Working Paper
Series

FINRISK
National Centre of Competence in Research
Financial Valuation and Risk Management

National Centre of Competence in Research
Financial Valuation and Risk Management

Working Paper No. 268

A latent factor model for ordinal data to Measure multivariate predictive ability of financial market movements

Philippe Huber Oliver Scaillet

Maria-Pia Victoria-Freser

First version: October 2006
Current version: October 2006

This research has been carried out within the NCCR FINRISK project on "Financial Econometrics for Risk Management"

A latent factor model for ordinal data to measure multivariate predictive ability of financial market movements

Philippe Huber[1], Olivier Scaillet[2] and Maria-Pia Victoria-Feser[3]

HEC - University of Geneva, Blv. Pont d'Arve 40,

CH-1211 Geneva, Switzerland

October 2005

[1]Partially supported by Swiss National Science Foundation, grant # 610-057883.99.

[2]Partially supported by the Swiss NCCR Finrisk, and member of International Center FAME.

[3]Partially supported by Swiss National Science Foundation, grants # 610-057883.99 and # PP001-106465 and by the Swiss NCCR Finrisk.

Abstract

In this paper we develop a structural equation model with latent variables in an ordinal setting which allows us to test broker-dealer predictive ability of financial market movements. We use a multivariate logit model in a latent factor framework, develop a tractable estimator based on a Laplace approximation, and show its consistency and asymptotic normality. Monte Carlo experiments reveal that both the estimation method and the testing procedure perform well in small samples. An empirical illustration is given for mid-term forecasts simultaneously made by two broker-dealers for several countries.

Key words: structural equation model, latent variable, generalised linear model, factor analysis, multinomial logit, forecasts, LAMLE, canonical correlation.

JEL Classification: C12, C13, C30, C51, C52, C53, G10.

MSC 2000: 62H12, 62H15, 62H20, 62H2, 62J12, 62H15, 62P20.

1 Introduction

Institutional investors make a large portion of overall trading volume in equity markets, and much of this trading activity is directed to brokerage houses who execute trades. In exchange for directed trades most of the brokerage houses provide so-called "soft dollars". Soft dollar arrangements are arrangements under which products or services other than execution of securities transactions are obtained by an institutional investor from or through a broker-dealer in exchange for the placement of his orders (see Blume 1993, Johnsen 1994 and Securities and Commission 1998 for detailed definition, history and law related to soft dollars). Typically under a soft dollar arrangement a brokerage firm rebate to an institutional client part of the non-execution-related commission paid, either by providing research to the institution, paying for third party research, or by buying research-related item, such as computer hardware and software, or magazine subscriptions. These arrangements are best thought as ways of subsidizing the research inputs that investors use to identify profitable trading opportunities. US regular surveys about the size of the soft dollars industry are conducted by Greenwich Associates. For example, their 2003 survey of 237 financial institutions indicates that soft dollar commissions totaled almost USD 1,005 million in 2003 up from USD 645 million in 2001. This represents about

USD 1 out of every USD 7 paid in commissions by those firms involved. Obviously soft dollars are costly, and it should be interesting for an institutional investor to determine from a statistical point of view whether these soft dollar inputs are worth being used (and indirectly paid for) or not. In this paper we aim at providing such a quantitative tool in the spirit of a structural equation modelling (SEM) with latent variables (see e.g. Aigner, Hsiao, Kapteyn, and Wansbek 1984 for an introduction). The model will take the form of a multivariate multinomial logit (MNL) with latent factors (see e.g. McFadden 1984 for an introduction).

The data at hand have been provided by the pension fund of the University of Geneva, and are historical data of financial forecasts from 2 broker-dealers about the mid-term evolution of the stock market in 5 countries and the bond market in 4 zones, respectively. These broker-dealers were asked each trimester during 6 years to provide their forecasts for each country in terms of market trends for the next 6 months. Each forecast is precisely defined as a given range for future variations of the different stock and bond indices, and the ranges correspond to strong bear, bear, neutral, bull, strong bull trends for the next 6 months. For our purpose they have been recorded on an ordinal scale from 1 to 5. In order to decide whether the forecasts are valid they should be confronted to the actual evolutions of the corresponding markets. The issue is to determine whether the forecasts made by the broker-dealers are in some sense "near" the realized market evolutions six months later. Recall that we are in a *multivariate* context since the forecasts concern different countries at the same time. Formally, in this paper, we aim at measuring (and testing for) the association between two random vectors, say $\mathbf{X}$ (the forecasts) and $\mathbf{Y}$ (the market realizations), whose size p is the same (4 or 5 countries), and whose entries consist of ordinal variables corresponding to the forecast and realized market states (values in $\{1, ..., 5\}$) for each country, respectively.

The paper is organized as follows. In Section 2 we review a well-known measure of association between two normal random vectors, namely the canonical correlation coefficient, and show that this coefficient has a natural interpretation in terms of SEM. This interpretation will prompt us to extend the SEM to the case of ordinal variables in Section 3, and use the corresponding association coefficient as a measure of predictive ability. To this end we develop a generalized linear latent variable model (GLLVM) with a logit link function in an ordinal setting. Estimation and asymptotic properties are investigated in Section 4. We rely on the so-called Laplace approximation (De Bruijn 1981) to get a tractable and fast estimation procedure of the latent variable model, and show the consistency and asymptotic normality of the resulting estimators. Section 5 is devoted to Monte Carlo experiments aimed at gauging the performance in small samples of the estimation method and the testing procedure. In particular we look at the probability coverage of confidence intervals based on a parametric bootstrap method. We gather the empirical results in Section 6, while technical details and proofs are relegated to an appendix.

2 Canonical correlation coefficient and SEM

In the context of continuous random vectors a well known measure of association under the normality assumption is the canonical correlation coefficient (see e.g. Mardia, Kent, and Bibby 1979). For a moment suppose that $(\mathbf{X}', \mathbf{Y}')'$ is distributed as a multivariate normal random variable with mean $(\boldsymbol{\mu}_X', \boldsymbol{\mu}_Y')'$ and covariance matrix

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{XX} & \boldsymbol{\Sigma}_{XY} \\ \boldsymbol{\Sigma}_{XY}' & \boldsymbol{\Sigma}_{YY} \end{pmatrix}.$$

The canonical correlation coefficient is then defined by

$$\rho_c = \frac{\mathbf{b}_X^{*\prime}\boldsymbol{\Sigma}_{XY}\mathbf{b}_Y^*}{\sqrt{\mathbf{b}_X^{*\prime}\boldsymbol{\Sigma}_{XX}\mathbf{b}_X^*\mathbf{b}_Y^{*\prime}\boldsymbol{\Sigma}_{YY}\mathbf{b}_Y^*}}, \tag{1}$$

where $\mathbf{b}_X^*$ and $\mathbf{b}_Y^*$ are the solutions to the maximization problem:

$$\max_{\mathbf{b}_X, \mathbf{b}_Y} \frac{\mathbf{b}_X' \boldsymbol{\Sigma}_{XY} \mathbf{b}_Y}{\sqrt{\mathbf{b}_X' \boldsymbol{\Sigma}_{XX} \mathbf{b}_X \mathbf{b}_Y' \boldsymbol{\Sigma}_{YY} \mathbf{b}_Y}}. \tag{2}$$

The canonical correlation coefficient ρ_c is actually the maximal correlation coefficient between any linear combinations of $\mathbf{X}$ and $\mathbf{Y}$.

