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

Daniel L. Oberski* Jeroen K. Vermunt

Tilburg University, 5000 LE Tilburg, The Netherlands

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.

*Daniel L. Oberski is postdoctoral researcher (email: doberski@tilburguniversity.edu); and Jeroen K. Vermunt is Full Professor (email: j.k.vermunt@uvt.nl), both at the Department of Methodology and Statistics, Tilburg University, The Netherlands. This research was supported by grant ——- of the Netherlands Organisation for Scientific Research (NWO).

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 and 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 and 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, the latent class model for binary data has proved key to describing prevalence and symptomatology of diseases and assessing the accuracy of diagnoses (Faraone and Tsuang, 1994), and to evaluating the sensitivity, specificity, and predictive validity of diagnostic tests in the absence of a gold standard (Walter and Irwig, 1988; Hui and Zhou, 1998; Garrett et al., 2002). For instance, Bartolucci and Forcina (2006) discussed an application to capture-recapture data for estimating the prevalence of HIV; Fergusson et al. (1994) studied the comorbidity of five problem behaviors in Christchurch adolescents; Bandeen-Roche et al. (1997) modeled severe disability as measured by five task disability outcomes and linked these to risk factors in a latent variable regression extension of the latent class model; Uebersax (1993) applied the latent class model to expert ratings of appropriateness of the medical procedure carotid endarterectomy to reduce the risk of strokes, and Uebersax and Grove (1993) to the measurement of liver metastasis in patients using three different medical imaging techniques.

Latent class models crucially assume "local independence": conditional on the latent class (finite mixture component), the responses should be independent. In practice this assumption may be violated. Local dependencies may arise when there are additional scientifically interesting dimensions, but also when there is a "nuisance" dependency. For instance, the "topic" of a document may not suffice to explain the number of times pairs of words occur together in it; responses on 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.

Violations of the local independence assumption in latent class models generally lead to bias in the outcomes of interest (Vacek, 1985; Torrance-Rynard and Walter, 1998; Albert and Dodd, 2004; Hadgu et al., 2005). Pepe and Janes (2007) therefore suggested that "careful justification of assumptions about the dependence between tests in diseased and nondiseased subjects is necessary", while Albert and Dodd (2008) suggested collecting additional gold standard information on a subset of subjects. These excellent suggestions may not always be practicable, however. A standard solution is then to increase the number of classes, so that the additional classes can represent (absorb) the dependencies. However, this solution is undesirable when the dependence is really a nuisance that is not of scientific interest: substantively uninterpretable latent classes may result. Interpretation in diagnostic testing, for instance, is much simpler if a two-class "diseased"/"nondiseased" model can be found. As an alternative to increasing the number of latent classes, Harper (1972); Hagenaars (1988); Espeland and Handelman (1989); Formann (1992); Qu et al. (1996); Bartolucci and Forcina (2006) and Reboussin et al. (2008) propose using a direct modeling approach, which involves including local dependence parameters into the latent class model. A practical problem, however, when using these local dependence latent class models is that one has to determine which local dependencies should included in the model.

This article introduces a procedure for detecting local dependencies among pairs of observed variables after fitting a (partial) local independence model. We propose to use "expected parameter change" (EPC) measures, which estimate the value that a restricted local dependence parameter would take on if it were freed in the model. Our proposal extends the direct modeling approach to the problem of local dependence by evaluating the substantive size and direction of possible local dependencies before introducing dependence parameters. 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). Instead of a statistical test of the hypothesis of no local dependence, however, the EPC measures provide approximately consistent estimates of the substantive size and direction of possible local dependencies. The method proposed may be seen as an extension of residuals-based measures of local dependence such as those proposed by Formann and Kohlmann (1996), Qu et al. (1996), Garrett and Zeger (2004), and Vermunt and Magidson (2005).

The article is organized as follows. Section 2 presents the latent class model for binary variables with possible local dependencies. The EPC_L and EPC_{GS} for such models are introduced in section 3. The 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 an i.i.d. 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 ξ can then be formulated (Formann, 1992) as

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

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

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

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 and 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).

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 $\boldsymbol{\psi} = \mathbf{0}$. 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 $\boldsymbol{\psi}$ 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 and identifiability

The maximum likelihood estimates $\hat{\boldsymbol{\theta}}_2$ under the restricted model can be found by maximizing 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. 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.

