DIW Berlin

German Institute
for Economic Research

Discussion Papers

291

Martin Spiess and
Gerhard Tutz

Alternative measures of the explanatory power of multivariate probit models with continuous or ordinal responses

Berlin, July 2002

Opinions expressed in this paper are those of the author and do not necessarily reflect views of the Institute.

DIW Berlin

German Institute
for Economic Research

Königin-Luise-Str. 5
14195 Berlin,
Germany

Phone +49-30-897 89-0
Fax +49-30-897 89-200

www.diw.de

ISSN 1619-4535

Alternative measures of the explanatory power of multivariate probit models with continuous or ordinal responses

Martin Spiess[a] and Gerhard Tutz[b]

Abstract

In this paper R^2-type measures of the explanatory power of multivariate linear and categorical probit models proposed in the literature are reviewed and their deficiencies are discussed. It is argued that a measure of the explanatory power should take into account the components which are explicitely modeled when a regression model is estimated while it should be indifferent to components not explicitely modeled. Based on this view three different measures for multivariate probit models are proposed. Results of a simulation study are presented designed to compare two measures in various situations and evaluate the BC_a bootstrap technique for testing the hypothesis that the corresponding measure is zero and to calculate approximate confidence intervals. The BC_a bootstrap technique turned out to work quite well for a wide range of situations, but may lead to misleading results if the true values of the corresponding measure is close to zero.

Key words: Pseudo-R^2; Measure of explanatory power; Multivariate probit model; Panel model; Simulation study; Bootstrap confidence intervals
JEL classification: C15, C33, C35, C52

[a]DIW Berlin, GSOEP, Koenigin-Luise-Str. 5, 14195 Berlin, Germany. Part of this work was supported by the programme 'Improving Human Potential (IHP) and the Socio-economic Knowledge Base — Enhancing Access to Research Infrastructures'.

[b]Department of Statistics, Ludwig–Maximilians–Universität Munich, München, Germany. Part of this work was supported by the Deutsche Forschungsgemeinschaft, SFB 386.

1 Introduction

For univariate linear regression models, the coefficient of determination is routinely printed out from most statistical software packages and can be interpreted as a measure of the explanatory power of the estimated model. However, in contrast to the univariate case, where the concepts of total and 'explained' variation are rather obvious, problems arise in the case of multivariate or panel models with continuous or ordered categorical responses.

Based on Hooper's R_T^2 and a pseudo-R^2 measure, the sample variant of which was proposed by McKelvey and Zavoina (1975) for univariate ordinal models, Spiess and Keller (1999) proposed a pseudo-R_T^2 measure to assess the explanatory power in panel or multivariate probit models. This measure takes values between zero and one. It is zero if all the regression parameters are zero and increases if the values of the model parameters increase. In a simulation study, Spiess (2001) found pseudo-R_T^2 to have the desirable property of having larger values if the modeled and estimated correlation structure is closer to the true correlation structure. However, in this simulation study all the covariates were generated independently over measurement points.

In the present paper it is shown that the above as well as some other measures proposed in the literature for multivariate linear models are functions not only of the systematic or covariate part of the model, the variance of the noise variable and the correlations between noise variables of different measurement points but also of the correlations between the covariates of the different measurement points. The latter components generally are not modeled when a regression model is estimated and are of no further interest. It is argued that a measure of the explanatory power should take into account all the components explicitely modeled when a regression model is estimated and at the same time should be indifferent to components not explicitely modeled. Based on this starting point three different measures of the exaplanatory power for multivariate probit models are proposed.

This paper is organized as follows. In Section 2 measures proposed in the literature for multivariate linear models are reviewed. In Section 3 their deficiencies are shown and discussed. Alternative measures for linear models to overcome these deficiencies are proposed in Section 4 and are generalized to multivariate probit or panel models with continuous or ordered categorical responses in Section 5. The results of a simulation study designed to compare two of the three proposed measures and to evaluate the *bias corrected and accelerated* (BC_a) bootstrap technique (Efron and Tibshirani, 1993) as a means to test the hypothesis that the true value of the corresponding measure is in fact zero or to calculate approximate confidence intervals are given in Section 6. The paper ends with some conluding remarks in Section 7. Finally, in the Appendix, the resampling method which was used in the simulation study is described.

2 R^2 for linear multivariate models

First, the population case is considered. Let $y = (y_1, \ldots, y_T)'$ denote the $(T \times 1)$ dimensional response or dependent variable and x the $(P \times 1)$ dimensional vector of random covariates where the first element is a degenerate random variable which equals one. Define $\Sigma_y = E(yy') - E(y)E(y)'$, $\Sigma_x = E(xx') - E(x)E(x)'$ and $\Sigma_{xy} = E(xy') - E(x)E(y)'$. The multivariate linear model considered has the form

$$y = \Pi' x + \epsilon, \tag{1}$$

where $\eta = \Pi' x$ is the linear predictor determined by the $(P \times T)$ parameter matrix Π and ϵ represents the $(T \times 1)$ dimensional noise variable with $\epsilon \sim N(0, \Sigma_\epsilon)$. Moreover, it is assumed that x and ϵ are independent.

Based on model (1), the matrix Σ_y may be written as

$$\Sigma_y = \Pi' \, \Sigma_x \, \Pi + \Sigma_\epsilon, \tag{2}$$

where we assume that Σ_y and Σ_ϵ are positive definite. Note that special cases of model (1) are the 'classical' multivariate regression model, where all the elements of Π are unrestricted and the covariates are often assumed to be normally distributed, a panel model where the regression parameters weighting the covariates are restricted to be identical over all T equations and a special case of the seemingly unrelated regressions (SUR) model, mainly used in econometrics. The seemingly unrelated regressions model is given by a specification where the rows of Π' consist of structural zeros in a way that, apart from the intercept, there are no common covariates for the T different equations and for each of the seemingly unrelated regressions one has the same number of observations. Furthermore, model (1) can be considered as the reduced form of a structural equations model.

If $T = 1$, the well-known coefficient of determination, R^2, is defined as the proportion of the variation 'explained' by the systematic or covariate part of the model and the total variation in the dependent variable given by the sum of 'explained' and variation in the noise variable. If ordinary least squares (OLS) or maximum likelihood (ML) estimation is used, its sample version, $\hat{\mathrm{R}}^2$, has the favorable property that $0 \leq \hat{\mathrm{R}}^2 \leq 1$. The extreme $\hat{\mathrm{R}}^2 = 0$ results if and only if the total variation in the dependent variable is equal to the residual variation and the extreme $\hat{\mathrm{R}}^2 = 1$ results if and only if the variation 'explained' by the covariate part of the model is equal to the total variation in the dependent variable. Furthermore, $\hat{\mathrm{R}}^2$ is monotonically related to the F-statistic for testing the hypothesis that the regression parameters are zero, and $\hat{\mathrm{R}}^2$ can be interpreted as the squared multiple correlation coefficient which gives the maximal correlation of the response variable and a linear combination of the covariates.

On the other hand, there are also some properties that suggest to interpret $\hat{\mathrm{R}}^2$ with some caution. A first point is that $\hat{\mathrm{R}}^2$ is sensitive to the definition of the response. For example, the log transformation of the response will result in a different $\hat{\mathrm{R}}^2$. A further point is that a value that may be judged to be high in one context may be considered to be low in another. Moreover, $\hat{\mathrm{R}}^2$ will never decrease if the number of covariates is increased, even if these additional variables have no explanatory power.

Sometimes $\hat{\mathrm{R}}^2$ is interpreted as a measure of the 'goodness-of-fit'. It should, however, be kept in mind that $\hat{\mathrm{R}}^2$ does not measure the goodness-of-fit of the distributional assumptions in some general sense. Rather it measures the reduction in variation if the chosen linear model is used to predict the responses as compared to using the general mean of the responses as prediction. Therefore, the measures considered in this paper will be referred to as measures of explanatory power instead.