Consider now the following structural equation model (SEM) for $\mathbf{X} = (X_1, ..., X_p)'$ and $\mathbf{Y} = (Y_1, ..., Y_p)'$:

$$\begin{aligned} X_j &= \alpha_{X_j} + \beta_{X_j} F_X + \epsilon_{X_j}, & j = 1, \ldots, p, \\ Y_j &= \alpha_{Y_j} + \beta_{Y_j} F_Y + \epsilon_{Y_j}, & j = 1, \ldots, p, \end{aligned} \tag{3}$$

with

$$\mathbf{F} = (F_X, F_Y)' \sim N(\mathbf{0}, \mathbf{R}), \qquad \mathbf{R} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \tag{4}$$

$$\boldsymbol{\epsilon} = \left(\epsilon_{X_1}, \ldots, \epsilon_{Y_p}\right)' \sim N(\mathbf{0}, \boldsymbol{\psi}), \qquad \boldsymbol{\psi} = \mathrm{diag}(\psi_{X_1}^2, \ldots, \psi_{Y_p}^2). \tag{5}$$

The model (3) can be interpreted as follows: given that the covariance structure in $\mathbf{X}$ (respectively $\mathbf{Y}$) is explained by the latent factor F_X (respectively F_Y) as in classical factor analysis, the relationship (i.e. correlation) between $\mathbf{X}$ and $\mathbf{Y}$ is summarized by the correlation parameter ρ. We may therefore ask the next question: what is the link between the latter and the standard canonical correlation coefficient ρ_c? The following proposition shows that ρ can be rewritten in an analogous form to (1), namely a correlation between linear combinations of $\mathbf{X}$ and $\mathbf{Y}$ but with a modified covariance matrix $\mathbf{\Sigma}^* = \mathbf{\Sigma} + \boldsymbol{\psi}$.

Proposition 1 *For the SEM (3) with (4) and (5), the correlation coefficient between the latent variables ρ is given*

$$\rho = \frac{\boldsymbol{\beta}'_X \mathbf{\Sigma}^*_{XY} \boldsymbol{\beta}_X}{\sqrt{\boldsymbol{\beta}'_X \mathbf{\Sigma}^*_{XX} \boldsymbol{\beta}_X \boldsymbol{\beta}'_Y \mathbf{\Sigma}^*_{YY} \boldsymbol{\beta}_Y}} \tag{6}$$

with $\boldsymbol{\beta}_X = \left(\beta_{X_1}, \ldots, \beta_{X_p}\right)'$, $\boldsymbol{\beta}_Y = \left(\beta_{Y_1}, \ldots, \beta_{Y_p}\right)'$ *and* $\mathbf{\Sigma}^* = \mathbf{\Sigma} + \boldsymbol{\psi}$.

Equation (6) can be interpreted as the correlation between two linear combinations of $\mathbf{X}$ and $\mathbf{Y}$ when the covariance structure is accounted for measurement error in the manifest variables via $\boldsymbol{\psi}$. The correlation coefficient ρ is thus different from the canonical correlation coefficient ρ_c by construction. The advantage of the former over the latter is that it can be easily generalized to the case of non normal variables. We suggest hereafter to extend the SEM (3) to the case of ordinal variables in order to derive an association measure valid in an ordinal setting . To our knowledge, such an association measure between multivariate variables outside the normal model has not been proposed yet. This association measure will constitute in our example an indicator of predictive ability of the broker-dealers.

Let us finally remark that the independence between the error terms in (5) is motivated by the fact that the latent factor structure $\mathbf{F}$ is assumed to describe entirely the correlation structure between the manifest variables $\mathbf{X}$ and $\mathbf{Y}$, or in other terms that the manifest variables are conditionally independent given the latent variables. In principle, we can always define a model with enough latent variables such that $\boldsymbol{\psi}$ remains diagonal as in classical factor analysis. Incorporating additional factors in (3) does not raise any difficulties (see below). Of course in that case, we loose the intuitive interpretation unveiled in Proposition 1 since the factor structure implies that the measure of association between $\mathbf{X}$ and $\mathbf{Y}$ has to be multidimensional instead of the scalar ρ. The measure of association cannot be reduced to a single parameter for higher factor dimension. Although this situation is certainly of practical interest, it will not occur in our empirical illustration.

3 SEM for predictive ability assessment

In this section we first generalize the SEM (3) to the case of ordinal variables with a multifactor structure before discussing its use in assessing predictive ability.

Let us introduce the $m_X \times 1$ random vector $\mathbf{F}_X$ and $m_Y \times 1$ random vector $\mathbf{F}_Y$ to build the factor structure $\mathbf{F} = (1, \mathbf{F}'_X, \mathbf{F}'_Y)'$ with $m_Y, m_X < p$. Using

$\mathbf{Z} = (\mathbf{X}', \mathbf{Y}')'$ the SEM (3) under (4) and (5) can be translated in a compact form as:

$$Z^{(l)}|\mathbf{F} \overset{\text{ind.}}{\sim} N\left(\boldsymbol{\lambda}_l'\mathbf{F}, \phi_l^2\right), \tag{7}$$

with $\boldsymbol{\lambda}_l = (\alpha_{X_l}, \boldsymbol{\beta}_{X_l}', \mathbf{0}')'$ and $\phi_l = \psi_{X_l}$, for $l = 1, \ldots, p$, while $\boldsymbol{\lambda}_l = (\alpha_{Y_{l-p}}, \mathbf{0}', \boldsymbol{\beta}_{Y_{l-p}}')'$ and $\phi_l = \psi_{Y_{l-p}}$, for $l = p+1, \ldots, 2p$. This implies $E\left[Z^{(l)}|\mathbf{F}\right] = \boldsymbol{\lambda}_l'\mathbf{F}$ and $\text{Var}\left[Z^{(l)}|\mathbf{F}\right] = \phi_l^2$. Note that the $\boldsymbol{\lambda}_l$ are called loadings.

We can generalize (7) to the family of exponential distributions with probability distribution function:

$$g_l(z^{(l)}|\mathbf{F}) = \exp\left\{\frac{u\left(\boldsymbol{\lambda}_l'\mathbf{F}\right) z^{(l)} - b\left(u\left(\boldsymbol{\lambda}_l'\mathbf{F}\right)\right)}{\phi_l} + c(z^{(l)}, \phi_l)\right\} \tag{8}$$

where $u\left(\boldsymbol{\lambda}_l'\mathbf{F}\right)$ is the so-called canonical parameter, $b\left(u\left(\boldsymbol{\lambda}_l'\mathbf{F}\right)\right)$ and $c(z^{(l)}, \phi_l)$ are specific functions whose form depends on the particular exponential distribution, and ϕ_l is a scale parameter (see McCullagh and Nelder 1989). Except for the normal case, the expectation $E\left[Z^{(l)}\middle|\,\mathbf{F}\right]$ is not a linear function of $\mathbf{F}$, but is linked to the linear predictor through a link function ν as

$$\nu\left(E\left[Z^{(l)}|\mathbf{F}\right]\right) = \boldsymbol{\lambda}_l'\mathbf{F}.$$

We further have that $u\left(\boldsymbol{\lambda}_l'\mathbf{F}\right) = \boldsymbol{\lambda}_l'\mathbf{F}$ when we choose the so-called canonical link function for ν. This model actually belongs to the class of Generalized Linear Latent Variables Model (GLLVM) which has been proposed by Moustaki (1996) and Moustaki and Knott (2000) under an assumption of independence between the Gaussian latent variables (diagonal $\mathbf{R}$). This type of modelling can be viewed as an extension of the usual Generalized Linear Models approach (McCullagh and Nelder 1989) to the latent factor framework.