Local identifiability is a crucial issue for the interpretation of results and the validity of asymptotic approximations (Forcina, 2008). The distribution F is said to be locally identifiable at the parameter $\boldsymbol{\theta}_2^0$ if there exists some neigborhood $\boldsymbol{\phi}$ of $\boldsymbol{\theta}_2^0$ such that

$$\begin{aligned} F_{\mathbf{Y}}(\mathbf{y}; \boldsymbol{\theta}_2^0) = F_{\mathbf{Y}}(\mathbf{y}; \boldsymbol{\theta}_2) \;\; \forall \mathbf{y} \in S_{\mathbf{Y}} \Leftrightarrow \\ \boldsymbol{\theta}_2 = \boldsymbol{\theta}_2^0, \forall \boldsymbol{\theta}_2 \in \boldsymbol{\phi} \subset \boldsymbol{\Theta}, \end{aligned} \tag{3}$$

where $S_{\mathbf{Y}}$ denotes the distributional support of $\mathbf{Y}$, and $\boldsymbol{\Theta}$ the set of all possible $\boldsymbol{\theta}_2$

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

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

(Bandeen-Roche et al., 1997, p. 1378; see also Huang and Bandeen-Roche, 2004). As noted by Goodman (1974) and shown by Catchpole and Morgan (1997, theorem 1), the model described in equation 1 will be identifiable if its Jacobian $\mathbf{S}$ of the expected response patterns with respect to the parameters is of full column rank (see also McHugh, 1956, theorem 1). The appendix gives the precise form of the Jacobian for the latent class model with possible local dependencies.

An obvious necessary condition for identifiability is that there are not more columns in $\mathbf{S}$ than independent rows, i.e. the number of parameters should not exceed the number of unique response patterns, $R - 1 \geq p$. However, this is not a sufficient condition, as evidenced by the unconstrained three-class model for four binary observed variables, which has one degree of freedom but is not identified.

According to Harper (1972, p. 58), a sufficient condition for identification of the latent class model containing all pairwise local dependencies is that $J \geq 2T + 2$. In practice fewer than $2T + 2$ items may suffice for local identifiability of local dependence parameters. Theorem 1 given in the appendix shows that whenever the local independence model is locally identifiable, as many local dependence parameters as exhaust the degrees of freedom are also locally identifiable.

As suggested by Forcina (2008, p. 5266), identifiability can be examined empirically by randomly sampling a large number of parameter sets and examining the rank of the expected information matrix for each set. Note that the rank of the expected information matrix is equal to the rank of the Jacobian (Formann, 1992). 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. Table 1 shows the results of applying this method to models with an increasing number of classes and variables. The table reports the number of local dependencies that can be identified, where a dash indicates that even the local independence model is not identifiable. These results show Theorem 1 of the appendix 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.

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 $\boldsymbol{\psi} = \mathbf{0}$. 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 (Saris 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.

Under the correctly specified model, let $\boldsymbol{\theta}^*$ be the value to which $\hat{\boldsymbol{\theta}}$ converges in probability as the sample size increases. Additionally, let the score $\mathbf{s}(\boldsymbol{\theta}^*) := \partial \ell / \partial \boldsymbol{\theta}$ evaluated at $\boldsymbol{\theta}^*$. Then the loglikelihood ℓ shown in equation 1 can be approximated by a Taylor expansion as

$$\ell \approx \hat{\ell} + \begin{bmatrix} \boldsymbol{\theta}_1^* - \hat{\boldsymbol{\theta}}_1 \\ \boldsymbol{\theta}_2^* - \hat{\boldsymbol{\theta}}_2 \end{bmatrix}' \begin{bmatrix} \mathbf{s}_1(\boldsymbol{\theta}^*) \\ \mathbf{s}_2(\boldsymbol{\theta}^*) \end{bmatrix} + \frac{1}{2} \begin{bmatrix} \boldsymbol{\theta}_1^* - \hat{\boldsymbol{\theta}}_1 \\ \boldsymbol{\theta}_2^* - \hat{\boldsymbol{\theta}}_2 \end{bmatrix}' \begin{bmatrix} \mathbf{I}_{Y11}^* & \mathbf{I}_{Y21}^* \\ \mathbf{I}_{Y12}^* & \mathbf{I}_{Y22}^* \end{bmatrix} \begin{bmatrix} \boldsymbol{\theta}_1^* - \hat{\boldsymbol{\theta}}_1 \\ \boldsymbol{\theta}_2^* - \hat{\boldsymbol{\theta}}_2 \end{bmatrix}, \quad (4)$$