In contrast to the univariate case, where the concepts of total and 'explained' variation are rather obvious, problems arise for $T > 1$. Then Σ_y and $\Pi' \Sigma_x \Pi$ are multivariate generalizations of the univariate concept of variance measures and there are various possibilities of defining a scalar measure of the explanatory power of a regression model. Two common measures of multivariate variation are the generalized variance, i.e. the determinant of the corresponding covariance matrix, and the total variation, i.e. the trace of the corresponding covariance matrix. Both have been used to construct measures of the explanatory power.

Let us first consider measures which are based on the generalized variance and their connection to correlation coefficients. Two measures which may be derived directly from (2) are the squared vector correlation coefficient

$$\rho_C^2 = \frac{|\Pi'\Sigma_x\Pi|}{|\Sigma_y|} = \left|\Sigma_y^{-1}\Pi'\Sigma_x\Pi\right|$$

and the measure

$$\rho_A^2 = 1 - \frac{|\Sigma_\epsilon|}{|\Sigma_y|} = 1 - \left|\Sigma_y^{-1}\Sigma_\epsilon\right| ,$$

where the positive square root of $q_A^2 = \left|\Sigma_y^{-1}\Sigma_\epsilon\right|$ is the so-called alienation coefficient (see e.g. Dhrymes, 1970; Hooper, 1959). Note that ρ_C^2 is a measure of the 'explained' generalized variance of the mutually dependent response variables and q_A^2 is a measure of the 'unexplained' generalized variance of the mutually dependent response variables which is also called the coefficient of simultaneous correlation. Unfortunately, in general ρ_C^2 and q_A^2 do not sum to unity. In univariate models both, ρ_C^2 and ρ_A^2, equal R^2 and thus may be seen as generalizations to the multivariate case.

Both measures can also be motivated by reference to canonical correlations. To see that both are functions of the canonical correlations between y and x, note that a determinant of a matrix can be written as the product of its eigenvalues, so that

$$\rho_C^2 = \prod_{t=1}^{T} \nu_t^2 \qquad \text{and} \qquad q_A^2 = \prod_{t=1}^{T} (1-\nu_t^2), \quad \text{i.e.} \quad \rho_A^2 = 1 - \prod_{t=1}^{T} (1-\nu_t^2)$$

where it can be shown that ν_t^2 and $1-\nu_t^2$ are the corresponding eigenvalues and ν_t^2 are the squared canonical correlations between the responses and the $P-1$ covariates, ignoring of course the constant term (Hooper, 1959).

From these relations it follows, that ρ_A^2 and ρ_C^2 are equal to one if all the canonical correlations are equal to one, representing complete dependence of y and x. Conversely, ρ_A^2 is zero if all the canonical correlations are equal to zero, representing complete independence of y and x. However, ρ_C^2 is zero if at least one of the canonical correlations equals zero. Therefore, if $P-1 < T$ then some of the eigenvalues vanish and ρ_C is equal to zero (Hooper, 1959), no matter how well in some other sense the covariate part 'explains' the variation in the dependent variable. Furthermore, the coefficients ρ_C^2 and q_A^2 tend to zero when T is large since both are represented as the product of quantities of a monotonically decreasing sequence between zero and one.

Therefore, instead of ρ_A^2 and ρ_C^2 one may consider

$$\sqrt[T]{\rho_C^2} = \sqrt[T]{\left|\Sigma_y^{-1}\Pi'\Sigma_x\Pi\right|} \qquad \text{and} \qquad 1-\sqrt[T]{q_A^2} = 1-\sqrt[T]{\left|\Sigma_y^{-1}\Sigma_\epsilon\right|},$$

which means taking the geometrical means of $1-\nu_t^2$ and ν_t^2, respectively. Again, however, $\sqrt[T]{\rho_C^2}$ and $\sqrt[T]{q_A^2}$ do not sum to unity (Hooper, 1959), which seems to be a requirement at least in the population. Therefore, we will exclude ρ_A^2, $1-\sqrt[T]{q_A^2}$, ρ_C^2 and $\sqrt[T]{\rho_C^2}$ from the further discussion.

Another possibility within the framework of canonical correlations is to consider the arithmetic mean of the squared canonical correlations rather than the geometric mean. Thus Hooper[1] (1959) proposed to use what he called the squared trace correlation which takes values between zero and one. The population version is given by

$$\begin{aligned} \rho_H^2 &= T^{-1}\sum_{t=1}^{T} \nu_t^2 \\ &= T^{-1}\mathrm{tr}(\Sigma_y^{-1}\Pi'\,\Sigma_x\,\Pi), \end{aligned}$$

where $\mathrm{tr}(A)$ is the trace of matrix A. Note that ρ_H^2 may also be derived from rewriting (2) in the form

$$\mathrm{I}_T = \Sigma_y^{-1}\Pi'\,\Sigma_x\,\Pi + \Sigma_y^{-1}\Sigma_\epsilon, \tag{3}$$

[1] In fact, Hooper (1959) considers a model where all the variables are centered. However, the results remain unchanged whether or not variables are centered (cf. Mardia, Kent and Bibby, 1995).

where I_T is the $(T \times T)$ identity matrix. By considering the trace on both sides of the equation one obtains

$$1 = T^{-1}\mathrm{tr}(\Sigma_y^{-1}\Pi'\,\Sigma_x\,\Pi) + T^{-1}\mathrm{tr}(\Sigma_y^{-1}\Sigma_\epsilon).$$

Thus ρ_H^2 as an extension of R^2 to the multivariate case may be seen as the arithmetic mean of components which reflect the ratio of explained to total variation. Another advantage of ρ_H^2 over ρ_A^2 and ρ_C^2 is that ρ_H^2 is unequal zero unless all the canonical correlations are zero, which will only be the case if the y's and x's are mutually independent. If $T = 1$ then ρ_H^2 is the square of the multiple correlation of the dependent variable with the covariates. The derivation of ρ_H^2 in the finite sample case is based on the use of unrestricted OLS regression parameter estimators.

Another measure which uses the trace as a measure of variability is Glahn's (1969) squared composite correlation coefficient

$$\rho_G^2 = \frac{\mathrm{tr}(\Pi'\,\Sigma_x\,\Pi)}{\mathrm{tr}(\Sigma_y)}.$$

From (2) one obtains

$$\mathrm{tr}(\Sigma_y) = \mathrm{tr}(\Pi'\,\Sigma_x\,\Pi) + \mathrm{tr}(\Sigma_\epsilon)$$

and

$$1 = \frac{\mathrm{tr}(\Pi'\,\Sigma_x\,\Pi)}{\mathrm{tr}(\Sigma_y)} + \frac{\mathrm{tr}(\Sigma_\epsilon)}{\mathrm{tr}(\Sigma_y)},$$

so that the first term on the right can be interpreted as that portion of variation 'explained' by the covariate part of the model and the second term on the right as the portion of variation in the noise variables. It is easily seen that $0 \leq \rho_G^2 \leq 1$ and ρ_G^2 increases if the portion of the variance in the noise variables decreases and vice versa. Furthermore, if $T = 1$ then ρ_G^2 is identical to ρ_H^2 and both may be interpreted as squared multiple correlations of the dependent variable and the covariates. As for ρ_H^2, the derivation of the sample version of ρ_G^2 assumes that the estimated residuals are uncorrelated with the covariates.