The conditional independence of the manifest variables $Z^{(l)}$ given the latent ones is again assumed, so that the conditional joint density of the manifest variables is $\prod_{l=1}^{2p} g_l(z^{(l)}|\mathbf{F})$ and their marginal joint distribution is

$$f(\mathbf{z}) = \int \left[\prod_{l=1}^{2p} g_l(z^{(l)}|\mathbf{v}) \right] h(\mathbf{v}) d\mathbf{v} \tag{9}$$

with $h(\mathbf{v})$ being the $N(\mathbf{0}, \mathbf{R})$ probability distribution function.

Because of the nature of the data at hand, we need to develop hereafter the case of ordinal variables, i.e. ordered categorical variables. Let $Z^{(l)}|\mathbf{F}$ follow a multinomial distribution with possible values (or categories) going from 1 to q_l. In the following we opt for a cumulative logit formulation (see Agresti 1990 for the advantages of this formulation over other ones, and Jöreskog and Moustaki 2001 for a comparison of different approaches in the framework of factor analysis with ordinal data) to account for the ordered nature of the categorical data. Let $P_{ls} = P\left[Z^{(l)} \leq s|\mathbf{F}\right]$, $s = 1, \ldots, q_l$, be the conditional cumulative distribution functions. The quantity $\log\left(P_{ls}/(1-P_{ls})\right)$ is the log-odds of falling into or below a category s versus falling above it for the manifest variable l. It is used in the logit link between the linear predictor and the conditional cumulative probability distribution:

$$\nu(P_{ls}) = \log\left(\frac{P_{ls}}{1-P_{ls}}\right) = \boldsymbol{\lambda}'_{ls}\mathbf{F}, \tag{10}$$

where $\boldsymbol{\lambda}_{ls} = (\alpha_{X_{ls}}, \boldsymbol{\beta}'_{X_l}, \mathbf{0}')'$ or $\boldsymbol{\lambda}_{ls} = (\alpha_{Y_{l-p,s}}, \mathbf{0}', \boldsymbol{\beta}'_{Y_{l-p}})'$ depending on the value of l. The subscript s in the α's indicates that each intercept depends not only on the manifest variable l but also on the category s. The constraint for each manifest variable that the p slope coefficients (the beta's) in $\boldsymbol{\lambda}_{ls}$ does not depend on the category s is known as the proportional-odds assumption, and essentially allows us to reduce the number of parameters to be estimated. The intercepts take the interpretation of thresholds and are monotonic in the sense that the lowest category receives the lowest threshold, and so on. They represent the log-odds of falling into or below category s when all latent variables are nil, while a given positive slope leads to an increase on the log-odds of falling into or below any category associated with a one unit increase in the corresponding latent variable. A positive slope indicates thus an increase in the odds themselves, and higher probabilities for the manifest variable to take low values. For identification purposes the highest threshold is set equal to infinity by convention, which means that we only need to estimate $q_l - 1$ threshold for the manifest variable l. With all these restrictions, the model is fully identifiable. In general, the thresholds can be assumed to be different for each manifest (ordinal) variable. However, when the ordinal variables are measured through questionnaires using the same measurement unit (for example percentages), we can constrain the thresholds to be equal for all manifest variables, i.e.

$$\boldsymbol{\lambda}_{ls} = \boldsymbol{\lambda}_s \quad \text{for all } l.$$

This is a suitable constraint for the analysis of our data (see section 6).

The scale parameter ϕ_l is here equal to 1, while the canonical parameter is not linear in the latent factors (since we do not use the canonical link function), but equal to

$$u\left(\boldsymbol{\lambda}_{ls}'\mathbf{F}\right) = \log\left(\frac{P_{ls}}{P_{l,s+1} - P_{ls}}\right),$$

while

$$b\left(u\left(\boldsymbol{\lambda}_{ls}'\mathbf{F}\right)\right) = \log\left(1+\exp\left(u\left(\boldsymbol{\lambda}_{ls}'\mathbf{F}\right)\right)\right) = \log\left(\frac{P_{l,s+1}}{P_{l,s+1}-P_{ls}}\right)$$

and $c(z^{(l)}, \phi_l) = 0$. The conditional distribution of the manifest variable is given by

$$\begin{aligned} g(z^{(l)}|\mathbf{F}) &= \prod_{s=1}^{q_l}(P_{ls}-P_{l,s-1})^{\iota(z^{(l)}=s)}, \\ &= \prod_{s=1}^{q_l-1}\left(\frac{P_{ls}}{P_{l,s+1}}\right)^{\iota(z^{(l)}\leq s)}\left(\frac{P_{l,s+1}-P_{ls}}{P_{l,s+1}}\right)^{\iota(z^{(l)}\leq s+1)-\iota(z^{(l)}\leq s)} \\ &= \exp\left(\sum_{s=1}^{q_l-1}\left[\iota(z^{(l)}\leq s)u\left(\boldsymbol{\lambda}_{ls}'\mathbf{F}\right)-\iota(z^{(l)}\leq s+1)b\left(u\left(\boldsymbol{\lambda}_{ls}'\mathbf{F}\right)\right)\right]\right) \end{aligned} \quad (11)$$

where $\iota(z^{(l)} = s) = 1$ if $z^{(l)} = s$ and 0 otherwise, and $\iota(z^{(l)} \leq s) = 1$ if $z^{(l)} \leq s$ and $\iota(z^{(l)} \leq s) = 0$ otherwise.

Note that instead of the logit function we could use a probit link. In practice however, the difference is very small since these two link functions are very close ($|\Phi(x)-\Psi(1.7x)| < 0.01, \forall x$, where Ψ is the logistic distribution function and Φ the normal cumulative distribution function, see e.g. Lord and Novick (1968). In the regression model (i.e. $\mathbf{F}$ is observed, and $\mathbf{Z}$ reduces to a univariate ordinal response variable), McCullagh and Nelder (1989) use the same approach to link the explanatory variable to the first moments of the response variable. Finally let us remark that a specification in terms of latent variables is a usual way to reduce the complexity of multinomial model calculation (see McFadden 1984, p. 1419), and achieve a relative parsimony in the modelling. This is even more relevant, if not inevitable, in a multivariate framework.

4 Estimation and asymptotic properties

Let $\mathbf{z} = [\mathbf{z}_1, \ldots, \mathbf{z}_n]'$, with $\mathbf{z}_i = [z_i^{(1)}, \ldots, z_i^{(2p)}]$, n the sample size and $2p$ the number of manifest variables. As the marginal distribution of the observed variable must be integrated out from the conditional distributions $g(z^{(l)}|\mathbf{F})$ given by (9), we use a Laplace approximation (see De Bruijn 1981) to approximate the likelihood function of the sample as it has been done in Huber, Ronchetti, and Victoria-Feser (2004) for other types of variables.

The Laplace approximation to integrals goes back to the original work of Laplace. This technique is widely used in mathematics; see e.g. De Bruijn (1981). In statistics, it has been used successfully to approximate posterior distributions in Bayesian statistics (see e.g. Tierney and Kadane 1986) and in relation to saddlepoint approximations (Field and Ronchetti 1990).

