Biometrika (2022), **00**, 0, *pp.* 1–8 https://doi.org/10.1093/biomet/asac031

Regression of exchangeable relational arrays

BY F. W. MARRS

Statistical Sciences, Los Alamos National Laboratory, P.O. Box 1663, Los Alamos, New Mexico 87545, U.S.A.
fmarrs3@lanl.gov

B. K. FOSDICK

Department of Statistics, Colorado State University, 102 Statistics Building, Fort Collins, Colorado 80523, U.S.A.
bailey.fosdick@coclostate.edu

AND T. H. MCCORMICK

Department of Statistics, University of Washington, Box 354322, Seattle, Washington 98195, U.S.A.
tylermc@uw.edu

SUMMARY

Relational arrays represent measures of association between pairs of actors, often in varied contexts or over time. Trade flows between countries, financial transactions between individuals, contact frequencies between school children in classrooms and dynamic protein-protein interactions are all examples of relational arrays. Elements of a relational array are often modelled as a linear function of observable covariates. Uncertainty estimates for regression coefficient estimators, and ideally the coefficient estimators themselves, must account for dependence between elements of the array, e.g., relations involving the same actor. Existing estimators of standard errors that recognize such relational dependence rely on estimating extremely complex, heterogeneous structure across actors. This paper develops a new class of parsimonious coefficient and standard error estimators for regressions of relational arrays. We leverage an exchangeability assumption to derive standard error estimators that pool information across actors, and are substantially more accurate than existing estimators in a variety of settings. This exchangeability assumption is pervasive in network and array models in the statistics literature, but not previously considered when adjusting for dependence in a regression setting with relational data. We demonstrate improvements in inference theoretically, via a simulation study, and by analysis of a dataset involving international trade.

Some key words: Array data; Dependent data; Generalized least squares; Weighted network.

1. INTRODUCTION

Entries in relational arrays quantify pairwise interactions between actors that may be of multiple types or observed over time. Examples include annual flows of migrants between countries (Aleskerov et al., 2017) and interactions between students over the course of a semester (Han et al., 2016). In economics, relational arrays are used to describe monetary transfers between individuals as part of informal insurance markets (see Bardham, 1984; Foster & Rosenzweig, 2001; Fafchamps, 2006; Attanasio et al., 2012; Banerjee et al., 2013). Other examples of data that are naturally represented as relational arrays include gene expressions (Zhang & Horvath, 2005) and international relations (Fagiolo et al., 2008).

A relational array $Y = (y_{ijr})$, where $i,j = 1,\dots,n, i \neq j, r = 1,\dots,R$, is composed of a series of R matrices of size $(n \times n)$, each of which describes the directed pairwise relationships among n actors of type r, e.g., time period r or relation context r. The diagonal elements of each matrix, for example y_{iir}, are assumed to be undefined, as we do not consider, e.g., international relations of a country with itself. The relationship from actor i to actor j may differ than that from j to i, such that $y_{ijr} \neq y_{jir}$ in general; however, the methods we propose extend to the symmetric relation case, see the Supplementary Material.

The primary goal in our setting is inference for linear regressions, exploring the effects of exogenous covariates on the values in the relational array, expressed as

$$y_{ijr} = \beta^{\mathrm{T}} x_{ijr} + \xi_{ijr} \quad (i,j = 1,\dots,m,\ i \neq j,\ r = 1,\dots,R), \tag{1}$$

where y_{ijr} is a continuous directed measure of the rth relation from actor i to actor j, x_{ijr} is a $(p \times 1)$ vector of covariates and ξ_{ijr} is an unobserved, scalar random error. For example, considering informal insurance markets, Fafchamps & Gubert (2007) examined how covariates such as geographical proximity and kinship relate to risk sharing relations after economic shocks.

A core challenge in making inference on β arises from the innate dependencies among error relations involving the same actor. For example, dependence often exists between trade relations involving the same country or between economic transfers originating from the same individual. This dependence may arise due to variation unaccounted for in the covariates, for example, from differences in production levels between nations or from individual differences in risk aversion. Standard regression techniques may lead to poor estimates of β and/or incorrect conclusions regarding the significance of the estimate of β. Approaches to account for error dependence in relational arrays have appeared in the statistics, biostatistics and econometrics literatures and can be characterized into two broad classes.