where $\mathbf{I}_Y^*$ is the observed information matrix at $\boldsymbol{\theta}^*$. As demonstrated in this equation, the information matrix is partitioned following the partition $\boldsymbol{\theta}' = (\boldsymbol{\theta}_1', \boldsymbol{\theta}_2')$.

To study what would happen if the restricted parameter vector $\boldsymbol{\theta}_1$ were freed, we find new estimates by maximizing ℓ (equation 1), this time with respect to both $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ (Sörbom, 1989, p. 373). This leads to the equality

$$\begin{bmatrix} \mathbf{s}_1(\boldsymbol{\theta}^*) \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{I}_{Y11}^* & \mathbf{I}_{Y21}^* \\ \mathbf{I}_{Y12}^* & \mathbf{I}_{Y22}^* \end{bmatrix} \begin{bmatrix} \boldsymbol{\theta}_1^* - \hat{\boldsymbol{\theta}}_1 \\ \boldsymbol{\theta}_2^* - \hat{\boldsymbol{\theta}}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}. \quad (5)$$

Note that $\mathbf{I}_Y^*$ cannot be obtained from the maximum likelihood solution as it depends on the unknown value $\boldsymbol{\theta}^*$. However, consistent estimates of the shift in parameter values if $\boldsymbol{\theta}_1$ were freed can be obtained from the restricted solution as the "expected parameter change" EPC $:= \hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1^* \approx -\hat{\mathbf{V}}^{-1}\mathbf{s}_1(\hat{\boldsymbol{\theta}})$, 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 $T = \mathbf{s}_1(\hat{\boldsymbol{\theta}})'\hat{\mathbf{V}}^{-1}\mathbf{s}_1(\hat{\boldsymbol{\theta}})$ which is distributed as $\chi^2_{\mathrm{rk}(\mathbf{S}_1)}$ under the null hypothesis.

Under the null hypothesis $\boldsymbol{\psi} = \mathbf{0}$, 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)

$$\begin{aligned} \mathrm{EPC}_L &= -\hat{\mathbf{V}}_L^{-1}\mathbf{s}_1(\hat{\boldsymbol{\theta}}) = -\hat{\mathbf{I}}_L^{-1}\,\mathbf{s}_1(\hat{\boldsymbol{\theta}}) \\ &= -(\hat{\mathbf{I}}_{L11} - \hat{\mathbf{I}}_{L12}\hat{\mathbf{I}}_{L22}^{-1}\hat{\mathbf{I}}_{L21})^{-1}\,\mathbf{s}_1(\hat{\boldsymbol{\theta}}), \end{aligned} \quad (6)$$

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 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 $T_L = \mathbf{s}_1(\hat{\boldsymbol{\theta}})'\hat{\mathbf{I}}_L^{-1}\mathbf{s}_1(\hat{\boldsymbol{\theta}})$, which under the null hypothesis has a chi square distribution with $\mathrm{rank}(\mathbf{S}_1)$ 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 and Silvey, 1958; Breusch and 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 and Glas, 2010).

The derivation of $\hat{\mathbf{V}}$ under the null hypothesis suggests that when $\boldsymbol{\psi} \neq \mathbf{0}$, the EPC_L is asymptotically biased. Under misspecified local independence, a "generalized", i.e. robust to misspecification, consistent estimate $\hat{\mathbf{V}}_{\mathrm{GS}}$ of $\mathbf{V}$ can be used (White, 1982). As shown by Boos (1992, p. 329),

