Modeling actor and partner effects in dyadic data when outcomes are categorical

Loeys, Tom & MOLENBERGHS, Geert (2013) Modeling actor and partner effects in dyadic data when outcomes are categorical. In: Psychological methods, 18 (2), p. 220-236.

DOI: 10.1037/a0030640
Handle: http://hdl.handle.net/1942/14740

Modeling Actor and Partner Effects in Dyadic Data When Outcomes are Categorical

Tom Loeys[1], Geert Molenberghs[2]

August 1, 2012

Affiliations:

[1] Department of Data Analysis, Faculty of Psychology and Educational Sciences, Universiteit Gent, Belgium (Email: tom.loeys@ugent.be)

[2] Interuniversity Institute for Biostatistics and Statistical Bioinformatics, Katholieke Universiteit Leuven, Belgium

Abstract

When two people interact in a relationship, the outcome of each person can be affected by both his or her own inputs and his or her partner's inputs. For Gaussian dyadic outcomes, linear mixed models taking into account the correlation within dyads, are frequently used to estimate actor's and partner's effects based on the actor-partner interdependence model. In this paper, we explore the potential of generalized linear mixed models (GLMMs) for the analysis of non-Gaussian dyadic outcomes. Several approximation techniques that are available in standard software packages for these GLMMs are investigated. Despite the different modeling options related to these different techniques, none of these have an overall satisfactory performance in estimating actor and partner effects and the within-dyad correlation, especially when the latter is negative and/or the number of dyads is small. An approach based on generalized estimating equations for the analysis of non-Gaussian dyadic data turns out to be an interesting alternative.

KEY WORDS: binary data, count data, dyadic data, generalized estimating equations, generalized linear mixed models, multilevel analysis

1 Introduction

Dyadic research has become immensely popular in the social and behavioral sciences. When two people interact in a relationship, the outcome of each person can be affected by both his or her own inputs and his or her partner's inputs. The Actor-Partner Interdependence Model (APIM) offers an appealing approach to model such dyadic behavior (Kenny, Kashy, & Cook, 2006). Indeed, it allows to simultaneously study the influence of a person's own predictor variable on his or her own outcome variable, which is called the *actor effect*, and on the outcome variable of the partner, which is called the *partner effect*, while allowing for non-independence in the two persons' responses. Typically, two types of dyads are considered. Dyads are called distinguishable when the two persons from all the dyads can be ordered in the same way (e.g., for hetero couples, persons within a dyad can be ordered by gender). Indistinguishable dyads occur when no ordering of persons exists within a dyad (like twins, for example). The left panel of Figure 1 shows a graphical presentation of the APIM with two distinguishable dyad members and an X and Y variable for each. The variables X_1 and X_2 represent the predictor variables of persons 1 and 2 of a dyad, respectively, whereas Y_1 and Y_2 represent the outcome variables for the two members. The model contains two actor effects a_1 and a_2 (represented by the horizontal arrows), and two partner effects p_{12} and p_{21} (represented by the diagonal arrows). The curved arrow on the left reflects the correlation between the predictor variables, while the one on the right represents the correlation between the error terms. An alternative but underutilized model (Ledermann & Kenny, 2012) to explore dyadic influences is the Common Fate Model (CFM). When a construct representing a common fate variable exists at the level of the dyad rather than at the individual level, the CFM is more appropriate (right panel of Figure 1). In contrast, self-referential or partner-referential measures that are expected to represent individual behaviors or attitudes are more suitable for the APIM. We will focus here on the APIM, which has clearly dominated the dyadic literature with more than 150 publications over the last three years (Kenny & Ledermann, 2012).

Multilevel modeling, also referred to as hierarchical modeling, has been shown to be a useful technique for the estimation of actor and partner effects in dyadic data (Kenny et al., 2006). In these multilevel models two different levels are distinguished: the lower level, or level 1, refers to the case of persons nested within a dyad. The lower-level unit is person, whereas the upper level, or level 2, is the dyad. Linear mixed models (LMM) are frequently used and well understood for the analysis of such dyadic data but their use is limited to (Gaussian) outcomes measured at the interval level. The analysis of non-Gaussian dyadic data on the other hand has received little attention in the literature so far. The generalized linear mixed model (GLMM), which is an extension of both the generalized linear model (GLM) (Nelder & Wedderburn, 1972) and the LMM, is potentially suitable for the analysis of clustered observations from the exponential dispersion family distribution (Agresti, 2000). McMahon, Pouget, and Tortu (2006) and Spain, Jackson, and Edmonds (2012) provided guidance on fitting the GLMM for dyadic data with binary outcomes but paid little attention to its properties.

We argue here that before proposing the GLMM as an appropriate approach for modeling dyadic data, two important issues need to be explored in more detail. First, while researchers in psychological or social sciences are often faced with clustered data (in educational measurement applications for example, when several test items are administered to students; in longitudinal studies when psychological measurements are repeatedly assessed over time, etc.), the cluster size of two when analyzing dyads is an important feature. Second, the possible negative correlation between observations within a dyad also warrants further exploration. Indeed, while in an item-response or longitudinal setting, measurements are typically positively correlated, negative correlations may occur within dyads. The strictness of parental supervision is one example of such negative correlation within dyads, where the more extreme in strictness one parent becomes, the more extreme in permissiveness the other parent is likely to become (Cook, 2001).

We therefore investigate in detail here the performance of multilevel modeling of non-Gaussian outcomes in a dyadic setting. We first discuss the traditional use of the GLMM and its interpretation, point to its limitations, and explore the potential of some other rather non-standard estimation techniques for these GLMMs. Next, we introduce the generalized estimating equations (GEE) approach (Liang and Zeger, 1986) as a viable alternative. While multilevel models have become immensely popular for the analyses of correlated data, GEE is relatively unused in the educational and behavioral sciences (Bauer & Sterba, 2011). GEE accommodates correlated outcome data too; but whereas multilevel models explicitly specify the joint distribution of the outcomes, GEE only models the univariate marginal expectations as a function of explanatory variables and empirically accounts for the presence of correlation in the data. Simulations for binary and count dyadic data are performed to compare, under a wide range of within-dyad correlations and for typical APIM sample sizes, the performance of the GEE-approach with different estimation techniques for the multilevel approach. Focus in these simulation studies lies on both the estimation of the actor and partner effects and the estimation of the within-dyad correlation. We end with an application of the different approaches to data from the Interdisciplinary Project for the Optimization of Separation Trajectories conducted in Flanders (IPOS) and present two illustrations. A first example illustrates the analysis of negatively correlated binary data and investigates the effect of actor's and partner's levels of feeling guilty during the break-up, on showing so-called forcing behavior or not during the postbreak-up negotiations in 29 ex-couples. The second example presents the analysis of positively correlated count data and explores in 33 ex-couples the effect of the actor's and partner's level of anxious attachment in their relationship with their ex-partner prior to the break-up on the number of unwanted pursuit behavior (UPB) perpetrations after separation.

2 Multilevel models

2.1 Linear mixed models

Let $\mathbf{Y}_i$ denote a 2-dimensional vector of measurements available for dyad $i = 1, \ldots, N$ with components Y_{i1} and Y_{i2} (e.g., the measurement for a male and female partner in a heterosexual couple[1]). Using LMMs for the APIM, two different formulations are typically considered (Kenny et al., 2006). The first one takes a hierarchical view (i.e., a multilevel approach) and specifies the so-called random-intercept model, with the random intercept capturing the correlation within a dyad:

$$Y_{ij} = \mathbf{x}_{ij}^t \boldsymbol{\beta} + b_i + \epsilon_{ij} \qquad j = 1, 2 \tag{1}$$

with $b_i \sim N(0, \tau)$ and $\epsilon_{ij} \sim N(0, \sigma^2)$. In APIM (1) $\mathbf{x}_{ij}$ is a vector of known covariates, typically including an actor's predictor variable x_{act}, a partner's predictor variable x_{par}, a distinguishing variable x_{dis} (like gender) in case of distinguishable dyads , and their interactions; $\boldsymbol{\beta}$ a vector of coefficients, called fixed effects, and b_i a random intercept. In a standard multilevel model, the assumption is made that the variance of the random effect is positive (i.e., $\tau \geq 0$).

The second formulation of the APIM takes a marginal view, i.e. it does not

[1] Throughout the manuscript we will assume distinguishable dyads but all models and estimation techniques that are presented can easily accommodate indistinguishable dyads as well.

incorporate random effects,

$$Y_{ij} = \mathbf{x}_{ij}^t \boldsymbol{\beta} + \epsilon_{ij} \qquad\qquad j = 1, 2 \tag{2}$$

but simply models the variance-covariance in the data[2]. Model (2) assumes in its most general form that the residuals ϵ_{ij} are bivariate normally zero-mean distributed with an unstructured variance-covariance

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}.$$

Here, ρ reflects the within-dyad correlation[3] and can take any value in the interval $[-1, +1]$.

It is important to note here that *marginal models* like (2) describe so-called population-averaged effects that refer to an averaging over dyads particular levels of predictors while dyad-specific models like (1) are *conditional models* that describe effects at the dyad level. However, since the marginal and conditional expectation of Y_{ij} are the same here (see Table 1), i.e. $E(Y_{ij}) = E(Y_{ij} \mid b_i)$, the parameters $\boldsymbol{\beta}$ in (1) and (2) share their interpretation. In other words, if on average *within dyads*, a 1-unit increase in the actor predictor x_{act} for example causes a shift of size β_1 for the actor's outcome Y (i.e. 'the conditional effect'), then this coefficient β_1 can also be interpreted as the effect on the population level, and the estimated *overall sample* means (i.e. 'the marginal effect') will also change with the same coefficient β_1 for such 1-unit increase. The marginal

[2]This method is sometimes referred to as the R-side covariance method.