The first set of approaches impose a parametric model on the errors. Specifically, they either use latent variables to model the array measurements as conditionally independent given the latent structure (Holland et al., 1983; Wang & Wong, 1987; Hoff et al., 2002; Li & Loken, 2002; Hoff, 2005) or model the error covariance structure directly subject to a set of simplifying assumptions (Hoff, 2011, 2015; Fosdick & Hoff, 2014). While these methods allow for possibly improved estimation of β and appropriate standard error estimators in the presence of relational dependence, the accuracy of inference on β depends on the extent to which the true error structure is consistent with the specified parametric model. In addition, many of these models are estimated in a Bayesian paradigm using Markov chain Monte Carlo approaches, which are commonly computationally expensive to estimate.

The second set of approaches to accounting for relational dependence relies heavily on empirical estimates of the error structure based on the regression residuals, first proposed by Fafchamps & Gubert (2007) and based on the spatial dependence work of Conley (1999). This framework is model agnostic, making as few assumptions as possible about the data generating process. One empirical approach estimates the regression coefficients using ordinary least squares, and then utilizes a sandwich covariance estimator, which is robust to a wide array of error structures, for the standard errors of the regression coefficients (Fafchamps & Gubert, 2007; Aronow et al., 2015). In finite samples, this estimator is hindered by the need to estimate a large number of covariance parameters with limited observations, see King & Roberts, 2015 for a discussion in other contexts, and is the reason why Wakefield (2013) suggests such estimators be labelled empirical rather than robust. We observe that standard errors from this empirical framework are often highly variable and are anticonservative.

In this work, we introduce an empirical estimation approach for relational arrays that incorporates an exchangeability assumption. This assumption is implicit in many of the model-based approaches discussed previously, and is a hallmark of Bayesian hierarchical models within, and outside, the relational context (Orbanz & Roy, 2015). Our key contribution is to define the covariance matrix, and a corresponding estimator, of the relational error array under exchangeability. Use of our parsimonious estimator produces superior estimates of β and its standard errors relative to existing approaches, and our estimator is easier to compute than existing Bayesian model-based and exchangeable bootstrapping (Menzel, 2017; Green & Shalizi, 2022) approaches. Reproduction code is available at `https://github.com/fmarrs3/netreg_public` and methods are implemented in the `R` package `netregR` (R Development Core Team, 2022).

2. Inference in relational regression

2.1. *Estimation of regression coefficients*

We employ a least squares framework to perform inference on β in the relational regression model in (1) (Aitkin, 1935). An unbiased estimator for β is the ordinary least squares estimator, $\hat{\beta} = (X^{\mathrm{T}}X)^{-1} X^{\mathrm{T}}y$, where X is an $\{Rn(n-1) \times p\}$ matrix of $(p \times 1)$ covariate vectors (x_{ijr}) and y is a vectorized representation of (y_{ijr}). The least squares estimator is the best linear unbiased estimator for β when the covariance matrix $\Omega = \mathrm{var}(y \mid X)$ is proportional to the identity matrix. Dependence is expected in relational data, e.g., between relations (i, j, r) and (i, k, r) that share actor i. If Ω were known, the best linear unbiased estimator for β is the generalized least squares estimator of Aitkin (1935),

$$\hat{\beta}_{\mathrm{GLS}} = (X^{\mathrm{T}}\Omega^{-1}X)^{-1}X^{\mathrm{T}}\Omega^{-1}y. \tag{2}$$

In practice, Ω is unknown and must be estimated. Given estimator $\hat{\Omega}$, alternating estimation of Ω with (2), replacing Ω with $\hat{\Omega}$ at each iteration, is termed feasible generalized least squares. When $\hat{\Omega}$ is consistent, feasible generalized least squares is asymptotically efficient for β (Greene, 2003; Hansen, 2015).

Regardless of whether the ordinary least squares estimator or (2) is used to estimate β, uncertainty estimates are required for inference. A common approach is to approximate the distribution of the β estimate as a multivariate normal random variable, and construct confidence intervals using an estimator of its variance: in the ordinary least squares setting,

$$\mathrm{var}(\hat{\beta} \mid X) = (X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}\Omega X(X^{\mathrm{T}}X)^{-1}, \tag{3}$$

and in the feasible generalized least squares setting,

$$\mathrm{var}(\hat{\beta}_{\mathrm{GLS}} \mid X) = (X^{\mathrm{T}}\widetilde{\Omega}^{-1}X)^{-1}X^{\mathrm{T}}\widetilde{\Omega}^{-1}\Omega\widetilde{\Omega}^{-1}X(X^{\mathrm{T}}\widetilde{\Omega}^{-1}X)^{-1}, \tag{4}$$

where $\widetilde{\Omega}$ is the final estimate of Ω from the generalized least squares procedure. Variance estimators are often constructed by substituting an estimator for Ω in (3) and (4), and are commonly termed sandwich estimators (Huber, 1967; White, 1980). Thus, inference for β requires an estimator for Ω, regardless of how β is estimated, and properties of the estimator of $\mathrm{var}(\hat{\beta} \mid X)$ depend strongly on the estimator of Ω.

