

Multi-market direction-of-change modeling using dependence ratios

Stanislav Anatolyev*
New Economic School

Studies in Nonlinear Dynamics & Econometrics, 2008

Abstract

We consider a multivariate dynamic model for the joint distribution of binary outcomes associated with directions-of-change for several markets or assets. The marginal distribution of each binary outcome follows a dynamic binary choice model, while the association structure is parameterized via possibly time varying dependence ratios. We illustrate the technique using daily stock index returns from three European markets, from three Baltic markets, and from two Chinese exchanges.

Key words: direction-of-change, binary choice, dependence ratio, stock returns.

*Address: Stanislav Anatolyev, New Economic School, Nakhimovsky Pr., 47, Moscow, 117418 Russia. E-mail: sanatoly@nes.ru. I would like to thank the Editor, an Associate Editor and two anonymous referees for very useful comments and suggestions.

1 Introduction

It has been long recognized that directional prediction of financial time series is as important as prediction of its levels, and helps investors form trading strategies, allocate funds efficiently and extract profits (e.g., Henriksson and Merton, 1981; Leitch and Tanner, 1991; Pesaran and Timmermann, 1995). There has been quite a bit of research which develops tools for testing for sign predictability and evaluation of directional forecasts (Henriksson and Merton, 1981; Breen, Glosten and Jagannathan, 1989; Pesaran and Timmermann 1992; Anatolyev and Gerko 2005; Christoffersen and Diebold, 2006; Chung and Hong, 2007). Applications may be found in Hartzmark (1991), Greer (2003) and Bekiros and Georgoutsos (2008), among many others. The question of how structural instability affects directional forecast evaluation is studied in Pesaran and Timmermann (2004). Lesser developed are tools for modeling directions-of-change, particularly in a multivariate context, although one can encounter univariate directional analyses in Rydberg and Shephard (2003), Christoffersen and Diebold (2006), Startz (2008), and Anatolyev and Gospodinov (2009).

Of course, one may use classical binary response analysis augmented by dynamic effects (e.g., Dueker, 2005; Kauppi and Saikkonen, 2008; Startz, 2008). Such models are widely used to primarily predict economy's expansions and recessions (e.g., Birchenhall, Jessen, Osborn, and Simpson, 1999; Dueker, 2005; Chauvet and Potter, 2005; Startz, 2008). While it is straightforward to carry these ideas over to modeling the directions-of-change in financial markets, it is often interesting, however, to analyze several markets simultaneously. The extension of univariate dynamic binary response models to this multivariate case can go along different routes. We take the route where the probability of an upward movement is modeled as an explicit parametric (such as logistic) function of observable history (de Jong and Woutersen, 2007; Kauppi and Saikkonen, 2008; Startz, 2008). Alternative routes such as the one with latent variables following their own dynamics (as in, e.g., Dueker, 2005) lead to computationally much more intensive estimation methods. Our main focus is on the issues of dependence across markets.

One way to model several directions-of-change is to describe the joint evolution of underlying returns based on a multivariate density (e.g., Engle, 2002; Bauwens and Laurent, 2005) possibly using copulas (e.g., Patton, 2006), and then convert it into the joint process followed by corresponding binary outcomes. This "indirect" approach should be classified as unnecessarily restrictive, because parameterizations of the joint distributions and its dynamics lead to numerous restrictions which may not be supported empirically with

confidence. Needless to say, such approach leads to a complicated likelihood function with a high-dimensional parameter vector. Also, it seems redundant to first describe the whole evolution of the joint distribution of several returns, and then reduce it to the evolution of its simple function. When modeling the directions-of-change directly, it is much easier to avoid unnecessary restrictions and describe fully the joint dynamics of direction indicators.

In this paper we propose the dynamic model for directions-of-change of several assets or markets which originates from a static multivariate binary choice model of Ekholm, Smith, and McDonald (1995) containing the description of binary marginals and so called dependence ratios, which are the ratios of joint success probabilities to products of marginal success probabilities for joint binary outcomes of all possible combinations of individual binary outcomes. The dependence ratios have a clear interpretation and a number of advantages over other association measures such as conditional odds ratios (Fitzmaurice and Laird, 1993), marginal log odds ratios (Glonck and McCullagh, 1995), a generalization of the two (Glonck, 1996), or copula-type representation (Tajar, Denuit, and Lambert, 2001). Certain interesting hypotheses and association models regarding the dependence structure are easily testable.

