

Tilburg University

## The Expected Parameter Change (EPC) for local dependence assessment in binary data latent class models

Oberski, Daniel L.; Vermunt, Jeroen K.

*Publication date:*  
2018

*Document Version*  
Publisher's PDF, also known as Version of record

[Link to publication in Tilburg University Research Portal](#)

*Citation for published version (APA):*  
Oberski, D. L., & Vermunt, J. K. (2018). *The Expected Parameter Change (EPC) for local dependence assessment in binary data latent class models.*

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# The Expected Parameter Change (EPC) for Local Dependence Assessment in Binary Data Latent Class Models

DL Oberski

JK Vermunt

## Abstract

Binary data latent class models crucially assume local independence, violations of which can seriously bias the results. We present two tools for monitoring local dependence in binary data latent class models: the “Expected Parameter Change” (EPC) and a generalized EPC, estimating the substantive size and direction of possible local dependencies. The asymptotic and finite sample behavior of the measures is studied, and two applications to the U.S. Census estimation of Hispanic ethnicity and medical experts’ ratings of x-rays demonstrate its value in arriving at a model that balances realism and parsimony.

R code implementing our proposal and including both example datasets is available online as supplementary material.

**KEY WORDS:** local independence, finite mixture model; diagnostic error; score test; generalized score

# 1 Introduction

The latent class model for binary data is a discrete finite mixture of binomials (Agresti, 2002), and has a wide range of applications in a diverse number of fields. In the social sciences, Hill & Kriesi (2001) classified patterns of longitudinal change in Swiss voters' support for car pollution abatement policies, while Johnson (1990) evaluated the measurement properties of alternative questions to measure ethnicity in the U.S. Census; in machine learning, the model has been used for classifying documents based on word events under the pseudonym "probabilistic latent semantic analysis" (Hofmann, 2001); and in education, Dayton & Macready (1988) analyzed how elementary school children's ability to correctly answer questions on addition, subtraction, multiplication, and division might indicate mastery of the subject. In the (bio)medical sciences, latent class analysis (LCA) for binary data has proved key to describing prevalence and symptomatology of diseases and assessing the accuracy of diagnoses (Faraone & Tsuang, 1994), and to evaluating the sensitivity, specificity, and predictive validity of diagnostic tests in the absence of a gold standard (Walter & Irwig, 1988; Hui & Zhou, 1998; Garrett et al., 2002).

The essential assumption in LCA is local independence. One way of viewing the local independence assumption is that it is assumed that, besides the latent class variable, there are no other unobserved variables influencing at least two indicators. For instance, one document "topic" may not suffice to explain the number of times pairs of words occur together in it; addition and subtraction test items may be more strongly associated to one another than to multiplication and division items; and a pair of radiologists might rate x-rays similarly if they trained in the same hospital. For more detailed explanations of how local dependence may arise, we refer to Biemer (2011, section 5.2). An important distinction to make is then whether this additional unobserved variable is of substantive interest or not (Oberski, 2013). For example, an educational researcher may, under cer-

tain circumstances, wish to distinguish between “higher” and “lower” arithmetic skills, whereas radiologist judgements’ dependence due to having trained in the same hospital is likely to be no more than a nuisance.

Local dependence is a potential problem because it can severely bias LCA outcomes of interest: estimates of classification error, class sizes, and the posterior classification of cases are all affected when local dependence is present (Vacek, 1985; Torrance-Rynard & Walter, 1998; Albert & Dodd, 2004). Hadgu et al. (2005, p. 610) argued that in the application of LCA to diagnostic tests, bias has serious clinical consequences such as an overrated degree of epidemic control, overtreatment, and unrecognized undertreatment. The degree to which such estimates will be biased depends on the size of the local dependence that is being ignored. Large local dependencies should not be ignored, whereas small local dependencies are not particularly consequential for the outcomes of interest.

Local dependence leads to LCA model misfit, to which the standard reaction is to increase the number of classes (McLachlan & Peel, 2000, chapter 6). The additional classes then represent (absorb) the dependencies. An alternative method is to model the additional variables explicitly (Hagenaars, 1993; Dendukuri et al., 2009). These methods are applicable when the dependence is of substantive interest, but introduce a problem of model interpretability when it is not (Yang & Becker, 1997; Oberski, 2013).

When local dependence is not of direct interest but represents a nuisance, a more interpretable model may be obtained by modeling the dependence directly. A variety of extended latent class models has been proposed to this end: additive probability models (Harper, 1972); loglinear (logistic) direct effects between indicators (Hagenaars, 1988; Formann, 1992); models with continuous random effects (Qu et al., 1996); and marginal models (Bartolucci & Forcina, 2006; Reboussin et al., 2008). These modeling approaches are in principle attractive when the local dependence is not of substantive interest, but suffer from two problems. First, not all local dependencies may be globally identifiable

from the data (Jones et al., 2010; Stanghellini & Vantaggi, 2013). Second, even when the local dependencies are locally identifiable in some part of the parameter space, the model estimates may become highly unstable; in particular, the substantive latent class variable and the nuisance local dependencies may become difficult to separate.

Because it is not usually desirable to model all possible local dependencies, the question arises as to *which* pairs of indicators should be modeled as dependent. Qu et al. (1996, 800-1) suggested plotting residual correlations with bootstrapped confidence intervals; Formann (2003, pp. 556-7) proposed a hypothesis test on the odds ratio in the residual bivariate crosstables, while Garrett & Zeger (2000, pp. 1063-4) discussed a Bayesian posterior predictive check on the log of that odds ratio; Sepúlveda et al. (2008) proposed a graphical method based on the log-odds ratios. Finally, Magidson & Vermunt (2004, pp. 9–11) proposed using “bivariate residuals” (BVR’s), the Pearson chi-square in the same crosstables (see also Vermunt & Magidson, 2005, pp. 72-4). An issue with the BVR, posterior predictive checks, or hypothesis tests is that they focus on the statistical significance of local dependence, while the substantive size of local dependencies is, in our view, the primary motivation for local dependence models: small dependencies are not likely to be relevant for the modeling goals whereas large dependencies should not go undetected. An additional issue with all of these measures is that their development has been *ad hoc*, in the sense that the connection between them and the local dependence model has remained unclear.

We introduce the “expected parameter change” (EPC) measure for detecting local dependencies in latent class models to resolve these problems. The EPC estimates the value that a restricted local dependence parameter would take on if it were freed in an alternative model. The EPC therefore directly evaluates the substantive size and direction of possible local dependencies. In addition, the EPC has a clear interpretation in terms of the local dependence model parameterization chosen. For example, when the alternative is

the loglinear latent class model with direct effects, the EPC has the interpretation of a conditional log-odds ratio. Other parameterizations are, however, also possible – including class-specific parameterizations as emphasized by Sepúlveda et al. (2008). In this sense the EPC is a generalization of (log) odds ratio residual measures. In addition, a hypothesis test on the EPC is equivalent to a score (“Lagrange multiplier”) test, which in turn was shown to be a generalization of the bivariate Pearson residual (BVR) by Oberski et al. (2013). The EPC therefore extends these existing local dependence measures while providing a rigorous framework for interpretation in terms of the alternative model.

We propose two variations of the expected parameter change: the  $EPC_L$  based on the expected information matrix, which is well known in structural equation modeling (Saris et al., 1987), and a novel “generalized”  $EPC_{GS}$ , based on an information matrix that can be expected to be more robust to model misspecification. The  $EPC_L$  is closely related to Rao’s classic efficient score test (Rao, 1948), while the  $EPC_{GS}$  is related to the generalized score test (White, 1982; Boos, 1992).

The article is organized as follows. Section 2 presents a local dependence latent class model for binary variables. The  $EPC_L$  and  $EPC_{GS}$  for such models are introduced in section 3. Asymptotic and sampling behavior of the  $EPC_L$  and  $EPC_{GS}$  under a range of simulation conditions are then evaluated in section 4. In sections 5 and 6, two real data applications from the literature, one in the social sciences and the other in diagnostic test assessment, demonstrate how these measures can aid in the detection of local dependence and yield different and more easily interpretable results.

## 2 Latent Class Model with local dependencies

Suppose a sample of size  $N$  is obtained on  $J$  observed binary variables, aggregated by the  $R$  response patterns into  $\mathbf{Y}$ . Let  $\mathbf{n}$  be the  $R$ -vector of observed response pattern

counts. The log-likelihood for the latent class model with  $T$  classes for the unobserved discrete variable  $\xi$  can then be formulated (Formann, 1992) as the linear-logistic (log linear) model,

$$\ell(\boldsymbol{\theta}) = \mathbf{n}' \log \Pr(\mathbf{Y}) = \mathbf{n}' \log \left[ \sum_{t=1}^T \Pr(\xi = t) \left( \frac{\exp(\boldsymbol{\eta}_t)}{\mathbf{1}'_R \exp(\boldsymbol{\eta}_t)} \right) \right], \quad (1)$$

where  $\log$  and  $\exp$  denote elementwise operations,  $\Pr(\xi = t) = \exp(\alpha_t) / \mathbf{1}'_T \exp(\boldsymbol{\alpha})$ , and

$$\boldsymbol{\eta}_t = \mathbf{X}_{(Y)}\boldsymbol{\tau} + \mathbf{X}_{(YY)}\boldsymbol{\psi} + \mathbf{X}_{(Y\xi_t)}\boldsymbol{\lambda}, \quad (2)$$

where  $\mathbf{X}_{(Y)}$ ,  $\mathbf{X}_{(YY)}$  and  $\mathbf{X}_{(Y\xi_t)}$  are design matrices for the observed variables' main effects  $\boldsymbol{\tau}$ , bivariate associations  $\boldsymbol{\psi}$ , and associations with the latent class variable  $\boldsymbol{\lambda}$ , respectively (Evers & Namboodiri, 1979). The vector  $\boldsymbol{\alpha}$  contains the logistic main effect parameters for the latent class proportions. This parameterization of the local dependence latent class model is similar to that adopted by Hagenaars (1988) and Formann (1992, section 4.3).