2.2. *Dyadic clustering estimator*

Fafchamps & Gubert (2007), Cameron et al. (2011), Aronow et al. (2015) and Tabord-Meehan (2018) proposed and described the properties of a flexible standard error estimator for relational regression that makes the sole assumption that two relations (i, j, r) and (k, l, s) are independent if (i, j) and (k, l) do not share an actor. This assumption implies that $\mathrm{cov}(y_{ijr}, y_{kls} \mid X) = \mathrm{cov}(\xi_{ijr}, \xi_{kls} \mid X) = 0$ for non-overlapping relation pairs, but places no restrictions on the covariance elements for pairs of relations that share an actor. Let Ω_{DC} denote the covariance matrix of ξ, subject to this non-overlapping pair independence assumption. Fafchamps & Gubert (2007) proposed estimating each nonzero entry of Ω_{DC} with a product of residuals, i.e., using $e_{ijr}e_{iks}$ to estimate $\mathrm{cov}(\xi_{ijr}, \xi_{iks})$, where $e_{ijr} = y_{ijr} - \hat{\beta}^{\mathrm{T}}x_{ijr}$. The estimator $\hat{\Omega}_{\mathrm{DC}}$ can be seen as that which takes the empirical covariance of the residuals defined by ee^{T}, where e is a vector of the set of residuals (e_{ijr}), and introduces zeros to enforce the non-overlapping pair independence assumption. Fafchamps & Gubert (2007) proposed a sandwich variance estimator for $\mathrm{var}(\hat{\beta} \mid X)$ in (3) based on $\hat{\Omega}_{\mathrm{DC}}$,

$$\hat{V}_{\mathrm{DC}} = (X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}\hat{\Omega}_{\mathrm{DC}}X(X^{\mathrm{T}}X)^{-1}. \tag{5}$$

We refer to $\hat{V}_{\mathrm{DC}}$ as the dyadic clustering estimator as it owes its derivation to the extensive literature on cluster-robust standard error estimators.

The dyadic clustering estimator in (5) has the attractive properties that it is asymptotically consistent under a wide range of error dependence structures and is fast to compute. However, $\hat{\Omega}_{\mathrm{DC}}$ estimates $O(R^2n^3)$ nonzero covariance elements separately based on $O(R^2n^2)$ dependent observations, and thus $\hat{V}_{\mathrm{DC}}$ is inherently quite variable. Only when there is extreme heterogeneity in the true covariance structure is the dyadic clustering method optimal and it will suffer a loss of efficiency otherwise. Lastly, $\hat{\Omega}_{\mathrm{DC}}$ is always singular and thus $\hat{\Omega}_{\mathrm{DC}}$ cannot be inverted for use in a feasible generalized least squares estimator of β.

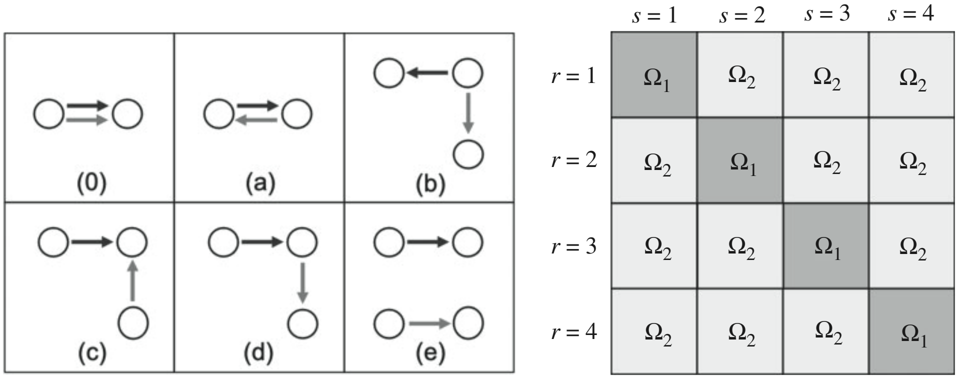

Fig. 1. Left: six distinguishable configurations of relation pairs (black and grey arrows) in an exchangeable relational model with unlabelled actors, corresponding to 12 covariance values (six each for $s = r$ and $s \neq r$). Right: block structure in Ω_E for $R = 4$, where blocks Ω_1 and Ω_2 correspond to $s = r$ and $s \neq r$, respectively.