In our dynamic version of the multivariate direction-of-change model, the marginal Bernoulli distributions are parameterized as functions of past directions for all markets involved in the analysis, while the dependence ratios may depend on past directions for the markets related to these dependence ratios. To illustrate the technique, we apply the model to three sets of daily stock market index returns: from three leading European markets, from three emerging markets of Baltic states, and from the two largest Chinese stock exchanges. The European and Chinese data show little dynamics but high dependence, while the Baltic data show a lot of dynamics but little dependence among the individual markets. We also demonstrate that our model is able to produce out-of-sample conditional directional forecasts of high quality which dominate forecasts produced by the “indirect” approach as well as those based on univariate models.

The paper is organized as follows. In section 2 we present a parameterization of the joint distribution of direction indicators using dependence ratios. In section 3 we develop a dynamic model and describe its properties. Section 4 contains applications to some developed and emerging markets. Section 5 concludes.

2 Direction indicators and dependence ratios

Suppose we have m markets with observable return x_i from market i , $i = 1, \dots, m$. Denote the direction indicator for market i by

$$I_i = \mathbb{I}\{x_i > c\}, \quad (1)$$

where $\mathbb{I}\{\cdot\}$ is the indicator function. Usually, we are interested in the case $c = 0$, but the constant c may in general be different from zero, especially when the markets are expected on average to grow or fall, and c may well be different across markets.

Denote by y the cell encoded as an $1 \times m$ vector of ones and zeros corresponding to “up” (success) and “down” (failure) directions-of-change, respectively, i.e. $y = (y_1, y_2, \dots, y_m) \in \{0, 1\}^m$. Let π^y be a cell probability corresponding to y , i.e. the joint probability mass of the vector $I = (I_1, I_2, \dots, I_m)'$. Denote $\mathbb{N}_1^y = \{i : y_i = I_i\}$ and $\mathbb{N}_2^y = \{i : y_i = 1 - I_i\}$, and also by m_1^y and m_2^y the cardinalities of \mathbb{N}_1^y and \mathbb{N}_2^y . Define the cell indicator for the cell y as

$$I^y = \mathbb{I}\{I = y'\} = \prod_{i=1}^m \mathbb{I}\{I_i = y_i\}.$$

Because $\mathbb{I}\{I_i = 1\} = I_i$ and $\mathbb{I}\{I_i = 0\} = 1 - I_i$, it is straightforward to see that

$$\begin{aligned} I^y &= \left(\prod_{k \in \mathbb{N}_1^y} I_k \right) \left(\prod_{i \in \mathbb{N}_2^y} (1 - I_i) \right) \\ &= \left(\prod_{k \in \mathbb{N}_1^y} I_k \right) \left(\sum_{j_1=0}^1 \dots \sum_{j_{m_2}=0}^1 \left((-1)^{\sum_{i=1}^{m_2} j_i} \right) \prod_{i \in \mathbb{N}_2^y} I_i^{j_i} \right). \end{aligned}$$

For example, in the case $m = 3$ and $y = (1, 0, 1)$ we have $\mathbb{N}_1^{(101)} = \{1, 3\}$, $\mathbb{N}_2^{(101)} = \{2\}$, $m_1^{(101)} = 2$, $m_2^{(101)} = 1$, $\pi^{(101)} = \Pr\{x_1 > c, x_2 \leq c, x_3 > c\}$, and $I^{(101)} = I_1(1 - I_2)I_3$.

Let us denote by ϕ_i the marginal (success) probability corresponding to market i , i.e.

$$\phi_i = \Pr\{x_i > c\} = E(I_i).$$

Denote by $\phi_{i_1 i_2 \dots i_\ell}$ the joint (success) probability corresponding to different markets i_1, i_2, \dots, i_ℓ , where $1 < \ell \leq m$ and $1 \leq i_j \leq m$ for all $j = 1, \dots, \ell$, i.e.

$$\phi_{i_1 i_2 \dots i_\ell} = \Pr\{x_{i_1} > c, x_{i_2} > c, \dots, x_{i_\ell} > c\}.$$

We factorize each joint probability $\phi_{i_1 i_2 \dots i_\ell}$ into the product of corresponding marginal probabilities $\phi_{i_1}, \phi_{i_2}, \dots, \phi_{i_\ell}$ and the *dependence ratio* $\gamma_{i_1 i_2 \dots i_\ell}$ for markets i_1, i_2, \dots, i_ℓ :

$$\phi_{i_1 i_2 \dots i_\ell} = \phi_{i_1} \phi_{i_2} \dots \phi_{i_\ell} \gamma_{i_1 i_2 \dots i_\ell}.$$

When the markets are independent, all dependence ratios equal unity. A dependence ratio greater than unity is indicative of positive association between the involved markets. The deviation of a dependence ratio from unity measures how far the corresponding joint success probability differs from that under independence. For example, $\gamma_{12} = 2$ means that the probability of $I_1 = I_2 = 1$ is twice that under independence; $\gamma_{123} = 1.5$ means that the probability of $I_1 = I_2 = I_3 = 1$ is 50% larger than that under independence; etc.