$$\begin{aligned}\hat{\mathbf{V}}_{\mathrm{GS}} &= (\mathbf{1}, -\hat{\mathbf{I}}_{Y12}\hat{\mathbf{I}}_{Y22}^{-1})\hat{\mathbf{D}}(\mathbf{1}, -\hat{\mathbf{I}}_{Y12}\hat{\mathbf{I}}_{Y22}^{-1})' \\ &= \hat{\mathbf{D}}_{11} - \hat{\mathbf{I}}_{Y12}\hat{\mathbf{I}}_{Y22}^{-1}\hat{\mathbf{D}}_{12}' - \hat{\mathbf{D}}_{12}\hat{\mathbf{I}}_{Y22}^{-1}\hat{\mathbf{I}}_{Y12}' + \hat{\mathbf{I}}_{Y12}\hat{\mathbf{I}}_{Y22}^{-1}\hat{\mathbf{D}}_{22}\hat{\mathbf{I}}_{Y22}^{-1}\hat{\mathbf{I}}_{Y12}',\end{aligned} \tag{7}$$

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" $\mathrm{EPC}_{\mathrm{GS}}$ is obtained as $\mathrm{EPC}_{\mathrm{GS}} = -\hat{\mathbf{V}}_{\mathrm{GS}}^{-1}\mathbf{s}_1(\hat{\boldsymbol{\theta}})$; the well known generalized score test (White, 1982) is $T_{\mathrm{GS}} = \mathbf{s}_1(\hat{\boldsymbol{\theta}})'\hat{\mathbf{V}}_{\mathrm{GS}}^{-1}\mathbf{s}_1(\hat{\boldsymbol{\theta}})$.

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. 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

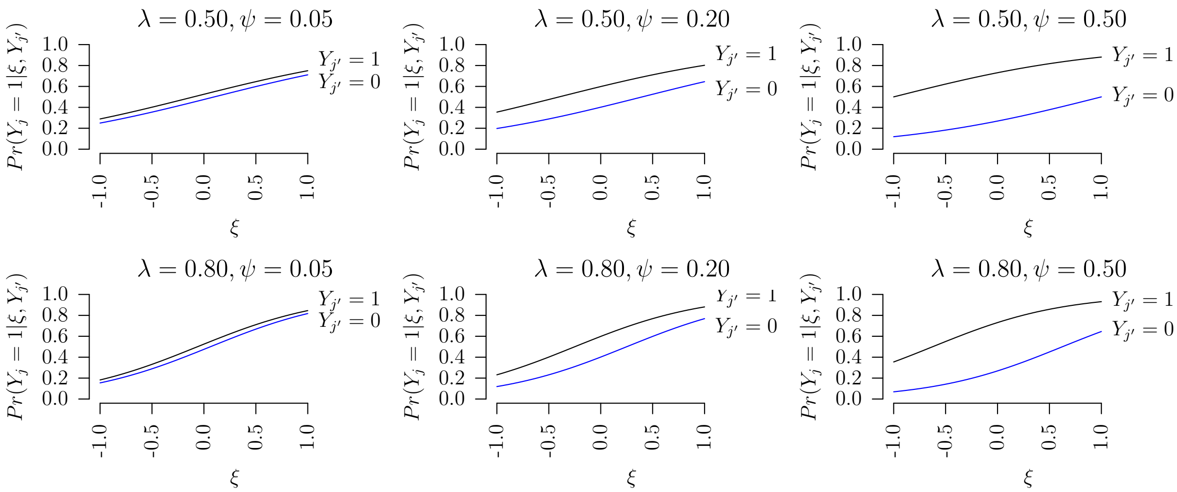
The population model is specified as a two-class model with five binary indicators and one local dependence between a pair of indicators. In our setup, all design matrices in equation 2 are chosen such that the columns sum to zero ("effect coding"). The intercepts $\boldsymbol{\tau} = \mathbf{0}$, the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

1. Local dependence size (ψ): -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);
2. Effect of latent variable on indicators (β): 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.

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



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 λ and the local dependence parameter ψ on the conditional probability for one observed variable. For illustrative purposes, many different values for the latent class variable ξ 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 λ 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 λ and the local dependence, ψ, between the two items. For illustrative purposes, the latent variable ξ 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 λ 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 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 percent-

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

	Local dependence (ψ)						
λ	-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 (ψ)						
λ	-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

age 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 relative bias appears to be within acceptable limits, showing that 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 400 times to yield a sampling distribution for EPC_L and EPC_{GS}.

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.