[3]More precisely it measures the residual intra-cluster correlation (ICC), i.e. the correlation between the measurement of the first person of the dyad and the measurement of the second person of the dyad that is left after accounting for the predictor effects in model (2).

variance-covariance matrix under model (1) has a compound symmetry structure with correlation equal to

$$\tau/(\sigma^2 + \tau), \tag{3}$$

and so in contrast to the marginal model formulation (2), the hierarchical formulation with the restriction $\tau \geq 0$ does not allow for negative correlation within dyads. The latter can be a serious restriction in dyadic settings where negative within-dyad correlations are not uncommon and therefore formulation (2) is typically preferred above formulation (1). If one takes a marginal view on model (1) though, negative values for τ are perfectly possible (Molenberghs & Verbeke, 2011)[4]. We will further refer to the latter approach as the 'unconstrained approach' as opposed to the more standard 'constrained approach'.

2.2 Generalized linear mixed models: a conditional approach

While the APIM was considered for the Gaussian outcomes in the previous section, we now focus on modeling dyadic binary and count data. Similar to model (1) for Gaussian outcomes, we consider the logistic-normal random intercept model (Snijders & Bosker, 1999) for binary dyadic data with a logit link[5]

[4]In such approach the conceptual interpretation of the random effect is abandoned and the hierarchical model approach merely used as a vehicle for estimation. The only restriction is that $\tau \geq -\sigma^2/2$ for the variance-covariance matrix to be positive definite.

[5]Other link functions like the probit or complementary log-log could be considered as well, but we will restrict attention to the logit-link here.

and assume no overdispersion:

$$\text{logit}[E(Y_{ij} \mid b_i)] = \mathbf{x}_{ij}^t \boldsymbol{\beta} + b_i \quad \text{with } b_i \sim N(0, \tau), \quad \tau > 0 \tag{4}$$

Unlike for LMMs (1), the marginal effects are different from the conditional effects here (see Table 1), unless the random intercept variance τ equals zero. Similar to LMMs (1), the outcomes Y_{i1} and Y_{i2} from dyad i are here too conditionally independent (i.e., given b_i) but are marginally non-negatively correlated (Table 1). The marginal correlation is not straightforward to calculate for Bernoulli outcomes, given the dependence of the variance on the mean. Pryseley, Tchonlafi, Verbeke, and Molenberghs (2011) derive an easy-to-calculate first-order approximation of the intra-cluster correlation (ICC)[6],

$$\rho \approx \frac{\tau}{\tau + \exp(\beta_0)(1 + \exp(-\beta_0))^2}, \tag{5}$$

where β_0 is the intercept of model (4) (assuming centered predictors). Observe that $\rho = 0$ when $\tau = 0$ and $\rho \to 1$ as $\tau \to +\infty$.

Next, we consider the Poisson-normal random-intercept model for count data (Agresti, 2000) with a log link

$$\log[E(Y_{ij} \mid b_i)] = \mathbf{x}_{ij}^t \boldsymbol{\beta} + b_i \quad \text{with } b_i \sim N(0, \tau), \quad \tau > 0. \tag{6}$$

The marginal effects of the explanatory variables are the same as the dyad-specific effects in model (6); Y_{i1} and Y_{i2} are marginally non-negatively correlated

[6]When Y_i is viewed as a realization of an underlying latent variable $\tilde{Y}_i$, another expression for the ICC - analogous to (3) - is given by $\tau/(\tau + \pi^2/3)$ with $\pi^2/3$ the variance of the standard logistic distribution. The latter expression should hence be viewed as the ICC at the underlying latent level as opposed to the observed.

when $\tau > 0$ (Table 1). Here too, the variance depends on the mean, but the ICC can be approximated (Pryseley et al., 2011) by

$$\rho = \frac{\exp\left(\beta_0 + \frac{1}{2}\tau\right)\left(\exp\tau - 1\right)}{1 + \exp\left(\beta_0 + \frac{1}{2}\tau\right)\left(\exp\tau - 1\right)} \tag{7}$$

where β_0 is the intercept of model (6) (assuming again no effect of the centered predictors). Again, $\rho = 0$ when $\tau = 0$ and $\rho \to 1$ as $\tau \to +\infty$. A limitation of the Poisson model is that the variance must equal the mean (Loeys, Moerkerke, De Smet & Buysse, 2012). However, count data often show overdispersion, with the variance exceeding the mean. As will be illustrated in a later example, the negative binomial distribution allows for such overdispersion. Molenberghs, Verbeke and Demétrio (2007) derive closed-form expressions for the ICC under the latter scenario.

In summary, we have illustrated that, similar to the LMM with random intercept, the GLMM with random intercept, leads to non-negative marginal correlations too when $\tau > 0$ is restricted to be positive. The similarity between marginal and conditional interpretation of the fixed effect parameters in the GLMM on the other hand depends on the link-function[7]. So far, we have only

[7]It can be shown in general that when the conditional mean is additive in a random effect on the log scale, the marginal mean equals the conditional mean plus a constant, such that slope parameters have the same interpretation in both formulations. No further distributional assumptions are needed in this case (Griswold & Zeger, 2004). When a logit or probit link is used with a normal random effect, the marginal mean parameters become attenuated by a factor which depends on parameters of the distribution of the covariates. For example for the binary case the marginal effect can be approximated by $\beta/\sqrt{c^2\tau + 1}$ with β the conditional effect from the logistic-normal model and $c = 16\sqrt{3}/(15\pi)$ (Heagerty, 1999).

discussed the counterpart of model (1) for non-Gaussian outcomes. In the next paragraph, we will see how we can get to a marginal formulation similar to (2) in the GLMM-framework.

2.3 Generalized linear mixed models: a marginal approach

Fitting GLMMs like models (4) and (6) proceeds by integrating over the random effects. Broadly speaking three different strategies have historically been considered to overcome the integration over the (normally) distributed random effects: (i) approximation of the integral using Gaussian quadrature, (ii) approximation of the integrand using Laplace's method, and (iii) a quasi-likelihood approach based on a linearized approximation. We refer the interested reader to Tuerlinckx, Rijmen, Verbeke, and De Boeck (2006) for an in-depth review and discussion of these different approximation methods. These three approximation techniques are available in standard software packages like SAS for example. We omit the technical details of (i) and (ii) here, but elaborate a bit further on (iii) as it will allow us to specify a marginalized GLMM. To explain the linearized approximation method, we can rewrite models (4) and (6) as

$$Y_{ij} = h(\mathbf{x}_{ij}^t\boldsymbol{\beta} + b_i) + \epsilon_{ij},$$

with h the inverse of the logit and log function, respectively. A first-order Taylor expansion around the estimated fixed effect and posterior mode of the random effect and further re-arrangement (Tuerlinck et al., 2006) lead to

$$Y_{ij} \approx \hat{\mu}_{ij} + v(\hat{\mu}_{ij})\mathbf{x}_{ij}^t(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) + v(\hat{\mu}_{ij})(b_i - \hat{b}_i) + \epsilon_{ij}, \tag{8}$$

with $v(\hat{\mu}_{ij})$ the approximate variance of the error term. It can be shown that (8) can be rewritten as a linear mixed model for pseudo data $\mathbf{Y}_i^*$ with fixed effects $\boldsymbol{\beta}$, random effects b_i and error terms $\boldsymbol{\epsilon}_i^*$. Therefore, estimation of $\boldsymbol{\beta}$, the fixed effect parameters, and the variance of b_i, can be obtained by iterating between updating the pseudo response and fitting the linear mixed model to the pseudo-data. This approach is referred to as penalized quasi-likelihood (PQL). The advantage of using such pseudo-likelihood approach is that it becomes possible to fit GLMMs without random effects and to take a marginal view with only residual association effects. We will further label this model as the marginalized GLMM. In contrast to the constrained random intercept models (4) and (6) for example, this marginalized GLMM allows - similar to model (2) for Gaussian outcomes - to model negatively correlated non-Gaussian outcomes. Alternatively, one may take a marginal view on the random intercept models (4) and (6) and give up the constraint $\tau > 0$ ('the unconstrained' random intercept approach). By doing so, one can allow for negatively correlated outcomes as well. In practice, it turns out that this is possible when the Laplace-approximation is used, but not under the Gaussian quadrature approximation (Pryseley et al., 2011).

3 Generalized Estimating Equations

Generalized estimating equations, as introduced by Liang and Zeger (1986), can be considered as an extension of the GLM that accommodates correlated outcome data too. It provides a general framework for the analysis of correlated continuous, ordinal, dichotomous, or count dependent data. GEE is often referred to as a marginal (or population-averaged) approach as opposed to the conditional approach exploited by multilevel models (Diggle, Heagerty, Liang, & Zeger, 2002). Whereas multilevel models explicitly specify the joint distribution of the outcomes, focus on modeling the dyad-specific expectation as a function of explanatory values, and allow one to disentangle the variability at the different levels, GEE is a moment-based method and only models the marginal expectations as a function of explanatory variables.

The GEE fitting algorithm can be described in four different steps that are repeated until convergence (Ghisletta & Spini, 2004).

1. A GLM is fitted assuming independence between observations. This GLM requires the specification of a link function that describes the linear relationship between the expected outcome and its predictors (e.g., the identity link for Gaussian data, the logit link for dichotomous or ordinal data and the log link for count data) and of the relationship between the mean μ and the variance, denoted $v(\mu)$.

2. Standardized residuals, contrasting the observed and expected (model-based) outcome, are calculated. Based on an assumed structure of the correlation matrix (such as independence or unstructured), a *working correlation* matrix C that characterizes the correlations among observations within dyads is computed using these standardized residuals. We suggest to use the unstructured working correlation structure here[8].

[8] An unstructured covariance matrix is no guarantee for a correct specification since the

3. An estimate of the covariance parameters is obtained from the assumed mean-variance association $v(\mu)$ and the working correlation matrix C.

4. Given the covariance estimate obtained in step 3, a set of estimating equations for the regression coefficients is solved[9].

5. The steps 2 to 4 result in an iterative scheme that switches between estimating the regression coefficients for fixed values of the covariance parameters, and estimating the covariance given the regression coefficients, and is continued until convergence occurs.

This scheme yields consistent estimators for the regression coefficients even if the correlational structure in step 2 was misspecified (but provided the linear relationship is correctly specified). These estimators are asymptotically multivariate normally distributed with a covariance matrix that can be consistently estimated (also in case the correlational structure was misspecified) by a so-called sandwich estimate (resulting in the *robust standard errors*).