Let $h : \mathbb{R}^m \rightarrow \mathbb{R}$ be a function which satisfies the following conditions: it is continuous and has a global maximum in $\hat{\mathbf{x}}$, its first and second derivatives exist in a neighborhood of $\hat{\mathbf{x}}$ and$\partial h(\hat{\mathbf{x}})/\partial \mathbf{x} = 0$ and $\mathbf{H}(\hat{\mathbf{x}}) = \partial^2 h(\hat{\mathbf{x}})/\partial x \partial x$, the Hessian matrix, is such that $-\mathbf{H}(\hat{\mathbf{x}})$ is positive definite. Moreover, $h(\mathbf{x})$ is sharply peaked in the neighborhood of $\hat{\mathbf{x}}$, i.e. two positive scalars b and c exist such that

$$h(\hat{\mathbf{x}}) \leq -b, \quad \text{if } |\hat{\mathbf{x}} - \mathbf{x}| \geq c.$$

Then,

$$\int e^{th(\mathbf{x})} \mathbf{dx} = (2\pi)^{m/2} \det(-\mathbf{H}(\hat{\mathbf{x}}))^{-1/2} t^{-1/2} \exp\left(th(\hat{\mathbf{x}})\right)\left(1 + O(t^{-1})\right), \quad t \to \infty. \tag{12}$$

Equation (12) is obtained by an expansion of $h(x)$ about its maximum $\hat{\mathbf{x}}$:

$$\begin{aligned}
\int e^{th(\mathbf{x})} \mathbf{dx} &\approx \int \exp\left(th(\hat{\mathbf{x}}) + t\frac{\partial}{\partial \mathbf{x}} h(\hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}}) + \frac{1}{2} t(\mathbf{x} - \hat{\mathbf{x}})\mathbf{H}(\hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})'\right) \mathbf{dx} \\
&= \exp\left(th(\hat{\mathbf{x}})\right) \int \exp\left(\frac{1}{2} t(\mathbf{x} - \hat{\mathbf{x}})\mathbf{H}(\hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})'\right) \mathbf{dx} \\
&= (2\pi)^{m/2} \det\left(-\mathbf{H}(\hat{\mathbf{x}})\right)^{-1/2} t^{-1/2} \exp\left(th(\hat{\mathbf{x}})\right).
\end{aligned}$$

Let $\boldsymbol{\lambda}$ denote the vector of all loadings and thresholds and $\mathbf{R}$ the correlation matrix of the latent factors, the approximated log-likelihood $\tilde{l}$ for a model with ordered multinomial distributed manifest variables is (see Appendix A.1)

$$
\begin{aligned}
\tilde{l}(\boldsymbol{\lambda}, \mathbf{R}|\mathbf{z}) \;=\; & \sum_{i=1}^{N} \Bigg(-\frac{1}{2}\log\,\det\Big((\boldsymbol{\Gamma}(\boldsymbol{\lambda}, \mathbf{R}, \widehat{\mathbf{F}}_i)\Big) - \frac{1}{2}\log\det(\mathbf{R}) \\
& + \sum_{l=1}^{2p}\sum_{s=1}^{q_l-1}\Big[\iota(z_i^{(l)} \le s)u(\boldsymbol{\lambda}_{ls}'\hat{\mathbf{F}}_i) \\
& -\iota(z_i^{(l)} \le s+1)\log(1+\exp u(\boldsymbol{\lambda}_{ls}'\hat{\mathbf{F}}_i))\Big] \\
& -\frac{\widehat{\mathbf{F}}_{i(2)}'\mathbf{R}^{-1}\widehat{\mathbf{F}}_{i(2)}}{2}\Bigg), \qquad (13)
\end{aligned}
$$

where $\boldsymbol{\Gamma}(\boldsymbol{\lambda}, \mathbf{R}, \widehat{\mathbf{F}}_i)$ is a correction matrix that comes from the Laplace approximation, $\hat{\mathbf{F}}_{i(2)} = [\hat{\mathbf{F}}_{iX}', \hat{\mathbf{F}}_{iY}']'$ and $\hat{\mathbf{F}}_i = [1, \hat{\mathbf{F}}_{i(2)}]$ is the estimator of the latent score for the i^{th} observation which is given by the implicit equation

$$
\begin{aligned}
\hat{\mathbf{F}}_{i(2)} \;:\; & = \hat{\mathbf{F}}_{i(2)}(\boldsymbol{\lambda}, \mathbf{R}, \mathbf{z}_i) = \sum_{l=1}^{2p}\sum_{s=1}^{q_l-1}\Big(\iota(z_i^{(l)} \le s)P_{l,s+1}(\boldsymbol{\lambda}_{ls}'\hat{\mathbf{F}}_i) \\
& -\iota(z_i^{(l)} \le s+1)P_{ls}(\boldsymbol{\lambda}_{ls}'\hat{\mathbf{F}}_i)\Big)\,\mathbf{R}\boldsymbol{\lambda}_{l(2)}, \qquad (14)
\end{aligned}
$$

where $\boldsymbol{\lambda}_{l(2)}$ is $\boldsymbol{\lambda}_{ls}$ without its first element.

Note that Huber, Ronchetti, and Victoria-Feser (2004) pointed out that $\hat{\mathbf{F}}_{i(2)}$ can be seen as the MLE of $\mathbf{F}_{i(2)}$. The Laplace Approximated MLE (LAMLE) of the models parameters are obtained from the optimization of $\tilde{l}$, whose derivatives can be computed explicitly, but are omitted here for the sake of space. Hereafter, we establish the consistency and asymptotic normality of the LAMLE $\widehat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta} = (\boldsymbol{\lambda}', \text{vech}(\mathbf{R})')'$ where vech($\mathbf{A}$) is the stack of the elements on and below the diagonal of $\mathbf{A}$.

Proposition 2 (Consistency) *Let $\boldsymbol{\theta} \in \Theta$. If Θ is compact,*

$$\widehat{\boldsymbol{\theta}} \xrightarrow{p} \boldsymbol{\theta}_0.$$

Note that the empirical approximated likelihood is here too complex to be shown to be concave in $\boldsymbol{\theta}$. Under concavity of the objective function, compactness can be replaced by the assumption of $\boldsymbol{\theta}_0$ being an element of the interior of a convex set Θ (see e.g. Theorem 2.7 of Newey and McFadden 1994).

Proposition 3 (asymptotic normality) *If Θ is compact, $\boldsymbol{\theta}_0 \in$ interior(Θ), $J_0 = E[\partial^2 \widetilde{l}(\boldsymbol{\theta}_0)/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}']$ is nonsingular,*

$$\sqrt{T}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow{d} N(0, \mathbf{J}_0^{-1}\mathbf{I}_0\mathbf{J}_0^{-1}),$$

with $\mathbf{I}_0 = E[\partial\widetilde{l}(\boldsymbol{\theta}_0)/\partial\boldsymbol{\theta}\partial\widetilde{l}(\boldsymbol{\theta}_0)/\partial\boldsymbol{\theta}']$.

We note that instead of using a Laplace approximation to approximate the integrals, we could use a Gauss-Hermite Quadrature procedure (Bock and Liberman (1970)), an adaptive Gauss-Hermite Quadrature procedure (as implemented in GLAMM in the STATA package), or Monte-Carlo methods. Huber, Ronchetti, and Victoria-Feser (2004) argue however that a Laplace approximation of the likelihood function is a better approach: the LAMLE are asymptotically unbiased and fast to compute. It should be stressed that alternative estimators have been proposed in the framework of Generalized Mixed Linear Models, such as McGilchrist (1994) best linear unbiased prediction (BLUP) based on the h-likelihood of Lee and Nelder (1996), or Green (1987) penalized quasi-likelihood (PQL) (see also Breslow and Clayton 1993). Huber, Ronchetti, and Victoria-Feser (2004) show that these estimators are all equal but different than the LAMLE. Finally, we could also in principle use a two-step approach as implemented in LISREL. In that approach, polychoric correlations (Muthén 1984, Poon and Lee 1987) between the manifest variables are first estimated and then used as sufficient statistics in the normal model (3) with (4) and (5). Huber, Ronchetti, and Victoria-Feser (2004) show that in the case of mixtures between normal and binary manifest variables, this procedure leads to biased estimators and incorrect inference.