A measure proposed by McElroy (1977) within the framework of seemingly unrelated regressions is Glahn's (1969) ρ_G^2 applied to the standardized model (1), i.e.

$$\Sigma_\epsilon^{-1/2}y = \Sigma_\epsilon^{-1/2}\Pi'x + \Sigma_\epsilon^{-1/2}\epsilon,$$

leading to the decomposition

$$\Sigma_\epsilon^{-1/2}\Sigma_y\Sigma_\epsilon^{-1/2} = \Sigma_\epsilon^{-1/2}\Pi'\,\Sigma_x\,\Pi\Sigma_\epsilon^{-1/2} + \mathrm{I}_T.$$

Taking the trace yields

$$\mathrm{tr}(\Sigma_\epsilon^{-1}\Sigma_y) = \mathrm{tr}(\Sigma_\epsilon^{-1}\Pi'\,\Sigma_x\,\Pi) + T$$

and

$$1 = \frac{\mathrm{tr}(\Sigma_\epsilon^{-1}\Pi'\,\Sigma_x\,\Pi)}{\mathrm{tr}(\Sigma_\epsilon^{-1}\Sigma_y)} + \frac{T}{\mathrm{tr}(\Sigma_\epsilon^{-1}\Sigma_y)}.$$

The corresponding measure proposed by McElroy (1977) has the form

$$\rho_M^2 = \frac{\mathrm{tr}(\Sigma_\epsilon^{-1}\Pi'\,\Sigma_x\,\Pi)}{\mathrm{tr}(\Sigma_\epsilon^{-1}\Sigma_y)}$$

with values between zero and one. Note that ρ_M^2 gives the portion of explained variation in the transformed, not in the original y. The properties of the sample version of ρ_M^2 are derived for the OLS estimator of the transformed model, which is a generalized least squares estimator in the not transformed model. It can be shown that this measure is monotonically related to a test statistic which investigates that all the regression parameters apart from intercepts are zero. Further, for $T = 1$, $\rho_H^2 = \rho_G^2 = \rho_M^2$.

Within the framework of structural equations systems, Carter and Nagar (1977) proposed a measure which is quite similar to the one proposed by McElroy (1977). In fact, the population versions of both measures are identical. The population version of the measure proposed by Carter and Nagar (1977) can be written as

$$\rho_{CN}^2 = \frac{\mathrm{tr}(\Sigma_\epsilon^{-1}\Pi'\,\Sigma_x\,\Pi)}{\mathrm{tr}(\Sigma_\epsilon^{-1}\Pi'\,\Sigma_x\,\Pi) + \mathrm{tr}(\mathrm{I}_T)} = \frac{\mathrm{tr}(\Sigma_\epsilon^{-1}\Pi'\,\Sigma_x\,\Pi)}{\mathrm{tr}(\Sigma_\epsilon^{-1}\Pi'\,\Sigma_x\,\Pi) + T}.$$

Note that, in contrast to McElroy‘s measure, $\Sigma_\epsilon^{-1}\Pi'\,\Sigma_x\,\Pi + \mathrm{I}_T$ instead of $\Sigma_\epsilon^{-1}\Sigma_y$ is used in the denominator. Both measures are identical in the population case. However, for the sample version, in contrast to McElroy (1977), Carter and Nagar (1977) do not restrict the parameter estimators to be OLS or GLS estimators, respectively, but consider consistent estimators of the parameters. Consequently, in finite samples, ρ_{CN}^2 and ρ_M^2 are not identical in general. For the sample version of their measure and certain estimators Carter and Nagar (1977) derive its asymptotic distribution to test the hypothesis that the population version is zero.

3 Deficiencies of the proposed measures

All of the above approaches have in common that, as in the univariate case, the variation of the covariate part of model (1) is considered as the part that ‘explains’ in some sense the variation in y, whereas the variances of the noise variables are considered as nuisance parameters. On the other hand, all the above measures are functions not only of the covariate part and the variances but also of the correlation matrix of the noise variables. This holds for the population versions as well as for the finite sample versions. Therefore, the values of the measures may change if the variances or correlations change, everything else being constant. Whereas the interpretation of the variances in the noise variables as measures of the unexplained part of the model seems clear, this is not the case for the correlations of the noise variables.

There is, however, another problem arising in the multivariate case. Most of the measures considered in the last section are functions of the off-diagonal elements of Σ_η as well, where $\Sigma_\eta = \Pi' \, \Sigma_x \, \Pi$ is the covariance matrix of the linear predictor $\eta = \Pi' x$. The exception is Glahn's ρ_G^2. Of course, the diagonal elements of Σ_η are just the variances of η_t where $\eta' = (\eta_1, \ldots, \eta_T)$ and can be interpreted as noted above. On the other hand, the off-diagonal elements lack a similar interpretation.

As a consequence of the above discussion, before defining or evaluating a measure of the explanatory power of a model, it has to be clarified how this measure should or should not change as a function of the various components. The answer to this question is closely linked to the problem of defining what is meant by 'explained' variation and which components of the model 'explain' something and which do not. The latter seems to be a necessary requirement for its interpretation with respect to the assessment of the chosen model.

To illustrate the foregoing discussion about the components of model (1) and their contribution to the various measures, let us consider the case $T = 2$. Let $\Pi = (\beta_{jt})$ be a $(P \times 2)$ matrix, $x = (x_1, \ldots, x_P)'$ a $(P \times 1)$ vector with $E(x) = 0$. Let

$$\Sigma_\epsilon = \begin{pmatrix} \sigma_{11}^2 & \sigma_{12} \\ \sigma_{12} & \sigma_{22}^2 \end{pmatrix}$$

be the (2×2) covariance matrix and define $\Sigma_\eta = \Pi' \, E(xx') \, \Pi = (a_{lm})$. Then

$$a_{11} = \sum_j \sum_k \beta_{j1}\beta_{k1} E(x_j x_k), \qquad a_{12} = \sum_j \sum_k \beta_{j1}\beta_{k2} E(x_j x_k),$$

$$a_{22} = \sum_j \sum_k \beta_{j2}\beta_{k2} E(x_j x_k)$$

and $a_{21} = a_{12}$. In this simple case, we obtain

$$\rho_H^2 = \frac{1}{2} \, \frac{(a_{22} + \sigma_{22}^2)a_{11} + (a_{11} + \sigma_{11}^2)a_{22} - 2(a_{12} + \sigma_{12})a_{12}}{(a_{22} + \sigma_{22}^2)(a_{11} + \sigma_{11}^2) - (a_{12} + \sigma_{12})^2} \quad ,$$

$$\rho_G^2 = \frac{a_{11} + a_{22}}{a_{11} + a_{22} + \sigma_{11}^2 + \sigma_{22}^2}$$

and

$$\rho_M^2 = \rho_{CN}^2 = \frac{\sigma_{22}^2 a_{11} + \sigma_{11}^2 a_{22} - 2\sigma_{12} a_{12}}{\sigma_{22}^2 a_{11} + \sigma_{11}^2 a_{22} - 2\sigma_{12} a_{12} + \sigma_{11}^2 \sigma_{22}^2 - \sigma_{12}^2} \quad .$$

The measure ρ_G^2 depends only on the diagonal elements of Σ_η and Σ_ϵ, whereas ρ_H^2, ρ_M^2 and ρ_{CN}^2 depend also on the off-diagonal elements of Σ_η and Σ_ϵ.

For simplicity, let $a_{11} = a_{22} = \sigma_{11}^2 = \sigma_{22}^2 = 1$ then $\rho = \sigma_{12}$ is a correlation, and one has

$$\rho_H^2 = \frac{2 - (a_{12} + \rho) a_{12}}{4 - (a_{12} + \rho)^2} \quad ,$$

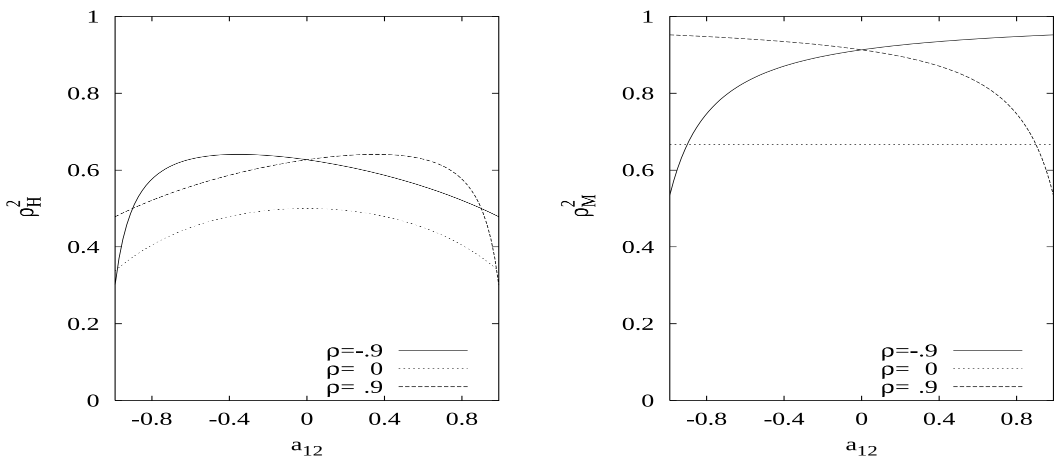
$$\rho_G^2 = \frac{1}{2}$$

and

$$\rho_M^2 = \rho_{CN}^2 = \frac{2 - 2\rho a_{12}}{2 - 2\rho a_{12} + 1 - \rho^2} \quad .$$