When the observed variables are “dummy-coded”, the loglinear local dependencies  $\boldsymbol{\psi}$  can be interpreted as log-odds ratios between pairs of items within the latent classes. The  $\boldsymbol{\psi}$  parameters should in general be interpreted as “direct effect” parameters; in effect, they are regression coefficients in the Poisson regression of within-class response pattern counts. Just as in regression analysis, when marginalized over all other response variables, the within-class log-odds dependency between a pair of items may differ from the direct effect parameter  $\boldsymbol{\psi}$ . Furthermore, even though  $\boldsymbol{\psi}$  is equal across classes, the marginal log-odds need not be. The model therefore does allow the marginal log-odds dependence to differ across classes to the extent allowed for by the model. This may be seen as a parsimonious way of modeling dependence. On the other hand, when the

goal of the analysis is to interpret the local dependence itself, marginal models such as those discussed by Bartolucci & Forcina (2006) and Reboussin et al. (2008) may be more appropriate.

The loglinear parameterization used here is equivalent to a parameterization in terms of probabilities when  $\boldsymbol{\psi} = \mathbf{0}$ , but has the advantage that probabilities below 0 and above 1 are not possible as sample estimates. For example, consider the  $T$ -class local independence model for  $K$  binary items,

$$\Pr(Y_1, \dots, Y_K) = \sum_{t=1}^T \left[ \Pr(\xi = t) \prod_{k=1}^K \Pr(Y_k | \xi = t) \right],$$

in which the joint probability given the latent class  $\prod_{k=1}^K \Pr(Y_k | \xi = t)$  may be reparameterized following the usual logistic formulation by taking

$$\ln \prod_{k=1}^K \Pr(Y_k | \xi = t) = \sum_{k=1}^K \ln \Pr(Y_k | \xi = t) = \sum_{k=1}^K (\tau_k x_{Y_k} + \lambda_{kt} x_{Y_k} x_{\xi}),$$

where  $x_{Y_k}$  is a design variable that depends on the value of  $Y_k$  and  $x_{\xi}$  a design variable depending on  $\xi$ . In dummy coding, for instance,  $x_{Y_k}$  will equal 1 if  $Y_k$  has the value 1, and 0 otherwise, whereas in effects coding, the corresponding values are +1 and -1. Note that the columns of the design matrices given in Equation (2) above are formed by  $x_{Y_k}$ ,  $x_{\xi}$ , and their products. However, the matrix formulation of the model given above is more flexible as it can also include “interactions” (local dependencies) between any number of variables, and can trivially be extended to include covariates predicting class membership. In what follows we therefore use this convenient formulation of the latent class model.

The  $p$ -vector of parameters  $\boldsymbol{\theta}$  can be defined as  $\boldsymbol{\theta}' = (\boldsymbol{\alpha}', \boldsymbol{\tau}', \boldsymbol{\lambda}', \boldsymbol{\psi}')$ . Thus, the full unconstrained model for binary variables has  $p = T - 1 + JT + \binom{J}{2}$  parameters. Typically, however, not all possible parameters are freed. The standard local independence latent



class model, for example, is obtained by setting  $\psi = \mathbf{0}$ . More generally, it is also possible to specify parameter restrictions of the form  $\mathbf{a}(\boldsymbol{\theta}) = \mathbf{0}$ . For the purposes of this paper, however, we will assume that the restrictions take the form of fixing some or all elements of  $\psi$  to a value (typically zero).

The parameter vector  $\boldsymbol{\theta}$  can then be partitioned into two parts: a part fixed to a value and a part corresponding to the  $p$  free parameters of the model. We will denote the fixed parameter vector by  $\boldsymbol{\theta}_1$  and the  $p$  free parameters by  $\boldsymbol{\theta}_2$ . In the typical latent class model assuming local independence  $\boldsymbol{\theta}_1 = \boldsymbol{\psi}$  and  $\boldsymbol{\theta}_2' = (\boldsymbol{\alpha}', \boldsymbol{\tau}', \boldsymbol{\lambda}')$ .

## 2.1 Estimation

The maximum likelihood estimates  $\hat{\boldsymbol{\theta}}_2$  under the restricted model can be found by maximizing the likelihood Equation (1) with respect to  $\boldsymbol{\theta}_2$  while keeping  $\boldsymbol{\theta}_1$  fixed at  $\hat{\boldsymbol{\theta}}_1$ . In the local independence model,  $\hat{\boldsymbol{\theta}}_1 = \hat{\boldsymbol{\psi}} = \mathbf{0}$ . Different methods of maximizing Equation (1) have been suggested in the literature, largely falling into the categories of expectation-maximization on the one hand (Dempster et al., 1977) and (quasi-) Newton optimization on the other. Either of these methods or a combination of them may be used to obtain the local maximum when it exists (Vermunt & Magidson, 2013). Since the optimization method used is inconsequential for our following discussion, we will simply assume that the maximum likelihood estimates  $\hat{\boldsymbol{\theta}}_2$  can be obtained by one or a combination of these methods.

## 2.2 Identifiability of local dependence parameters

Local identifiability is a crucial issue for the interpretation of latent class model results and the validity of asymptotic approximations (Forcina, 2008). A model is said to be locally identifiable in an open neighborhood when, within this neighborhood, there is one unique

set of parameter values that can generate a given likelihood (e.g. Skrondal & Rabe-Hesketh, 2004, chapter 5).

Goodman (1974) noted that the latent class model will be locally identifiable if the Jacobian,  $\mathbf{S}$ , of the expected response patterns with respect to the parameters is of full column rank (see also McHugh 1956, Theorem 1; Catchpole & Morgan 1997, Theorem 1; Bandeen-Roche et al. 1997, p. 1378; Huang & Bandeen-Roche 2004). The appendix gives the precise form of this Jacobian for the latent class model in Equation (1) with possible local dependencies. A necessary but not sufficient condition for identifiability is that there should be a nonzero number of degrees of freedom. That is, the number of parameters (columns of  $\mathbf{S}$ ) should not exceed the number of unique patterns (rows of  $\mathbf{S}$ ), i.e.,  $p \leq R - 1$ .

Local identifiability of dependencies is of especial importance for the EPC measures. The expected parameter change measures to be developed here are valid only when the single hypothetical local dependency under investigation would be identifiable from the data in the neighborhood of the maximum-likelihood solution. In general, additional parameters are not necessarily identifiable even when there are positive degrees of freedom. Notwithstanding this general situation, however, for the class-independent local dependencies considered in model (2), it can be proved that identification is possible as long as there are positive degrees of freedom (Theorem 1).

**Theorem 1.** *Consider the model in Equation (2) with local dependencies fixed to zero,  $\psi = \mathbf{0}$ . Assume this model is locally identifiable and the number of degrees of freedom is positive,  $d > 0$ . Then a model including at most  $d$  free elements of  $\psi$  is locally identifiable.*

*Proof.* The proof can be found in Appendix A. □

Theorem 1 is useful for the development of the EPC measures, since it greatly simplifies the definition of situations in which EPC measures are appropriate.

Table 1: Number of identifiable local dependence parameters out of total possible.

	<i>Number of observed variables (J)</i>			
	<i>J = 3</i>	<i>J = 4</i>	<i>J = 5</i>	<i>J = 6</i>
<i>Number of classes (T)</i>				
<i>T = 2</i>	0/3	6/6	10/10	15/15
<i>T = 3</i>	-	-	10/10	15/15
<i>T = 4</i>	-	-	8/10	15/15
<i>T = 5</i>	-	-	-	15/15
<i>T = 6</i>	-	-	-	15/15

To demonstrate this result, Table 1 shows local identifiability of models with an increasing number of classes and variables assessed by the method of Forcina (2008, p. 5266). Identifiability is examined empirically by randomly sampling a large number of parameter sets and examining the rank of the expected information matrix for each set. If for each of these random points the information matrix is numerically of full rank, then the model is locally identified with probability close to one (Formann, 1992). Shown are the number of local dependencies (elements of  $\psi$ ) that can be identified, where a dash indicates that even the local independence model is not identifiable. Table 1 shows Theorem 1 in action: for instance, since the local independence model with four classes and five response variables is identifiable and has eight degrees of freedom, exactly eight out of the ten pairwise local dependencies are identifiable.

As shown in model (2), we only consider loglinear local dependencies that are constant over (do not interact with) classes. For loglinear local dependency parameters that may differ over classes, as are considered in graphical models, identifiability conditions are less straightforward. Stanghellini & Vantaggi (2013) provide such conditions for the two-class model, as well as the subspace of parameters in which local identifiability occurs. Jones et al. (2010) investigated identifiability of class-dependent parameters for a particular set of models with covariates. In such cases the derivation of the EPC measures given below will also be applicable, but identification needs to be assessed more carefully.

### 3 Expected Parameter Change (EPC)

Our approach to monitoring possible local dependencies in latent class analysis sets out from the observation that local dependencies that have not been parameterized will constitute model misspecifications in the restriction  $\psi = \mathbf{0}$ . Assuming the local dependencies would be identifiable from the data if parameterized, the expected parameter change (EPC) is an approximately consistent estimate of local dependence misspecifications that can be obtained after fitting the restricted model. In this section we derive the EPC and the closely related score test for detecting local dependencies, following the literature on the EPC for structural equation models (Sarlis et al., 1987; Sörbom, 1989), and on generalized score tests (Boos, 1992). The appendix provides the first and second derivative matrices used in this section.

After estimation of the local independence latent class model, sample estimates  $\hat{\theta}$  are obtained that we will assume converge in probability to a population value  $\theta^*$  as sample size increases, i.e.  $\hat{\theta} \rightarrow \theta^*$ . These estimates can be seen as having been obtained under the model described in the previous section, but with the restriction that all local dependencies are exactly zero,  $\psi = \mathbf{0}$ : this local independence model is the null model.