The results of Proposition 3 could in principle be used for inference when the sample sizes are large. When this is not that case, it is more suitable to use other techniques. For the correlation estimator $\hat{\rho}$, we propose to use the transformation function η introduced by Fisher (1915) that stabilizes the variance of the estimator:

$$\eta(\rho) = \tanh^{-1}(\rho) = \frac{1}{2}\log\left(\frac{1+\rho}{1-\rho}\right),$$

and η is approximately normal

$$\eta(\rho) \sim \mathcal{N}\left(\nu_\rho, \frac{1}{n-3}\right),$$

with $\nu_\rho = \tanh^{-1}(\rho) + \frac{\rho}{2(n-1)}$. A discussion about the Fisher transformation can be found in Efron (1982). In practice, we compute the variance of $\hat{\eta}$ which is simply $(n-3)^{-1}$, calculate its confidence interval, and transform it back to a confidence interval for $\hat{\rho}$.

For the other parameter estimators, we use a two-step approach based on a parametric bootstrap: first, we calculate the estimators $\hat{\boldsymbol{\theta}}$ from the observed sample and then, we generate 1000 new samples under the estimated distributions to get new estimators $\hat{\boldsymbol{\theta}}^*_k$, where $k = 1, \ldots, 1000$. We find the biases and endpoints of the confidence intervals for $\hat{\boldsymbol{\theta}}$ using a bias-corrected acceleration (BC_a) technique as described in Efron (1987), Efron and Tibshirani (1993), and Shao and Tu (1995).

5 Monte Carlo experiments

In order to evaluate the performance of our model and our estimator in finite samples, we have performed a simulation study[1]. We consider the model (8) with $p = 10$, equal thresholds and parameter values given in Tables 1 to 3. With these two sets of parameters, we also choose different values for the correlation coefficient, namely $\rho = -0.5, 0$, and 0.5. The first set of parameters ($S1$) was chosen to match one of the real examples analyzed in Section 6, and the other ($S2$) to reflect what should be sensible in practice for other cases, i.e., a conservative attitude implying large probabilities associated to small or no changes. For each set of parameters, we simulated 500 samples of size $n = 30$ and computed the LAMLE of $\boldsymbol{\lambda}$ and the predictive criterion ρ. The distribution of the sample bias estimates is presented for each estimator in the form of boxplots in Figures 1 to 3 for a correlation $\rho = 0.5$ (for the other values of ρ, we found similar results). Figure 4 shows the boxplots for the estimated correlation $\hat{\rho}$ under the parameter set $S1$ (for the other set, the results are similar). We can see that even for a relatively small sample size (given the size of the model), the performance is very good in that there is no apparent bias for all parameters, including ρ.

We have also studied the small sample performance of the probability

[1] The code is available from the authors upon request.

coverage of 95% confidence intervals for $\hat{\rho}$ computed with the Fisher transformation, and found a probability coverage of 84.9%.

6 Data Analysis

The database contains the forecasts (in terms of trends) of two broker-dealers A and B about the mid-term (6 months) evolution of the stock market in five different countries (Switzerland, Germany, France, Great Britain and USA) for A and the bond market in four zones (Switzerland, Euro Zone, Great Britain and USA) for B. The trends are precisely defined as corresponding to a given future variation x with: x <-10%, -10%< x <-5%, -5% < x < 5%, 5% < x < 10%, 10% < x, for stock markets, and x <-0.25%, -0.25%< x <-0.10%, -0.10% < x < 0.10%, 0.10% < x < 0.25%, 0.25% < x, for bond markets. In both cases, we compare the forecasts to the actual returns of the corresponding markets six months later. The sample starts in July 1997 and finishes in April 2003 with one forecast every quarter (22 observations).

The estimated loadings for both broker-dealers are given in Tables 4 and 5 with biases and 95% confidence intervals, all computed via a parametric bootstrap. The estimated correlation between both latent variables for broker-dealers A and B is given in Table 6. The scores of the latent variables for both broker-dealers are displayed in Figures 5 and 6.

The study of the correlation estimates (see Table 6) indicates that the broker-dealer forecasts match the actual market evolution. The correlation for both broker-dealers are significantly positive. We can therefore conclude that the forecasts are accurate to some extent. Alternatively, we can look at the latent scores $\widehat{F}_{X_i}$ and $\widehat{F}_{Y_i}$ and see graphically how they evolve. For broker-dealer A, they are given in Figure 5 and for broker-dealer B in Figure 6. For both broker-dealers, the evolution of the two lines (predicted and actual) is pretty similar, thus reflecting the fact that the predictions on the five stock markets and on the four bond markets are in phase with the actual evolution of the indices. This reflects again the ability of the dealer-brokers to predict market evolution relatively accurately.

Tables 4 and 5 present the estimated loadings for dealer-brokers A and B, respectively. They give another type of information about the behavior of the broker-dealers. Indeed, the correlation reflects the ability of the dealer-broker to predict the changes in markets directions, but not necessarily the size of the changes. The latter can be inferred from the loadings because they act as a multiplicative factor of the latent variables. In other words, the latent variables give the directions of the market moves, while the loadings give the (average) sizes of these moves. In Tables 4 and 5 we can see that the loadings for the actual markets are systematically and significantly higher than the corresponding loadings related to the forecasts. This difference is certainly due to the fact that the forecasts are in general too conservative: although the direction of the movements are correctly predicted, their size is underestimated in all markets by the broker-dealers.

A Appendix

A.1 Development of the LAMLE for ordered multinomial distributed manifest variables

Let us write the marginal distribution of $\mathbf{x}_i$ as

$$f_{\boldsymbol{\lambda}}(\mathbf{z}_i) = \int e^{p \cdot Q(\boldsymbol{\lambda}, \mathbf{R}, \mathbf{F}, \mathbf{z}_i)} \mathbf{dF}, \tag{15}$$

where

$$\begin{aligned} Q(\boldsymbol{\lambda}, \mathbf{R}, \mathbf{F}, \mathbf{z}_i) &= \frac{1}{p} \Bigg[\sum_{l=1}^{p} \sum_{s=1}^{m_l - 1} \Big[\iota(z_i^{(l)} \leq s) u(\boldsymbol{\lambda}'_{ls} \mathbf{F}) \\ &\quad - \iota(z_i^{(l)} \leq s+1) \log(1 + \exp(u(\boldsymbol{\lambda}'_{ls} \mathbf{F}))) \Big] \\ &\quad - \frac{\mathbf{F}'_{(2)} \mathbf{R}^{-1} \mathbf{F}_{(2)}}{2} - \frac{q}{2} \log(2\pi) \Bigg]. \end{aligned} \tag{16}$$

with $\mathbf{F}_{(2)} = [\mathbf{F}'_X, \mathbf{F}'_Y]'$ and $\mathbf{F} = [1, \mathbf{F}_{(2)}]$. We use the q-dimensional Laplace approximation to eliminate the integral from the density (15) (see De Bruijn 1981 or Tierney and Kadane 1986, pp. 82-86) and obtain

$$f_{\boldsymbol{\lambda}}(\mathbf{z}_i) = \left(\frac{2\pi}{p}\right)^{q/2} \left(\det(-\mathbf{U}(\hat{\mathbf{F}}_i))\right)^{-1/2} \exp\left(pQ(\boldsymbol{\lambda}, \mathbf{R}, \hat{\mathbf{F}}_i, \mathbf{z}_i)\right) (1 + O(p^{-1})), \tag{17}$$

where

$$\mathbf{U}(\hat{\mathbf{F}}_i) = \frac{\partial^2 Q(\boldsymbol{\lambda}, \mathbf{R}, \mathbf{F}_i, \mathbf{z}_i)}{\partial \mathbf{F} \partial \mathbf{F}'}\bigg|_{\mathbf{F}=\hat{\mathbf{F}}_i} = -\frac{1}{p}\boldsymbol{\Gamma}(\boldsymbol{\lambda}, \mathbf{R}, \hat{\mathbf{F}}_i), \tag{18}$$