Figure 1: *Variation of Hooper's ρ_H^2 and McElroy's ρ_M^2 against a_{12} ($-.99 \le a_{12} \le .99$) for several values of ρ ($\rho = -.9, 0, .9$).*

Hooper's ρ_H^2 McElroys's ρ_M^2



It is evident from the above formulas and Figure 1 that ρ_H^2, ρ_M^2 and ρ_{CN}^2 are not only functions of ρ and a_{12}, but that generally the influence of, say, ρ depends on the value of a_{12} and vice versa. Consequently, the interpretation of the measures ρ_H^2, ρ_M^2 and ρ_{CN}^2 even in the population is not clear. Comparing ρ_H^2 and ρ_M^2 or ρ_{CN}^2 it is evident that they behave quite differently to changes in ρ and a_{12}.

To summarize, the values of ρ_H^2, ρ_M^2 and ρ_{CN}^2 depend on components of model (1) whose contributions to a measure of the explanatory power of the model is not clear. The exception is the measure proposed by Glahn (1969), ρ_G^2, which is a function of only those components that are interpreted in univariate models as the variation 'explained' by the model and 'unexplained' variation. Correlations between the noise variables are ignored.

Now, turning to the sample versions, the above properties of course carry over. As a consequence, everything else being constant, the values of $\hat{\rho}_H^2$, $\hat{\rho}_M^2$ and $\hat{\rho}_{CN}^2$ may change from one sample to the other if, e.g., the off-diagonal elements of $\widehat{\Sigma}_y$ do not change but those of $\widehat{\Sigma}_\eta$ and $\widehat{\Sigma}_\epsilon$ do. So, using $\hat{\rho}_H^2$, $\hat{\rho}_M^2$ and $\hat{\rho}_{CN}^2$ as descriptive measures, comparisons over different samples may be problematic at least. Again, $\hat{\rho}_G^2$ is robust with respect to changes in the corresponding off-diagonal elements.

On the other hand, with the exception of Carter and Nagars' (1977) $\hat{\rho}_{CN}^2$, all the above measures and their properties are derived by using OLS-estimators, or at least esimators for which the estimated residuals are uncorrelated with the covariates, for the regression parameters either in the original or in a transformed model. If other estimators, e.g. estimators that take into account a priori restrictions, are used, then the finite sample orthogonality property of the corresponding measures break down. On the other hand, $\hat{\rho}_{CN}^2$ suffers from the fact that generally there is no neat finite sample interpretation.

Furthermore, $\hat{\rho}_H^2$ and $\hat{\rho}_G^2$ do not take into account whether or not a certain correlation structure is modeled and estimated. For example, multivariate regression models allow the noise variables in (1) to be correlated. At the same time, however, for the general multivariate regression model given in (1) ordinary least squares, generalized least squares and maximum likelihood estimation lead to the same estimator for Π, which does not depend on the covariance matrix Σ_ϵ of the noise variables (Mardia, Kent and Bibby, 1995, pp. 173–174). On the other hand, $\hat{\rho}_M^2$ and $\hat{\rho}_{CN}^2$ are functions — although considered as nuisance — of the estimated correlation structure, but, as has been discussed for the population versions, the effect of different correlation structures depends — everything else being constant — on the off diagonal elements of Σ_η.

4 Alternative measures

As argued above, when defining a measure of the explanatory power of model (1) it should be clarified what is to be understood as the explanatory power in terms of the components involved. Following the notion of a measure of the explanatory power in univariate models, the role of the diagonal elements of Σ_η and Σ_ϵ are clear: The larger the values of $\text{vecdiag}(\Sigma_\eta)$, where $\text{vecdiag}(A)$ means extracting the diagonal elements of A into a column vector, the larger the portion of variation 'explained' by the covariate part of the model and vice versa. Correspondingly, the larger the values of $\text{vecdiag}(\Sigma_\epsilon)$ the larger the portion of variation 'unexplained' by the covariate part of the model and vice versa.

When model (1) is estimated using OLS, GLS or ML approaches, then parameters of the (conditional) mean of $y|\eta$ and the (conditional) covariance matrix of $y|\eta$ are estimated and often some are subject to restrictions. To be more precise, we consider the case where the covariance matrix is estimated by restricting the structure of the correlation matrix. The estimated parameters are the correlation structure parameters and, if identifiable, variances of the noise variables, where the latter can be interpreted as the 'unexplained' variation in the corresponding y_t's. For example, the correlation matrix $R_\epsilon = V^{-1/2}\Sigma_\epsilon V^{-1/2}$, where $V = \text{Diag}(\Sigma_\epsilon)$ and $\text{Diag}(A)$ denotes a diagonal matrix with diagonal elements identical to those of A, may be restricted to be an equicorrelation matrix without fixing the correlation coefficient. If independence is assumed, then the correlation structure parameter of the assumed equicorrelation matrix is restricted to zero. On the other extreme, the structure of R_ϵ may be completely unrestricted, leading to the estimation of $T(T-1)/2$ correlation coefficients.

The off-diagonal elements of Σ_η, on the other hand, are not explicitely modeled, although they enter implicitely if parameters are estimated. Furthermore, all the covariates assumed to have a significant effect on y_t should enter the corresponding equation. Hence, following Glahn (1969), from Σ_η only the corresponding diagonal element is considered important.

Again following Glahn (1969), by variation in the y_t's we mean the variation reflected in the diagonal elements of Σ_y. So, up to now, three important components are identified, i.e. $\text{vecdiag}(\Sigma_y)$, $\text{vecdiag}(\Sigma_\eta)$ and $\text{vecdiag}(\Sigma_\epsilon)$. Note that none of the above quantities capture any 'cross equation' dependencies. To take into account the modeled correlation structure, which by the above arguments is considered as part of the structure captured by the model to 'explain' the variation in the y_t's, and at the same time define a measure for the explanatory power in accordance with the above principle of considering the variation equation by equation, we use the squared multiple correlation of each tth noise variable with all remaining $T-1$ noise variables which is given by $\varrho^2_{t|t'} = 1-(r^t)^{-1}$, where r^t is the tth diagonal element of R_ϵ^{-1}. Since ϵ is normally distributed, the multiple correlation coefficient is the correlation between ϵ_t and the (linear) regression function of ϵ_t on the remaining $\epsilon_{t'}$'s and the squared multiple correlation coefficient measures the fraction of reduction in the variation of ϵ_t obtained by conditioning on all the remaining $\epsilon_{t'}$'s (e.g. Muirhead, 1982). Furthermore, it can be shown that the squared multiple correlation coefficient is identical to the well known coefficient of determination in the univariate linear regression model.

The choice of linear regression models of ϵ_t on the remaining $\epsilon_{t'}$'s equation by equation is justified by the normality assumption of the noise variables. In that case correlations equal to zero imply independence of the variables and $\varrho^2_{t|t'} = 0$. Further note that $\varrho^2_{t|t'}$ is maximal if R_ϵ is the 'true' correlation matrix.