	Local dependence (ψ)					
	$\psi = -0.05$		$\psi = -0.20$		$\psi = -0.50$	
	Loading (λ)					
No. obs.	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

	$\psi = +0.05$		$\psi = +0.20$		$\psi = +0.50$	
No. obs.	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

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 ψ 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 ψ 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 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 the EPC_{GS}. Even for very 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 much better than the EPC_L both in the population and in finite samples. Overall the bias in the EPC_{GS} can be viewed as acceptable for the purpose of detecting the substantive size of local dependencies.

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.

	Local dependence (ψ)					
	$\psi = -0.05$		$\psi = -0.20$		$\psi = -0.50$	
	Loading (λ)					
No. obs.	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
	$\psi = +0.05$		$\psi = +0.20$		$\psi = +0.50$	
No. obs.	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

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: whether Spanish was spoken at home during childhood ("Language-interview") and Hispanic origin ("Origin-interview"). In a subsequent reinterview, two additional indicators of ethnicity were obtained: Hispanic ancestry ("Ancestry-reinterview") and a repetition of the "Origin-interview" measure ("Origin-reinterview"). We analyze the group of 9485 respondents not born in a Hispanic country. Of interest are false positive and false negative rates for the alternative question formulations.

Johnson (1990) 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 fit a model with two separate two-class latent variables corresponding to the two measurement occasions, which improved the deviance to 3.1 with 4 degrees of freedom ($p = 0.54$; BIC equals -33.5). The false negative rates were, respectively, 0.19 and 0.08 for Ancestry-reinterview and Origin-reinterview, measuring the first latent variable, and respectively 0.17 and 0.22 for Origin-interview and Language-interview, which measure the second latent variable. False positive rates were below 0.01 for all variables. Since the latent variables represent a nuisance dependency due to the measurement occasion, however (p. 64), these false positive and false negative rates are difficult to interpret in sociological terms.

Table 6: Local dependencies between indicators of Hispanic ethnicity in the Census.

Local dependence			EPC_L	T_L	EPC_{GS}	T_{GS}	$\tilde{\psi}$	Wald
Ancestry-re	$\leftrightarrow$	Language-in	0.92	5.0	1.45	7.9	0.22	0.2
Ancestry-re	$\leftrightarrow$	Origin-in	-0.76	2.5	-1.23	4.1	-0.23	0.1
Ancestry-re	$\leftrightarrow$	Origin-re	2.94	45.6	1.32	20.5	1.82	18.7
Language-in	$\leftrightarrow$	Origin-in	4.14	97.1	1.59	37.2	3.52	53.4
Language-in	$\leftrightarrow$	Origin-re	-1.08	7.9	-1.76	12.8	1.33	7.1
Origin-in	$\leftrightarrow$	Origin-re	1.10	6.1	2.20	12.2	0.52	0.3

Table 6 shows the results of calculating the EPC_L, EPC_{GS}, and the corresponding score statistics after fitting the two-class model. For reference, the column $\tilde{\psi}$ reports the estimate of the local dependence parameter concerned when it is freed, and the last column gives the corresponding Wald test. Table 6 clearly indicates the local dependencies that led Johnson (1990) to fit a two-dimensional model: the two local dependencies between observed variables measured at the same occasion have large score statistics and large and positive EPC_L and EPC_{GS} estimates.

Based on Table 6, a reasonable alternative model is the two-class model including both local dependencies – 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 ψ, the false negative rates can be interpreted as being with respect to a common latent class variable that might be labeled "Hispanic ethnicity".

The false negative rates, i.e. respondents "incorrectly" reporting a non-Hispanic ethnicity, estimated under the model allowing for local dependence within measurement occasions are 0.38, 0.30, 0.33, and 0.29, for Ancestry-reinterview, Origin-reinterview, Origin-interview, and Language-interview, respectively. The respective estimated false positive rates were 0.001, 0.001, 0.002, and 0.005. This suggests, in accordance with expectations, that the pure repetition of Origin has similar measurement properties on both occasions. This more easily interpretable model would lead to two new conclusions for the U.S. Census: 1) Origin may be the better measure 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

In their discussion of local dependencies in diagnostic testing, Qu et al. (1996, pp. 804-6) discuss a dataset due to Espeland and Handelman (1989) consisting of the ratings five dentists gave to dental x-rays that may show incipient caries. Each rating is a binary observed variable, and two latent classes represent true caries state. Qu et al. (1996, pp. 804-6) suggested that the two-class model taking into account local dependencies is easier to interpret than a four class model that was used in the earlier literature. These authors discussed an interesting alternative approach to taking local dependency into account, whereby the dependencies are parameterized as arising from a continuous random effect variable. This model is more parsimonious than freeing all local dependencies at once but also assumes that the dependencies are all in the same direction.