$$\begin{aligned}\boldsymbol{\Gamma}(\boldsymbol{\lambda}, \mathbf{R}, \hat{\mathbf{F}}_i) &= \sum_{l=1}^{p}\sum_{s=1}^{q_l-1}\left[\iota(z_i^{(l)} \leq s+1)\eta(\boldsymbol{\lambda}_{ls}'\hat{\mathbf{F}}_i) - \iota(z_i^{(l)} \leq s)\eta(\boldsymbol{\lambda}_{ls+1}'\hat{\mathbf{F}}_i))\right]\boldsymbol{\lambda}_{l(2)}\boldsymbol{\lambda}_{l(2)}' \\ &\quad +\mathbf{R}^{-1},\end{aligned} \tag{19}$$

where $\boldsymbol{\lambda}_{l(2)}$ is $\boldsymbol{\lambda}_{ls}$ without its first element, and

$$\eta(\boldsymbol{\lambda}_{ls}'\hat{\mathbf{F}}_i) = \frac{\partial P_{ls}(\boldsymbol{\lambda}_{ls}'\hat{\mathbf{F}}_i)}{\partial(\boldsymbol{\lambda}_{ls}'\hat{\mathbf{F}}_i)} = \frac{1}{1+\exp(\boldsymbol{\lambda}_{ls}'\hat{\mathbf{F}}_i)}P_{ls}(\boldsymbol{\lambda}_{ls}'\hat{\mathbf{F}}_i).$$

Note that $\hat{\mathbf{F}}_i$, the MLE of $\mathbf{F}_i$ as mentioned is Section 4, is the optimum of $Q(\boldsymbol{\lambda}, \mathbf{R}, \mathbf{F}, \mathbf{z}_i)$, i.e. the root of $\partial Q(\boldsymbol{\lambda}, \mathbf{R}, \mathbf{F}, \mathbf{z}_i)/\partial \mathbf{F} = 0$, and is defined through the iterative equation (14).

Putting (16), (18) and (19) into the distribution function (17) for the whole sample gives the approximated log-likelihood (13).

A.2 Proof of Proposition 1

We can write (3) as

$$\mathbf{Z} = \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} = \boldsymbol{\alpha} + \boldsymbol{\Lambda}\mathbf{F} + \boldsymbol{\epsilon}, \tag{20}$$

where $\boldsymbol{\alpha} = \left(\alpha_{X_1}, \ldots, \alpha_{Y_p}\right)'$, and where $\boldsymbol{\Lambda}$ is a $(2p) \times 2$ matrix with

$$\boldsymbol{\Lambda} = \begin{pmatrix} \boldsymbol{\beta}_X & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\beta}_Y \end{pmatrix},$$

so that

$$\mathbf{F} = \left(\boldsymbol{\Lambda}'\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{\Lambda}' \left(\mathbf{Z} - \boldsymbol{\alpha} - \boldsymbol{\epsilon}\right).$$

This yields

$$\text{Var}(\mathbf{F}) = \left(\boldsymbol{\Lambda}'\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{\Lambda}' \left(\boldsymbol{\Sigma} + \boldsymbol{\psi}\right) \boldsymbol{\Lambda} \left(\boldsymbol{\Lambda}'\boldsymbol{\Lambda}\right)^{-1} = \mathbf{R},$$

which gives Equation (6) after an identification term by term.

A.3 Proof of Proposition 2

a) Under our distributional assumption for the latent variables, $\boldsymbol{\theta}_0 = (\boldsymbol{\lambda}_0^{'}, \text{vech}(\mathbf{R}_0)^{'})'$ is identified. b) Recall that $b\left(u\left(\boldsymbol{\lambda}_{ls}'\mathbf{F}\right)\right) = \log\left(1 + \exp\left(u\left(\boldsymbol{\lambda}_{ls}'\mathbf{F}\right)\right)\right)$. Hence $|b\left(u\left(\boldsymbol{\lambda}_{ls}'\mathbf{F}\right)\right)| \leq \log 2$ for $\boldsymbol{\lambda}_{ls}'\mathbf{F} \leq 0$, and $|b\left(u\left(\boldsymbol{\lambda}_{ls}'\mathbf{F}\right)\right)| \leq \|\boldsymbol{\lambda}_{ls}\|\|\mathbf{F}\|$ for $\boldsymbol{\lambda}_{ls}'\mathbf{F} > 0$ (using $\log(1+v) < \log(v)$ for $v > 0$). Besides $|\log\det\boldsymbol{\Gamma}(\boldsymbol{\lambda}, \mathbf{R}, \mathbf{F})| \leq C$ and $|\iota(z^{(l)} \leq s)\boldsymbol{\lambda}_{ls}'\mathbf{F}| \leq |\iota(z^{(l)} \leq s)|\|\boldsymbol{\lambda}_{ls}\|\|\mathbf{F}\|$. Using the definition of $\mathbf{F}$, we deduce that $E[\|\mathbf{F}\|] < \infty$, and thus $E[|\widetilde{l}(\boldsymbol{\theta})|] < \infty$.

Combining a) and b) establishes that $E[\widetilde{l}(\boldsymbol{\theta})]$ has a unique maximum at $\boldsymbol{\theta}_0$

(see e.g. Lemma 2.2 of Newey and McFadden 1994). Since the data are i.i.d., Θ is compact, $\widetilde{l}(\boldsymbol{\theta})$ is continuous at each $\boldsymbol{\theta}$ with probability one, and there is a function $d(\mathbf{F})$ with $|\widetilde{l}(\boldsymbol{\theta})| \leq d(\mathbf{F})$ such that $E[d(\mathbf{F})] < \infty$ (cf. b)), we deduce that $E[\widetilde{l}(\boldsymbol{\theta})]$ is continuous and that the empirical approximated likelihood converges uniformly in probability to that quantity (see e.g. Lemma 2.4 of Newey and McFadden 1994). Therefore the conditions of Theorem 2.1 of Newey and McFadden (1994)) are fulfilled, and we get the stated result.

A.4 Proof of Proposition 3

By computing the explicit expression of $\partial^2 \widetilde{l}(\boldsymbol{\theta})/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'$, we can show that we can find a function $d(\mathbf{F})$ with $|\partial^2 \widetilde{l}(\boldsymbol{\theta}_0)/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'| \leq d(\mathbf{F})$ such that $E[d(\mathbf{F})] < \infty$ (as in b) of the previous proof). Hence, since the data are i.i.d., Θ is compact, and $\partial^2 \widetilde{l}(\boldsymbol{\theta}_0)/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'$ is continuous at each $\boldsymbol{\theta}$ with probability one, we deduce that $E[\partial^2 \widetilde{l}(\boldsymbol{\theta}_0)/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}']$ is continuous and that the Hessian of the empirical approximated likelihood converges uniformly in probability to that quantity (see e.g. Lemma 2.4 of Newey and McFadden 1994). The asymptotic normality results then from the Lindberg-Levy Central Limit Theorem and Theorem 3.1 of Newey and McFadden (1994).

References

Agresti, A. (1990). *Categorical Data Analysis*. New York: Wiley.

Aigner, D. J., C. Hsiao, A. Kapteyn, and T. Wansbek (1984). Latent variables models in econometrics. In Z. Griliches and M. Intriligator (Eds.), *Handbook of Econometrics*, Volume II. Amsterdam: North-Holland.

Blume, M. (1993). Soft dollars and the brokerage industry. *Financial Analysts Journal 49*, 36–44.

Bock, R. D. and M. Liberman (1970). Fitting a response model for n dichotomously scored items. *Psychometrika 35*, 179–197.