Ekhholm, Smith, and McDonald (1995) show that there is explicit, albeit nonlinear, mapping from the collection of ϕ_i for all i and $\phi_{i_1 i_2 \dots i_\ell}$ and $\gamma_{i_1 i_2 \dots i_\ell}$ for all combinations i_1, i_2, \dots, i_ℓ , to the collection of π^y for all $y \in \{0, 1\}^m$. It is this collection of π^y that enters the likelihood function. The mapping can be written as follows:

$$\begin{aligned} \pi^y &= E(I^y) \\ &= \sum_{j_1=0}^1 \dots \sum_{j_{m_2}=0}^1 \left((-1)^{\sum_{i=1}^{m_2} j_i} \right) E \left[\prod_{k \in \mathbb{N}_1^y} \prod_{i \in \mathbb{N}_2^y} I_k I_i^{j_i} \right]. \end{aligned} \quad (2)$$

The expectation of the product at the end of this formula has the form

$$\phi_{i_1} \phi_{i_2} \dots \phi_{i_\ell} \gamma_{i_1 i_2 \dots i_\ell}$$

for an appropriate combination i_1, i_2, \dots, i_ℓ . For example, when $m = 3$, there are three dependence ratios of order 2, $\gamma_{12}, \gamma_{13}, \gamma_{23}$, and one dependence ratio of order 3, γ_{123} ; the mapping looks as follows:

$$\begin{aligned} \pi^{(111)} &= \phi_1 \phi_2 \phi_3 \gamma_{123} \\ \pi^{(110)} &= \phi_1 \phi_2 (\gamma_{12} - \phi_3 \gamma_{123}) \\ \pi^{(101)} &= \phi_1 \phi_3 (\gamma_{13} - \phi_2 \gamma_{123}) \\ \pi^{(011)} &= \phi_2 \phi_3 (\gamma_{23} - \phi_1 \gamma_{123}) \\ \pi^{(100)} &= \phi_1 (1 - \phi_2 \gamma_{12} - \phi_3 \gamma_{13} + \phi_2 \phi_3 \gamma_{123}) \\ \pi^{(010)} &= \phi_2 (1 - \phi_1 \gamma_{12} - \phi_3 \gamma_{23} + \phi_1 \phi_3 \gamma_{123}) \\ \pi^{(001)} &= \phi_3 (1 - \phi_1 \gamma_{13} - \phi_2 \gamma_{23} + \phi_1 \phi_2 \gamma_{123}) \\ \pi^{(000)} &= 1 - \phi_1 - \phi_2 - \phi_3 + \phi_1 \phi_2 \gamma_{12} + \phi_1 \phi_3 \gamma_{13} + \phi_2 \phi_3 \gamma_{23} - \phi_1 \phi_2 \phi_3 \gamma_{123} \end{aligned}$$

Now, let previously considered variables additionally indexed by t correspond to observation t ($t = 1, 2, \dots, T$), containing the vector of returns

$x_t = (x_{1,t}, x_{2,t}, \dots, x_{m,t})'$ and associated indicator vector $I_t = (I_{1,t}, I_{2,t}, \dots, I_{m,t})'$. Note that Ekholm, Smith, and McDonald (1995) and Ekholm, McDonald and Smith (2000) parameterize the marginal probabilities as some functions of exogenous variables characterizing corresponding IID units. In our context, “units” are time periods having a different nature, so we keep the marginal probabilities constant for now and postpone parameterizations to the next section. The total loglikelihood is then written as

$$\ell(x_1, x_2, \dots, x_T; \varphi) = \sum_{t=1}^T \left(\sum_{y \in \{0,1\}^m} I_t^y \log \pi^y \right) \quad (3)$$

and is to be maximized with respect to $2^m - 1$ parameters in the collection

$$\varphi = (\phi_1, \phi_2, \dots, \phi_m, \gamma_{12}, \dots, \gamma_{m-1,m}, \gamma_{123}, \dots, \gamma_{m-2,m-1,m}, \dots, \gamma_{12\dots m}).$$

The maximum likelihood theory suggests that under suitable conditions, the maximum likelihood estimator $\hat{\varphi}$ of φ is consistent, asymptotically normal and asymptotically efficient, with an information matrix consistently estimated in a straightforward way (see, e.g., Gouriéroux and Monfort, 1981). One notable condition in this context is that no marginal success probability is exactly zero or one (i.e. some markets must grow faster and some slower than at rate c), which will hold if the constant c in (1) is chosen judiciously.

