.045	0.922	-0.027	0.025	0.285	0.954	500
S.1	TALE: Stage 1	-0.012	0.048	0.050	0.946	-0.051	0.028	0.305	0.956	500
	TALE: Stage 2	-0.015	0.046	0.048	0.932	-0.021	0.025	0.285	0.954	500
	TALE: Stage 3	0.001	0.045	0.045	0.934	-0.021	0.025	0.285	0.951	500
	TALE: Stage 4	0.001	0.045	0.045	0.934	-0.001	0.024	0.273	0.955	500
	Tetrad logit	-0.005	0.060	0.060	0.970					500
	JML	-0.017	0.056	0.058	0.903	-0.092	0.033	0.362	0.968	483
	JML + RBC	-0.001	0.055	0.055	0.936	-0.047	0.032	0.355	0.956	483
S.2	TALE: Stage 1	-0.015	0.059	0.061	0.958	-0.112	0.058	0.571	0.958	500
	TALE: Stage 2	-0.018	0.055	0.058	0.944	-0.046	0.043	0.412	0.955	500
	TALE: Stage 3	-0.002	0.054	0.054	0.956	-0.046	0.043	0.412	0.952	500
	TALE: Stage 4	-0.002	0.054	0.054	0.956	-0.002	0.030	0.332	0.954	500
	Tetrad logit	-0.009	0.075	0.076	0.968					500
	JML	-0.017	0.071	0.072	0.910	-0.142	0.041	0.462	0.980	334
	JML + RBC	0.001	0.069	0.069	0.907	-0.073	0.038	0.448	0.962	334
S.3	TALE: Stage 1	-0.019	0.074	0.076	0.962	-0.375	0.154	1.753	0.952	500
	TALE: Stage 2	-0.020	0.071	0.074	0.940	-0.132	0.110	1.026	0.958	500
	TALE: Stage 3	-0.003	0.070	0.070	0.944	-0.132	0.110	1.026	0.956	500
	TALE: Stage 4	-0.003	0.070	0.070	0.944	-0.003	0.036	0.415	0.953	500
	Tetrad logit	-0.017	0.050	0.053	0.956					500
	JML	-0.016	0.046	0.048	0.914	-0.059	0.028	0.302	0.961	479
	JML + RBC	0.000	0.045	0.045	0.952	-0.029	0.027	0.299	0.953	479
S.4	TALE: Stage 1	-0.012	0.048	0.049	0.956	-0.081	0.060	0.540	0.954	500
	TALE: Stage 2	-0.015	0.045	0.048	0.946	-0.030	0.044	0.372	0.952	500
	TALE: Stage 3	0.001	0.044	0.044	0.962	-0.030	0.044	0.372	0.950	500
	TALE: Stage 4	0.001	0.044	0.044	0.962	0.000	0.026	0.284	0.952	500
	Tetrad logit	-0.029	0.062	0.068	0.950					500
	JML	-0.015	0.057	0.059	0.893	-0.092	0.035	0.377	0.967	347
	JML + RBC	0.002	0.056	0.056	0.919	-0.048	0.033	0.370	0.954	347
S.5	TALE: Stage 1	-0.016	0.063	0.065	0.950	-0.240	0.117	1.360	0.947	500
	TALE: Stage 2	-0.018	0.058	0.060	0.930	-0.098	0.100	0.913	0.951	500
	TALE: Stage 3	-0.001	0.057	0.056	0.940	-0.098	0.100	0.913	0.949	500
	TALE: Stage 4	-0.001	0.057	0.056	0.940	-0.001	0.031	0.349	0.951	500
	Tetrad logit	-0.031	0.078	0.084	0.962					500
	JML	-0.004	0.065	0.065	0.875	-0.141	0.040	0.463	0.975	88
	JML + RBC	0.014	0.064	0.065	0.909	-0.075	0.037	0.449	0.958	88
S.6	TALE: Stage 1	-0.018	0.076	0.078	0.948	-0.657	0.219	2.598	0.924	500
	TALE: Stage 2	-0.018	0.070	0.073	0.926	-0.323	0.202	2.307	0.947	500
	TALE: Stage 3	-0.001	0.069	0.069	0.932	-0.323	0.202	2.307	0.945	500
	TALE: Stage 4	-0.001	0.069	0.069	0.932	-0.002	0.037	0.429	0.951	500

Notes: Entries, including 95% coverage rates, are based on successful replications among 500 Monte Carlo replications with $n = 150$, where success means that the computation completed and returned estimates. JML denotes joint maximum likelihood; JML + RBC is the residualized bias-correction benchmark. TALE Stages 1–4 are triad anchoring, stabilized dyadic refinement, bias correction for β, and bias-reduced effect profiling, respectively. Stage 3 retains the Stage-2 A^n, while Stage 4 retains the Stage-3 β. Bias, SD, RMSE, and 95% Cov. denote Monte Carlo mean bias, standard deviation, root mean squared error, and coverage. For A^n, Bias and SD use the average error $n^{-1}\sum_i(\widehat{A}_i - A_i)$, RMSE uses $\{n^{-1}\sum_i(\widehat{A}_i - A_i)^2\}^{1/2}$, and coverage is averaged across agents.

Table B.6: Monte Carlo Results for (γ_0, γ_1)

	γ_0				γ_1			
Design	Bias	SD	RMSE	95% Cov.	Bias	SD	RMSE	95% Cov.
S.1	0.000	0.035	0.035	0.956	0.000	0.026	0.026	0.972
S.2	-0.001	0.038	0.038	0.966	0.000	0.030	0.030	0.964
S.3	-0.002	0.045	0.045	0.958	0.002	0.034	0.034	0.978
S.4	0.000	0.038	0.038	0.952	-0.002	0.029	0.030	0.962
S.5	0.000	0.040	0.040	0.966	0.001	0.033	0.033	0.962
S.6	-0.001	0.042	0.042	0.964	0.002	0.035	0.035	0.968

Notes: Entries, including 95% coverage rates, are based on successful replications among 500 Monte Carlo replications with $n = 150$. The estimates of (γ_0, γ_1) are obtained by projecting the final TALE agent-effect estimates $\widehat{A}^{n,\mathrm{TALE}}$ on $(1, X_i)$. Bias, SD, RMSE, and 95% Cov. denote Monte Carlo mean bias, standard deviation, root mean squared error, and coverage.