Breslow, N. E. and D. G. Clayton (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association 88*, 9–25.

De Bruijn, N. G. (1981). *Asymptotic Methods in Analysis* (third ed.). New York: Dover Publications.

Efron, B. (1982). Transformation theory: How normal is a family of distributions? (Ed. Corr: V10 p1032). *The Annals of Statistics 10*, 323–339.

Efron, B. (1987). Better bootstrap confidence intervals (C/R: P186-200). *Journal of the American Statistical Association 82*, 171–185.

Efron, B. and R. Tibshirani (1993). *An Introduction to the Bootstrap.*

Chapman and Hall.

Field, C. A. and E. Ronchetti (1990). *Small Sample Asymptotics.* Lectures Notes-Monograph Series. Haywood (CA): Institute of Mathematical Statistics.

Fisher, R. A. (1915). Frequency distribution of the values of the correlation coeeficient in samples from an indefinitely large population. *Biometrika 10*, 507–521.

Green, P. J. (1987). Penalized likelihood for general semi-parametric regression models. *International Statistical Review 55*, 245–259.

Huber, P., E. Ronchetti, and M.-P. Victoria-Feser (2004). Estimation of generalized latent trait models. *Journal of the Royal Statistical Society, Series B 66*, 893–908.

Johnsen, B. (1994). Property rights to investment research: The agency costs of soft dollar brokerage. *The Yale Journal of Regulations 11*, 57–73.

Jöreskog, K. G. and I. Moustaki (2001). Factor analysis for ordinal variables: a comparison of three approaches. *Multivariate Behavioural Research 36*, 347–387.

Lee, Y. and J. A. Nelder (1996). Hierarchical generalized linear models. *Journal of the Royal Statistical Society, Series B 58*, 619–678.

Lord, F. M. and M. E. Novick (1968). *Statistical Theories of Mental Test Scores*. Addison-Wesley.

Mardia, K. V., J. T. Kent, and J. M. Bibby (1979). *Multivariate Analysis*. London: Academic Press.

McCullagh, P. and J. A. Nelder (1989). *Generalized Linear Models*. London: Chapman and Hall. Second edition.

McFadden, D. (1984). Econometric analysis of qualitative response models. In Z. Griliches and M. Intriligator (Eds.), *Handbook of Econometrics*, Volume II. Amsterdam: North-Holland.

McGilchrist, C. A. (1994). Estimation in generalized mixed models. *Journal of the Royal Statistical Society, Series B 56*(1), 61–69.

Moustaki, I. (1996). A latent trait and a latent class model for mixed observed variables. *British Journal of Mathematical and Statistical Psychology 49*, 313–334.

Moustaki, I. and M. Knott (2000). Generalized latent trait models. *Psychometrika 65*, 391–411.

Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical and continuous latent variables indicators. *Psychometrika 49*, 115–132.

Newey, W. and D. McFadden (1994). Large sample estimation and hypothesis testing. In R. F. Engle and D. McFadden (Eds.), *Handbook of Econometrics*, Volume IV. Amsterdam: North-Holland.

Poon, W.-Y. and S.-Y. Lee (1987). Maximum likelihood estimation of multivariate polyserial and polychoric correlation coefficients (corr: V53 p301). *Psychometrika 52*, 409–430.

Securities and E. Commission (1998). Inspection report on the soft dollar practices of broker-dealers. *Investment Advisers and Mutual Funds*.

Shao, J. and D. Tu (1995). *The Jackknife and Bootstrap.* Springer-Verlag.

Tierney, L. and J. B. Kadane (1986). Accurate approximations for posterior moments and marginal densities. *Journal of the American Statistical Association 81*, 82–86.

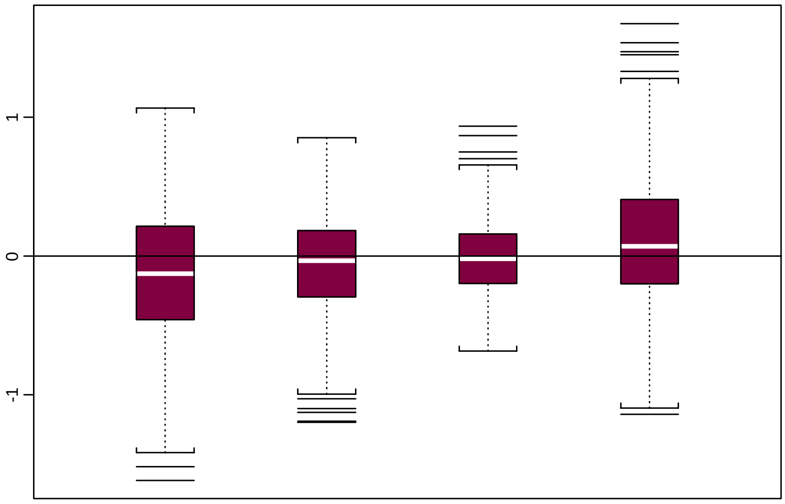

Figure 1: Distributions of threshold estimates for simulation $S1$ with $\rho = 0.5$.

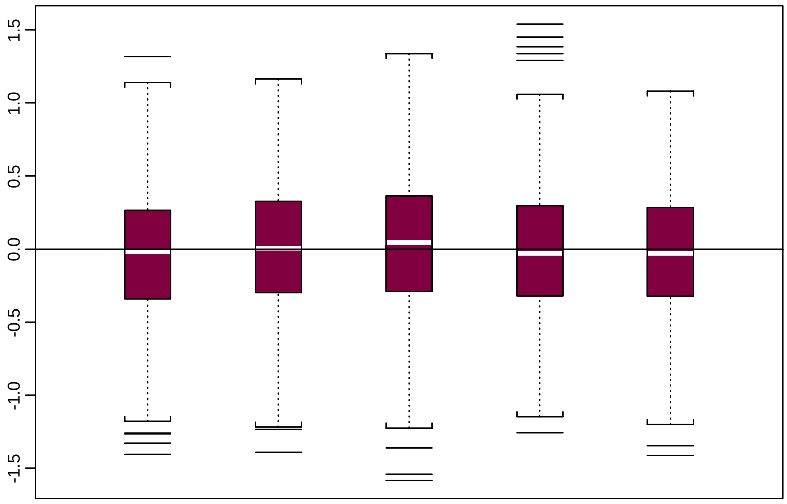

Figure 2: Distributions of loading estimates (first latent variable) for simulation $S1$ with $\rho = 0.5$.

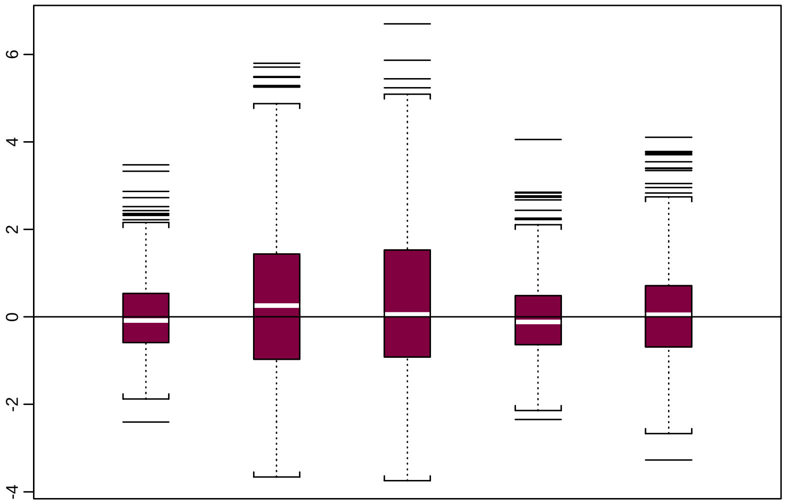

Figure 3: Distributions of loading estimates (second latent variable) for simulation $S1$ with $\rho = 0.5$.