However, instead of model (1), we consider a scaled version of it, and differing from the measures proposed by McElroy (1977) and Carter and Nagar (1977), we start with the decomposition of the covariance matrix of the scaled y_t's with scaling matrix $V^{1/2}$ where $V = \mathrm{Diag}(\Sigma_\epsilon)$, i.e.

$$V^{-1/2}\Sigma_y V^{-1/2} = V^{-1/2}\Pi'\Sigma_x\Pi V^{-1/2} + R,$$

where the subscript ϵ is suppressed. We then decompose, equation by equation, the transformed variation in the noise variable which is equal to one within each equation, into variation 'explained' by the correlation structure of the tth noise variable and the $T-1$ remaining noise variables, i.e. $\varrho^2_{t|t'}$, and the 'unexplained' variation, i.e. $1-\varrho^2_{t|t'}$. As the population version of a measure of the explanatory power we define

$$\rho_1^2 = \frac{\mathrm{tr}(V^{-1}\Pi'\Sigma_x\Pi) + \mathrm{tr}(I_T - \mathrm{Diag}(R^{-1})^{-1})}{\mathrm{tr}(V^{-1}\Pi'\Sigma_x\Pi) + T},$$

where $\mathrm{tr}(I_T - \mathrm{Diag}(R^{-1})^{-1}) = \mathrm{tr}(\mathrm{Diag}(\varrho^2_{1|1'}, \ldots, \varrho^2_{T|T'}))$ is the portion 'explained' by the correlation structure and $\mathrm{tr}(\mathrm{Diag}(R^{-1})^{-1})$ is the portion 'not explained'. In decomposing the variation of the transformed y_t's into variation 'explained' and 'unexplained' by the model, we follow the idea behind Carter and Nagars' (1977) measure. Note that ρ_1^2 is invariant with respect to changes of scale in any of the y_t's. However, this measure may be simplified to yield

$$\rho_1^2 = \frac{\sum_t \sigma^{-2}_{\epsilon_t}\sigma^2_{\eta_t} + \sum_t \varrho^2_{t|t'}}{\sum_t \sigma^{-2}_{\epsilon_t}\sigma^2_{\eta_t} + T} \tag{4}$$

and can be interpreted as the proportion of the variation 'explained' in all T equations to the total variation given by the sum over all T equations. Obviously, $0 \le \rho_1^2 < 1$ and, if besides the constant at least one covariate is included, $\rho_1^2 = 0$ iff $\Pi = 0$ and $R = \mathrm{I}_T$.

Similar measures can be defined, e.g. if one is interested in componentwise contributions to an overall ρ^2, then

$$\rho_2^2 = T^{-1}\sum_t \frac{\sigma^{-2}_{\epsilon_t}\sigma^2_{\eta_t} + \varrho^2_{t|t'}}{\sigma^{-2}_{\epsilon_t}\sigma^2_{\eta_t} + 1}, \tag{5}$$

may be prefered, which is just the unweighted mean of the T values of the corresponding measure of the explanatory power.

To derive a measure for the explanatory power which is closer in spirit to Hoopers' (1959) measure, let

$$M = \mathrm{Diag}(\sigma^{-2}_{\epsilon_1}\sigma^2_{\eta_1}, \ldots, \sigma^{-2}_{\epsilon_T}\sigma^2_{\eta_T})$$

and

$$\varrho^2_{y_t|y_{t'}} = \frac{\sigma^2_{\epsilon_t} r'_{t,t'}(M_{t'} + R_{t'})^{-1} r_{t,t'}}{\sigma^2_{y_t}},$$

where for a $(T \times T)$ matrix A, $A_{t'}$ is that part of A where the tth row and column has been discarded and $r'_{t,t'}$ is the vector of correlations of the tth noise variable with the remaining $T-1$ noise variables. The measure $\varrho^2_{y_t|y_{t'}}$ can be interpreted as the squared multiple correlation of y_t with the remaining $y_{t'}$'s after having removed the off diagonal elements of Σ_η, i.e. only allowing for the correlations between the noise variables to 'explain' the dependence on the remaining response variables. As a measure of the explanatory power we can define

$$\begin{aligned}
\rho_3^2 &= T^{-1}\mathrm{tr}((M+R)^{-1}M) \\
&= T^{-1}\sum_t \left(\frac{\sigma^2_{y_t}}{\sigma^2_{\epsilon_t}} - \frac{\sigma^2_{y_t}}{\sigma^2_{\epsilon_t}}\varrho^2_{y_t|y_{t'}}\right)^{-1} \frac{\sigma^2_{\eta_t}}{\sigma^2_{\epsilon_t}} \\
&= T^{-1}\sum_t \frac{\sigma^2_{\eta_t}\sigma^{-2}_{\epsilon_t}}{(\sigma^2_{\eta_t}\sigma^{-2}_{\epsilon_t}+1)(1-\varrho^2_{y_t|y_{t'}})}, \qquad (6)
\end{aligned}$$

where

$$\left(\frac{\sigma^2_{y_t}}{\sigma^2_{\epsilon_t}} - \frac{\sigma^2_{y_t}}{\sigma^2_{\epsilon_t}}\varrho^2_{y_t|y_{t'}}\right)^{-1}$$

is the tth diagonal element of $(M+R)^{-1}$. The measure ρ_3^2 may be preferred over ρ_1^2 and ρ_2^2 if it is intended to assign more weight to the covariate-part of the model.

It can be shown that $0 \leq \rho_3^2 \leq \rho_2^2 < 1$ and, if besides the constant at least one covariate is included, that $\rho_2^2 = 0$ iff both, $\Pi = 0$ and $R = \mathrm{I}_T$, whereas $\rho_3^2 = 0$ iff $\Pi = 0$. Furthermore, if all correlations are zero, then $\rho_2^2 = \rho_3^2$.

Generally, larger values of $\mathrm{vecdiag}(\Sigma_\eta)$ and $\varrho^2_{t|t'}$ or $\varrho^2_{y_t|y_{t'}}$, respectively, lead to larger values of ρ_1^2, ρ_2^2 and ρ_3^2. Conversely, larger values of $\mathrm{vecdiag}(\Sigma_\epsilon)$ lead to smaller values. If $T=1$, then $\rho_1^2 = \rho_2^2 = \rho_3^2 = \rho_H^2 = \rho_G^2 = \rho_M^2 = \rho_{CN}^2$.

5 A generalization to multivariate probit models

Now consider an ordered categorical model where each observable scalar response variable z_t takes on a value $k \in \{0, \ldots, K\}$. A threshold model is assumed, where

$$y^* = \Pi' x + \epsilon, \qquad y^* = (y_1^*, \ldots, y_T^*)' \qquad (7)$$

and

$$z_t = k \Longleftrightarrow \theta_k < y_t^* \leq \theta_{k+1},$$

where the z_t's are observed variables but the y_t^*'s are not, and the θ_j are threshold values, which are for ease of presentation assumed to be identical for all t and

$\theta_0 = -\infty$ and $\theta_{K+1} = +\infty$. Since not all threshold parameters and the intercept are identifiable, a restriction on these parameters has to be chosen when this model is estimated. However, the derivations that follow do not depend on a particular choice of the restriction.

For univariate probit models with ordinal responses, where $T = 1$, McKelvey and Zavoina (1975) proposed a measure to assess the explanatory power, the population version of which can be written as

$$\rho^2_{MZ} = \frac{\sigma_\epsilon^{-1}\sigma_\eta^2}{\sigma_\epsilon^{-1}\sigma_\eta^2 + 1},$$

where the index t is omitted.

Following Hoopers (1959) approach, Spiess and Keller (1999) proposed a generalization of McKelvey and Zavoina's (1975) measure to multivariate probit models with ordered categorical responses. For the general model, the population version of their measure can be written as

$$\rho^2_{SK} = T^{-1}\text{tr}((V^{-1/2}\Sigma_\eta V^{-1/2} + R)^{-1}(V^{-1/2}\Sigma_\eta V^{-1/2})). \tag{8}$$

Comparing this measure with ρ_3^2, we see that the difference between the two is that the off-diagonal elements are ignored in ρ_3^2. In fact, ρ^2_{SK} shares with Hooper's (1959) measure the disadvantage that it depends on correlations between covariates and cross products of the regression parameters of the different measurement points.