Consider the alternative model that one local bivariate dependence parameter, i.e. some element  $\psi$  of the vector  $\psi$ , is nonzero. Hypothetically this additional parameter  $\psi$  could be included in the model as a free parameter, and the model re-estimated. More generally, the part of the parameter vector that is free in both the null and this alternative model is denoted  $\theta_2$ , while the additional part under consideration such as the local dependence parameter(s) are collected in the vector  $\theta_1$ . The hypothetical parameter estimates that would be obtained under this alternative model are denoted as  $\tilde{\theta}$  here. Using the standard device of a sequence of local alternatives (e.g., Maydeu-Olivares & Joe 2005, p. 721; Cameron & Trivedi 2005, p. 248), the alternative model estimates, as the sample

size increases, also converge to the population value  $\theta^*$ , that is,  $\tilde{\theta} \rightarrow \theta^*$ . This assumption, common to derivations of the asymptotic distribution of Wald and score statistics, can be informally stated as assuming that the model is “not too misspecified”, relative to the sample size (e.g., Kolenikov et al., 2010, p. 6). Violation of this assumption may occur when the local dependence in question is large and the sample size small, an issue that will be investigated in the simulation study of Section 4.

We now examine the loglikelihood value obtained under the null model, from the point of view of the restriction of interest that the local dependence equals zero. Using a Taylor expansion of the log likelihood at the maximum-likelihood solution,

$$\ell \approx \hat{\ell} + \begin{bmatrix} \theta_1^* - \hat{\theta}_1 \\ \theta_2^* - \hat{\theta}_2 \end{bmatrix}' \begin{bmatrix} s_1(\theta^*) \\ s_2(\theta^*) \end{bmatrix} + \frac{1}{2} \begin{bmatrix} \theta_1^* - \hat{\theta}_1 \\ \theta_2^* - \hat{\theta}_2 \end{bmatrix}' \begin{bmatrix} \mathbf{I}_{Y11}^* & \mathbf{I}_{Y21}^* \\ \mathbf{I}_{Y12}^* & \mathbf{I}_{Y22}^* \end{bmatrix} \begin{bmatrix} \theta_1^* - \hat{\theta}_1 \\ \theta_2^* - \hat{\theta}_2 \end{bmatrix}, \quad (3)$$

where  $s(\theta^*) = \partial\ell/\partial\theta$  is the score vector and  $\mathbf{I}_Y^*$  is the observed information matrix, both evaluated at the population value  $\theta^*$ . As demonstrated in Equation (3), both are partitioned into parameters  $\theta_2$  included in both the null and alternative models, and parameters  $\theta_1$  that are being considered as possibly different from their restricted solution,  $\theta' = (\theta_1', \theta_2')$ .

To study what would happen if the restricted parameter vector  $\theta_1$  were freed, we find new estimates by maximizing  $\ell$  (Equation (1)), this time with respect to both  $\theta_1$  and  $\theta_2$  (Sörbom, 1989, p. 373). This leads to the equality

$$\begin{bmatrix} s_1(\theta^*) \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{I}_{Y11}^* & \mathbf{I}_{Y21}^* \\ \mathbf{I}_{Y12}^* & \mathbf{I}_{Y22}^* \end{bmatrix} \begin{bmatrix} \theta_1^* - \hat{\theta}_1 \\ \theta_2^* - \hat{\theta}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}. \quad (4)$$

Note that  $\mathbf{I}_Y^*$  cannot be obtained from the maximum likelihood solution as it depends on the unknown value  $\theta^*$ . However, consistent estimates of the shift in parameter values

if  $\theta_1$  were freed can be obtained from the restricted solution as the “expected parameter change”  $\text{EPC} = \hat{\theta}_1 - \theta_1^* \approx \hat{\mathbf{V}}^{-1} \mathbf{s}_1(\hat{\theta})$ , where  $\hat{\mathbf{V}}$  is consistent estimate of  $\mathbf{I}_Y^*$  evaluated at the restricted solution. This implies that  $\hat{\mathbf{V}}$  consistently estimates the variance of the score vector  $\mathbf{s}_1$ , so that a score statistic can be obtained as  $T = \mathbf{s}_1(\hat{\theta})' \hat{\mathbf{V}}^{-1} \mathbf{s}_1(\hat{\theta})$  which is distributed as  $\chi_{\text{rk}(\mathbf{S}_1)}^2$  under the null hypothesis.

Under the null hypothesis  $\psi = \mathbf{0}$ , the information matrix  $\mathbf{I}_Y^*$  is consistently estimated by the expected information matrix evaluated at the restricted solution  $\hat{\mathbf{I}}_L$ , so that (Rao, 1948)

$$\begin{aligned} \text{EPC}_L &= \hat{\mathbf{V}}_L^{-1} \mathbf{s}_1(\hat{\theta}) = \hat{\mathbf{I}}_L^{-1} \mathbf{s}_1(\hat{\theta}) \\ &= (\hat{\mathbf{I}}_{L11} - \hat{\mathbf{I}}_{L12} \hat{\mathbf{I}}_{L22}^{-1} \hat{\mathbf{I}}_{L21})^{-1} \mathbf{s}_1(\hat{\theta}), \end{aligned} \tag{5}$$

where the last step, following from the rules for inverting a partitioned non-singular matrix, is computationally convenient since the partition  $\hat{\mathbf{I}}_{L22}$  of the information matrix corresponding to the free parameters will usually already be at hand in latent class modeling software. The  $\text{EPC}_L$  defined above is popular in the field of structural equation modeling (Saris et al., 1987). Rao (1948)’s efficient score statistic can be obtained as  $T_L = \mathbf{s}_1(\hat{\theta})' \hat{\mathbf{I}}_L^{-1} \mathbf{s}_1(\hat{\theta})$ , which under the null hypothesis has a chi square distribution with  $\text{rank}(\mathbf{S}_1)$  degrees of freedom. The efficient score statistic is known in the structural equation modeling literature as the “modification index” (MI) (Sörbom, 1989), and in the econometrics literature as the Lagrange multiplier test (Aitchison & Silvey, 1958; Breusch & Pagan, 1980). By the same argument of consistency under the null hypothesis, the expected information matrix  $\hat{\mathbf{I}}_L$  can be replaced by the observed information evaluated at the restricted solution,  $\hat{\mathbf{I}}_Y$  (see Glas, 1999; van der Linden & Glas, 2010).

The derivation of  $\hat{\mathbf{V}}$  under the null hypothesis suggests that when  $\psi \neq \mathbf{0}$ , the  $\text{EPC}_L$  is asymptotically biased. Under misspecified local independence, a “generalized”, i.e. robust to misspecification, consistent estimate  $\hat{\mathbf{V}}_{\text{GS}}$  of  $\mathbf{V}$  can be used (White, 1982). As

shown by Boos (1992, p. 329),

$$\begin{aligned}\hat{\mathbf{V}}_{\text{GS}} &= (\mathbf{1}, -\hat{\mathbf{I}}_{Y12}\hat{\mathbf{I}}_{Y22}^{-1})\hat{\mathbf{D}}(\mathbf{1}, -\hat{\mathbf{I}}_{Y12}\hat{\mathbf{I}}_{Y22}^{-1})' \\ &= \hat{\mathbf{D}}_{11} - \hat{\mathbf{I}}_{Y12}\hat{\mathbf{I}}_{Y22}^{-1}\hat{\mathbf{D}}'_{12} - \hat{\mathbf{D}}_{12}\hat{\mathbf{I}}_{Y22}^{-1}\hat{\mathbf{I}}'_{Y12} + \hat{\mathbf{I}}_{Y12}\hat{\mathbf{I}}_{Y22}^{-1}\hat{\mathbf{D}}_{22}\hat{\mathbf{I}}_{Y22}^{-1}\hat{\mathbf{I}}'_{Y12},\end{aligned}\tag{6}$$

where  $\mathbf{D}$  is the outer product matrix of first derivatives of the log-likelihood (see appendix) and  $\hat{\mathbf{I}}_Y$  and  $\hat{\mathbf{D}}$  denote quantities evaluated at the sample estimates  $\hat{\boldsymbol{\theta}}$  under the restricted model. A “generalized expected parameter change”  $\text{EPC}_{\text{GS}}$  is obtained as  $\text{EPC}_{\text{GS}} = \hat{\mathbf{V}}_{\text{GS}}^{-1}\mathbf{s}_1(\hat{\boldsymbol{\theta}})$ ; the well known generalized score test (White, 1982) is  $T_{\text{GS}} = \mathbf{s}_1(\hat{\boldsymbol{\theta}})'\hat{\mathbf{V}}_{\text{GS}}^{-1}\mathbf{s}_1(\hat{\boldsymbol{\theta}})$ .

### 3.1 Form of the $\text{EPC}_L$ for a local dependence parameter

So far we have given the expected parameter change statistics in generality. That is, the equations given above may, in fact, be applied to any restricted parameter, not only to local dependencies. To gain more insight into the use of these statistics for the detection of local dependence, we give here the form of the EPC for detecting local dependencies between two variables  $Y_j$  and  $Y_{j'}$  in the parameterization of Equation (2).

As noted above, the  $\text{EPC}_L$  is defined as minus the derivative of the log-likelihood with respect to the local dependence parameter, premultiplied by the inverse information matrix under the alternative, both evaluated at the restricted solution,

$$\text{EPC}_L = (\mathbf{I}_L|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}) \left( \frac{\partial \ell}{\partial \psi_{jj'}} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \right),\tag{7}$$

where  $\boldsymbol{\theta} = (\boldsymbol{\alpha}', \boldsymbol{\tau}', \boldsymbol{\lambda}', \psi_{jj'})'$ . In the derivatives involved in the information matrix  $\mathbf{I}_L$  and  $\partial \ell / \partial \psi_{jj'}$ , a key role is played by the design matrices  $\mathbf{X}_{(Y)}$ ,  $\mathbf{X}_{(Y\xi_t)}$ , and  $\mathbf{X}_{(YY)}$ . For example, the derivative  $\partial \ell / \partial \psi_{jj'}$  is determined by the corresponding column of the design matrix  $\mathbf{X}_{(YY)}$ : with dummy coding, this column will be a vector that equals 1 for all

patterns in which both  $Y_j$  and  $Y_{j'}$  equal 1, and 0 otherwise; using effects coding, the corresponding values will be 1 and -1.

The derivative with respect to a hypothetical local dependence parameter between two variables is closely related to the raw residual in the bivariate crosstable between these two variables. Since in the current model all residuals in the bivariate crosstable will be equal to each other in absolute value, we shall consider  $r_{11}$ , that is, the observed minus the expected number of observations in which the two variables  $Y_j$  and  $Y_{j'}$  are both equal to 1, i.e.,  $r_{11} = n_{11} - \hat{\mu}_{11}$ , where  $n_{kl}$  and  $\hat{\mu}_{kl}$  respectively denote the observed and expected number of observations in cell  $(k, l)$  of the crosstable.