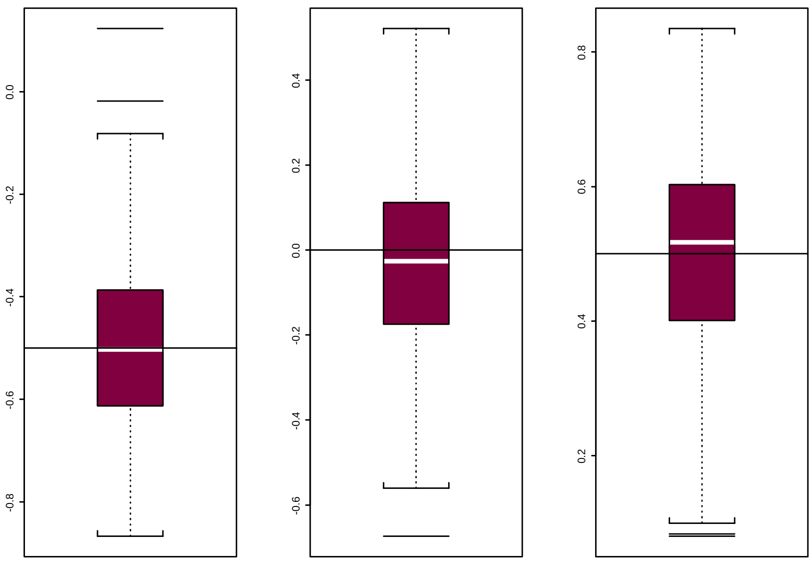

Figure 4: Distributions of correlation estimates for simulation $S1$ with $\rho = 0.5$.

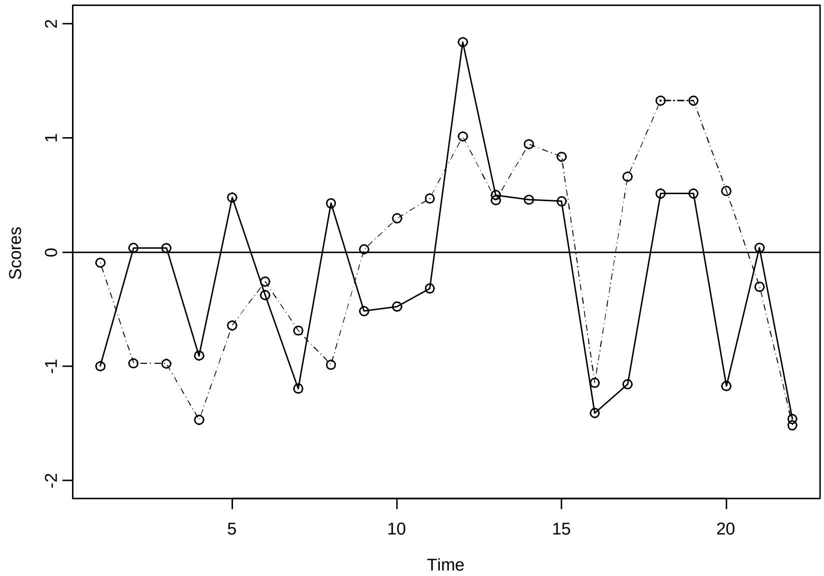

Figure 5: Estimated scores for broker-dealer A. The plain line is the forecast and the dotted line the actual level for the stock market.

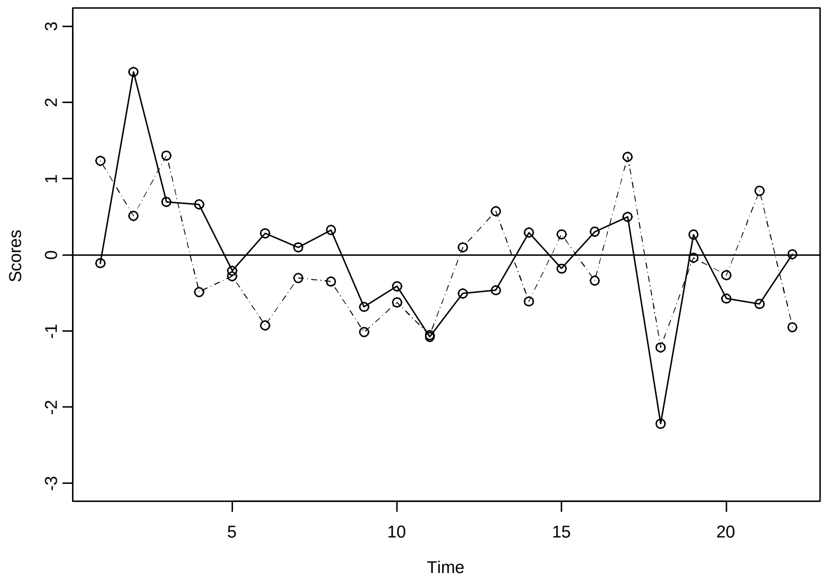

Figure 6: Estimated scores for broker-dealer B. The plain line is the forecast and the dotted line the actual level for the bond market.

Thresholds	Cumulative probabilities
-4.60	0.01
-2.94	0.05
0.85	0.70
4.60	0.99

Table 1: Thresholds for simulation $S1$.

Thresholds	Cumulative probabilities
-2.19	0.10
-1.39	0.20
1.39	0.80
2.19	0.90

Table 2: Thresholds for simulation $S2$.

Latent 1	Latent 2
1.60	0.00
1.75	0.00
1.70	0.00
1.30	0.00
1.50	0.00
0.00	5.00
0.00	9.00
0.00	9.00
0.00	5.00
0.00	6.00

Table 3: Loadings for simulation $S1$ and $S2$.

	Predicted				Observed			
Market	Estimator	Bias	$l_{0.95}$	$u_{0.95}$	Estimator	Bias	$l_{0.95}$	$u_{0.95}$
CH	1.596	-0.078	0.441	3.071	5.557	-0.280	3.716	10.069
D	1.762	-0.063	0.456	3.197	8.925	-1.091	5.399	17.079
F	1.725	-0.066	0.411	3.407	8.982	-1.102	5.493	17.709
UK	1.286	-0.035	0.207	2.724	4.980	-0.162	3.165	8.386
USA	1.506	-0.103	0.302	3.023	6.222	-0.442	4.178	11.628

Table 4: Estimated loadings for broker-dealer A. The biases, lower ($l_{0.95}$) and upper ($u_{0.95}$) confidence bounds were computed with a parametric bootstrap.

	Predicted				Observed			
Market	Estimator	Bias	$l_{0.95}$	$u_{0.95}$	Estimator	Bias	$l_{0.95}$	$u_{0.95}$
CH	0.632	0.038	-0.985	1.824	2.971	-0.082	1.836	5.608
EU	0.886	-0.029	-0.912	2.200	5.869	-1.049	2.554	10.014
UK	0.544	0.029	-0.930	1.732	5.419	-0.897	2.662	9.234
USA	1.219	-0.105	-0.921	2.665	7.180	-1.647	3.327	14.059

Table 5: Estimated loadings for broker-dealer B. The biases, lower ($l_{0.95}$) and upper ($u_{0.95}$) confidence bounds were computed with a parametric bootstrap.

Broker-dealer A				Broker-dealer B			
Estimator	Bias	$l_{0.95}$	$u_{0.95}$	Estimator	Bias	$l_{0.95}$	$u_{0.95}$
0.398	-0.034	0.013	0.722	0.320	0.069	0.027	0.691

Table 6: Estimated correlations between both latent variables.