Note that if maximization of (3) is performed with respect to the collection of π^y , $y \in \{0, 1\}^m$ subject to the constraint $\sum_{y \in \{0,1\}^m} \pi^y = 1$, the resulting ML estimates of π^y are equal to the empirical probabilities for corresponding cells. The ML estimates of marginal success probabilities and dependence ratios can be recovered using the mapping (2), which is, however, very hard to do even for small m (not to mention that to compute standard errors is even harder) due to a highly nonlinear nature of (2). Therefore, it is preferable to maximize (3) with respect to the parameters in φ in the first place.

As pointed out in Ekholm, Smith, and McDonald (1995) and Ekholm, McDonald and Smith (2000), using the dependence ratios has several advantages over using other association structures such as conditional odds ratios frequently used in statistical processing of binary data (for example, in Fitzmaurice and Laird, 1993).

First, the interpretation of dependence ratios does not depend on how many markets are in analysis, while that of conditional odds ratios does. This means that adding a new market to a set of markets under analysis does not change the old dependence ratios and their interpretation. For example, γ_{12} is unrelated to the value of m (provided that $m \geq 2$), while the conditional

odds ratio for markets 1 and 2 evidently depends on m in the conditional set ($y_3 = 0, \dots, y_m = 0$).

Second, when $m > 2$, no explicit mapping from the collection of marginals and conditional odds ratios to the collection of π^y exists, while it does exist in case of dependence ratios. This greatly facilitates maximum likelihood methods.

Third, the method described permits handling missing data in a straightforward way, while this is not the case when the association parameters are conditional. In analyzing several financial markets it sometimes happens that some markets are closed (because of, for example, national holidays) when others are open, which results in missing data. If the “missing at random” assumption can be accepted, for the observation t having missing $y_{i_1^*}, \dots, y_{i_\ell^*}$ we replace in the loglikelihood $\sum_y I_t^y \log \pi^y$ by $\sum_y I_t^y \log \sum_{i_1^*, \dots, i_\ell^*} \pi^{y|y_{i_1^*}^* \dots i_\ell^*}$, where $\pi^{y|y_{i_1^*}^* \dots i_\ell^*}$ signifies the cell probability corresponding to y given the particular combination $y_{i_1^*}^* \dots i_\ell^*$.

There are nonetheless two minor disadvantages in using dependence ratios to using conditional odds ratios. First, the dependence ratio is not invariant to the exchange of 1 and 0, i.e. to what we call the success and what we call the failure. If we switch codes 1 and 0 for a couple of markets, the corresponding conditional log odds ratio will only change its sign, while the dependence ratio will take on a new value according to a nonlinear mapping that also involves the marginals. However, if the direction of the market does not possess symmetry (indeed, the perception of bull and bear markets is different), this may be considered as an advantage rather than a drawback. Second, the maximum likelihood estimates of marginal distribution parameters and conditional odds ratios have a property of asymptotic block-diagonality of the asymptotic variance matrix, a property not shared by dependence ratios. However, this block diagonality, being a pleasant property, is not critical at all.

Beside conditional odds ratios, there exist yet other measures of association between binary variables which also are less advantageous than the dependence ratios. The marginal log odds ratios (Glonek and McCullagh, 1995), which are contrasts of log odds ratios in various joint distributions, are more interpretable than conditional odds ratios, but also lack the existence of an explicit mapping to the collection of π^y . The copula-type representation for bivariate binary data proposed by Tajar, Denuit and Lambert (2001) seems unnecessarily complex and not generalizable to more than two binary variables in a straightforward way.

It is also important to realize that the parameterization of the above model

is complete in the sense that no assumptions are made beyond the description of the full joint distribution of all m binary outcomes. An alternative, “indirect” approach where one fully describes the joint distribution of underlying returns and subsequently converts it into the joint distribution of corresponding binary outcomes, should be classified as unnecessarily restrictive, because parameterizations of the marginal distributions and the dependence structure of returns themselves necessarily leads to numerous restrictions, not to mention excessive complicatedness of this approach.

Ekhholm, Smith, and McDonald (1995) suggest a number of hypotheses about certain configurations of the dependence ratios. These hypotheses impose restrictions on dependence ratios $\gamma_{i_1 i_2 \dots i_\ell}$ and thus on joint probabilities $\phi_{i_1 i_2 \dots i_\ell}$ without restricting marginal probabilities ϕ_1, \dots, ϕ_m . The most restrictive hypothesis is that of independence across markets which assumes that all dependence ratios are equal to unity. Another one is a hypothesis of horizontal homogeneity which assumes equal dependence within all subsets of markets of equal size. For example, when $m = 3$, this implies two restrictions $\gamma_{12} = \gamma_{23} = \gamma_{13}$; when $m = 4$, this implies eight restrictions $\gamma_{12} = \gamma_{13} = \gamma_{14} = \gamma_{23} = \gamma_{24} = \gamma_{34}$ together with $\gamma_{123} = \gamma_{124} = \gamma_{134} = \gamma_{234}$. When horizontal homogeneity holds and $m > 2$, the hypothesis of vertical homogeneity assumes that the effect on the dependence ratio of expanding the subset of markets by one is the same irrespective of the size of the initial subset. For example, when $m = 3$, this hypothesis implies one additional restriction $\gamma_{123} = \gamma_{12}^2$; when $m = 4$, it implies two additional restrictions $\gamma_{1234} = \gamma_{123}^{3/2} = \gamma_{12}^3$.