Like the GLMM, the GEE-approach can easily deal with a wide range of outcome types such as binary, categorical, count, or interval data. Unlike the constrained GLMM with a random intercept though, the correlation of out-

covariance structure may further depend, for example, on certain covariates. Assuming independent observations within dyads on the other hand, and hence the choice for an independence working correlation matrix, may lead to some small gain in efficiency in estimating the actor- and partner-effects provided the independence assumption truly holds.

[9]The ith dyad contributes a three-way product involving the partial derivative of μ_i with respect to the regression parameter, times the inverse of the dyad's variance-covariance, times the difference between the dyad's responses and their mean (see details in Appendix A1).

comes within a dyad is not restricted to be positive. The GEE-approach does not make full distributional assumptions (only the mean-variance relationship), and no likelihood-based methods as in the GLMM can be used for testing actor and partner effects for example. Instead, parameter testing can be based on Wald statistics constructed with the asymptotic normality of the estimators together with their estimated covariance matrix. A criticism often made is that the sandwich variance estimate of GEE may underestimate the variability in the parameter estimates when the number of clusters (dyads in this particular case) is small (McCaffrey & Bell, 2006), resulting in tests that have greater than nominal type 1 error rates (i.e., too liberal tests). Rotnitzky and Jewell (1990) describe an alternative procedure for testing effects of predictors, the so-called *score test*. The test statistic for this score test is based on the generalized estimating 'score-like' equations[10] that are solved to produce parameter estimates for the GEE model. Finally we note here that, while GLMMs explicitly specify the correlation, the *unstructured* working correlation (as suggested in step 2) in the GEE-approach is only a device to support estimation of the regression parameters, and no standard errors are given along these working correlations. The resulting correlations should therefore only be interpreted informally (Molenberghs & Verbeke, 2005)[11].

[10]Loosely speaking these score like equations are of similar form as the score equations derived for GLM, and the principle of the GEE score test is the same as the likelihood-based score test. More technical details can be found in Appendix A1.

[11]When the association structure is of primary interest, one should turn to some extensions of GEE. Examples of the latter are second-order extensions of GEE (GEE2) that include the marginal pairwise association as well, or alternating logistic regressions that use conditional

4 Simulations

In this section, we compare the performance of five different approaches to the estimation of actor and partner effects in the APIM and the estimation of the within-dyad correlation for Bernoulli or Poisson dyadic outcomes, which are either positively or negatively correlated, with GEEs or GLMMs:

(1) a GEE-approach with p-values for tests of fixed-effect parameters based on a robust Wald test, and using an unstructured working correlation matrix;

(2) the same GEE-approach as in (1) but with p-values based on the score test;

(3) a GLMM with a random intercept (RI), a constrained variance component ($\tau > 0$), and computation based on adaptive Gaussian quadrature;

(4) a GLMM with RI, an unconstrained variance component, and computation based on Laplace approximation;

(5) a marginalized GLMM and computation based on linearized approximation (pseudo-likelihood methods).

All simulations were performed in SAS version 9.2 and used the GENMOD procedure for (1) and (2) (with TYPE3 option in the MODEL statement for the latter), the NLMIXED procedure for (3) (with default method= adaptive Gaussian quadrature), and the GLIMMIX procedure for (4) and (5) (with method=LAPLACE and option NOBOUND for (4), and method=RSPL for

probability ideas (Molenberghs & Verbeke, 2005).

(5))[12]. A literature review of studies using the APIM revealed that sample sizes typically ranged from 30 to 300 dyads (first quartile=60, median=100, and third quartile=150)[13], but even dyadic sample sizes as small as 12 were recently reported (Tambling, Johnson, & Johnson, 2012). We therefore considered number of dyads equal to 10, 30, 60, 100, 150, or 300 in the simulation study. Results from each simulation setting are based on 2000 repetitions. It should be noted though that in case of convergence issues for a particular estimation method, estimates were not included for that approach. Such non-convergence occurred in about 15% of the cases for the marginalized GLMM when the ICC was positive or negative, and for the constrained GLMM with RI when the ICC was negative (both for small and large samples).

4.1 Correlated Bernoulli outcomes

For the simulation settings with a positive ICC, responses Y_{ij} were generated from a Bernoulli distribution with probability p_{ij} following the APIM with fixed effects for the actor's and partner's predictor and a distinguishing variable, and a random intercept:

$$\text{logit}(p_{ij} \mid b_i) = \beta_0 + \beta_{act}x_{act,ij} + \beta_{par}x_{par,ij} + \beta_{dis}x_{dis,ij} + b_i, \quad j = 1, 2, \quad (9)$$

with $x_{dis,ij}$ coded as 1 if $j = 1$ and -1 if $j = 2$, actor and partner predictors x_{act} and x_{par} generated from a standard bivariate normal distribution with cor-

[12] One may use the GLIMMIX procedure for (3) as well using the method=QUAD option.

[13] Special thanks to Robert Wickham from the university of Houston for sharing his database on the use of the APIM.

relation 0.50[14]. We set $\beta_0, \beta_{act}, \beta_{par}$ and β_{dis} equal to zero, while values of τ were chosen such that the ICCs were approximately equal to 0.30, 0.15, or 0.05. For the simulation setting with a negative intra-cluster correlation, we relied on Leisch, Weingessel, and Hornik (1988) who show how to simulate multivariate binary distributions with a given correlation structure from a multivariate normal distribution. By dichotomizing the normal variates and the appropriate choice of the correlation between normal variates, one can obtain the required marginal and pairwise probabilities. We generated binary dyadic data with marginal probabilities 0.5 and 0.5 for Y_{i1} and Y_{i2} (i.e., no actor and partner effect of the standard bivariate normal distributed x_{act} and x_{par} with correlation 0.50, and no effect of the distinguishing variable x_{dis}) and joint probability 0.175, 0.2125, and 0.2375, leading to ICCs of −0.3, −0.15, and −0.05, respectively.

For the GEE-approaches and marginalized GLMM-approach, the following working model is assumed:

$$\text{logit}(p_{ij}) = \beta_0 + \beta_{act}x_{act,ij} + \beta_{par}x_{par,ij} + \beta_{dis}x_{dis,ij}, \quad j = 1, 2, \tag{10}$$

while for the GLMM-approaches with RI, model (9) is assumed.

To ensure equal marginal and conditional parameter effects, data were first generated under the null hypothesis of no actor and no partner effect. By doing so, the size of the test of $\beta_{act} = 0$ (β_{par}=0, respectively) at the nominal 5% level under each of the five approaches can easily be assessed (note that with 2000

[14]Smaller correlations between predictors were considered in this setting, as well as in all settings described further, but did not reveal major differences from the results presented.

simulations, the standard error on the estimated size is about 0.5%, and empirical type 1 errors for appropriate tests are therefore expected to lie between 4% and 6%). Empirical type 1 errors for the test of no actor effect are presented in Figure 2 (results for the test of no partner effect were very similar). While under the GEE-approach, the robust Wald test tends to be too liberal when the numbers of dyads is extremely small, the performance of the score test is satisfactory under all settings considered (slightly conservative for small number of dyads in some cases). Both the constrained and unconstrained RI-model yield a too conservative test under positive ICC settings when the number of dyads is small. With increasing negative values of the ICC, the constrained RI-model (which will then typically force the random-intercept variance to be zero) yields a much too conservative test. The unconstrained RI-approach jumps from too conservative tests for small samples to too liberal tests for larger samples when the ICC is negative. The marginalized GLMM-approach performs relatively well in terms of type 1 error, both under positive and negative ICC scenarios, except when the number of dyads is small.

Overall, we conclude that under the null hypothesis the marginal approaches perform better than the conditional approaches in preserving the type 1 error, and we proceed for now with the former only to explore the performance in estimating the residual ICC. The upper panel of Figure 3 shows the median of the estimated ICC under both marginal approaches for the six values we considered for the ICC. Although the 'standard' GEE-approach does not formally aim to estimate the ICC, its estimate obtained from the unstructured working correlation is very informative and recovering the ICC well. In contrast, there is - regardless of the sample size - indication of a serious negative bias for the ICC estimate from the marginalized GLMM for increasing absolute values of the ICC.

Next, data were generated following model (9) assuming an effect of x_{act} and x_{par} ($\beta_1 = \log 1.5 \approx 0.405$ and $\beta_2 = \log 0.75 \approx -0.287$, respectively). Note that the marginal effects of x_{act} and x_{par} have no longer the same value as the conditional effects, but can be approximated by $\beta_1/\sqrt{c^2\tau + 1}$ and $\beta_2/\sqrt{c^2\tau + 1}$ respectively, with $c = 16\sqrt{3}/(15\pi)$ (Molenberghs & Verbeke, 2005). Because of the approximation techniques that are used for GLMMs, estimates of non-zero fixed effect in the GLMMs are known to be frequently biased[15]. The upper left panel of Figure 4 presents the mean of the estimated actor and partner effects for the scenario where the ICC equals 0.15. We found no evidence of severe bias for either the marginal or conditional effects using the five approaches, except when the sample size is extremely small. The upper middle panel of Figure 4 shows the power to detect the actor effect at the nominal 5% significance level. Not surpisingly, we find the GEE-Wald test to have highest power at lower sample sizes (as the test was seen to be too liberal). The robust score test from the GEE-approach is performing well as compared to the multilevel approaches. Finally, to shed some light on the performance of the approximation formula (5)

[15]Breslow and Lin (1995) studied the 'worst case' scenario of binary responses in a matched-pairs design and found the asymptotic bias in the pseudo-likelihood estimator of β to be of the order of $| \tau |$. The bias for the Laplace estimator is of smaller order, while adaptive quadrature leads to nearly unbiased estimated (Pinheiro & Chao, 2006).

for the ICC in the RI-model, the estimated within-dyad correlations under the 5 approaches are presented in the upper right panel of Figure 4, illustrating once more the excellent performance of the GEE-approach in recovering the ICC.

4.2 Correlated Poisson outcomes

For the simulation settings with a positive ICC, responses Y_{ij} were generated from a Poisson distribution with mean μ_{ij} following the APIM with RI

$$\log(\mu_{ij} \mid b_i) = \beta_0 + \beta_{act} x_{act,ij} + \beta_{par} x_{par,ij} + \beta_{dis} x_{dis,ij} + b_i, \quad j = 1,2 \quad (11)$$

with $\beta_0 = \beta_{act} = \beta_{par} = \beta_{dis} = 0$, $b_i \sim N(0, \tau)$ and x_{act}, x_{par} and x_{dis} as before. The number of dyads i considered was again 10, 30, 60, 100, 150, or 300. Values of τ were chosen such that ICCs approximately equal to 0.30, 0.15, or 0.05 were obtained.