From the general form of the derivatives, given in the Appendix, it can then be shown that using effects coding,

$$\begin{aligned}
\frac{\partial \ell}{\partial \psi_{jj'}} &= \mathbf{n}' \sum_{t \in 1..T} \Pr(\xi = t | \mathbf{Y}) [\mathbf{x}_{(y_j y_{j'})} - \mathbf{x}'_{(y_j y_{j'})} \Pr(\mathbf{Y} | \xi)] \\
&= \mathbf{n}' [\mathbf{x}_{(y_j y_{j'})} - \mathbf{x}'_{(y_j y_{j'})} \Pr(\mathbf{Y})] \\
&= \sum_{k=l} (n_{kl} - \hat{\mu}_{kl}) - \sum_{k \neq l} (n_{kl} - \hat{\mu}_{kl}) = 4r_{11}
\end{aligned} \tag{8}$$

where the second step is due to the fact that  $\psi_{jj'}$  is class-independent, and the last step follows because the off-diagonal residuals have a sign opposite to the diagonal residuals. If dummy coding is chosen instead of effect coding,  $\partial \ell / \partial \psi_{jj'} = r_{11}$ . The precise form of the information matrix  $\mathbf{I}_L$  is given in the Appendix.

There is, therefore, a close relationship between the residual in the bivariate crosstable and the expected parameter change. The EPC equals the raw residual “divided by” its variance under the alternative model. This finding is in close correspondence with the finding of Oberski et al. (2013) that the score test is closely related to the bivariate residuals.



## 4 Asymptotic and finite sample evaluation of expected parameter change

In this section we evaluate both the asymptotic and sampling performance of the suggested  $EPC_L$  and  $EPC_{GS}$  statistics for detecting relevant local dependencies. The goal of this investigation is to evaluate the feasibility of the EPC as a measure of the size of possible local dependencies.

EPC measures when applied to population data should approximate the true size of the local dependence. EPC measures calculated on samples should be sample estimates of the population EPC. Under different conditions, we examine:

- To what extent the *population* EPC corresponds to the true local dependence;
- To what extent the average *sample* EPC corresponds to the population EPC.

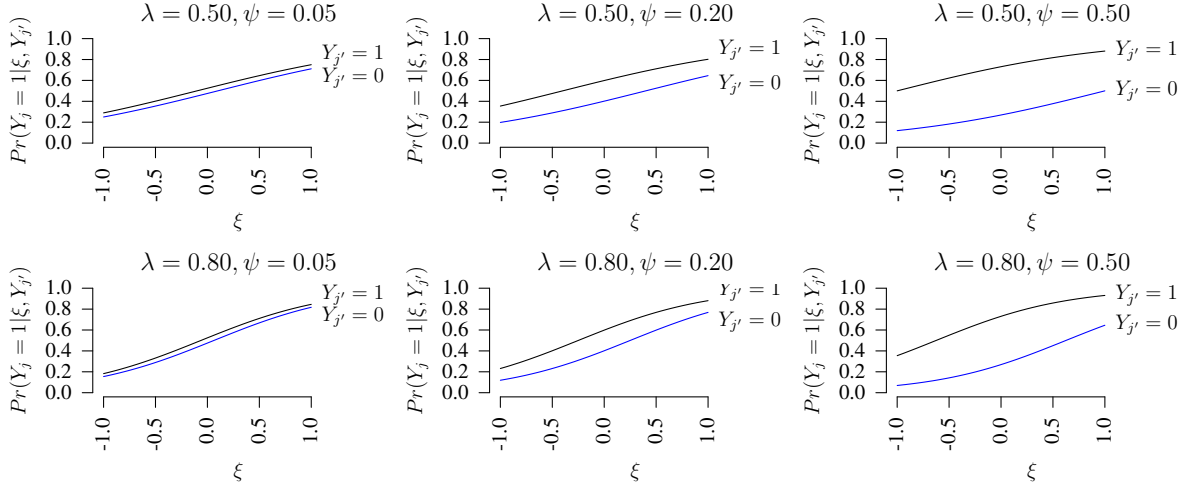
By performing both a population and a finite sample analysis, we can separate errors due to the approximation inherent in the EPC on the one hand from errors due to sampling fluctuations on the other.

### 4.1 Setup

The population model is specified as a two-class model with five binary indicators and one local dependence between a pair of indicators. In our setup, all design matrices in equation (2) are chosen such that the columns sum to zero (“effect coding”). The intercepts  $\tau = \mathbf{0}$ , the latent class intercept  $\alpha = 0.20$ , and the “loadings” and bivariate local dependence are varied across conditions:

1. Local dependence size ( $\psi$ ): -0.50 (high-negative), -0.20 (middle-negative), -0.05 (low-negative), 0 (none), +0.05 (low-positive), +0.20 (middle-positive), 0.50 (high-positive);

Figure 1: Effect of local dependence. Shown is the conditional probability that an observed variable  $Y_j = 1$  given the latent class variable  $\xi$  and a different observed variable  $Y_{j'}$  ( $j \neq j'$ ), for six conditions. (Only conditions with positive slopes are shown here.)



2. Effect of latent variable on indicators ( $\lambda$ ): 0.5 (medium-low), 0.8 (high).

A subsequent Monte Carlo simulation crosses these 14 conditions with sample size,

3 Sample size (nobs): 128, 256, 512, 1024, 2048.

We therefore examine the sampling performance of the two statistics for 70 conditions in total.

To give the reader an idea of the implications of these conditions, Figure 1 shows the effect of choosing different combinations of the slope parameter  $\lambda$  and the local dependence parameter  $\psi$  on the conditional probability for one observed variable. For illustrative purposes, many different values for the latent class variable  $\xi$  are plotted; in practice there will be only  $T$  points along the horizontal axis. The Figure shows that  $\psi = 0.05$  constitutes a rather small local dependency, while choosing  $\psi = 0.5$  has a very large effect on the implied conditional probability. This effect is larger in absolute terms when the slope parameter  $\lambda$  is small.

To illustrate the implications of these conditions, Figure 1 depicts how the class-specific response probability for variable  $Y_j$  is affected by the value of a different variable  $Y_{j'}$  for particular values of  $\lambda$  and the local dependence,  $\psi$ , between the two items. For illustrative purposes, the latent variable  $\xi$  is treated as continuous, but in fact it takes on only the values 0 and 1. It can be seen that  $\psi = 0.05$  constitutes a rather small local dependency (lines are close to one another), while choosing  $\psi = 0.5$  has a very large effect on the implied class-specific response probabilities. This effect is larger in absolute terms when the slope parameter  $\lambda$  is small.

## 4.2 Asymptotic performance

We will first evaluate the asymptotic performance of the  $EPC_L$  and  $EPC_{GS}$  obtained from the  $H_0$  model which omits the local dependence. For this purpose, we compute maximum likelihood estimates under the  $H_0$  model using the population proportions under the  $H_1$  model as data. Since this amounts to minimizing the Kullback-Leibler distance, we refer to this model as the “KL-model”. The KL-model provides the asymptotic value (as the sample size approaches infinity) of the EPC and score statistic given  $H_1$ .

The top parts of Tables 2 and 3 show the obtained  $EPC_L$  and  $EPC_{GS}$  values under the different conditions. It can be seen that when there is no misspecification, i.e. when the true local dependence parameter is zero, both EPC’s will also estimate zero. When there is a small misspecification of -0.05 or +0.05, both EPC’s have population values that are very close to the true local dependence. The top part of Table 2 shows that with larger local dependencies in absolute value, the population  $EPC_L$  is a biased estimate of the true local dependence parameter. The percentage relative bias in the  $EPC_L$  is shown in the bottom part of Table 2. Local dependencies of +0.2 and +0.5 cause larger asymptotic biases than their negative counterparts. Under the condition with lower slopes and the largest positive misspecification, the  $EPC_L$  is an 338% overestimate of the true local

Table 2: Population  $EPC_L$  statistics under the 14 simulation conditions.

		Local dependence ( $\psi$ )						
$\lambda$		-0.5	-0.2	-0.05	0	0.05	0.2	0.5
0.5		-0.374	-0.165	-0.047	-0.000	0.054	0.313	2.190
0.8		-0.329	-0.159	-0.047	-0.000	0.054	0.277	1.425
		Percent relative bias in $EPC_L$						
0.5		-25	-17	-6	-	9	56	338
0.8		-34	-20	-6	-	7	39	185

Table 3: Population  $EPC_{GS}$  statistics under the 14 simulation conditions.

		Local dependence ( $\psi$ )						
$\lambda$		-0.5	-0.2	-0.05	0	0.05	0.2	0.5
0.5		-0.403	-0.186	-0.050	-0.000	0.050	0.181	0.694
0.8		-0.439	-0.209	-0.051	-0.000	0.048	0.167	0.344
		Percent relative bias in $EPC_{GS}$						
0.5		-19	-7	-1	-	-1	-10	39
0.8		-12	5	3	-	-4	-17	-31

dependence. In contrast, with negative local dependencies, the  $EPC_L$  is underestimated in absolute terms.

Table 3 shows the population  $EPC_{GS}$  estimates (top part) as well as the percentage bias relative to the true population local dependence (bottom part). The table shows that the relative asymptotic bias in the  $EPC_{GS}$  is uniformly much lower than that in the  $EPC_L$ : on average it is 60% lower. Overall the  $EPC_{GS}$  has much better asymptotic performance.

### 4.3 Finite sample performance

In finite samples, sampling fluctuations in the score and the  $\mathbf{V}$  matrix will influence the EPC's as well. We therefore performed a Monte Carlo simulation to evaluate the sampling behavior of these statistics. From each of the 70 populations, a sample of  $N$  observations was drawn and the  $EPC_L$  and  $EPC_{GS}$  were calculated. This process was replicated

Table 4: Monte Carlo simulation results: median  $EPC_L$  statistics over 400 replications for each condition. For comparison, the bottom rows provide population values obtained from the KL-model.