Although ρ^2_{SK} is written as a function of V and σ^2_η, in the multivariate probit model with ordinal responses it is not possible to estimate all variances and regression parameters independently from each other. Therefore, instead of model (7), we consider the transformed model

$$V^{-1/2}y^* = B'x + V^{-1/2}\epsilon, \tag{9}$$

where $B' = V^{-1/2}\Pi'$ leading to $B'\Sigma_x B = V^{-1/2}\Sigma_\eta V^{-1/2}$, which leaves the value of ρ^2_{SK} unchanged.

Now, given the derivations in section 4, and the idea behind McKelvey and Zavoinas' (1975) and Spiess and Kellers' (1999) measure, alternative measures of the explanatory power in multivariate probit models may be defined. The population versions we propose are exactly those given in section 4.

Given consistent estimators or correctly preassigned values of the regression and correlation structure parameters, we define the finite sample versions, i.e. the estimators of the corresponding population versions as

$$\hat{\rho}_1^2 = \frac{\sum_t \hat{\sigma}^2_{\eta_t^*} + \sum_t \hat{\varrho}^2_{t|t'}}{\sum_t \hat{\sigma}^2_{\eta_t^*} + T}, \tag{10}$$

$$\hat{\rho}_2^2 = T^{-1}\sum_t \frac{\hat{\sigma}^2_{\eta_t^*} + \hat{\varrho}^2_{t|t'}}{\hat{\sigma}^2_{\eta_t^*} + 1}, \tag{11}$$

and

$$\hat{\rho}_3^2 \;=\; T^{-1}\sum_t \frac{\hat{\sigma}^2_{\eta_t^*}}{(\hat{\sigma}^2_{\eta_t^*}+1)(1-\hat{\varrho}^2_{y_t|y_{t'}})}, \qquad (12)$$

respectively, where $\hat{\sigma}^2_{\eta_t^*}$ is the tth diagonal element of $\widehat{B}'\Sigma_x\widehat{B}$, $\hat{\varrho}^2_{t|t'}$ is the estimated squared multiple correlation given the estimated correlation matrix, $\widehat{R}_\epsilon$, of the noise variables in the underlying linear model and $\hat{\varrho}_{y_t|y_{t'}}$ is given by

$$\hat{\varrho}^2_{y_t|y_{t'}} = \frac{\hat{r}'_{t,t'}(\hat{M}_{t'}+\hat{R}_{t'})^{-1}\hat{r}_{t,t'}}{\hat{\sigma}^2_{\eta_t^*}+1},$$

where $\hat{r}'_{t,t'}$, $\hat{M}_{t'}$ and $\hat{R}_{t'}$ are the estimates of the corresponding quantities described in section 4.

It must be noted, however, that the proposed measures should be interpreted with some caution. First, they share with the coefficient of determination the limitations already discussed in section 2. Second, for a given sample, generally, there is no orthogonal decomposition of the total variation into variation 'explained' and not 'explained' by the model. This orthogonality only holds in the limit.

6 A simulation study: Comparisons, tests and confidence intervals

In this section we present results of a simulation study designed to compare two of the proposed measures, namely $\hat{\rho}_1^2$ and $\hat{\rho}_3^2$, and to evaluate the *bias corrected and accelerated* (BC_a) bootstrap technique (Efron and Tibshirani, 1993) as a means to test hypotheses and to calculate confidence intervals.

Instead of the categorical multivariate model described in section 5, however, we simulated and estimated a linear model since in order to calculate the proposed measures and to make asymptotic statements, only the parameter estimates, their estimated covariance matrix and their asymptotic distribution is necessary, where it is assumed that the corresponding estimators are consistent.

The model simulated is

$$y_n = X_n\beta + \epsilon_n,$$

where y_n is a (3×1)-vector of responses, X_n is a (3×3)-matrix of two covariates and a column of ones, $\epsilon_n \sim N(0,R)$ is a (3×1)-vector of noise variables distributed independently from the covariates. The first covariate was drawn using a (pseudo) random number generator to simulate a uniformly distributed continuous variable with variance equal one. The second was generated according to the Gamma distribution with variance and mean 0.5. Note that this model is just a special case of the more general model (1). The simulated covariance matrix was a correlation matrix which was either the identity matrix or a matrix $R(\gamma)$ corresponding to a stationary autoregressive process of order one with parameter $\gamma \in \{0.3\ 0.5\ 0.8\}$. N was set to 100 and the number of simulated data sets was 500 in each case considered.

The model was estimated using a variant of the iterative method described in Spiess and Keller (1999) adapted to the linear case where all the parameters are estimated simultaneously using the true values as starting values. However, since the proposed measures are based on the transformed model

$$\sigma_\epsilon^{-1} y_n = X_n \sigma_\epsilon^{-1} \beta + \epsilon_n^*,$$

we estimated $\beta^* = \beta/\sigma_\epsilon$ and the correlation structure parameter γ. The latter depends on the assumed correlation structure. Three different structures were considered: Equicorrelation, AR(1) and an unrestricted structure. Two versions of an equicorrelation matrix are considered: The identity matrix and a matrix with off-diagonal elements not restricted to zero. If independence was assumed, only β^* was estimated. For the equicorrelation structure, one has $r_{t,t'} = \gamma$ for all $t \neq t'$, for the AR(1) structure $r_{t,t'} = \gamma^{|t-t'|}$ for all $t \neq t'$. In these cases γ is a scalar. If an unrestricted structure is assumed, then γ is a (3×1) vector containing the off-diagonal elements.

Given consistent estimates of the interesting parameters, $\hat{\rho}_1^2$ and $\hat{\rho}_3^2$ were calculated. The first two tables show the results, i.e. the means $\bar{\hat{\rho}}_1^2$ and $\bar{\hat{\rho}}_1^2$, for two different values of the correlation structure parameter, $\gamma = 0.3$ (Table 1) and $\gamma = 0.8$ (Table 2), given an AR(1) structure and various values for the regression parameters. Models are estimated assuming independence, an equicorrelation, an AR(1) and an unrestricted correlation structure.

Obviously, the values of ρ_1^2 exceed those of ρ_3^2 in all cases considered. Further, the means of the estimated statistics over the simulated data sets, $\bar{\hat{\rho}}_1^2$ and $\bar{\hat{\rho}}_3^2$, respectively, are larger if the assumed correlation structure is closer to the truth: The corresponding values are smallest if independence is assumed, they get larger if an equicorrelation structure is assumed and are close to the true values if an AR(1) or an unrestricted correlation structure is assumed. The differences assuming an AR(1) structure and an unrestricted structure are rather negligible. The differences in the means of the estimates are larger for medium values of ρ_1^2 and ρ_3^2, respectively. Furthermore, these differences are larger if $\bar{\hat{\rho}}_1^2$ is considered as compared to $\bar{\hat{\rho}}_3^2$, which is to be expected, since ρ_3^2 gives more weight to the covariate part of the model. Comparing corresponding entries in Table 1 and 2 it can be seen that the values of both measures increase if the value of γ increases. The general conclusions from these two tables are the same if a model with AR(1) structure and $\gamma = .5$ (not reported) is considered.

To test the hypothesis that the R^2-type measure under consideration is zero or to calculate a confidence interval, the BC_a bootstrap technique is considered. In a first stage $j = 1, \ldots, J$ samples, generally 2000, are drawn with replacement

Table 1: *Mean values of estimated* ρ_1^2 *and* ρ_3^2 ($\bar{\hat{\rho}}_1^2$ *and* $\bar{\hat{\rho}}_3^2$) *assuming an independence* (Ind)*, equicorrelation* (Equi)*, AR(1)* (AR(1)) *or unrestricted* (Free) *correlation structure over* 500 *simulated data sets. Simulated model:* $N = 100$, $\beta^* = (\beta_0\ \beta_1\ \beta_2)'$ *with* $\beta_0 = 0$ *and correlation structure AR(1) with* $\gamma = .3$.