For the simulation setting with a negative ICC, we extend the approach of Leisch et al. (1988) and show how to simulate multivariate Poisson distributions with a given correlation structure. We first generate samples from a bivariate standard normal distribution with correlation ρ_N. Whereas before the Gaussian random variables where dichotomized to yield binary events of 0 or 1, they will now be discretized into M different states to yield counts of $0, 1, 2, \ldots, M$. Precisely, we want to generate counts Y_{ij} that have count probabilities $\Pr(Y_{ij} = k) = p_{ijk}$. Samples are generated by discretizing a 2-dimensional normal random variable U by setting $Y_{ij} = k$ if $\gamma_{ij,k} < U_{ij} \leq \gamma_{ij,k+1}$, with $\gamma_{ij,k} = \Phi^{-1}(\Pr(Y_{ij} < k))$ for each $k = 1, 2, \ldots, M$. It can be shown that the value of ρ_N is uniquely determined by the value of the desired correlation between Poisson outcomes.

We generated bivariate Poisson outcomes with marginal means equal to 2 for Y_{i1} and Y_{i2} (i.e., no effect of x_{act}, x_{par} and x_{dis}) and pairwise correlation equal to -0.30, -0.15, and -0.05, respectively.

For the GEE-approaches and marginalized GLMM-approach, the following working model was assumed

$$\log(\mu_{ij}) = \beta_0 + \beta_{act}x_{act,ij} + \beta_{par}x_{par,ij} + \beta_{dis}x_{dis,ij}, \quad j = 1, 2, \tag{12}$$

while for the GLMM-approaches with RI, model (11) was assumed.

As data were generated first under the null hypothesis of no effect of X here too, we can again assess the size of the test of $\beta_{act} = 0$ ($\beta_{par} = 0$, respectively) at the nominal 5% level. The empirical sizes for the test of no actor effect are presented in Figure 5. While under the GEE-approach, the robust Wald test is far too liberal when the numbers of dyads is small - even more pronounced than in the Bernoulli setting - the performance of the score test is satisfactory under all settings considered, except for some conservatism in very small samples. The constrained RI-model yield a too conservative test under positive ICC settings when the number of dyads is small. With increasing negative values of the ICC, the constrained RI-model yields a way too liberal test. The unconstrained RI-approach again jumps from too conservative tests for small samples to too liberal tests for larger samples when the ICC is negative. The marginalized GLMM-approach tends to perform well in all settings (except for extremely small sample sizes). Interestingly when comparing the marginal approaches in their performance to estimate the ICC, we observe similar findings as for the Bernoulli outcomes (lower panel of Figure 3).

Next, data were generated following model (11) assuming an effect of x_{act} and x_{par} ($\beta_1 = \log 1.25 \approx 0.223$ and $\beta_2 = \log 0.85 \approx -0.163$, respectively). As shown before, the marginal effect of x_{act} and x_{par} are the same as the conditional effects for this setting. The lower left panel of Figure 4 presents the mean of the estimated actor and partner effects for the scenario where the ICC equals 0.30. We found no evidence of any bias for any of the 5 approaches, except when the sample size is extremely small. The lower middle panel of Figure 4 shows the power to detect the actor effect at the nominal 5% significance level. Not surpisingly, we find again the GEE-Wald test to have highest power at lower sample sizes (as the test was seen to be too liberal). In contrast to the setting with Bernoulli outcomes, the robust score test from the GEE-approach is performing slightly worse now in terms of power as compared to the multilevel approaches. The ICC is again well recovered from the working correlation in the GEE-approach (the lower right panel of Figure 4), better than by the approximation (7) for the GLMMs with RI and than the marginalized GLMM.

5 Examples

The two studies presented below are subsamples of the Interdisciplinary Project for the Optimization of Separation Trajectories conducted in Flanders (IPOS; www.scheidingsonderzoek.be), which is a cooperation of psychologists, lawyers, and economists from Ghent University and the University of Leuven. This research project carried out a large-scale recruitment of formerly married partners.

All couples who divorced between March 2008 and March 2009 in four major courts in Flanders were systematically approached in the waiting room to participate in a study on divorce (N = 8896). The individual respondents (i.e., not both ex-partners) willing to participate (N = 3921; response rate = 44.1%) were subsequently contacted for an interview in view of a computerized survey. To reduce the survey's length and lessen the burden on the respondents, the survey was divided into a basic intake assessment assigned to each respondent, and three different questionnaire packages (measuring emotions, parent-child relationships, or ex-partner relationships) which were randomly distributed among the participants. As the recruitment strategy did not directly target the ex-partners simultaneously, only dyadic data from about 30 ex-couples were part of the same sample for each of the questionnaire packages. Therefore results presented below should merely be seen as an illustration of the different approaches.

5.1 Correlated binary data: forcing behavior or not during negotiations in ex-couples

The first example explores the effect of feeling guilty on negotiation behavior. Negotiation behavior was assessed with the Dutch Test for Conflict Handling (DUTCH, De Dreu, Evers, Beersma, Kluwer, & Nauta, 2001). One of the subsscales of the DUTCH measures forcing behavior (e.g., "I fight for a good outcome for myself."), measured with four items to be answered on a 5-point Likert scale from 1 (*totally disagree*) to 5 (*totally agree*). For illustration purposes, participants with an average score higher than three were artificially classified here as showing forcing behavior (denoted as $Y = 1$). Out of the 29 ex-couples in total, there was one couple where both ex-partners showed forcing behavior, six cases where only the male partner showed forcing behavior, nine cases where only the female partner showed forcing behavior and 13 couples were none of the partners showed forcing behavior. Guilt was assessed with the Guilt in Separation Scale (Wietzker, Buysse, Loeys, & Brondeel, 2012) and is computed as the mean of 10 items (e.g. "I am responsible for his/her misery."), measured on a 7-point Likert scale from 1 (*never*) to 7 (*always*). Throughout the analysis we will use the mean values of guilt, with the person's own score denoted as GUILTA and his or her partner's score as GUILTP. In addition, we use gender as the distinguishing variable in the couple, denoted as SEX and effect coded as 1 for men and -1 for women.

We used both marginal approaches and conditional approaches to explore the impact of feeling guilty on forcing behavior. For the GEE-approach and the marginalized multilevel approach, we specified the following linear relation on the logit scale between showing forcing behavior and feeling guilty (as there was no evidence of gender-specific actor or partner effects no additional interactions were considered):

$$\text{logit}[E(Y_{ij})] = \beta_0 + \beta_1 * GUILTA_{ij} + \beta_2 * GUILTP_{ij} + \beta_3 * SEX_{ij}. \quad (13)$$

For the conditional multilevel approach, we considered the following RI-model

$$\text{logit}[E(Y_{ij} \mid b_i)] = \beta_0 + \beta_1 * GUILTA_{ij} + \beta_2 * GUILTP_{ij} + \beta_3 * SEX_{ij} + b_i, \quad (14)$$

with $b_i \sim N(0, \tau)$ and τ either constrained (using adaptive Gaussian quadrature for optimization) or not (using Laplace approximation for optimization). Results are presented in Table 2 (corresponding SAS-code can be found in Appendix A2). The following trends are observed: Feeling guilty is associated with a decrease of the forcing behavior, while guilt emotions of the partner show a reverse effect. Overall, men show less forcing behavior than women. We repeat the different interpretation of the conditional and marginal models here. For example, from the constrained RI-approach, we estimate that within a dyad, a one-unit increase in the guilty score of the partner, corresponds to an increase of exp(0.69) of the odds of showing forcing behavior. Marginally, we estimate with the GEE-approach that such increase is associated with an increase of exp(0.58) of that odds in the sample of ex-couples. It's worth noting here that the estimated actor and partner effects under the marginalized multilevel approach are substantially different from the marginal effects under the GEE-approach, as are the significance of the effects. This might be attributed to poor convergence of the GLMM for these particular data. The estimated intra-dyad correlation from the working GEE-correlation matrix equals -0.18. Given this indication of negative correlation in forcing behavior between ex-partners, it is therefore not surprising that the constrained RI-model (14) resulted in an estimated zero random effect variance, and some conservatism in the estimated standard errors of the predictors.

5.2 Correlated count data: the number of unwanted pursuit behaviors in ex-couples

In the second example, we focus on a sample of 33 ex-couples who responded to an adapted version of the Relational Pursuit-Pursuer Short Form (RP-PSF; Cupach & Spitzberg, 2004) used to assess the extent of UPB-perpetrations the participant showed towards the ex-partner since the break-up. The total of 28 RP-PSF items (ranging from 'leaving unwanted gifts' to 'threatening to hurt yourself'), each measured on a 5-point Likert scale from 0 (*never*) to 4 (*over 5 times*), was used as an overall index of perpetration (with higher scores indicating higher levels of perpetrations). A participant who answered *never* to all these 28 UPB-items will have an UPB-outcome equal to 0; a participant who answered *over 5 times* to 'leaving unwanted gifts' and *never* to all other items will have an UPB-total equal to 4 for example; while a participant who answered *over 5 times* to all items will have the maximum score of 112. While many predictors for the UPB-outcome were measured, we limit our attention here to the impact of the actor's and partner's level of anxious attachment in their relationship with their ex-partner before the break-up, which was measured using a total of five anxious attachment items (e.g., 'My desire to be very close sometimes scared my ex-partner away') from an adapted Experience in Close Relationships Scale-Short form (ECR-S; Wei, Russell, Mallinckrodt, & Vogel, 2007). Throughout the analysis we will use the mean-centered values of anxious attachment, with the person's own score denoted as ANXA and his or her partner's score as ANXP.

Figure 6 shows the right-skewed distribution of the observed number of UPB-perpetrations. Such count data are frequently modeled using the Poisson distribution, but the corresponding predicted frequencies in Figure 6 clearly reveal lack-of-fit here. The negative binomial distribution, relaxing the Poisson-assumption of equality of the mean and the variance, yields a much better fit and will further be assumed.