Table 7 shows the EPC's and score statistics for the local dependence parameters. The EPC's and score statistics are large for five bivariate local dependencies. The local dependency between dentists three and four is negative, violating the random effects model's assumption that all dependencies are in the same direction.

Table 7: Local dependencies between five dentists' x-ray ratings for caries.

Dentist dependence	EPC_L	T_L	EPC_{GS}	T_{GS}	$\tilde{\psi}$	Wald
1 ↔ 2	0.32	3.1	0.35	3.4	0.33	3.8
1 ↔ 3	1.04	34.0	0.97	31.6	1.07	40.5
1 ↔ 4	0.59	13.1	0.59	13.1	0.61	13.6
1 ↔ 5	0.47	2.7	0.44	2.6	0.42	2.3
2 ↔ 3	0.56	6.8	0.53	6.4	0.56	8.6
2 ↔ 4	0.23	1.8	0.22	1.7	0.25	1.8
2 ↔ 5	0.63	16.4	0.48	12.6	0.54	18.7
3 ↔ 4	-0.30	2.7	-0.35	3.2	-0.37	2.8
3 ↔ 5	0.76	5.1	0.55	3.7	0.70	6.0
4 ↔ 5	0.42	3.5	0.27	2.3	0.46	3.7

The deviance for the fully restricted two-class model equals 129.9 with 20 degrees of freedom ($p < 10^{-10}$) and the BIC is -35.4, while Qu et al. (1996, p. 805)'s random effects model including two additional equality restrictions has a much improved deviance of 15.8 with 12 degrees of freedom (bootstrap $p = 0.38$) and BIC equal to -83.4. An even more parsimonious model can be found in a stepwise manner by using the EPC's and score statistics shown in Table 7. We freed the large local dependencies between dentists one and two, one and five, two and four, two and five, and three and five. This model has a deviance of 28.4 with 15 degrees of freedom (bootstrap $p = 0.07$) and a BIC of -95.5. The BIC therefore favors this model over the random effects latent class model. The estimates of specificity, sensitivity, and prevalence are similar for these two models, never differing more than 0.03 in absolute value. There is one exception, however: under Qu et al.'s final model, the specificity of dentist four is estimated at the boundary of unity and the sensitivity at 0.68, while these estimates are 0.997 and 0.57 respectively under the final local dependence model arrived at using the EPC's and score statistics.

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 accomodated. 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 an experimental version of the standard latent class modeling software Latent GOLD, which allows for the above extensions (Vermunt and Magidson, 2005). The online supplement provides R code (R Core Team, 2012) for the applications.

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 $\boldsymbol{\tau}, \boldsymbol{\lambda}$, or $\boldsymbol{\psi}$ is obtained as

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

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

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

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

$$\mathbf{s} := \frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \sum_{i=1}^{N} \frac{\partial \ell_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \mathbf{S}(\boldsymbol{\theta})'\mathbf{n}. \tag{A.3}$$

Define the observed and expected information matrices as

$$\mathbf{I}_Y := -\frac{\partial \mathbf{s}}{\partial \boldsymbol{\theta}'} = -\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}, \tag{A.4}$$

$$\mathbf{I}_L := E_L(\mathbf{I}_Y) = \sum_{r=1}^{R} \hat{n}_r \mathbf{S}_r(\boldsymbol{\theta})' \mathbf{S}_r(\boldsymbol{\theta}), \tag{A.5}$$

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

$$\mathbf{D} := \sum_{r=1}^{R} n_r \mathbf{S}_r(\boldsymbol{\theta})' \mathbf{S}_r(\boldsymbol{\theta}). \tag{A.6}$$

The form of the Jacobian in equation A.1 can be used to determine identifiability.

Theorem 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 A.1 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 A.1, (ii) and (iii) guarantee that the columns of $\mathbf{S}_{\text{new}}$ are also independent of those in $\mathbf{S}_{\boldsymbol{\theta}_2}$. □

The theorem suggests that when the local independence model is identifiable and the number of local dependencies ψ 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.* Wiley-Interscience, New York.