More intricate hypotheses implying existence of latent factors underlying the joint distribution of indicators are described in Ekhholm, McDonald and Smith (2000). One of their association models, called “a latent binary factor”, seems to be relevant to financial markets. This model states that all m indicators are independent conditional on a realization of a single unobservable binary factor, say $b \in \{0, 1\}$. This imposes the following constraints on the dependence structure: there is horizontal homogeneity, and each ℓ^{th} order ($\ell = 2, \dots, m$) dependence ratio equals $(v + (1 - v)w^\ell) / (v + (1 - v)w)^\ell$, where $v = E(b) \in (0, 1)$ and $w = E(I_i | b = 0) / E(I_i | b = 1) \in (0, 1)$ are two “deep” parameters. For example, if $m = 3$, this means that $\gamma_{12} = \gamma_{23} = \gamma_{13} = (v + (1 - v)w^2) / (v + (1 - v)w)^2$ and $\gamma_{123} = (v + (1 - v)w^3) / (v + (1 - v)w)^3$. The factor b may be interpreted as an indicator reflecting fundamental changes in all economies associated with the involved markets and driving the returns in such a way that all “residual” movements are independent across the markets. The parameters v and w are interpretable from their definitions.

3 Adding dynamics

There is sufficient evidence that the directions of returns are predictable from past history. Christoffersen and Diebold (2006), Linton and Whang (2007) and Anatolyev and Gospodinov (2009) find convincing evidence of sign predictability of US stock returns, while Anatolyev (2008) discovers significant directional predictability in some Eastern European stock markets. Christoffersen and Diebold (2006) show that directional predictability may be induced by volatility dynamics alone even when the returns are conditionally mean independent. Thus, even efficient markets may well be characterized by directional predictability.

As noted in the previous section, in Ekholm, Smith, and McDonald (1995) and Ekholm, McDonald and Smith (2000) marginal probabilities and dependence ratios are handled separately. In particular, the authors suggest setting each marginal probability to some function (such as logistic) of associated covariates. We also follow the approach of separate modeling, but, because in our context the “units” indexed by t are time periods rather than IID individual objects, we parameterize the marginal probabilities and possibly dependence ratios as functions of the history of trades. We include past directions-of-change in all m markets as variables driving the evolution of marginal probabilities, and the impact of past indicators is allowed to be different for different markets:

$$\phi_{i,t} = \Lambda(\theta_{i,t}), \quad i = 1, \dots, m, \quad (4)$$

where

$$\theta_t = \omega + \sum_{j=1}^p \Gamma_j \theta_{t-j} + \sum_{j=1}^q \Delta_j I_{t-j}, \quad (5)$$

where $\Lambda(\cdot)$ is the logistic function

$$\Lambda(z) = \frac{\exp(z)}{1 + \exp(z)},$$

$\theta_t = (\theta_{1,t}, \theta_{2,t}, \dots, \theta_{m,t})'$, ω is an $m \times 1$ vector of intercepts, each Γ_j is an $m \times m$ matrix of autoregressive parameters, and each Δ_j is an $m \times m$ matrix of loadings on past directions-of-change. Such specification for determination of marginal probabilities is similar to that in Mosconi and Seri (2006) whose latent regression in a bivariate probit analysis includes a finite number of past indicators as driving variables. In the latent processes approach, a vector of the latent variable and explanatory variables is set to follow a VAR process as in Dueker (2005). We extend these ideas and in addition include past values of

θ_t on the right side of (5). In fact, our specification is a two-way generalization of the binary autoregressive model first proposed in Cox (1981). Even though past indicators are binary variables, the equation (5) is able to generate a rich set of values for θ_t and thus for marginal probabilities. The properties of θ_t are familiar from the multivariate ARMA literature. The equation (5) is a parsimonious representation of an infinite order autoregressive structure $\theta_t = \tilde{\omega} + \Psi(L) I_t$ for $\Psi(L) = \Gamma(L)^{-1} \Delta(L)$, where $\Gamma(z) = I_m - \sum_{j=1}^p \Gamma_j z^j$ and $\Delta(L) = \sum_{j=1}^q \Delta_j z^j$ are finite order matrix polynomials assumed to be of full rank, and L is the lag operator. The usual stationarity condition requires that the roots of $\det(\Gamma(z))$ lie outside the unit circle.