3. Standard errors under exchangeability

3.1. *Exchangeability in relational models*

A common modelling assumption for relational and array structured errors is exchangeability. Defined by de Finetti for a univariate sequence of random variables, exchangeability was generalized to array data by D. N. Hoover in his 1979 Princeton Institute for Advanced Study preprint and Aldous (1981). The errors in a relational data model are jointly exchangeable if the probability distribution of the error array $\Xi = (\xi_{ijr})$ is invariant under any simultaneous permutation of the rows and columns, and secondary permutation of the third dimension. Mathematically, this means that $\mathrm{pr}(\Xi) = \mathrm{pr}\{\Pi(\Xi)\}$, where $\Pi(\Xi) = \{\xi_{\pi(i)\pi(j)\nu(r)}\}$ is the error array with its indices reordered according to permutation operators π and ν. Intuitively, exchangeability in the regression context means that the observed covariates are sufficiently informative such that the labels of the rows and columns in the error array are uninformative. Similarly, the ordering of the third dimension of the error array is uninformative to its distribution. This assumption may be appropriate when the third dimension of the array represents different contexts of observations, such as economic trade sectors, that have no inherent ordering, or when the third dimension represents time periods, but the bulk of the temporal variation is accounted for in the covariates. Many of the conditionally independent parametric latent variable models cited in § 1 have this joint exchangeability property (Hoff, 2008; Bickel & Chen, 2009).

3.2. *Impact of exchangeability on the covariance structure*

Li & Loken (2002) and Hoff (2005) described several particular random effects models for $R = 1$. The corresponding error covariance matrices have different entries depending on the model, yet all covariance matrices have at most six unique entries. A key contribution of this paper is to formalize and extend this observation, showing that any jointly exchangeable model for relational array Ξ results in an Ω of the same form, with at most six unique terms when $R = 1$ and at most 12 unique terms when $R > 1$.

Proposition 1. *If a probability model for a directed relational array Ξ is jointly exchangeable and has finite second moments, then the covariance matrix of Ξ contains at most* 12 *unique values.*

The 12 (possibly) unique entries in Ω correspond to the 12 distinguishable configurations of relation pairs (i, j, r) and (k, l, s) with unlabelled actors. The 12 configurations can be separated into two sets of six identical configurations of relations (i, j) and (k, l) with unlabelled actors, as depicted in Fig. 1, where each set corresponds to $r = s$ and $r \neq s$. A proof is provided in the Supplementary Material.

3.3. *Covariance matrices of exchangeable relational arrays*

Similar to the dyadic clustering estimator, we assume that non-overlapping relation pairs are independent, such that $\mathrm{cov}(\xi_{kls}, \xi_{ijr}) = 0$ for any s and r when (i, j, k, l) are distinct. This assumption sets two of the 12 parameters in Ω to zero. We introduce a new class of covariance matrices that contain 10 possibly nonzero entries, $\phi_0^{(\eta)}, \phi_a^{(\eta)}, \phi_b^{(\eta)}, \phi_c^{(\eta)}, \phi_d^{(\eta)}$ $(\eta = 1, 2)$, associated with Fig. 1(a). The separation of covariances by $r = s$ and $r \neq s$ implies that Ω consists of blocks of Ω_1 and Ω_2, each consisting of five nonzero terms for $r = s$ $(\eta = 1)$ and $r \neq s$ $(\eta = 2)$, respectively; see Fig. 1(b). We define an exchangeable covariance matrix as any covariance matrix of this form and denote it Ω_E.