No. obs.	Local dependence ( $\psi$ )					
	$\psi = -0.05$		$\psi = -0.20$		$\psi = -0.50$	
	Loading ( $\lambda$ )					
	0.5	0.8	0.5	0.8	0.5	0.8
128	-0.053	-0.054	-0.164	-0.163	-0.377	-0.330
256	-0.040	-0.046	-0.157	-0.159	-0.374	-0.333
512	-0.055	-0.045	-0.166	-0.158	-0.380	-0.332
1024	-0.041	-0.047	-0.164	-0.160	-0.378	-0.328
2048	-0.045	-0.051	-0.163	-0.162	-0.376	-0.330
Population	-0.047	-0.047	-0.165	-0.159	-0.374	-0.329

No. obs.	Local dependence ( $\psi$ )					
	$\psi = +0.05$		$\psi = +0.20$		$\psi = +0.50$	
	Loading ( $\lambda$ )					
	0.5	0.8	0.5	0.8	0.5	0.8
128	0.032	0.030	0.212	0.235	1.235	0.993
256	0.053	0.045	0.278	0.282	1.695	1.199
512	0.052	0.049	0.282	0.292	1.962	1.330
1024	0.049	0.051	0.294	0.271	2.079	1.358
2048	0.055	0.057	0.302	0.276	2.110	1.351
Population	0.054	0.054	0.313	0.277	2.190	1.425

400 times to yield a sampling distribution for  $EPC_L$  and  $EPC_{GS}$ .

Table 4 shows the median  $EPC_L$  estimates over each of the 400 samples in each of the conditions. For the  $EPC_L$  to be unbiased with respect to the true local dependence, these values should correspond to the size of the  $\psi$  local dependence parameter shown in the table headers. Considering the population bias reproduced in the rows marked “population”, we would not expect unbiasedness with respect to  $\psi$  in general. Except in conditions with sample sizes 128 and 256, the median sample estimates in Table 4 are close to the population values.

With a small sample size of  $N = 128$ , the sample estimates of the  $EPC_L$  are biased with respect to the population values. Paradoxically, the small-sample estimates can be

Table 5: Monte Carlo simulation results: median  $EPC_{GS}$  statistics over 400 replications for each condition. For comparison, the bottom rows provide population values obtained from the KL-model.

No. obs.	Local dependence ( $\psi$ )					
	$\psi = -0.05$		$\psi = -0.20$		$\psi = -0.50$	
	Loading ( $\lambda$ )					
	0.5	0.8	0.5	0.8	0.5	0.8
128	-0.052	-0.053	-0.155	-0.196	-0.330	-0.424
256	-0.040	-0.049	-0.168	-0.207	-0.350	-0.432
512	-0.057	-0.051	-0.185	-0.202	-0.375	-0.439
1024	-0.044	-0.052	-0.186	-0.208	-0.388	-0.438
2048	-0.048	-0.055	-0.183	-0.211	-0.396	-0.440
Population	-0.050	-0.051	-0.186	-0.209	-0.403	-0.439

No. obs.	Local dependence ( $\psi$ )					
	$\psi = +0.05$		$\psi = +0.20$		$\psi = +0.50$	
	Loading ( $\lambda$ )					
	0.5	0.8	0.5	0.8	0.5	0.8
128	0.027	0.027	0.102	0.136	0.305	0.214
256	0.047	0.042	0.130	0.158	0.468	0.263
512	0.046	0.045	0.152	0.164	0.605	0.298
1024	0.044	0.046	0.165	0.162	0.619	0.321
2048	0.049	0.050	0.171	0.166	0.670	0.326
Population	0.050	0.048	0.181	0.167	0.694	0.344

closer to the true misspecification than the population values are (see for example the conditions with  $\psi = +0.20$  and  $\psi = +0.50$ ). As expected, increasing the sample size brings the median  $EPC_L$  closer to the population value. It is clear that the conditions with the larger slopes perform much better than those with lower slopes, both in the population and in finite samples. With five indicators and true slopes equal to 0.8, the  $EPC_L$  provides reasonable estimates in all conditions. Whether this condition is satisfied cannot be verified from a given restricted sample solution, since the restriction itself may bias the loading estimates.

Table 5 shows the Monte Carlo simulation results for  $EPC_{GS}$ . Even for small sample sizes, the median  $EPC_{GS}$  over simulated samples is close to the population  $EPC_{GS}$ . The

sample  $EPC_{GS}$  estimates are close to the true local dependence parameters. The  $EPC_{GS}$  clearly performs better than the  $EPC_L$  both in the population and in finite samples overall. An exception occurs in the case of  $\psi = 0.20$  with small samples<sup>1</sup>: here the average  $EPC_L$  values are closer to 0.20 than are the average  $EPC_{GS}$  values, which are downwards biased. This occurs due to the interplay of an upwards bias in the asymptotic value of  $EPC_L$  combined with a downwards small-sample bias. These approximate bias canceling effects only appear to occur in a few circumstances and in practice one may not want to rely on their occurrence, however. Overall the bias in the  $EPC_{GS}$  can be viewed as acceptable for the purpose of detecting the substantive size of local dependencies.

The  $EPC_{GS}$  appears preferable to  $EPC_L$ . However, if in practice the observed information matrix is close to singular,  $EPC_{GS}$  will not be computable so that  $EPC_L$  may be an alternative in those cases. On the other hand, when the model is very large, the expected information matrix, which involves all possible response patterns, may require a prohibitively large amount of computer resources; in such cases  $EPC_L$  may not be computable.

## **5 Application 1: Measurement of Hispanic ethnicity in the U.S. Census**

Johnson (1990) performed a latent class analysis of four indicators of Hispanic ethnicity in the U.S. Census. For 9701 respondents to the 1986 National Content Test, two indicators were obtained during an initial interview (at time point  $t = 1$ ): whether Spanish was spoken at home during childhood (“Language <sub>$t=1$</sub> ”) and Hispanic origin (“Origin <sub>$t=1$</sub> ”). In a subsequent reinterview ( $t = 2$ ), two additional indicators of ethnicity were obtained: Hispanic ancestry (“Ancestry <sub>$t=2$</sub> ”) and a repetition of the “Origin” measure (“Origin <sub>$t=2$</sub> ”).

---

<sup>1</sup>We thank an anonymous reviewer for pointing out this finding.

Following Johnson (1990), we analyze the group of 9485 respondents not born in an Hispanic country. Of interest are false positive rates,  $\Pr(Y = \text{Yes}|\xi = \text{No})$ , and false negative rates,  $\Pr(Y = \text{No}|\xi = \text{Yes})$ , for the alternative question formulations.

Johnson (1990) first fitted a two-class model to these data, yielding a deviance of 103.6 with 6 degrees of freedom ( $p < 10^{-5}$ ), and a Bayesian Information Criterion (BIC) of 48.7. The two class model's lack of fit to the data led the authors to subsequently fit a model with two separate two-class latent variables corresponding to the two measurement occasions. In terms of probabilities their model can be written

$$\begin{aligned} \Pr(\text{Language}_{t=1}, \text{Origin}_{t=1}, \text{Ancestry}_{t=2}, \text{Origin}_{t=2}) = \\ \sum_{\{j,k\} \in \{1,2\} \times \{1,2\}} [\Pr(\xi_{t=1} = j, \xi_{t=2} = k) \Pr(\text{Language}_{t=1} | \xi_{t=1} = j) \Pr(\text{Origin}_{t=1} | \xi_{t=1} = j) \times \\ \Pr(\text{Ancestry}_{t=2} | \xi_{t=2} = k) \Pr(\text{Origin}_{t=2} | \xi_{t=2} = k)]. \quad (9) \end{aligned}$$

It can be seen in Equation (9) that instead of one single  $\xi$  variable, two latent discrete variables  $\xi_{t=1}$  and  $\xi_{t=2}$  are defined. Crucially, the conditional probability of an item at a time point only depends on the latent variable corresponding to that time point. Conditional independence given the time-specific latent variable is assumed. The relationship between the two latent variables  $\Pr(\xi_{t=1} = j, \xi_{t=2} = k)$  is freely estimated, but could equally well be viewed as a set of four class proportions. An alternative way of viewing this model is therefore as a highly restricted four-class model, where each of the four classes corresponds to a cell in the cross-table of the two latent variables (Hagenaars, 1988). Due to these restrictions the model parameters are identifiable.

Johnson (1990)'s final analysis is model (9) applied to the Hispanic ethnicity data. This indeed improved the deviance to 3.1 with 4 degrees of freedom ( $p = 0.54$ ; BIC equals -33.5). Conditional probability estimates based on the two-variable model are



Table 6: Conditional probability estimates in Johnson’s (1990) two-variable model. Shown are the conditional probabilities for each item given its corresponding latent variable ( $\xi_{t=1}$  or  $\xi_{t=2}$ ).

Class #	Ancestry <sub>t=2</sub>		Language <sub>t=1</sub>		Origin <sub>t=1</sub>		Origin <sub>t=2</sub>	
	No	Yes	No	Yes	No	Yes	No	Yes
1	0.999	0.001	0.995	0.005	0.998	0.002	0.999	0.001
2	0.186	0.814	0.171	0.829	0.218	0.782	0.075	0.925

shown in Table 6. For conciseness, the probabilities in Table 6 are conditional on the latent class variable corresponding to the item in question (i.e.,  $\xi_{t=1}$  or  $\xi_{t=2}$ ).

The latent variables represent a nuisance dependency due to the measurement occasion (also remarked by the original authors, p. 64). The resulting conditional probabilities from model (9) in Table 6 are therefore difficult to interpret in sociological terms, because we obtain measurement properties of each item as a measurement of the specific time point,  $\Pr(Y_{t=1,2}|\xi_{t=1,2})$ , but not the false positive and false negative rates of interest,  $\Pr(Y_t|\xi)$ . Instead of the multiple latent variable model (9), it may therefore be preferable to fit a model with a single dichotomous latent class variable but that does account for the time dependence between items.