Past returns in place of past indicators potentially may also be used as driving variables in (5). However, in our applications we have discovered a much better explanatory power of past indicators (cf. Kauppi and Saikkonen, 2008). Kauppi and Saikkonen (2008) and Startz (2008) in a univariate context suggest including lagged marginal probabilities $\phi_{i,t}$ as regressors instead of lagged $\theta_{i,t}$. Another possibility is including current and lagged values of volatility measures as in Christoffersen and Diebold (2006), Christoffersen, Diebold, Mariano, Tay, and Tse (2007) and Anatolyev and Gospodinov (2009); this requires availability of realized volatility data or supplementary models for volatility.

We refer to the specification (4)–(5) as the MGARL(p, q) model where the acronym MGARL stands for “multivariate generalized autoregressive logit”. Of course, the normal CDF $\Phi(\cdot)$ may be used in place of the logistic function $\Lambda(\cdot)$, in which case we have the MGARP model, the last letter standing for “probit”. Logit specifications can be found in the literature stressing financial applications (e.g., Foresi and Peracchi, 1995; Rydberg and Shephard, 2003; Christoffersen and Diebold, 2006; Anatolyev and Gospodinov, 2009), while probit is more favored in macroeconomic applications (e.g., Chauvet and Potter, 2005; Dueker, 2005; Kauppi and Saikkonen, 2008) as well as in microeconomic empirical examples (e.g., Mosconi and Seri, 2006). The assumption of the logistic distribution (4) as well as the dynamic equation (5) are testable; see, for example, a review in Gouriéroux (2000).

Let us first assume that all $2^m - m - 1$ dependence ratios do not depend on time. Then the dynamics of joint probabilities is automatically determined by the evolution of marginal probabilities. The dependence ratios change their interpretation: they are now indicative of the degree of *conditional* dependence among the involved markets given past realizations in all m markets. As there are $2^m - m - 1$ dependence ratios, in total there are $(m + pm^2 + qm^2) + (2^m -$

$m - 1$) parameters in the collection

$$\psi = (\omega, \Gamma_1, \dots, \Gamma_p, \Delta_1, \dots, \Delta_q, \gamma_{12}, \dots, \gamma_{m-1,m}, \gamma_{123}, \dots, \gamma_{m-2,m-1,m}, \dots, \gamma_{12\dots m}).$$

The loglikelihood function equals

$$\ell(x_1, x_2, \dots, x_T; \psi) = \sum_{t=1}^T \left(\sum_{y \in \{0,1\}^m} I_t^y \log \pi_t^y \right),$$

and is to be maximized with respect to ψ . The maximum likelihood theory suggests that under suitable conditions, the maximum likelihood estimator $\hat{\psi}$ of ψ is consistent and asymptotically normal, with an information matrix consistently estimated in a straightforward way. The “suitable conditions” are likely to include the requirements for a correct specification and stationarity. The exact formulation of such conditions is outside the scope of this paper; the closest reference is de Jong and Woutersen (2007).

The dependence ratios too may be parameterized as functions of past data. It is reasonable to assume that they depend on past directions only from those markets that are related to these dependence ratios. Two natural flexible specifications are

$$\gamma_{i_1 i_2 \dots i_\ell, t} = \kappa_{i_1 i_2 \dots i_\ell} \exp \left(\sum_{i \in \{i_1, i_2, \dots, i_\ell\}} \sum_{j=1}^r \lambda_{i,j} I_{i,t-j} \right) \quad (6)$$

and

$$\gamma_{i_1 i_2 \dots i_\ell, t} = \kappa_{i_1 i_2 \dots i_\ell} + \sum_{i \in \{i_1, i_2, \dots, i_\ell\}} \sum_{j=1}^r \lambda_{i,j} I_{i,t-j}, \quad (7)$$

where $1 < \ell \leq m$ and $1 \leq i_j \leq m$ for all $j = 1, \dots, \ell$. The vector of parameters ψ then contains all $\kappa_{i_1 i_2 \dots i_\ell}$ and $\lambda_{i,j}$ in place of all $\gamma_{i_1 i_2 \dots i_\ell}$. The λ -parameters indicate how the degree of dependence varies with directions of past movements, i.e. whether bullish or bearish markets are more dependent. The dynamics of joint probabilities is determined by both evolution of marginal probabilities and evolution of dependence ratios. The generalization of homogeneity hypotheses in case when the dependence ratios are time varying is straightforward, but the latent binary factor association model does not seem to be.