4. Exchangeable estimator definition and evaluation

4.1. *Exchangeable covariance estimator*

Consider relational regression models with Ω of the exchangeable form Ω_E. The proposed exchangeable estimator of $\text{var}(\hat{\beta} \mid X)$ is then

$$\hat{V}_E = (X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}\hat{\Omega}_E X(X^{\mathrm{T}}X)^{-1}, \qquad \hat{\Omega}_E = \sum_{\eta=1}^{2}\sum_{u=0}^{d}\hat{\phi}_u^{(\eta)}\mathcal{S}_u^{(\eta)},$$

where $\mathcal{S}_u^{(\eta)}$ denotes the $\{Rn(n-1) \times Rn(n-1)\}$ binary matrix with 1s in the entries corresponding to relation pairs of type $(u = 0, a, b, c, d; \eta = 1, 2)$, as defined in Fig. 1. We propose estimating the 10 parameters in Ω by averaging the residual products that share the same index configurations, corresponding to (a)–(d) in Fig. 1. For example, the estimate of $\text{cov}(\xi_{kls}, \xi_{ijr})$, corresponding to $u = b$ and $\eta = 2$, is

$$\hat{\phi}_b^{(2)} = \binom{R}{2}^{-1}\frac{1}{n(n-1)(n-2)}\sum_{r\neq s}\sum_{i}\sum_{j\neq i} e_{ijr}\left(\sum_{k\neq j} e_{iks} - e_{ijs}\right).$$

The remaining nine estimators for $(s = 0, a, \ldots, e; \eta = 1, 2)$ are defined analogously, and $\hat{\Omega}_E$ may be interpreted as the projection of $\hat{\Omega}_{\mathrm{DC}}$ into the vector space over symmetric matrices of the form of Ω_E.

4.2. *Comparison of the exchangeable estimator with dyadic clustering*

It is intuitive that the moment-based exchangeable estimator is consistent, and more efficient than the dyadic clustering estimator, whenever the exchangeability assumption is satisfied. One might expect the highly parameterized dyadic clustering estimator to trade-off high variance for reduced bias. However, we derive the result that the dyadic clustering estimator is biased downwards, and that this bias is larger than twice the bias of the exchangeable estimator. One concludes that a trade-off for the robustness of the dyadic clustering estimator is anticonservatism. The proof of Theorem 1 is provided in the Supplementary Material.

Theorem 1. *Consider error vector ξ and normally distributed covariate vector x with zero means, exchangeable covariance matrices and bounded fourth moments, where*

$$y_{ij} = \beta_1 + x_{ij}\beta_2 + \xi_{ij}.$$

Then, the dyadic clustering estimator for $\text{var}(\hat{\beta}_2)$ *is biased downwards,*

$$n^2\text{Bias}(\hat{V}_{\mathrm{DC}}) + O(n^{-1/2}) \leqslant -2n^2|\text{Bias}(\hat{V}_E)| + O(n^{-1/2}) \leqslant 0,$$

noting that both $n^2\text{Bias}(\hat{V}_{\mathrm{DC}})$ *and* $|n^2\text{Bias}(\hat{V}_E)|$ *are* $O(1)$.

We conducted a simulation study to compare the bias and 95% confidence interval coverage when using the exchangeable and dyadic clustering estimators. We simulated from a model with three covariates, one each of binary, positive real and real valued, with exchangeable and nonexchangeable error models, for $R = 1$ and $n = 20, 40, 80, 160, 320$. Figure 2 shows the estimated mean coverage and middle 95% of coverages across various X realizations. In all settings, the estimated mean coverage of the exchangeable estimator

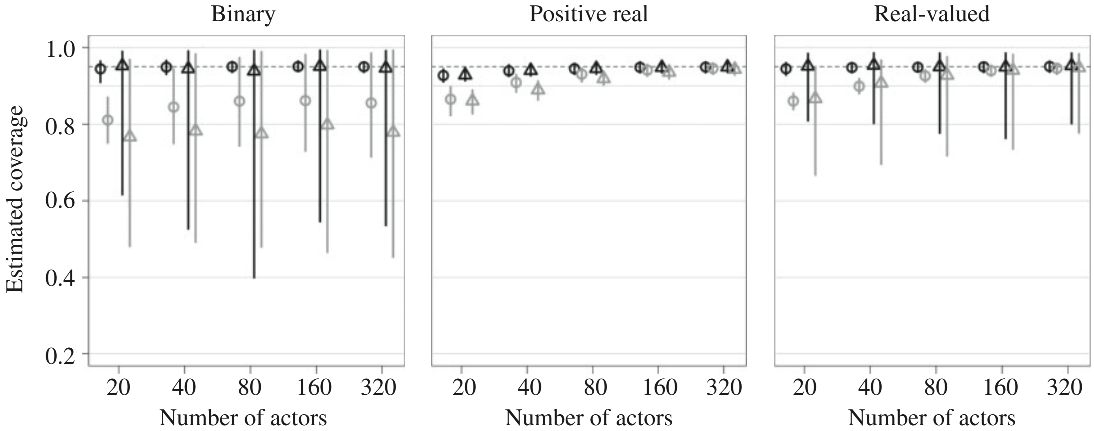

Fig. 2. Estimated probability that the true coefficient is in the 95% confidence interval for each of three covariates (binary, positive and real valued) when the errors are generated from exchangeable (circles) and nonexchangeable (triangles) models. Points denote mean estimated coverage and lines represent the middle 95% of coverages for exchangeable (black) and dyadic clustering (grey) estimators.