The question is now whether, starting from the independence model, the score and EPC measures discussed above could have succeeded in detecting the relevant dependencies. Table 7 shows the results of calculating the  $EPC_L$ ,  $EPC_{GS}$ , and the corresponding score statistics after fitting the two-class independence model. The dependence between items measured at the same time point is clearly indicated as the primary source of misfit. Moreover, the other pairs of items exhibit negative dependence. Such negative dependence is commonly thought to occur when there is multidimensionality among items measuring different latent variables (e.g., Yen, 1984, p. 127). Here, however, the multidimensionality in question is not of substantive interest. It merely represents extraneous factors of the measurement occasion which are not the focus of the investigation. An

Table 7: Expected parameter changes (EPC’s) and score tests (T) for local dependence between indicators of Hispanic ethnicity in the Census.

Local dependence			$EPC_L$	$T_L$	$EPC_{GS}$	$T_{GS}$
Language <sub>t=1</sub>	↔	Ancestry <sub>t=2</sub>	-0.92	5.0	-1.45	7.9
Origin <sub>t=1</sub>	↔	Ancestry <sub>t=2</sub>	-1.08	7.9	-1.76	12.8
Origin <sub>t=2</sub>	↔	Ancestry <sub>t=2</sub>	4.14	97.1	1.59	37.2
Origin <sub>t=1</sub>	↔	Language <sub>t=1</sub>	2.94	45.6	1.32	20.5
Origin <sub>t=2</sub>	↔	Language <sub>t=1</sub>	-0.76	2.5	-1.23	4.1
Origin <sub>t=2</sub>	↔	Origin <sub>t=1</sub>	-1.10	6.1	-2.20	12.2

Table 8: Conditional probability estimates in final model including local dependencies.

Class		Ancestry <sub>t=2</sub>		Language <sub>t=1</sub>		Origin <sub>t=1</sub>		Origin <sub>t=2</sub>	
#	Size	No	Yes	No	Yes	No	Yes	No	Yes
1	0.976	0.999	0.001	0.995	0.005	0.998	0.002	0.999	0.001
2	0.024	0.383	0.617	0.288	0.712	0.329	0.671	0.299	0.701

alternative to the multidimensional model is therefore a model that frees the two large and positive local dependence parameters,  $Origin_{t=2} \leftrightarrow Ancestry_{t=2}$  and  $Origin_{t=1} \leftrightarrow Language_{t=1}$ . Such a model retains its interpretability as a measurement model for Hispanic ethnicity while also accounting for time dependence.

Based on the dependencies indicated as substantively relevant in Table 7, we fit the loglinear latent class model with two classes and two loglinear local dependencies. The class sizes and conditional probability estimates from this model are shown in Table 8. Contrary to those in Table 6, the conditional probabilities in Table 8 can be interpreted as sensitivity and specificity estimates. This model produces identical expected frequencies and deviance to the multidimensional model chosen by Johnson (1990), and is therefore equivalent to it. Crucially, however, the false negative rates of interest differ considerably. Since the nuisance dependencies due to measurement occasions are absorbed by the local dependence parameters  $\psi$ , the false negative rates can be interpreted as being with respect to a common latent class variable that might be labeled “Hispanic ethnicity”.

Table 9: Local dependence estimates in final model including two log-odds local dependencies.

Local dependence			Est.	Wald
Origin <sub>t=2</sub>	↔	Ancestry <sub>t=2</sub>	3.166	50.2
Origin <sub>t=1</sub>	↔	Language <sub>t=1</sub>	1.677	19.3

Table 9 shows the loglinear dependence parameter estimates obtained from the two-class dependence model. The Wald test under this model indicates the same as the score test under the independence model: that these local dependencies differ significantly from zero. These estimates can be interpreted as the log-odds ratio between two items conditional on the latent class. In this particular model, these log-odds ratios are also the marginal log-odds ratios within classes over all possible response patterns. It can be seen that the local dependencies are rather strong, especially at time  $t = 2$ .

The final model results in Table 8 show that Origin<sub>t=1</sub> and Origin<sub>t=2</sub> have very similar measurement properties. This appears reasonable given that we are dealing with the same measure at two different time points. In contrast, Table 6 shows large differences between the measurement properties of the same question at different time points. The final model, thus, yields results that are easier to interpret than Johnson’s (1990) model, but does account for the local dependencies due to measurement occasion. It leads to two new conclusions for the U.S. Census: 1) Considering the false positive and negative rates in Table 8, Origin and Language may be the better measures of ethnicity, where the choice of measurement occasion is inconsequential; 2) the false negative rates in all indicators are considerable, meaning that the number of U.S. residents of Hispanic ethnicity is likely to be underestimated.

## 6 Application 2: Dentistry x-ray ratings

Espeland & Handelman (1989) used latent class modeling to explain 3869 ratings given

by five dentists ( $y_1$ – $y_5$ ) to x-rays that may show incipient caries (1) or not (0). Each rating is a binary observed variable, and two latent classes represent true caries state. Fitting the two-class local independence model, which we call  $M_0$ , to these data yields a badly fitting model with a deviance ( $L^2$ ) of 129.9 on 20 degrees of freedom (bootstrap  $p < 10^{-10}$ ) and a BIC of -35.4. These authors then suggested to increase the number of classes to four.

Qu et al. (1996, pp. 804-6) re-analyzed these data, and argued that the two-class model taking into account local dependencies is easier to interpret than the four class model suggested by Espeland & Handelman (1989). Qu et al. (1996, p. 799) introduced an alternative approach to taking local dependency into account, whereby the dependencies are parameterized as arising from a Gaussian random effect variable, whose effects are allowed to differ over classes. A reformulation of their model can be obtained by modifying Equation (2) as

$$\boldsymbol{\eta}_t = \mathbf{X}_{(Y)}\boldsymbol{\tau} + \mathbf{X}_{(Y\xi_t)}\boldsymbol{\lambda} + \mathbf{b}_t \cdot \eta, \quad (10)$$

where  $\eta$  is a Gaussian latent factor (random effects) variable,  $\eta \sim N(0, 1)$ , the latent variables  $\xi$  and  $\eta$  are assumed to be independent, and  $\mathbf{b}_t$  are vectors of class-specific factor loadings. That is, instead of using direct loglinear effect parameters to model the local dependencies, a continuous latent factor on which all items load is posited. Parameter estimates can be obtained through numerical integration (Vermunt & Magidson, 2013).

Qu et al. (1996, p. 805) fit their random effects model under the restriction that the loadings for  $y_1$  and  $y_2$ , the first two items, are equal, i.e.,  $\mathbf{b}_{(1,t)} = \mathbf{b}_{(2,t)}$ . This model,  $M_{\text{Qu}}$ , appears to fit the data quite well, with a deviance of 15.8 on 12 degrees of freedom (bootstrap  $p = 0.38$ ). The BIC is equal to -83.4, so that the improvement in model fit appears to somewhat outweigh the added complexity of this model.

While it appears to fit the data well, the random effects model  $M_{\text{Qu}}$  has two problems.

First, the maximum-likelihood solution given by Qu et al. (1996) involves boundary estimates that prevent the estimation of asymptotic standard errors. When standard errors are estimated using nonparametric bootstrapping, four (out of eight) factor loadings do not differ statistically significantly from zero. There thus appears to be considerable room for model simplification and additional parsimony. Second, even under the already rather complex  $M_{Qu}$ , a score test for the local dependence parameters reveals one strong and statistically significant residual dependence between  $y_3$  and  $y_4$  ( $T_L = 10.5$ ,  $EPC_L = -1.22$ ). This indicates that the random effects model cannot completely account for all local dependencies.

It is in principle possible to formulate a model that includes local dependence parameters as well as a random effect, but it would be preferable to find a more parsimonious model using loglinear local dependence parameters that also fits the data well. We now demonstrate that the EPC measures together with the score tests can be used to find a model that does not have the disadvantages of the random effects model, but that does account for local dependence while also retaining the easier-to-interpret two-class solution.

Under the local independence model  $M_0$ , we calculated EPC measures and score tests, shown in Table 10. It can be seen that EPC's and score statistics are large for the five bivariate local dependencies in rows 2, 3, 5, 7, and 9 of the Table.

Based on the EPC and score test values under  $M_0$  in Table 10, we proceed to formulate a model in which the loglinear local dependence parameters corresponding to rows 2, 3, 5, 7, and 9 of the Table are freed. This model, which we call  $M_1$ , has 15 degrees of freedom, a deviance of 35.7 ( $p = 0.0019$ ), and a BIC of -88.2. Although the BIC would prefer this model over the random effects model, there still appears to be model misfit. The largest  $EPC_L$  (0.93) and score test ( $T_L = 8.8$ ) are those for the fixed local dependence between  $y_1$  and  $y_5$  (row 4 in the Table). Moreover, the free local dependence parameter between  $y_1$  and  $y_4$  (row 3 in the Table) is estimated at a small (0.16) and not statistically significant

Table 10: EPC and score tests for loglinear local dependence parameters between five dentists' x-ray ratings for caries under the local independence model  $M_0$ .

Number	Bivariate dependence	$EPC_L$	$T_L$	$EPC_{GS}$	$T_{GS}$
1	1 ↔ 2	0.32	3.1	0.35	3.4
2	1 ↔ 3	1.04	34.0	0.97	31.6
3	1 ↔ 4	0.59	13.1	0.59	13.1
4	1 ↔ 5	0.47	2.7	0.44	2.6
5	2 ↔ 3	0.56	6.8	0.53	6.4
6	2 ↔ 4	0.23	1.8	0.22	1.7
7	2 ↔ 5	0.63	16.4	0.48	12.6
8	3 ↔ 4	-0.30	2.7	-0.35	3.2
9	3 ↔ 5	0.76	5.1	0.55	3.7
10	4 ↔ 5	0.42	3.5	0.27	2.3

value ( $p = 0.59$ ) under  $M_1$ .

Our final model ( $M_2$ ) therefore fixed the  $y_1 \leftrightarrow y_4$  dependence parameter to zero and freed the  $y_1 \leftrightarrow y_5$  dependence, as suggested by the EPC measures obtained under  $M_1$ . The final model  $M_2$  has a deviance of 28.4 with 15 degrees of freedom (bootstrap  $p = 0.07$ ) and a BIC of -95.5. This model therefore appears to fit the data well, and is strongly preferred by BIC over  $M_{Qu}$ ,  $M_0$ , and  $M_1$ .