					$\bar{\hat{\rho}}_1^2$					$\bar{\hat{\rho}}_3^2$	
β_1	β_2	ρ_1^2	Ind	Equi	AR(1)	Free	ρ_3^2	Ind	Equi	AR(1)	Free
0	0	.115	.004	.083	.113	.126	0	.004	.004	.004	.004
.15	−.15	.153	.043	.124	.153	.165	.048	.043	.047	.049	.049
.35	−.1	.218	.118	.194	.221	.232	.129	.118	.127	.130	.131
.5	−.63	.463	.397	.448	.466	.474	.411	.393	.405	.409	.411
.7	−.7	.553	.500	.542	.557	.564	.511	.496	.507	.510	.512
1	−1	.705	.672	.699	.709	.713	.676	.667	.673	.675	.675
2	−2	.902	.891	.900	.903	.905	.890	.889	.889	.889	.889

Table 2: *Mean values of estimated* ρ_1^2 *and* ρ_3^2 ($\bar{\hat{\rho}}_1^2$ *and* $\bar{\hat{\rho}}_3^2$) *assuming an independence* (Ind)*, equicorrelation* (Equi)*, AR(1)* (AR(1)) *or unrestricted* (Free) *correlation structure over* 500 *simulated data sets. Simulated model:* $N = 100$, $\beta^* = (\beta_1\ \beta_2\ \beta_3)$‘*, with* $\beta_1 = 0$*, correlation structure AR(1) with* $\gamma = .8$.

					$\bar{\hat{\rho}}_1^2$					$\bar{\hat{\rho}}_3^2$	
β_2	β_3	ρ_1^2	Ind	Equi	AR(1)	Free	ρ_3^2	Ind	Equi	AR(1)	Free
0	0	.687	.004	.636	.686	.690	0	.004	.004	.004	.004
.15	−.15	.700	.043	.704	.700	.704	.122	.043	.108	.122	.124
.5	−.63	.810	.398	.782	.811	.814	.555	.394	.546	.554	.554
.7	−.7	.842	.502	.819	.843	.845	.628	.498	.624	.628	.628

to calculate $\hat{\rho}_{1,j}^2$ and $\hat{\rho}_{3,j}^2$. Each bootstrap sample consists of N matrices X_{n^*} drawn independently with replacement from the original sample. Since, generally, estimation of the multivariate probit model is expensive especially if categorical responses are present, instead of sampling the corresponding response variables and estimating the model parameters, parameter values were drawn from their estimated asymptotic distribution. A description of the resampling method is given in the Appendix. A description of the BC_a technique is given in Efron and Tibshirani (1993, pp. 184–186). Note, however, that the BC_a method presupposes that a normalizing transformation exists (Efron, 1987). See also Davison and Hinkley (1997) or Hjorth (1994) for a discussion of the bootstrap technique.

To test the hypothesis that the R^2-type measure under consideration is zero, one-sided confidence intervals are calculated, using the BC_a technique. The true regression parameter values were set to zero and the true correlation matrix was set to the identity matrix. Note that in this case $\rho_1^2 = \rho_3^2 = 0$. For $J = 2000$ bootstrap replications for each of 500 simulated data sets with $N = 100$ and correctly assuming independence, using the BC_a technique, the portion of values larger than the calculated 95% quantile was .046 and for the 90% quantile was .084 for both measures. The results changed only slightly when $J = 8000$ bootstrap replications were used.

Table 3: *Non-Coverage of the true values of* ρ_1^2 *and* ρ_3^2*:* BC_a *technique over* 500 *simulated data sets and* 2000 *bootstrap replications each. Error rates in the lower* (l_1^{05}, l_3^{05}) *and upper* (l_1^{95}, l_3^{95}) *tails (intended error rate in each tail:* .05*) and central confidence intervals* (CI1, CI3) *with intended coverage* .9. *Simulated model:* $N = 100$ *with AR(1) and* $\gamma = .3$ *and* $\gamma = .5$*, respectively. Estimation: AR(1).*

γ	β_1	β_1	ρ_1^2	l_1^{05}	l_1^{95}	CI1	ρ_3^2	l_3^{05}	l_3^{95}	CI3
.3	0	0	.115	.038	.120	.158	0			
.3	.05	.01	.117	.042	.124	.166	.003	.094	.166	.26
.3	.15	.15	.153	.048	.106	.154	.048	.07	.042	.112
.3	.35	−.1	.219	.05	.096	.146	.129	.056	.056	.112
.3	.5	−.7	.491	.058	.046	.104	.443	.084	.056	.14
.3	1	−1	.705	.062	.04	.102	.676	.09	.052	.142
.3	2	−2	.902	.068	.046	.114	.89	.094	.052	.146
.5	0	0	.3	.036	.07	.106	0			
.5	.15	−.3	.371	.036	.054	.09	.135	.07	.052	.122
.5	.5	−.63	.575	.058	.038	.096	.446	.08	.062	.142
.5	1.5	−1	.835	.06	.03	.09	.780	.078	.04	.118

In a next step, approximate 90% central confidence intervals were calculated. The results for a model with various values of the regression parameters, an AR(1) structure with $\gamma = .3$ and $\gamma = .5$, respectively, and assuming an AR(1) structure are given in Table 3. The results in Table 3 suggest, that the error rates for the approximate confidence intervals using the BC_a technique are in an acceptable range for most of the cases considered. However, there is a high error rate for ρ_3^2 if its value is close to zero (.003) in which case the distribution of $\hat{\rho}_3^2$ is highly skewed to the right. The general result did not change if the number of bootstrap replicates were raised up to 8000. If the error rates in the lower and upper tails are considered, a similar picure emerges: For smaller true values of ρ_1^2 and ρ_3^2 coverage rates tend to be less symmetrical. Again, in these cases distributions are highly skewed.

Table 4 shows the results if models were simulated with $R = \mathrm{I}_T$ and assuming independence. Again, the general results are replicated for this model. With respect to the error rates of the confidence intervals the BC_a technique worked quite well in most of the cases considered. However, for true values of ρ_1^2 and ρ_3^2 very close to zero the technique leads to very high error rates. Very similar results were found for a model with a simulated AR(1) structure and $\gamma = .8$.

Table 4: *Non-Coverage of the true values of ρ_1^2 and ρ_3^2: BC_a technique over* 500 *simulated data sets and* 2000 *bootstrap replications each. Error rates in the lower* (l_1^{05}, l_3^{05}) *and upper* (l_1^{95}, l_3^{95}) *tails (intended error rate in each tail:* .05*) and central confidence intervals* (CI1, CI3) *with intended coverage* .9*. Simulated model: $N = 100$ with $R = I_T$. Estimation: Independence.*

γ	β_1	β_1	ρ_1^2	l_1^{05}	l_1^{95}	CI1	ρ_3^2	l_3^{05}	l_3^{95}	CI3
0	.05	−.1	.012	.044	.19	.234	.012	.078	.092	.17
0	.15	−.3	.101	.05	.094	.144	.101	.072	.048	.12
0	.3	−.3	.153	.046	.082	.128	.153	.068	.054	.122
0	.8	−.63	.51	.068	.044	.112	.51	.08	.04	.12

7 Conclusions

In this paper R^2-type measures that have been proposed in the literature are reviewed and their deficiencies are discussed. Based on the notion that the values of a measure of the explanatory power should change as a function only of those parts of the model which are explicitely modeled, alternative measuers are proposed which are functions of the variances of the linear predictors, the modeled correlation structure parameters and the variances of the noise variables. Estimates of these measures can easily be calculated given estimates of the model parameters. The measures proposed differ with respect to whether the mean of pointwise contributions or an overall contribution to the explanatory power of the model is considered (ρ_2^2 and ρ_3^2 vs. ρ_1^2) or with respect to the weight given to the correlation structure versus the covariate part of the model (ρ_1^2 and ρ_2^2 vs. ρ_3^2). In contrast to previously proposed measures, they all have in common that correlations between covariates over measurement points or, if panel models are considered points in time, have no effect on the value of the corresponding measure.