is closer to the nominal 0.95 level than the dyadic clustering estimator. This difference is most pronounced for the binary covariate, where there is reduced signal to noise relative to the other covariates. The average bias of the dyadic clustering estimator is typically more than four times that of the exchangeable estimator under the exchangeable error model, confirming Theorem 1 and driving the poorer coverage performance.

5. Patterns in international trade

We demonstrate the implications of using our exchangeable estimator in a study of international trade among 58 countries over $R = T = 20$ years. These data were previously analysed and made available by Westveld & Hoff (2011). Following Tinbergen (1962), Ward & Hoff (2007) and Westveld & Hoff (2011) we use a modified gravity mean model to represent log yearly trade between each pair of countries as a linear function of seven covariates in years 1981–2000. Westveld & Hoff (2011) proposed a model, which we refer to as the mixed effects model, which explicitly decomposes the regression error term ϵ_{ijt} for each time period and pair of actors into time-dependent sender and receiver effects, resulting in 13 error covariance parameters that are estimated using a latent variable representation and Bayesian Markov chain Monte Carlo methodology. We propose estimating the gravity mean model using feasible generalized least squares, assuming that the errors are jointly exchangeable. As noted in § 4.1, the proposed approach estimates 10 error covariance parameters.

We compared the exchangeable and mixed effects approaches, and ordinary least squares as a baseline, in an out-of-sample prediction study. Here we estimated the regression coefficients using the first K years of trade data for $K = 4, \ldots, 19$ and used the estimates to predict trade values in the following year. Figure 3 provides the coefficient of determination, R^2, for the three procedures when predicting trade flows in years 5 through 20. There is a median increase in R^2 of about 10% (30%) when using the proposed exchangeable approach relative to the mixed effects approach (ordinary least squares). The proposed approach performs better than the other approaches for all time periods, although the gap in performance decreases as K increases. These results suggest that the more parsimonious exchangeable approach represents the data better than the mixed effects model, and yet, the exchangeable approach runs in a small fraction of the time of the mixed effects approach. See the Supplementary Material for additional details.

We compared the coefficients of the ordinary least squares, mixed effects and exchangeable approaches. The R^2 between the ordinary least squares and mixed effects coefficients is about 0.46, while the exchangeable and mixed effects coefficients have an R^2 of 0.78. About 40% of the ordinary least squares coefficients, using dyadic clustering standard errors, were significantly different from the mixed effects coefficients, while only 1% of the exchangeable coefficients were significantly different from the mixed