Although the local dependence parameters are not of scientific interest in this application, it may aid understanding of the loglinear local dependence model to examine these values. Table 11 provides the estimates of the local dependence parameters under  $M_2$ , together with their corresponding Wald tests. In addition to these parameter estimates, it is possible to compute the within-class log-odds ratio between a pair of items, marginalized over all other variables (we thank an anonymous reviewer for this suggestion). These marginal log-odds ratios are given in the final two columns of Table 11. Because the locally dependent item pairs overlap, the  $\psi$  parameters no longer correspond to the marginal dependence within classes, as can be seen in the Table. Furthermore, the marginalized local dependencies differ over classes even though the loglinear dependence parameters  $\psi$

Table 11: Final local dependence latent class model ( $M_2$ ) results. Estimates of log-odds dependencies ( $\hat{\psi}$ ) between pairs of items with Wald tests under  $M_2$ . Also shown is the implied marginal dependence between pairs of items within classes.

Number	Bivariate dependence	Log-odds ratios within classes			
		Conditional		Marginal	
		Est. ( $\hat{\psi}$ )	Wald	Class 1	Class 2
1	1 ↔ 2	-		0.25	0.68
2	1 ↔ 3	1.38	49.7	1.48	1.66
3	1 ↔ 4	-		0	0
4	1 ↔ 5	0.74	60.4	0.84	1.25
5	2 ↔ 3	1.29	8.0	1.40	1.57
6	2 ↔ 4	-		0	0
7	2 ↔ 5	0.73	34.4	0.79	1.21
8	3 ↔ 4	-		0	0
9	3 ↔ 5	1.38	59.0	1.58	1.85
10	4 ↔ 5	-		0	0

-: Fixed to zero.

do not. Finally, it can be seen that the marginalized within-class dependence between  $y_1$  and  $y_2$  (row 1 of Table 11) is nonzero even though the loglinear dependence parameter is fixed to zero. This shows that the loglinear dependence model can account for certain, possibly class-dependent, marginal local dependencies in a rather parsimonious manner.

The estimates of specificity, sensitivity, and prevalence under  $M_2$  are given in Table 12. It can be seen that specificity values are quite high, with few false negatives, except for dentist #5 (0.68). At the same time, dentist #5 appears to have a rather high sensitivity (0.82) of x-ray judgments compared with his or her colleagues, who appear to err on the side of non-detection of caries. The estimated values under  $M_2$  in Table 12 differ from those under  $M_{Qu}$  especially for dentist #4:  $M_{Qu}$  would estimate this dentist's sensitivity at 0.68, whereas Table 12 shows that under the  $M_2$  model this estimate is 0.57. If this model is indeed to be preferred, it would appear that this dentist's performance with regard to false negatives is even worse than was previously thought.

Table 12: Local dependence latent class model results. Prevalence (class sizes), and sensitivity,  $\Pr(y = \text{Yes}|\xi = 2)$ , and specificity,  $\Pr(y = \text{No}|\xi = 1)$ , estimates for the five dentists’ judgments of x-rays for caries under  $M_2$ .

Class		Dentist 1		Dentist 2		Dentist 3		Dentist 4		Dentist 5	
#	Size	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
1	0.79	0.99	0.01	0.88	0.12	0.96	0.04	1.00	0.00	0.68	0.32
2	0.21	0.63	0.37	0.39	0.61	0.53	0.47	0.43	0.57	0.18	0.82

## 7 Conclusion

We have shown how the  $EPC_L$  and  $EPC_{GS}$  can aid in the detection of local dependence when the commonly made local independence assumption in latent class analysis of binary data does not hold. The asymptotic and finite sample properties of these measures appear adequate for this purpose. Applications to two real datasets previously analyzed by other authors demonstrated the advantage of this approach in trading off model realism and parsimony, and showed that different and more easily interpretable results can be obtained.

Extensions to polytomous data are possible in our framework by adjusting the relevant design matrices. Unless additional restrictions are imposed, the local dependence parameter for a pair of variables will then become multivariate. Class-specific and trivariate local dependencies can likewise be accommodated, as can latent class models including covariates. Finally, the  $EPC_L$  and  $EPC_{GS}$  could be applied to other parameters than local dependencies. For example, Glas (1999) suggested examining item bias (direct effects of covariates on response variables) in item response models. Based on our findings, the  $EPC_L$ ,  $EPC_{GS}$ , and corresponding score statistics have been implemented in the standard latent class modeling software Latent GOLD 5, which allows for the above extensions (Vermunt & Magidson, 2013, pp. 133–4). The online supplement provides R code (R Core Team, 2012) for the applications.

Although Section 2 developed the EPC measures for restrictions in a general maximum-



likelihood framework with misspecification, our simulation and applications have been limited to the investigation of class-independent loglinear local dependence parameters with exactly one misspecification. An evaluation of the performance of the EPC measures for such parameters, as well as in the case of more than one misspecification, remain topics for future investigation.

When conditional dependencies are among non-overlapping pairs of items, the direct loglinear effect parameters will equal the marginal local dependencies between pairs of items conditional on the latent class. When pairs of items overlap, this is no longer the case, but marginal dependence could be parameterized using the marginal model formulation of Bartolucci & Forcina (2006) and Reboussin et al. (2008). This will be advantageous when the goal is to interpret the local dependence parameters, or when the base model from which EPC's are calculated already includes local dependencies itself. Development of score tests and EPC measures for marginal models is therefore another interesting topic for future study.

## **A Information matrices, Jacobian, and identification of the locally dependent latent class model**

This appendix defines the information matrices, Jacobian, and outer product matrix for (partially) locally dependent latent class models used in the derivation of the EPC. We also provide a theorem giving conditions under which the local dependence parameters are locally identifiable.

By applying the rules of vector differentiation to model (1), the Jacobian of the patternwise likelihood vector with respect to one of the parameter vectors  $\tau$ ,  $\lambda$ , or  $\psi$  is

obtained as

$$\mathbf{S}_{(\cdot)} = \frac{\partial \log \Pr(\mathbf{Y})}{\partial (\cdot)} = \sum_{t=1}^T [\mathbf{1}' \otimes \Pr(\xi = t | \mathbf{Y} = \mathbf{y}) \circ (\mathbf{X}_{(\cdot)} - E_R[\mathbf{X}_{(\cdot)}])], \quad (11)$$

where  $\circ$  denotes the elementwise (“Hadamard”) product, the kronecker product  $\otimes$  here serves to duplicate the posterior probabilities columnwise,  $\mathbf{X}_{(\cdot)}$  is the design matrix corresponding to either  $\boldsymbol{\tau}$ ,  $\boldsymbol{\lambda}$ , or  $\boldsymbol{\psi}$ , and  $E_R[\mathbf{X}_{(\cdot)}]$  is a matrix with  $R$  rows, in which each row equals  $\mathbf{X}'_{(\cdot)} \Pr(\mathbf{Y} = \mathbf{y} | \boldsymbol{\xi})$ . For a two-class model with effect coding, the Jacobian with respect to the latent class intercept parameter is

$$\mathbf{S}_\alpha = \frac{\partial \log \Pr(\mathbf{Y})}{\partial \alpha} = 2[\Pr(\xi = 1 | \mathbf{Y} = \mathbf{y}) - \Pr(\xi = 1)]. \quad (12)$$

That is, the Jacobian depends on the change in the latent class classification before and after observation of  $\mathbf{Y}$ . This change therefore plays a large role in the determinant of the outer product of the patternwise score vectors used below.

Using obvious notation for the full Jacobian  $\mathbf{S}(\boldsymbol{\theta})$ , the gradient ( $p$ -score vector) over all response patterns will equal

$$\mathbf{s} = \frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \sum_{i=1}^N \frac{\partial \ell_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \mathbf{S}(\boldsymbol{\theta})' \mathbf{n}. \quad (13)$$

Define the observed and expected information matrices as

$$\mathbf{I}_Y = -\frac{\partial \mathbf{s}}{\partial \boldsymbol{\theta}'} = -\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}, \quad (14)$$

$$\mathbf{I}_L = E_L(\mathbf{I}_Y) = \sum_{r=1}^R \hat{n}_r \mathbf{S}_r(\boldsymbol{\theta})' \mathbf{S}_r(\boldsymbol{\theta}), \quad (15)$$

where  $\hat{n}_r = n \cdot \Pr(\mathbf{Y} = \mathbf{y}_r)$  and the outer product matrix as

$$\mathbf{D} = \sum_{r=1}^R n_r \mathbf{S}_r(\boldsymbol{\theta})' \mathbf{S}_r(\boldsymbol{\theta}). \quad (16)$$

The form of the Jacobian in equation (11) can be used to determine identifiability.

**Lemma 1.** *Assume that  $\mathbf{S}_{\boldsymbol{\theta}_2}$  is of full column rank. Let  $\mathbf{X}_{\text{new}}$  denote a design matrix such that:*

- (i.)  $\mathbf{X}_{\text{new}}$  is of full column rank;
- (ii.) The number of columns in  $\mathbf{X}_{\text{new}}$  is smaller than or equal to the number of degrees of freedom  $df = R - 1 - \text{rk}(\mathbf{S}_{\boldsymbol{\theta}_2})$ ;
- (iii.) The columns of  $\mathbf{X}_{\text{new}}$  are linearly independent of the columns of the design matrix  $\mathbf{X}_{\boldsymbol{\theta}_2}$  corresponding to the parameters  $\boldsymbol{\theta}_2$ ;
- (iv.)  $\mathbf{X}_{\text{new},t} = \mathbf{X}_{\text{new}}$  for all  $t \in \{1..T\}$ .

Then the parameters  $\boldsymbol{\theta}_{\text{new}}$  corresponding to  $\mathbf{X}_{\text{new}}$  in model (1) are locally identifiable.