Aitchison, J. and Silvey, S. (1958). Maximum-likelihood estimation of parameters subject to restraints. *The Annals of Mathematical Statistics*, 29(3):813–828.

Albert, P. and 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.

Albert, P. and Dodd, L. (2008). On estimating diagnostic accuracy from studies with multiple raters and partial gold standard evaluation. *Journal of the American Statistical Association*, 103(481):61–73.

Bandeen-Roche, K., Miglioretti, D., Zeger, S., and Rathouz, P. (1997). Latent variable regression for multiple discrete outcomes. *Journal of the American Statistical Association*, 92(440):1375–1386.

Bartolucci, F. and 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.

Boos, D. (1992). On generalized score tests. *The American Statistician*, 46(4):327–333.

Breusch, T. and Pagan, A. (1980). The lagrange multiplier test and its applications to model specification in econometrics. *The Review of Economic Studies*, 47(1):239–253.

Catchpole, E. and Morgan, B. (1997). Detecting parameter redundancy. *Biometrika*, 84(1):187–196.

Dayton, C. and Macready, G. (1988). Concomitant-variable latent-class models. *Journal of the American Statistical Association*, 83(401):173–178.

Dempster, A., Laird, N., and Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 1–38.

Espeland, M. and Handelman, S. (1989). Using latent class models to characterize and assess relative error in discrete measurements. *Biometrics*, pages 587–599.

Evers, M. and Namboodiri, N. (1979). On the design matrix strategy in the analysis of categorical data. *Sociological methodology*, 10:86–111.

Faraone, S. and Tsuang, M. (1994). Measuring diagnostic accuracy in. *Am J Psychiatry*, 1(51):651.

Fergusson, D., Horwood, L., and Lynskey, M. (1994). The comorbidities of adolescent problem behaviors: A latent class model. *Journal of abnormal child psychology*, 22(3):339–354.

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. and Kohlmann, T. (1996). Latent class analysis in medical research. *Statistical methods in medical research*, 5(2):179–211.

Garrett, E., Eaton, W., and 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. and Zeger, S. (2004). 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., and 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.

Harper, D. (1972). Local dependence latent structure models. *Psychometrika*, 37(1):53–59.

Hill, J. and 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. and 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. and 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.

McHugh, R. (1956). Efficient estimation and local identification in latent class analysis. *Psychometrika*, 21(4):331–347.

Pepe, M. and Janes, H. (2007). Insights into latent class analysis of diagnostic test performance. *Biostatistics*, 8(2):474–484.

Qu, Y., Tan, M., and Kutner, M. (1996). Random effects models in latent class analysis for evaluating accuracy of diagnostic tests. *Biometrics*, pages 797–810.

R Core Team (2012). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Rao, C. R. (1948). Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. In *Proceedings of the Cambridge Philosophical Society*, volume 44, pages 50–57. Cambridge Univ Press.

Reboussin, B., Ip, E., and 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., and Sörbom, D. (1987). The detection and correction of specification errors in structural equation models. *Sociological Methodology*, 17:105–129.

Sörbom, D. (1989). Model modification. *Psychometrika*, 54(3):371–384.

Torrance-Rynard, V. and Walter, S. (1998). Effects of dependent errors in the assessment of diagnostic test performance. *Statistics in Medicine*, 16(19):2157–2175.

Uebersax, J. (1993). Statistical modeling of expert ratings on medical treatment appropriateness. *Journal of the American Statistical Association*, 88(422):421–427.

Uebersax, J. and Grove, W. (1993). A latent trait finite mixture model for the analysis of rating agreement. *Biometrics*, pages 823–835.

Vacek, P. (1985). The effect of conditional dependence on the evaluation of diagnostic tests. *Biometrics*, pages 959–968.

van der Linden, W. and Glas, C. (2010). Statistical tests of conditional independence between responses and/or response times on test items. *Psychometrika*, 75(1).

Vermunt, J. K. and Magidson, J. (2005). Technical guide for latent GOLD 4.0: Basic and advanced. *Belmont Massachusetts: Statistical Innovations Inc.*

Walter, S. and 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*, pages 1–25.