Irrespective of whether there is dynamics in the model or not, the predicted probabilities for market i conditional on all the other $m - 1$ indicators in the same period can be constructed as

$$\phi_{i,t|i} = \frac{\Pr \{I_{1,t}, \dots, I_{i-1,t}, 1, I_{i+1,t}, \dots, I_{m,t} | I_{t-1}, I_{t-2}, \dots\}}{\sum_{y_i \in \{0,1\}} \Pr \{I_{1,t}, \dots, I_{i-1,t}, y_i, I_{i+1,t}, \dots, I_{m,t} | I_{t-1}, I_{t-2}, \dots\}}. \quad (8)$$

This formula contains only values of π_t^y for $y = (I_{1,t}, \dots, I_{i-1,t}, 1, I_{i+1,t}, \dots, I_{m,t})$ and for $y = (I_{1,t}, \dots, I_{i-1,t}, 0, I_{i+1,t}, \dots, I_{m,t})$. The denominator can also be interpreted as π_t^y for $y = (I_{1,t}, \dots, I_{i-1,t}, I_{i+1,t}, \dots, I_{m,t})$ relative to the whole set of markets less the i^{th} market. For example, if $m = 3$, the predicted conditional probabilities for the first market are $\phi_{1,t|-1} = \phi_{1,t}\gamma_{123,t}/\gamma_{23,t}$ if $I_{2,t} = I_{3,t} = 1$; $\phi_{1,t|-1} = \phi_{1,t}(\gamma_{12,t} - \phi_{3,t}\gamma_{123,t}) / (1 - \phi_{3,t}\gamma_{23,t})$ if $I_{2,t} = 1, I_{3,t} = 0$; etc. Similarly one can construct predicted conditional joint probabilities. Of course, the feasible forecasts $\hat{\phi}_{i,t|-i}$ require replacing the unknown parameters by their estimates from $\hat{\psi}$.

4 Empirical analysis of various stock markets

To illustrate the technique, we use three sets of stock market data, two of them with 3 indexes (so that $m = 3$) and one with 2 indexes (so that $m = 2$). The first set includes daily indexes from the developed European markets: DAX (Deutsche Börse AG), CAC (Paris Bourse) and FTSE (London Stock Exchange) from the beginning of 1991 to the end of 2000 (2542 days in total). The second set includes daily indexes from the young emerging markets of Baltic states: TALSE (Tallinn Stock Exchange) of Estonia, VILSE (Vilnius Stock Exchange) of Lithuania, and RIGSE (Riga Stock Exchange) of Latvia from January 2000 to January 2005 inclusive (1302 days in total). The third set includes daily indexes from two Chinese stock exchanges, Shanghai a and Shenzhen a, from October 1992 to January 2005 inclusive (3007 days in total). For each data set, the series under analysis x_t is a collection of logarithmic returns for three or two indexes. Some summary statistics of the data are collected in Table 1.

Throughout, we set $c = 0$, and use the logistic function for the link. For simplicity, we handle missing data (on days when a stock exchange in one country did not work while a stock exchange in another country in the set worked) by imputing returns by dividing the total return by a number of days with missing data (for indicators, this effectively means replicating the sign of return for the following day). All procedures are written in GAUSS. The loglikelihood maximization is performed using the `cml` (constrained maximum likelihood) procedure. The beginning-of-sample values for dynamically changing variables are set to sample means, for example, $\theta_1 = \left(I_m - \sum_{j=1}^p \Gamma_j\right)^{-1} \left(\omega + (\sum_{j=1}^q \Delta_j)(T^{-1} \sum_{t=1}^T I_t)\right)$. In choosing lag orders p and q , we fix r at zero and apply the general-to-specific methodology where, starting from maximal orders $p = 1$ and $q = 3$ we successively remove

more distant lags unless at least one parameter in matrix Γ_j and at least one parameter in matrix Δ_j are significant at the 5% level. After having fixed p and q , we set $r = 2$ and remove lags of indicators unless t-ratios for all λ -parameters left exceed unity. Additional experimentation shows that the final configurations of p , q and r are insensitive to the order of removal of distant lags.

Table 1. Summary statistics for the return data from the European, Baltic and Chinese stock markets.

	Mean, $\times 10^{-4}$	StDev, $\times 10^{-2}$	Ups	Downs
DAX	6.12	1.20	1367	1174
CAC	5.43	1.20	1347	1194
FTSE	4.24	0.91	1312	1229
TALSE	9.23	1.09	725	576
VILSE	8.84	0.88	725	576
RIGSE	10.60	1.73	729	572
Shanghai a	1.84	2.69	1484	1522
Shenzhen a	0.27	2.43	1513	1493

Notes: “Mean” and “StDev” stand for the mean and standard deviation of log returns; “Ups” and “Downs” – numbers of ups and downs.