*Proof.* To show local identifiability, it suffices to show that  $\mathbf{S}_{\text{new}}$  is of full column rank and its columns linearly independent of those in  $\mathbf{S}_{\boldsymbol{\theta}_2}$  (Goodman, 1974). Since  $\mathbf{X}_{\text{new}}$  is not class-specific by (iv), equation (11) reduces to  $\mathbf{S}_{\text{new}} = \mathbf{X}_{\text{new}} - E_R(\mathbf{X}_{\text{new}})$ , so that  $\text{rk}(\mathbf{S}_{\text{new}}) = \text{rk}(\mathbf{X}_{\text{new}})$ , implying full column rank by (i). Furthermore, by assumption  $\text{rk}(\mathbf{S}_{\boldsymbol{\theta}_2}) = \text{rk}(\mathbf{X}_{\boldsymbol{\theta}_2})$ , so that by equation (11), (ii) and (iii) guarantee that the columns of  $\mathbf{S}_{\text{new}}$  are also independent of those in  $\mathbf{S}_{\boldsymbol{\theta}_2}$ .  $\square$

The proof of Theorem 1 follows directly from the Lemma and model (2). It suggests that when the local independence model is identifiable and the number of local dependencies  $\psi$  to be freed does not exceed the degrees of freedom, these additional parameters will also be identifiable.

## B Supplemental materials

**R code:** Provides S4 classes to perform latent class analysis for binary variables with local dependencies and obtain the  $EPC_L$  and  $EPC_{GS}$  and score tests. Includes both data sets used as examples in the article. (GNU zipped tar file)

## References

- Agresti, A. (2002). *Categorical data analysis, 2nd ed.* New York: Wiley-Interscience.
- Aitchison, J., & Silvey, S. (1958). Maximum-likelihood estimation of parameters subject to restraints. *The Annals of Mathematical Statistics*, 29(3), 813–828.
- Albert, P., & Dodd, L. (2004). A cautionary note on the robustness of latent class models for estimating diagnostic error without a gold standard. *Biometrics*, 60(2), 427–435.
- Bandeen-Roche, K., Miglioretti, D., Zeger, S., & Rathouz, P. (1997). Latent variable regression for multiple discrete outcomes. *Journal of the American Statistical Association*, 92(440), 1375–1386.
- Bartolucci, F., & Forcina, A. (2006). A class of latent marginal models for capture–recapture data with continuous covariates. *Journal of the American Statistical Association*, 101(474), 786–794.
- Biemer, P. (2011). *Latent class analysis of survey error.* New York: Wiley.
- Boos, D. (1992). On generalized score tests. *The American Statistician*, 46(4), 327–333.
- Breusch, T., & Pagan, A. (1980). The lagrange multiplier test and its applications to model specification in econometrics. *The Review of Economic Studies*, 47(1), 239–253.

- Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: methods and applications*. Cambridge university press.
- Catchpole, E., & Morgan, B. (1997). Detecting parameter redundancy. *Biometrika*, *84*(1), 187–196.
- Dayton, C., & Macready, G. (1988). Concomitant-variable latent-class models. *Journal of the American Statistical Association*, *83*(401), 173–178.
- Dempster, A., Laird, N., & Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, *39*(1), 1–38.
- Dendukuri, N., Hagdu, A., & Wang, L. (2009). Modeling conditional dependence between diagnostic tests: A multiple latent variable model. *Statistics in Medicine*, *28*, 441–461.
- Espeland, M., & Handelman, S. (1989). Using latent class models to characterize and assess relative error in discrete measurements. *Biometrics*, *45*(2), 587–599.
- Evers, M., & Namboodiri, N. (1979). On the design matrix strategy in the analysis of categorical data. *Sociological Methodology*, *10*, 86–111.
- Faraone, S., & Tsuang, M. (1994). Measuring diagnostic accuracy in. *Am J Psychiatry*, *1*(51), 651.
- Forcina, A. (2008). Identifiability of extended latent class models with individual covariates. *Computational Statistics & Data Analysis*, *52*(12), 5263–5268.
- Formann, A. (1992). Linear logistic latent class analysis for polytomous data. *Journal of the American Statistical Association*, *87*(418), 476–486.

- Formann, A. (2003). Latent class model diagnosis from a frequentist point of view. *Biometrics*, 59(1), 189–196.
- Garrett, E., Eaton, W., & Zeger, S. (2002). Methods for evaluating the performance of diagnostic tests in the absence of a gold standard: a latent class model approach. *Statistics in Medicine*, 21(9), 1289–1307.
- Garrett, E., & Zeger, S. (2000). Latent class model diagnosis. *Biometrics*, 56(4), 1055–1067.
- Glas, C. (1999). Modification indices for the 2-PL and the nominal response model. *Psychometrika*, 64(3), 273–294.
- Goodman, L. (1974). Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika*, 61(2), 215.
- Hadgu, A., Dendukuri, N., & Hilden, J. (2005). Evaluation of nucleic acid amplification tests in the absence of a perfect gold-standard test: a review of the statistical and epidemiologic issues. *Epidemiology*, 16(5), 604–612.
- Hagenaars, J. A. P. (1988). Latent structure models with direct effects between indicators local dependence models. *Sociological Methods & Research*, 16(3), 379–405.
- Hagenaars, J. A. P. (1993). *Loglinear models with latent variables*. Sage.
- Harper, D. (1972). Local dependence latent structure models. *Psychometrika*, 37(1), 53–59.
- Hill, J., & Kriesi, H. (2001). Classification by opinion-changing behavior: A mixture model approach. *Political Analysis*, 9(4), 301–324.

- Hofmann, T. (2001). Unsupervised learning by probabilistic latent semantic analysis. *Machine Learning*, 42(1), 177–196.
- Huang, G., & Bandeen-Roche, K. (2004). Building an identifiable latent class model with covariate effects on underlying and measured variables. *Psychometrika*, 69(1), 5–32.
- Hui, S., & Zhou, X. (1998). Evaluation of diagnostic tests without gold standards. *Statistical methods in medical research*, 7(4), 354–370.
- Johnson, R. (1990). Measurement of hispanic ethnicity in the us census: An evaluation based on latent-class analysis. *Journal of the American Statistical Association*, 85(409), 58–65.
- Jones, G., Johnson, W., Hanson, T., & Christensen, R. (2010). Identifiability of models for multiple diagnostic testing in the absence of a gold standard. *Biometrics*, 66, 855–863.
- Kolenikov, S., Steinley, D., & Thombs, L. (2010). *Statistics in the social sciences*. New York: John Wiley & Sons.
- Magidson, J., & Vermunt, J. K. (2004). Latent class models. In D. Kaplan (Ed.), *The sage handbook of quantitative methodology for the social sciences* (pp. 175–198). Thousand Oaks, California: Sage Publications, Inc.
- Maydeu-Olivares, A., & Joe, H. (2005). Limited-and full-information estimation and goodness-of-fit testing in  $2^n$  contingency tables. *Journal of the American Statistical Association*, 100(471), 1009–1020.
- McHugh, R. (1956). Efficient estimation and local identification in latent class analysis. *Psychometrika*, 21(4), 331–347.
- McLachlan, G., & Peel, D. (2000). *Finite mixture models* (Vol. 299). Wiley-Interscience.

- Oberski, D. (2013). Local dependence in latent class models: application to voting in elections. In E. Brentari & M. Carpita (Eds.), *Advances in latent variables [sis 2013 conference proceedings]*. Milan, Italy: Vita e Pensiero.
- Oberski, D., Van Kollenburg, G., & Vermunt, J. (2013). A Monte Carlo evaluation of three methods to detect local dependence in binary data latent class models. *Advances in Data Analysis and Classification*, 7(3). doi: 10.1007/s11634-013-0146-2
- Qu, Y., Tan, M., & Kutner, M. (1996). Random effects models in latent class analysis for evaluating accuracy of diagnostic tests. *Biometrics*, 52(3), 797–810.
- R Core Team. (2012). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from <http://www.R-project.org/> (ISBN 3-900051-07-0)
- Rao, C. R. (1948). Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. *Proceedings of the Cambridge Philosophical Society*, 44(1), 50–57.
- Reboussin, B., Ip, E., & Wolfson, M. (2008). Locally dependent latent class models with covariates: an application to under-age drinking in the usa. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 171(4), 877–897.
- Saris, W., Satorra, A., & Sörbom, D. (1987). The detection and correction of specification errors in structural equation models. *Sociological Methodology*, 17, 105–129.
- Sepúlveda, R., Vicente-Villardón, J. L., & Galindo, M. P. (2008). The biplot as a diagnostic tool of local dependence in latent class models: a medical application. *Statistics in Medicine*, 27, 1855–1869.



- Skrondal, A., & Rabe-Hesketh, S. (2004). *Generalized latent variable modeling : multilevel, longitudinal, and structural equation models*. Boca Raton, FL: Chapman & Hall/CRC.
- Sörbom, D. (1989). Model modification. *Psychometrika*, *54*(3), 371–384.
- Stanghellini, E., & Vantaggi, B. (2013). Identification of discrete concentration graph models with one hidden binary variable. *Bernoulli*, DOI: 10.3150/12-BEJ435.
- Torrance-Rynard, V., & Walter, S. (1998). Effects of dependent errors in the assessment of diagnostic test performance. *Statistics in Medicine*, *16*(19), 2157–2175.
- Vacek, P. (1985). The effect of conditional dependence on the evaluation of diagnostic tests. *Biometrics*, *41*(4), 959–968.
- van der Linden, W., & Glas, C. (2010). Statistical tests of conditional independence between responses and/or response times on test items. *Psychometrika*, *75*(1).
- Vermunt, J. K., & Magidson, J. (2005). Technical guide for latent GOLD 4.0: Basic and advanced. *Belmont Massachusetts: Statistical Innovations Inc.*
- Vermunt, J. K., & Magidson, J. (2013). *Technical guide for latent GOLD 5.0: Basic and advanced*. Belmont, MA: Statistical Innovations Inc.
- Walter, S., & Irwig, L. (1988). Estimation of test error rates, disease prevalence and relative risk from misclassified data: a review. *Journal of clinical epidemiology*, *41*(9), 923–937.
- White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica: Journal of the Econometric Society*, *51*(1), 1–25.

Yang, I., & Becker, M. (1997). Latent variable modeling of diagnostic accuracy. *Biometrics*, 53(3), 948–958.

Yen, W. (1984). Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. *Applied Psychological Measurement*, 8(2), 125–145.