As in Example 1, we used both marginal and conditional approaches to explore the impact of anxious attachment on the number of UPBs. For the GEE-approach and the marginalized multilevel-approach, we specified the following linear relation on the logarithmic scale between the expected number of UPBs and its predictors:

$$\begin{aligned}\log[E(UPB_{ij})] &= \beta_0 + \beta_1 * ANXA_{ij} + \beta_2 * ANXP_{ij} + \beta_3 * SEX_{ij} \\ &\quad + \beta_4 * ANXP_{ij} * SEX_{ij} + \beta_5 * ANXP_{ij} * SEX_{ij}, \end{aligned} \quad (15)$$

while for the conditional multilevel-approach, we considered the following a RI-model

$$\begin{aligned}\log[E(UPB_{ij} \mid b_i)] &= \beta_0 + \beta_1 * ANXA_{ij} + \beta_2 * ANXP_{ij} + \beta_3 * SEX_{ij} \quad (16) \\ &\quad + \beta_4 * ANXP_{ij} * SEX_{ij} + \beta_5 * ANXP_{ij} * SEX_{ij} + b_i, \end{aligned}$$

with $b_i \sim N(0, \tau)$. Estimated parameters under different estimation methods are presented in Table 3 (the corresponding SAS code can be found in Appendix A3). The estimated ICC from the working correlation in GEE equals 0.07. Because of the linearity of the random effect on the log-scale, the conditional and marginal approaches lead to the same interpretation of parameters. Differences between the five approaches are smaller than in Example 1 (except again for the marginalized multilevel approach). Both for male and female actors, we observe an increase in the expected number of UPBs for increasing levels of the anxious attachment level of the actor. In contrast, while increasing anxious attachment levels of the male partner before the break-up is associated with an increase in the number of UPBs in women, the reverse trend is observed in men.

6 Discussion

While LMMs have frequently been used to model Gaussian dyadic outcomes, we have shown in this paper that GLMMs might not always be the best option to model non-Gaussian dyadic outcomes. This becomes especially true when the correlation between outcomes in a dyad is negative and/or the sample size is small. We explored the performance of different estimation techniques within the GLMM-framework, along with their potential to allow for negative ICCs, but found none of these to be overall satisfactory. While the marginalized GLMM performed relatively well with respect to estimating actor and partner effects in settings with either negative or positive within-dyad correlation, the latter is poorly estimated under such an approach. The GEE-approach, which is relatively unused within the social sciences[16], offers an interesting alternative in

[16]One of the reasons for being less popular might be that GEE is only valid under missing completely at random (MCAR) and covariate-dependent missingness, while GLMM is valid under the less restrictive missing at random (MAR) assumption. While it is common practice to exclude couples with the outcome of one the partners missing, it is possible to opt for an analysis where also incomplete information is used. In a likelihood context, this renders the

this context. The robust Wald test of GEE turned out to perform well, except for (extremely) small samples where the score test can be used instead.

Besides the LMM-framework, the structural equation modeling (SEM) framework is frequently used to analyze Gaussian dyadic data too (Newsom, 2002). In contrast to the LMM-framework, the SEM-framework easily allows for formal tests of goodness of fit[17], for mediational models (see Ledermann, Macho, & Kenny, 2011 for mediation in dyadic data) and for latent variables. When dealing with binary or ordinal response scales, SEM typically assumes that these data represent categorizations of underlying continuous variables [18]. The relationships of these underlying continuous variables are captured in a polychoric correlation matrix, and (robust) weighted least square estimation could be used for the parameters of the marginal model matching the GEE-model. Such an approach will be theoretically reasonable only in some cases. While for attitude items, the researcher may be more interested in the relationships among the continuous underlying latent variables than in the relationship beween the observed *agree* and *strongly agree* responses on the items; it may be difficult for

resulting analysis valid under the assumption that the missing data are missing at random. However, for dyadic data analysis, missingness on both the outcome and covariates for one member of the dyad, will not allow to estimate partner effects (for the other member of the dyad) anyway, unless one is willing to use an imputation-based method.

[17]It should be noted that using SEM to estimate the APIM with distinguishable dyads allowing for 'gender-specific' actor and 'gender-specific' partner effects is a saturated model and so it has zero degrees of freedom and no measures of fit can be computed.

[18]More precisely, limited information methods make this 'underlying continuous variable' assumption. Full information methods alternatively model the entire multivariate categorical distribution of the observed variables.

other variables like current drug user (*yes* or *no*) to conceive them as realizations of an underlying continuous variable. Moreover, for count outcomes such an approach does not work either. Interestingly, Muthén, du Toit, and Spisic (1997) compared the performance of robust WLS and (the second order extension of) GEE for binary outcomes in a longitudinal simulation setting, and found superior behavior of GEE in settings with 200 or 400 observation units, especially when the prevalence of the outcome is small, but more comparable behavior in larger samples that are less frequently seen in a dyadic context though.

On the other hand, the flexibility of SEM to deal with latent variables should not be neglected. Within the dyadic modeling world this might not only be an important asset for the APIM discussed here, but even more so for estimation in the CFM (right panel of Figure 1). The latter is indeed most easily seen as a latent variable dyadic model, and SEM the most natural framework to disentangle variability at the dyadic and at the subject level. While Gonzalez and Griffin (2002) showed how the CFM with distinguishable dyads can be casted within the multilevel framework too with the common-fate variables conceptualized as random intercepts, the CFM can not be tackled with the marginal approach taken by GEE.

To conclude, this article has shown the merits of the GEE-approach for estimating actor and partner effects in the wide range of typical dyadic sample sizes. Indeed, by expanding the types of data that can easily be analyzed with the APIM, its straightforward allowance for both positive and negative within-dyad correlations and its ease of implementation, GEE can add significantly to the toolbox of relationship researchers everywhere. As a rule thumb we suggest the use of the GEE score test for drawing inference about the actor and partner effects when the number of dyads is smaller than 50, while the robust Wald test may be used for larger samples. Multilevel approaches with a random intercept capturing the correlation within dyads are not recommended for estimating actor and partner effects from categorical dyadic data, especially when the sample size is small or the within-dyad correlation is negative. The marginalized multilevel approach on the other hand typically works well for estimating actor and partner effects, but shows serious negative bias in estimating the ICC. Although the GEE-approach treats the within-dyad correlation as nuisance, its estimate for the ICC from the unstructured working correlation turns out to be informative here. If formal inference about the ICC is needed though, GEE-extensions allowing for this are available (Molenberghs & Verbeke, 2005), but these are less commonly available in standard software packages.

References

Agresti, A. (2000). *Categorical Data Analysis*, NJ: Wiley.

Bauer, D. J., & Sterba, S. K. (2011). Fitting multilevel models with ordinal outcomes: performance of alternative specifications and methods of estimation. *Psychological Methods, 16,* 373-390.

Breslow, N. E., & Lin, X. (1995). Bias correction in generalised linear mixed models with a single component of dispersion. *Biometrika, 82,* 81-91.

Cook, W. L. (2001). Interpersonal influence in families: A social relations model analysis. *Child Development, 72,* 1179-1197.

Cupach, W. R., & Spitzberg, B. H. (2004). *The dark side of relationship pursuit. From attraction to obsession and stalking.* Mahwah, NJ: Lawrence Erlbaum Associates.

De Dreu, C. K. W., Evers, A., Beersma, B., Kluwer, E. S., & Nauta, A. (2001). A theory-based measure of conflict management strategies in the workplace. *Journal of Organizational Behavior, 22(6),* 645-668.

Diggle, P. J., Heagerty, P., Liang, K.-Y., & Zeger, S. L. (2002). *Longitudinal data analysis.* Oxford, UK: Oxford University Press.

Ghisletta, P., & Spini, D. (2004). An introduction to generalized estimating equations and an application to assess selectivity effects in a longitudinal study on very old individuals. *Journal of Educational and Behavioral Statistics, 29,* 421-437.

Gonzalez, R., & Griffin, D. (2002). Modeling the personality of dyads and groups. *Journal of Personality, 70,* 901-924.

Griswold, M. E., & Zeger, S. L. (2004). On marginalized multilevel models and their computation. *Johns Hopkins University, Department of Biostatistics Working Paper #99.*

Heagerty, P. J. (1999). Marginally specified logistic-normal models for longitudinal binary data *Biometrics, 55,* 688–698

Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis.* New York: Guilford Press.

Kenny, D. A., & Ledermann, T. (2012). Bibliography of actor-partner interdependence model. Retrieved from http://davidakenny.net/doc/apimbiblio.pdf

Ledermann, T., & Kenny, D. A. (2012). The common fate model for dyadic data: variations of a theoretically important but underutilized model. *Journal of Family Psychology, 26,* 140-148.

Ledermann, T., Macho, S., & Kenny, D. A. (2011). Assessing mediation in dyadic data using the actor-partner interdependence model. *Structural Equation Modeling: A Multidisciplinary Journal, 18,* 595-612.

Leisch, F., Weingessel, A., & Hornik, K. (1988). On the generation of correlated artificial binary data. *Journal of Statistical Planning and Inference, 20,* 131-154.

Liang, K.-Y., & Zeger, S. L. (1986). Longitudinal Data Analysis using Generalized Linear Models. *Biometrika, 73,* 13-22.

Loeys, T., Moerkerke, B., De Smet, O., & Buysse, A. (2012) The analysis of zero-inflated count data: beyond zero-inflated Poisson regression. *British Journal of Mathematical and Statistical Psychology, 65,* 163-180.

McCaffrey, D., & Bell, R. M. (2006). Improved hypothesis testing for coefficients in generalized estimating equations with small samples of clusters. *Statistics in Medicine, 25,* 4081-4098.

McMahon, J. M., Pouget, E. R., & Tortu, S. (2006). A guide for multilevel modeling of dyadic data with binary outcomes using SAS PROC NLMIXED. *Computational Statistics and Data Analysis, 50,* 3663-3680.

Molenberghs, G., & Verbeke, G. (2005). *Models for discrete longitudinal data.* New York: Springer.