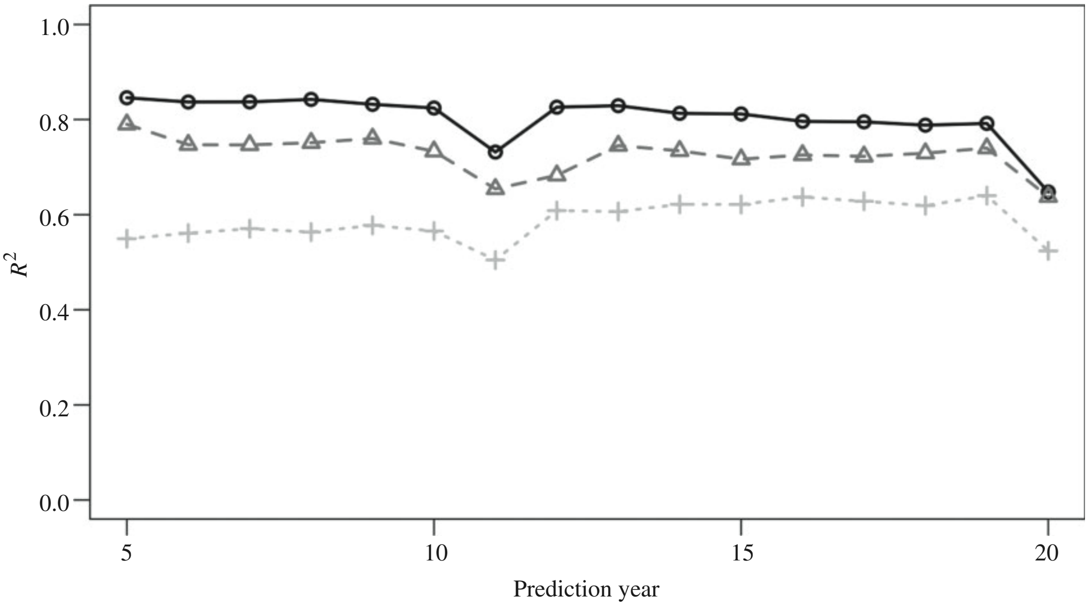

Fig. 3. The coefficient of determination R^2 when predicting one-year-ahead trade flows from exchangeable (black circle), mixed effects (grey triangle) and ordinary least squares (light grey plus).

effects coefficients. Finally, the ordinary least squares coefficients have standard errors that are, on average, over 1.5 times the standard errors of the exchangeable coefficients, with exchangeable standard errors. Together with the one-year-ahead prediction results, the coefficient and standard error comparisons suggest that the proposed approach can revise ordinary least squares coefficients in the direction of a higher fidelity model, giving more precise estimates of the coefficients, while requiring few modelling decisions and with limited runtime penalty. Based on the success of the exchangeable approach in the trade data analysis and simulation study, we recommend that researchers use the feasible generalized least squares estimator of the coefficient vector β as demonstrated here, unless an unbiased estimator of the coefficient vector is specifically desired.

6. Discussion

The proposed exchangeable estimator leverages exchangeability for maximal symmetry and parsimony in the covariance matrix of relational array Y. The exchangeability assumption may not be appropriate when the true error covariances are substantially heterogeneous. We propose using a permutation test based on the dyadic clustering estimator for testing the hypothesis of exchangeable errors. The procedure consists of generating a null distribution of $\hat{V}_{\text{DC}}$ in (5) by randomly permuting the residual array in a manner consistent with exchangeability. If the observed estimator is extreme relative to the null distribution, this suggests that the errors are nonexchangeable. Details and simulations are available in the Supplementary Material.

Supplementary material

The Supplementary Material contains details of the estimator and analyses, proofs and a proposed test for exchangeability.

References

Aitkin, A. (1935). On least squares and linear combination of observations. *Proc. R. Soc. Edin.* **55**, 42–8.

Aldous, D. J. (1981). Representations for partially exchangeable arrays of random variables. *J. Mult. Anal.* **11**, 581–98.

Aleskerov, F., Meshcheryakova, N., Rezyapova, A. & Shvydun, S. (2017). Network analysis of international migration. In *Models, Algorithms, and Technologies for Network Analysis*, V. A. Kalyagin, A. I. Nikolaev, P. M. Pardalos & O. A. Prokopyev, eds. Cham: Springer, pp. 175–85.

Aronow, P. M., Samii, C. & Assenova, V. A. (2015). Cluster–robust variance estimation for dyadic data. *Polit. Anal.* **23**, 564–77.

Attanasio, O., Barr, A., Cardenas, J. C., Genicot, G. & Meghir, C. (2012). Risk pooling, risk preferences, and social networks. *Am. Econ. J.* **4**, 134–67.

Banerjee, A., Chandrasekhar, A. G., Duflo, E. & Jackson, M. O. (2013). The diffusion of microfinance. *Science* **341**, 363–73.

Bardham, P. (1984). *Land, Labor and Rural Poverty*. New York: Oxford University Press.

Bickel, P. J. & Chen, A. (2009). A nonparametric view of network models and Newman–Girvan and other modularities. *Proc. Nat. Acad. Sci. USA* **106**, 21068–73.

Cameron, A. C., Gelbach, J. B. & Miller, D. L. (2011). Robust inference with multiway clustering. *J. Bus. Econ. Statist.* **29**, 238–49 .

Conley, T. G. (1999). GMM estimation with cross sectional dependence. *J. Economet.* **92**, 1–45.

Fafchamps, M. (2006). Development and social capital. *J. Develop. Studies* **42**, 1180–98.

Fafchamps, M. & Gubert, F. (2007). The formation of risk sharing networks. *J. Develop. Econ.* **83**, 326–50.