The results of the simulation study suggest that the two measures considered, ρ_1^2 and ρ_3^2, in fact reflect the explanatory power as defined above: The values of both increase if the absolute true values of the regression parameters or the correlations increase. Furthermore, the estimated values increase if the modeled and estimated correlation matrix comes closer to the true correlation matrix. The results of the simulation study further suggests, that the BC_a bootstrap technique which can easily be implemented given estimates of the model parameters, their estimated covariance matrix and (asymptotic) distribution, leads to acceptable results if the null hypothesis that the corresponding measure is zero is tested given this null hypothesis is in fact true. However, more simulations would be helpful to see e.g. whether this result also holds with differentially distributed covariates or more measurement points. Furthermore, the BC_a bootstrap technique worked quite well for a wide range of possible values of the corresponding measure if approximate confidence intervals are calculated. However, if the distributions of the estimates are highly skewed as e.g. if the corresponding measure is close to zero, then the technique may lead to unacceptable high error rates. At least in these cases it seems worth to consider other bootstrap methods, in particular double bootstrap techniques, as an alternative. However, the naive double bootstrap seems to be too costly to be applied if simulations are to be run. Therefore, given the non-standard resampling method used here and described in the Appendix, a technique should be developed to reduce the computational burden. This technique should then be evaluated with the help of simulations.

Appendix

Description of the resampling scheme

As in most problems, the estimators considered, $\hat{\theta}$, are assumed to be consistent and asymptotically normally distributed. To account for the fact that the asymptotic covariance matrix is estimated, for the jth bootstrap sample, a value for the covariance matrix, $\Sigma^*_{j,\hat{\theta}}$, is drawn from a Wishart distribution in the first step. Then, noise variables are simulated from the normal distribution with expected value zero and covariance matrix equal to the estimated correlation matrix $\hat{R}$. From the corresponding values a bootstrap correlation matrix, R^*_j, is calculated. The bootstrap covariance matrix of the estimator $\hat{\theta}$, $\Sigma^*_{j,\hat{\theta}}$, is then transformed to a bootstrap covariance matrix, $\Sigma^*_{j,\hat{\vartheta}}$, of the estimator $\hat{\beta}$, the estimator of the parameters of the systematic part of the model, and the bootstrap correlations, the off-diagonal elements of R^*_j. Then, given the bootstrap correlations and this bootstrap covariance matrix, parameters of the systematic part of the model are simulated from the corresponding conditional distribution. If independence is assumed, i.e. $R = \mathrm{I}_T$, then the parameters of the systematic part may be simulated simply by using the simulated covariance matrix of the first step. Bootstrap versions of ρ_1^2 and ρ_3^2 are then easily calculated using the bootstrap sample of covariates.

More specifically parameters are drawn according to the following scheme:

1. Draw $l = 1, \ldots, NT$ rowvectors w_l of dimension K independently from each other according to the $N(0, \hat{\Sigma}_{\hat{\theta}})$ distribution, where K is the number of estimated parameters, i.e. the dimension of $\hat{\beta}^*$ plus the number of estimated correlation structure parameters, and $\hat{\Sigma}_{\hat{\theta}}$ is the estimated covariance matrix of these estimators. Calculate $\Sigma^*_{j,\hat{\theta}} = (NT)^{-1}(w_1, \ldots, w_{NT})(w_1, \ldots, w_{NT})'$.

2. If it is not assumed that independence holds, then draw $m = 1, \ldots, N$ rowvectors v_m of dimension T independently from each other according to the $N(0, \hat{R})$ distribution, where $\hat{R}$ is the estimated correlation matrix, depending on the assumed correlation structure. Calculate the $(T \times T)$ jth bootstrap sample correlation matrix R^*_j of v_m, $m = 1, \ldots, N$. Let r^*_j be the vector of lower triangular elements of R^*_j.

3. If it is not assumed that independence holds, then calculate $\Sigma^*_{j,\hat{\vartheta}} = G'\Sigma^*_{j,\hat{\theta}}G$, where $G = \frac{\partial\vartheta}{\partial\theta}\big|_{\hat{\theta}}$, $\theta = (\beta', \gamma')'$ and $\vartheta = (\beta', r')'$ where r is the vector of the lower triangular elements of R, the assumed model of the correlation structure, and is a function of the correlation structure parameter γ. For the Equicorrelation structure we have $r = (\gamma, \gamma, \gamma)$, for the AR(1) structure $r = (\gamma, \gamma^2, \gamma)$ and for the unrestricted structure, with $\gamma = (\gamma_1, \gamma_2, \gamma_3)'$, $r = (\gamma_1, \gamma_2, \gamma_3)$. In most cases, however, $\Sigma^*_{j,\hat{\vartheta}}$ will be singular. Let

$$\Sigma^*_{j,\hat{\vartheta}} = \begin{pmatrix} \Sigma^*_{j,\hat{\beta}} & \Sigma^*_{j,\hat{\beta}\hat{r}} \\ \Sigma^*_{j,\hat{r}\hat{\beta}} & \Sigma^*_{j,\hat{r}} \end{pmatrix}.$$

4. Draw β^* according to the (conditional) $N(\hat{\beta} + \Sigma^*_{j,\hat{r}\hat{\beta}}\Sigma^{*-}_{j,\hat{r}}(r^*_j - \hat{r}), \Sigma^*_{\hat{\beta}|\hat{r}})$ distribution, where $\hat{r}$ ist the vector of lower triangular elements of $\hat{R}$. The matrix $\Sigma^*_{\hat{\beta}|\hat{r}}$ is given by

$$\Sigma^*_{\hat{\beta}|\hat{r}} = \Sigma^*_{j,\hat{\beta}} - \Sigma^*_{j,\hat{r}\hat{\beta}}\Sigma^{*-}_{j,\hat{r}}\Sigma^*_{j,\hat{\beta}\hat{r}}.$$

   If independence is assumed, then draw β^* according to the $N(\hat{\beta}, \Sigma^*_{\hat{\beta}})$ distribution.

5. Calculate ρ^{2*}_1 and ρ^{2*}_3.

References

Carter, R.A.L. & Nagar, A.L. (1977). Coefficients of Correlation for Simultaneous Equation Systems. *Journal of Econometrics, 6*, 39–50.

Davison, A.C. & Hinkley, D.V. (1997). *Bootstrap methods and their application.* Cambridge: Cambridge University Press.

Dhrymes, P.J. (1970). *Econometrics. Statistical Foundations and Applications.* New York: Harper & Row.

Efron, B. (1987). Better Bootstrap Confidence Intervals. *Journal of the American Statistical Association, 82(397),* 171–185.

Efron, B. & Tibshirani, R.J. (1993). *An Introduction to the Bootstrap.* New York: Chapman and Hall.

Glahn, H.R. (1969). Some Relationships Derived From Canonical Correlation Theory. *Econometrica, 37(2),* 252–256.

Hjorth, J.S.U. (1994). *Computer Intensive Statistical Methods. Validation model selection and bootstrap.* London: Chapman & Hall.

Hooper, J.W. (1959). Simultaneous equations and canonical correlation theory. *Econometrica, 27,* 245–256.

Mardia, K.V., Kent, J.T. & Bibby, J.M. (1995, 10th printing). *Multivariate Analysis.* London: Academic Press.

McElroy, M.B. (1977). Goodness of fit for seemingly unrelated regressions. Glahn's $R^2_{y.x}$ and Hooper's $\bar{r}^2$. *Journal of Econometrics, 6*, 381–387.

McKelvey, R.D. & Zavoina, W. (1975). A statistical model for the analysis of ordinal level dependent variables. *Journal of Mathematical Sociology, 4,* 103–120.

Muirhead, R.J. (1982). *Aspects of Multivariate Statistical Theory.* New York: John Wiley & Sons.

Spiess, M. & Keller, F. (1999). A mixed approach and a distribution-free multiple imputation technique for the estimation of a multivariate probit model with missing values. *British Journal of Mathematical and Statistical Psychology, 52,* 1–17.

Spiess, M. (2001). Evaluation of a pseudo-R^2 measure for panel probit models. *British Journal of Mathematical and Statistical Psychology, 54,* 325–333.

Verbeek, M. (2000). *A Guide To Modern Econometrics.* Chichester: John Wiley & Sons.