Molenberghs, G., Verbeke, G., & Demétrio, C. G. B. (2007). An extended random effects approach to modeling repeated, overdispersed count data. *Lifetime Data Analysis, 13,* 513-531.

Molenberghs, G., & Verbeke, G. (2011). A note on a hierarchical interpretation for negative variance components. *Statistical Modelling, 11,* 389-408.

Muthén, B. O., du Toit, S. H. C., & Spisic, D. (1997). Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes. Unpublished manuscript.

Nelder, J. A., & Wedderburn, R. W. M. (1972). Generalized linear models. *Journal of the Royal Statistical Society, Series A, 135,* 370-384.

Newsom, J. T. (2002). A multilevel structural equation model for dyadic data. *Structural Equation Modeling: A Multidisciplinary Journal, 9,* 431-447.

Pinheiro, J. C., & Chao, E. C. (2006). Efficient Laplacian and Adaptive Gaussian Quadrature Algorithms for Multilevel Generalized Linear Mixed Models. *Journal of Computational and Graphical Statistics, 15,* 5881.

Pryseley, A., Tchonlafi, C., Verbeke, G., & Molenberghs, G. (2011). Estimating negative variance components from Gaussian and non-Gaussian data: a mixed model approach. *Computational Statistics and Data Analysis, 55,* 1071-1085.

Rotnitzky, A., & Jewell, N. P. (1990). Hypothesis testing of regression parameters in semiparametric generalized linear models for cluster data. *Biometrika, 77,* 485-489.

Snijders, T., & Bosker, R. (1999). *Multilevel Analysis. An introduction to basic and advanced multilevel modeling.* Sage: London.

Spain, S. M., Jackson, J. J., & Edmonds, G. W. (2012). Extending the actor-partner interdependence model for binary outcomes: a multilevel logistic approach. *Personal Relationships.* DOI: 10.1111/j.1475-6811.2011.01371.x

Tambling, R. B., Johnson, S. K., & Johnson, L. N. (2011). Analyzing dyadic data from small samples: a pooled regression actor-partner interdependence model approach. *Counseling Outcome Research and Evalulation, 2,* 101-114.

Tuerlinckx, F., Rijmen, F., Verbeke, G., & De Boeck, P. (2006). Statistical inference in generalized linear mixed models: a review. *British Journal of Mathematical and Statistical Psychology, 59,* 225-255.

Wei, M., Russell, D. W., Mallinckrodt, B., & Vogel, D. L. (2007). The experiences in close relationship scale (ECR)-short form: reliability, validity, and factor structure. *Journal of Personality Assessment, 88,* 187-204.

Wietzker, A., Buysse, A., Loeys, T., & Brondeel, R. (2012). Easing the Conscience-Feeling Guilty Makes People Cooperate in Divorce Negotiations. *Journal of Social and Personal Relationships, 29,* 324-336.

Appendix A1.

Generalized Estimating Equations

Let $\mathbf{Y}_i$ denote a 2-dimensional vector of measurements (with components Y_{i1} and Y_{i2}) available for dyad $i = 1, \ldots, N$. The marginal mean $\mu_{ij} = E(Y_{ij})$ is related to the explanatory variables $\mathbf{x}_{ij}^t$ by the following expression

$$g(\mu_{ij}) = \mathbf{x}_{ij}^t \boldsymbol{\beta}$$

with g a known link function. The variance $\mathbf{V}_i = \mathrm{Var}(\mathbf{Y}_i)$ equals $\phi \mathbf{A}_i^{1/2} \mathbf{R_i} \mathbf{A_i}^{1/2}$, where the matrix $\mathbf{A_i}$ is a diagonal matrix containing the variance function of the model (with the variance function $v(\cdot)$ describing how the variances along the diagonal of $\mathbf{A_i}$ depend on the mean), ϕ a possibly unknown scale parameter (equal to 1 if no overdispersion is assumed). $\mathbf{R}_i$ a correlation matrix whose structure is generally unknown, but for which a working correlation matrix $\mathbf{C}(\alpha)$ is assumed under the GEE-approach. The Generalized Estimating Equation of Liang and Zeger (1986) for estimating the vector of regression parameters $\boldsymbol{\beta}$ is then given by

$$\mathbf{S}(\boldsymbol{\beta}) = \sum_{i=1}^{N} \mathbf{D}_i^t \mathbf{V}_i^{-1} \left(\mathbf{Y}_i - \boldsymbol{\mu}_i(\boldsymbol{\beta})\right) = 0$$

where $\mathbf{D}_i = \partial \boldsymbol{\mu}_i / \partial \boldsymbol{\beta}$.

The robust Wald test

The *model-based* estimator of the covariance of $\hat{\boldsymbol{\beta}}$ is given by $\boldsymbol{\Sigma}_m(\hat{\boldsymbol{\beta}}) = \mathbf{I}_0^{-1}$, where

$$\mathbf{I}_0 = \sum_{i=1}^{N} \frac{\partial \boldsymbol{\mu}_i^t}{\partial \boldsymbol{\beta}} \mathbf{V}_i^{-1} \frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}}$$

which is the GEE equivalent of the inverse of the Fisher information matrix that is often used in generalized linear models as an estimator for the covariance estimate of the maximum likelihood estimator of $\boldsymbol{\beta}$.

The *robust* (or sandwich) estimator of the covariance of $\hat{\boldsymbol{\beta}}$ is given by $\boldsymbol{\Sigma}_r(\hat{\boldsymbol{\beta}}) = \mathbf{I}_0^{-1}\mathbf{I}_1\mathbf{I}_0^{-1}$, with

$$\mathbf{I}_1 = \sum_{i=1}^{N} \frac{\partial \boldsymbol{\mu}_i^t}{\partial \boldsymbol{\beta}} \mathbf{V}_i^{-1} \mathrm{Cov}(\mathbf{Y}_i) \mathbf{V}_i^{-1} \frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}}$$

where $\mathrm{Cov}(\mathbf{Y}_i)$ is estimated by $\left(\mathbf{Y}_i - \boldsymbol{\mu}_i(\hat{\boldsymbol{\beta}})\right)\left(\mathbf{Y}_i - \boldsymbol{\mu}_i(\hat{\boldsymbol{\beta}})\right)^t$.

The score test

Let $\boldsymbol{\beta}_1$ denote a subset of the parameter vector $\boldsymbol{\beta}$ and consider testing $H_0: \boldsymbol{\beta}_1 = 0$. Further, let $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)$ and $\tilde{\boldsymbol{\beta}} = (0, \tilde{\boldsymbol{\beta}}_2)$ be the regression parameter vector resulting from solving the GEE in the restricted parameter space under H_0. The score test statistic is then given by $S(\tilde{\boldsymbol{\beta}})^t \boldsymbol{\Sigma}_m \boldsymbol{\Sigma}_r^{-1} \boldsymbol{\Sigma}_m S(\tilde{\boldsymbol{\beta}})$.

Appendix A2.

```
* GEE with binary outcome FORCING2*;
proc genmod data=couple1 descending;
class GENDER ID;
model FORCING2=GUILT_A GUILT_P SEX/D=binomial link=logit type3;
repeated subject=ID/type=un withinsubject=GENDER corrw;
run;

* constrained random intercept *;
proc nlmixed data=couple1;
parms beta0=-1.01 beta1=-0.61 beta2=0.59 beta3=-0.85 s2u=1;
eta=beta0+beta1*GUILT_A+beta2*GUILT_P+beta3*SEX+u;
mu=exp(eta)/(1+exp(eta));
model FORCING2~binary(mu);
random u~normal(0,s2u) subject=ID;
run;

/* Alternatively one may use the GLIMMIX procedure
proc glimmix data=couple1 method=quad;
model FORCING2=GUILT_A GUILT_P SEX/dist=bin link=logit s;
random intercept/subject=ID;
run;
*/

* unconstrained random intercept *;
proc glimmix data=couple1 method=laplace nobound;
model FORCING2=GUILT_A GUILT_P SEX/dist=bin link=logit s;
random intercept/subject=ID;
run;

* marginalized multilevel model *;
proc glimmix data=couple1 method=RSPL;
model FORCING2=GUILT_A GUILT_P SEX/dist=bin link=logit s;
random _residual_/subject=ID type=un VCORR;
run;
```

Appendix A3.

```
* GEE with count outcome UPB *;
proc genmod data=couple2;
class ID GENDER;
model UPB=ANX_A ANX_P SEX SEX*ANX_A SEX*ANX_P/D=nb link=log TYPE3;
REPEATED SUBJECT=ID/type=UN withinsubject=GENDER corrw;
run;

* constrained random intercept *;
proc nlmixed data=couple2;
parms b0=0, b1=0, b2=0, b3=0, b4=0, b5=0, k=4,s2u=0.1;
linp =b0+b1*ANX_A+b2*ANX_P+b3*SEX+b4*SEX*ANX_A+b5*SEX*ANX_P+u;
mu = exp(linp);
p  = 1/(1+mu*k);
model UPB ~ negbin(1/k,p);
random u~normal(0,s2u) subject=ID;
run;

/* Alternatively one may use the GLIMMIX procedure
proc glimmix data=couple2 method=quad;
model UPB = ANX_A ANX_P SEX SEX*ANX_A SEX*ANX_P / dist=negbin s;
random intercept/subject=ID;
run;
*/

* unconstrained random intercept *;
proc glimmix data=couple2 method=laplace nobound;
model UPB = ANX_A ANX_P SEX SEX*ANX_A SEX*ANX_P / dist=negbin s;
random intercept/subject=ID;
run;

* marginalized multilevel model *;
proc glimmix data=couple2 method=RSPL;
model UPB = ANX_A ANX_P SEX SEX*ANX_A SEX*ANX_P / dist=negbin s;
random _residual_/subject=ID type=un VCORR;
run;
```

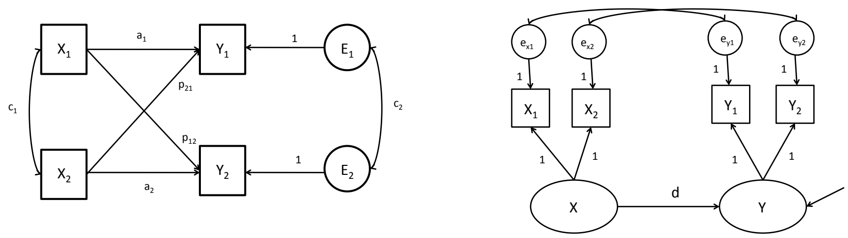

Figure 1: Left panel: The Actor Partner Interdependence Model for distinguishable dyads where a is the actor effect and p is the partner effect. Right panel: The Common Fate Model where d is direct effect.