Fagiolo, G., Reyes, J. & Schiavo, S. (2008). On the topological properties of the world trade web: a weighted network analysis. *Phys. A* **387**, 3868–73.

Fosdick, B. K. & Hoff, P. D. (2014). Separable factor analysis with applications to mortality data. *Ann. Appl. Statist.* **8**, 120–47.

Foster, A. D. & Rosenzweig, M. R. (2001). Imperfect commitment, altruism, and the family: evidence from transfer behavior in low-income rural areas. *Rev. Econ. Statist.* **83**, 389–407.

Green, A. & Shalizi, C. R. (2022). Bootstrapping exchangeable random graphs. *arXiv:* 1711.00813v2.

Greene, W. H. (2003). *Econometric Analysis*, 5th ed. Englewood Cliffs, NJ: Prentice Hall.

Han, G., McCubbins, O. & Paulsen, T. (2016). Using social network analysis to measure student collaboration in an undergraduate capstone course. *NACTA J.* **60**, 176.

Hansen, B. E. (2015). *Econometrics*. Available at http://www.ssc.wisc.edu/ bhansen/econometrics/. Last accessed May 2022.

Hoff, P. D. (2005). Bilinear mixed-effects models for dyadic data. *J. Am. Statist. Assoc.* **100**, 286–95.

Hoff, P. D. (2008). Modeling homophily and stochastic equivalence in symmetric relational data. In *Proc. 20th Int. Conf. Neural Info. Proces. Syst.*, pp. 657–64. Red Hook, NY: Curran Associates.

Hoff, P. D. (2011). Separable covariance arrays via the Tucker product, with applications to multivariate relational data. *Bayesian Anal.* **6**, 179–96.

Hoff, P. D. (2015). Multilinear tensor regression for longitudinal relational data. *Ann. Appl. Statist.* **9**, 1169–93.

Hoff, P. D., Raftery, A. E. & Handcock, M. S. (2002). Latent space approaches to social network analysis. *J. Am. Statist. Assoc.* **97**, 1090–8.

Holland, P. W., Laskey, K. B. & Leinhardt, S. (1983). Stochastic blockmodels: first steps. *Social Networks* **5**, 109–37.

Huber, P. J. (1967). The behavior of maximum likelihood estimates under nonstandard conditions. In *Proc. 5th Berkeley Symp. Math. Statist. Prob.*, vol. 1, pp. 221–33. Berkeley, CA: University of California Press.

King, G. & Roberts, M. E. (2015). How robust standard errors expose methodological problems they do not fix, and what to do about it. *Polit. Anal.* **23**, 159–79.

Li, H. & Loken, E. (2002). A unified theory of statistical analysis and inference for variance component models for dyadic data. *Statist. Sinica* **12**, 519–35.

Menzel, K. (2017). Bootstrap with clustering in two or more dimensions. *arXiv:* 1703.03043v3.

Orbanz, P. & Roy, D. M. (2015). Bayesian models of graphs, arrays and other exchangeable random structures. *IEEE Trans. Pat. Anal. Mach. Intel.* **37**, 437–61.

R Development Core Team (2022). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0, http://www.R-project.org.

Tabord-Meehan, M. (2018). Inference with dyadic data: asymptotic behavior of the dyadic-robust *t*-statistic. *J. Bus. Econ. Statist.* **37**, 1–10.

Tinbergen, J. (1962). *Shaping the World Economy: Suggestions for an International Economic Policy*. New York: Twentieth Century Fund.

Wakefield, J. (2013). *Bayesian and Frequentist Regression Methods*. New York: Springer.

Wang, Y. J. & Wong, G. Y. (1987). Stochastic blockmodels for directed graphs. *J. Am. Statist. Assoc.* **82**, 8–19.

Ward, M. D. & Hoff, P. D. (2007). Persistent patterns of international commerce. *J. Peace Res.* **44**, 157–75.

Westveld, A. H. & Hoff, P. D. (2011). A mixed effects model for longitudinal relational and network data, with applications to international trade and conflict. *Ann. Appl. Statist.* **5**, 843–72.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* **48**, 817–38.

Zhang, B. & Horvath, S. (2005). A general framework for weighted gene co-expression network analysis. *Statist. Appl.. Genet. Molec. Biol.* **4**, doi: 10.2202/1544-6115.1128.

[*Received on* 28 *January* 2021. *Editorial decision on* 11 *May* 2022]