outcome	Gaussian	Bernoulli	Poisson
link	identitiy	logistic	logarithm
$E(Y_{ij})$	$\mathbf{x}_{ij}^t\boldsymbol{\beta}$	$E\left[\frac{\exp(\mathbf{x}_{ij}^t\boldsymbol{\beta}+b_i)}{1+\exp(\mathbf{x}_{ij}^t\boldsymbol{\beta}+b_i)}\right]$	$\exp(\mathbf{x}_{ij}^t\boldsymbol{\beta}+\tau/2)$
		$\neq \frac{\exp(\mathbf{x}_{ij}^t\boldsymbol{\beta})}{1+\exp(\mathbf{x}_{ij}^t\boldsymbol{\beta})}$	
$\mathrm{covar}(Y_{i1}, Y_{i2})$	τ	$\mathrm{cov}\left[\mathrm{expit}(\mathbf{x}_{i1}^t\boldsymbol{\beta}+b_i), \mathrm{expit}(\mathbf{x}_{i2}^t\boldsymbol{\beta}+b_i)\right]$	$\exp(\mathbf{x}_{i1}^t+\mathbf{x}_{i2}^t)\boldsymbol{\beta}\ \left[\exp(\tau)\left(\exp(\tau)-1\right)\right]$

Table 1: Random intercept models (with random intercept $b_i \sim N(0,\tau)$): marginal mean and covariance.

	GEE		Multilevel		
	WALD	SCORE	RI CONSTR.	RI UNCONSTR.	MARGINALIZED
GUIA	-0.61 (0.38)		-0.74 (0.38)	-0.53 (0.30)	-2.20 (1.02)
	p=0.103	p=0.043	p=0.058	p=0.093	p=0.671
GUIP	0.58 (0.27)		0.69 (0.38)	0.73 (0.30)	0.82 (0.44)
	p=0.030	p=0.032	p=0.076	p=0.022	p=0.074
SEX	-0.85 (0.41)		-1.11 (0.65)	-0.67 (0.40)	-1.65 (3.20)
	p=0.039	p=0.034	p=0.100	p=0.104	p=0.609

Table 2: Example 1: the effect of feeling guilty on forcing behavior: comparison of five estimation/modeling methods. Estimated parameters (with standard errors) and corresponding p-values are presented.

	GEE		Multilevel		
	WALD	SCORE	RI CONSTR.	RI UNCONSTR.	MARGINALIZED
ANXA	0.122 (0.021)		0.123 (0.042)	0.125 (0.039)	0.181 (0.032)
	$p \leq 0.001$	p=0.005	p=0.007	p=0.004	$p \leq 0.001$
ANXP	-0.035 (0.022)		-0.042 (0.044)	-0.039 (0.041)	-0.085 (0.032)
	p=0.120	p=0.298	p=0.350	p=0.343	p=0.012
SEX	-0.510 (0.226)		-0.540 (0.293)	-0.549 (0.275)	-0.706 (0.218)
	p=0.024	p=0.062	p=0.075	p=0.055	p=0.003
SEX*ANXA	0.041 (0.018)		0.052 (0.043)	0.036 (0.039)	0.093 (0.033)
	p=0.027	p=0.121	p=0.234	p=0.368	p=0.008
SEX*ANXP	-0.074 (0.024)		-0.074 (0.042)	-0.075 (0.041)	-0.135 (0.033)
	p=0.002	p=0.048	p=0.091	p=0.076	$p \leq 0.001$

Table 3: Example 2: the effect of anxious attachment on unwanted pursuit behavior: comparison of five estimation/modeling methods. Estimated parameters (with standard errors) and corresponding p-values are presented.

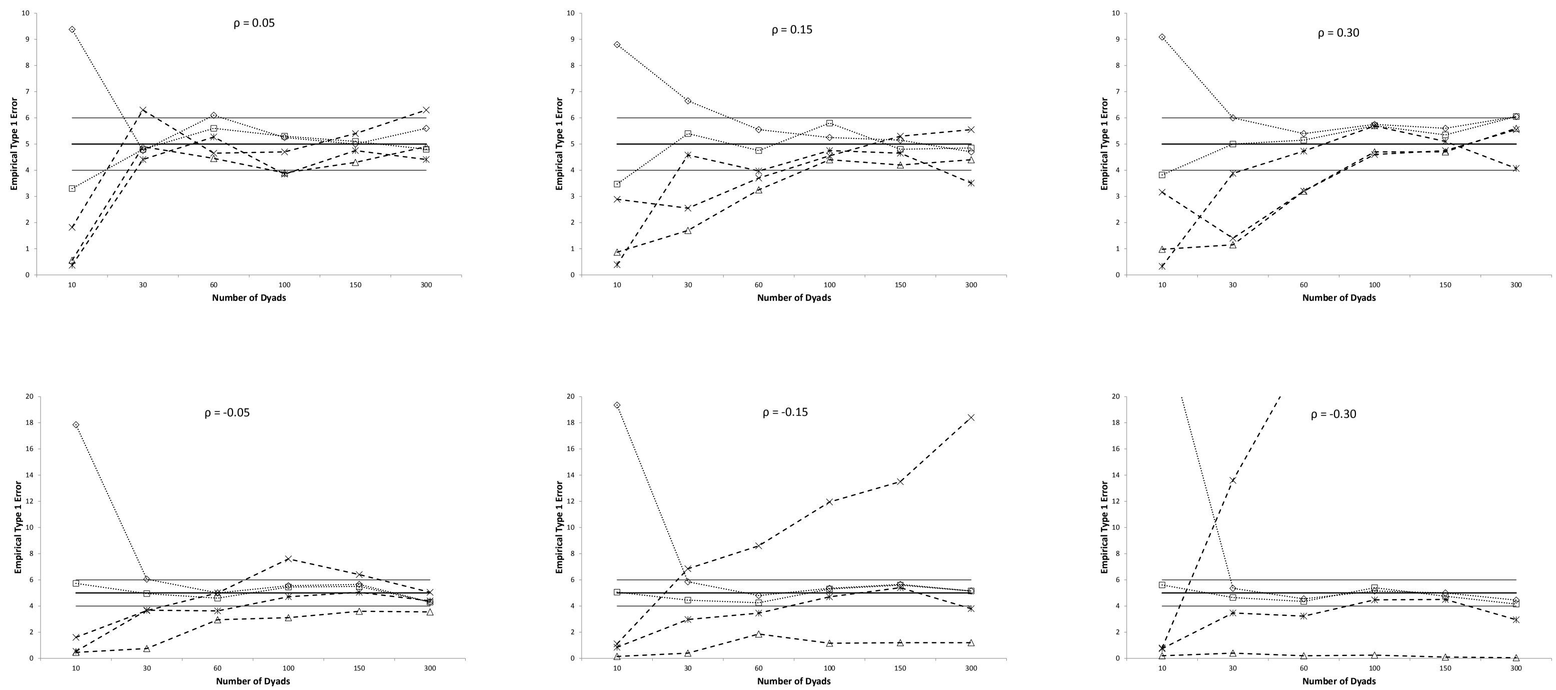

Figure 2: Performance under the APIM for binary outcome with GEE (dotted lines) and GLMM (dashed lines) under the null hypothesis of no actor effect: empirical type 1 error of the test for an actor effect at the nominal 5% level of the GEE-Wald ($\diamond$), GEE-Score ($\square$), Multilevel Constrained Random Intercept ($\triangle$), Multilevel Unconstrained Random Intercept ($\times$) and Multilevel Marginalized ($*$). Note that under the scenario $\rho = -0.30$ the type 1 error exceeds 20% for the GEE-Wald test when $N = 10$ and for the Multilevel Unconstrained Random Intercept Approach when $N \geq 60$.

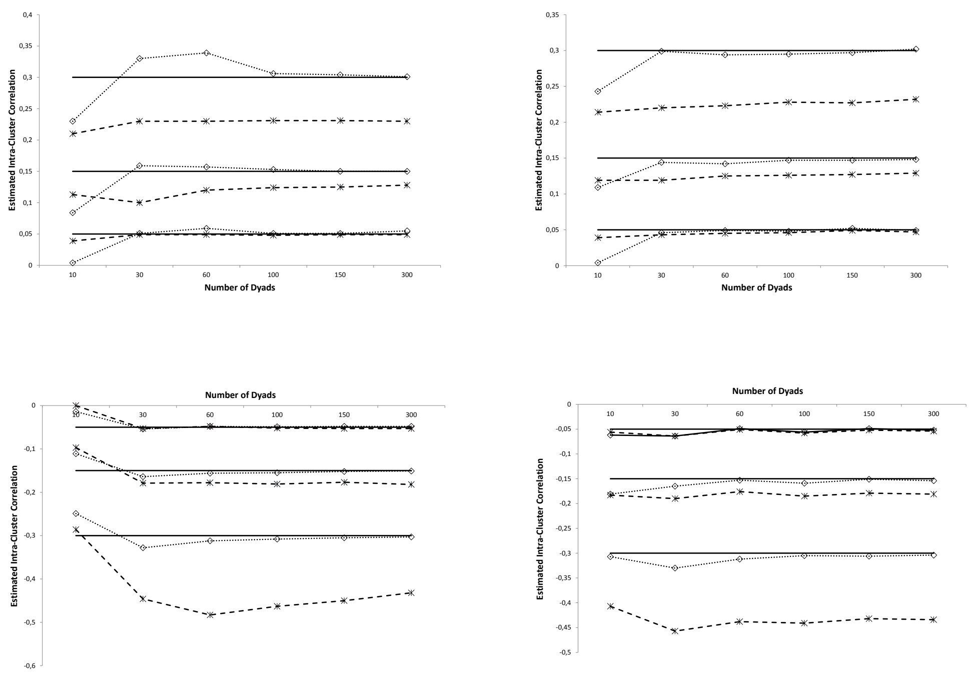

Figure 3: Estimation of the ICC (true ICC=solid line) in the APIM. Each of the 4 panels presents a setting with a low, medium and high ICC (0.05, 0.15 and 0.30 in absolute value, respectively). The upper panels consider positive ICCs and the lower panels negative ICCs. The left and right panels correspond to settings with Bernoulli and count outcomes, respectively. Estimated ICCs are based on the working correlation under the GEE-approach (dotted line with $\diamond$) and on the correlation of the pseudo-residuals under the marginalized multilevel approach (dashed line with $*$). Note that the GEE-Wald and GEE-score approach yield the same ICC.

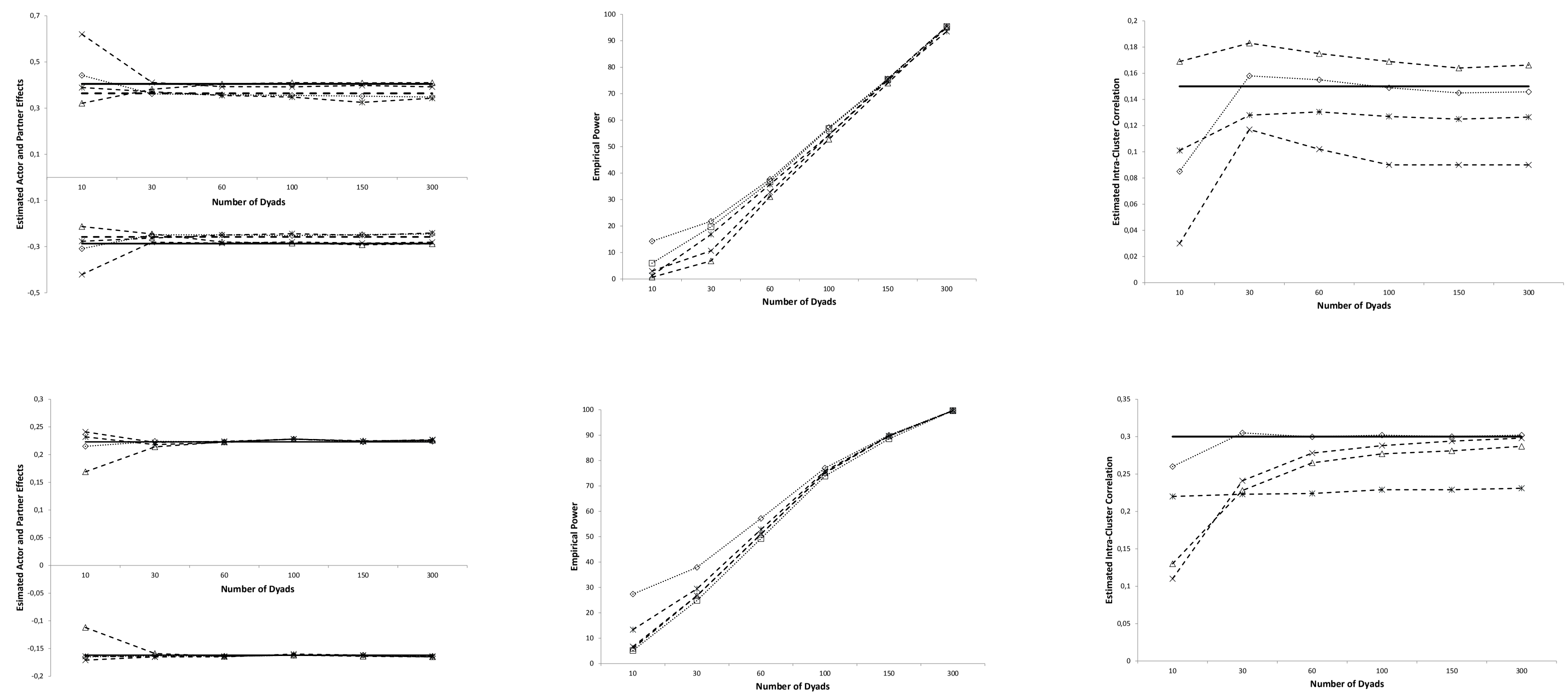

Figure 4: Performance under the APIM for Bernoulli outcome (upper panel) and count outcome (lower panel) in the presence of an actor and partner effect for the GEE-Wald (◇), GEE-Score (□), Multilevel Constrained Random Intercept (△), Multilevel Unconstrained Random Intercept (×) and Multilevel Marginalized (∗) approach. The left panel shows the mean estimated actor (positive) and partner (negative) effect (solid black represents the true marginal effect, while the dashed line represents the conditional effect). The middle panel presents the empirical power for the test of no actor effect as a function of sample size. The right panel presents the estimated ICC (the black solid line corresponds to the true ICC). Note that the GEE-Wald and GEE-score approach yield the same estimated parameters and the same estimated ICC.

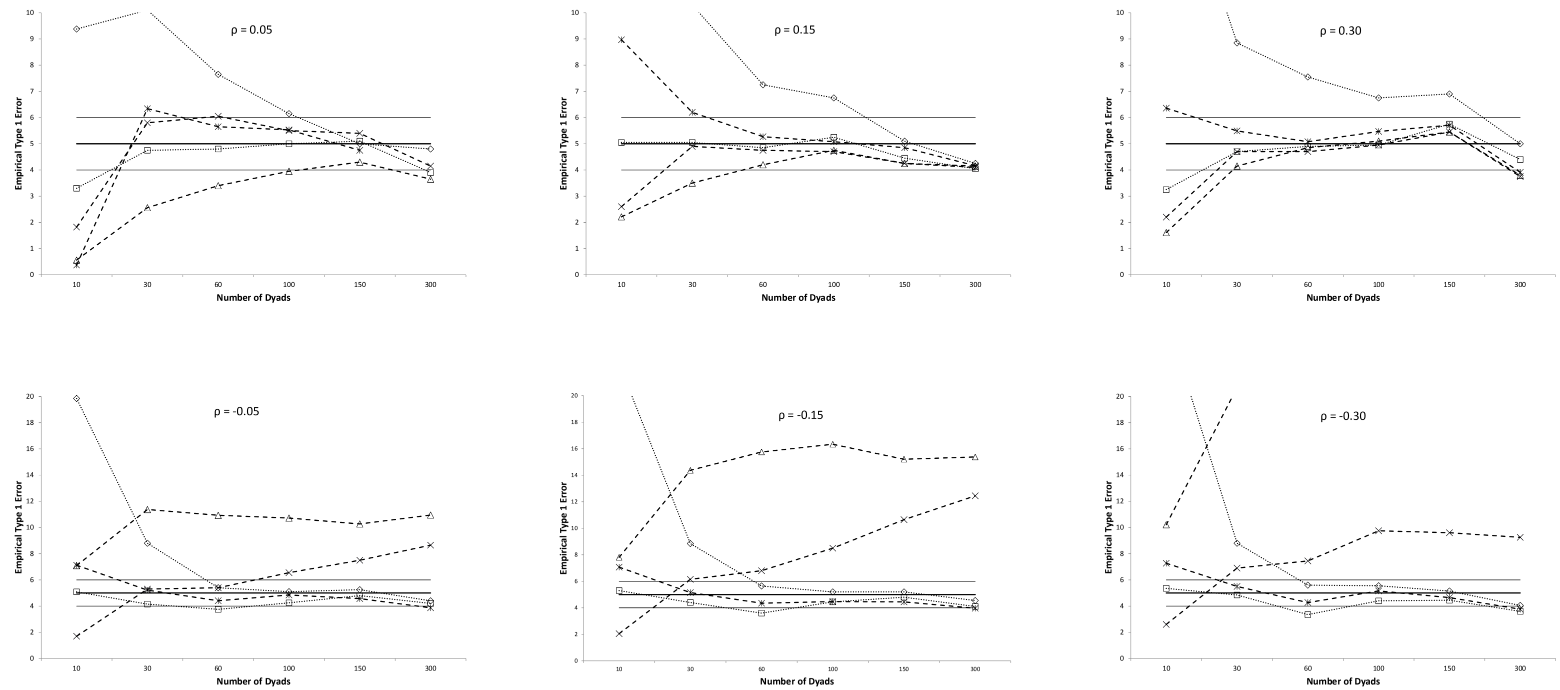

Figure 5: Performance under the APIM for count outcome with GEE (dotted lines) and GLMM (dashed lines) under the null hypothesis of no actor effect: empirical type 1 error of the test for an actor effect at the nominal 5% level of the GEE-Wald ($\diamond$), GEE-Score ($\Box$), Multilevel Constrained Random Intercept ($\triangle$), Multilevel Unconstrained Random Intercept ($\times$) and Multilevel Marginalized ($*$). Note that the type 1 error often exceeds 10% for the GEE-Wald test when N is small for positive ρ and 20% for negative ρ. Similarly the type 1 error exceeds 20% for the Multilevel Unconstrained Random Intercept Approach when $N \geq 30$ under the scenario $\rho = -0.30$.

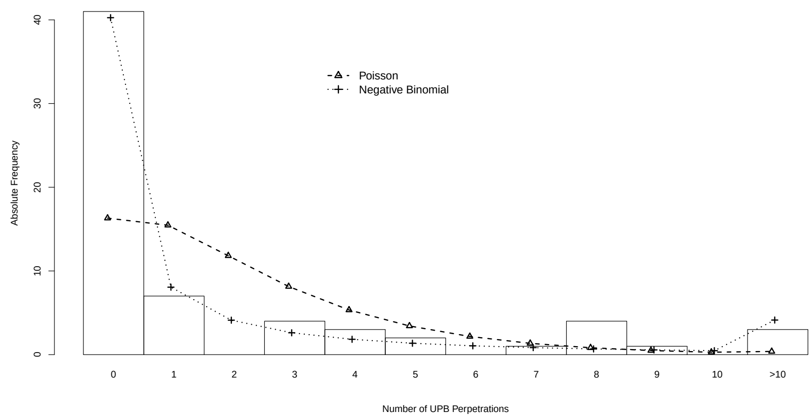

Figure 6: The observed distribution of the number of unwanted pursuit behaviors in the 33 ex-couples (with the expected distribution under a Poisson and negative binomial distribution).