The estimation results are reported in Tables 2a, 2b, and 2c. Note first that for the European and Chinese markets there is no statistically significant dynamics in marginal success probabilities. This partly reflects the efficiency of the developed European markets. The point estimates of ω_i and their signs reflect the extent to which the markets are bullish or bearish (cf. Table 1). In contrast to the European and Chinese ones, the inefficient Baltic markets reveal severe serial dependence in marginal success probabilities: the coefficients at lagged θ_t are quite close to unity (two of the three being within two standard deviations from 1), and whole two lags of past indicators are statistically significant at the 5% significance level. Interestingly, all off-diagonal elements of Δ_1 and Δ_2 are statistically insignificant; moreover, they are jointly insignificant (the LR test statistic is 7.47 which is far below conventional critical values for the $\chi^2(12)$ distribution). Therefore, we have imposed their equality to zero and show only the (reestimated) diagonal elements δ_1 and δ_2 of Δ_1 and Δ_2 . This means that last period indicators realized in the other markets have influence on the marginal probability of a given market only indirectly through the term $\Gamma_1\theta_{t-1}$.

Table 2a. Estimates of parameters of the MGARL model for the European stock markets.

	DAX	CAC	FTSE
ω	0.152 0.040	0.120 0.040	0.065 0.040

Notes: Standard errors are below point estimates.

Table 2b. Estimates of parameters of the MGARL model for the Chinese stock markets.

	Shanghai	Shenzhen
ω	-0.024 0.037	0.013 0.036

Notes: Standard errors are below point estimates.

Table 2c. Estimates of parameters of the MGARL model for the Baltic stock markets.

	TALSE	VILSE	RIGSE
ω	0.068 0.042	-0.359 0.123	-0.039 0.026
Γ_1	$\begin{pmatrix} 0.920 \\ 0.064 \\ 0.245 \\ 0.129 \\ 0.114 \\ 0.050 \end{pmatrix}$	$\begin{pmatrix} 0.079 \\ 0.038 \\ 0.648 \\ 0.093 \\ -0.062 \\ 0.030 \end{pmatrix}$	$\begin{pmatrix} -0.297 \\ 0.139 \\ 0.648 \\ 0.370 \\ 1.012 \\ 0.075 \end{pmatrix}$
δ_1	$\begin{pmatrix} 0.449 \\ 0.104 \end{pmatrix}$	$\begin{pmatrix} 0.414 \\ 0.117 \end{pmatrix}$	$\begin{pmatrix} -0.164 \\ 0.103 \end{pmatrix}$
δ_2	$\begin{pmatrix} -0.441 \\ 0.105 \end{pmatrix}$	$\begin{pmatrix} -0.005 \\ 0.143 \end{pmatrix}$	$\begin{pmatrix} 0.204 \\ 0.110 \end{pmatrix}$

Notes: The row vectors δ_1 and δ_2 contain diagonal elements of Δ_1 and Δ_2 , all off-diagonal elements being set to zero. Standard errors are below point estimates.

Tables 3a and 3b contain estimates of dependence ratios for the European and Baltic markets. The dependence ratios are constant in both data sets; no significant predictability is attained using flexible specifications (6) or (7). For the European markets the dependence ratios of order two are quite high, and signify that the probability of two markets both going up is 30% to 40% larger than this probability would be under independence. The dependence ratio of order three is around 2, i.e. the probability of the three markets all going up is twice as large as this probability under independence. These figures reflect a high degree of integration of trades in these leading European markets. In contrast, for the Baltic markets the dependence ratios of order two are close to unity, if at all statistically significantly different from it. These

figures are a reflection of a low degree of integration of trades in the emerging stock markets, despite the fact that they are geographically close and were established nearly at the same time under similar economic conditions. Table 3c contains estimates of the dependence ratio for the Chinese markets. The dependence ratio is not statistically significantly far from a constant; using the flexible specification (7) yields slightly significant dependence which we nevertheless show. As the results indicate, even higher than in the case of European markets is the dependence ratio of order two for the two Chinese exchanges, which is of no surprise. This figure implies that the probability of the two exchanges both going up is 72% larger than this probability would be under independence. Also, the dependence ratio tends to be negatively associated with past upward movements in both exchanges. The impact of the Shanghai return is first positive, but in the next period it is annihilated and the long-run effect is slightly negative.

Table 3a. Estimates of dependence ratios in the MGARL model for the European stock markets.

DAX-CAC	DAX-FTSE	CAC-FTSE	DAX-CAC-FTSE
1.348 0.020	1.319 0.020	1.413 0.022	2.058 0.055

Notes: Standard errors are below point estimates.

Table 3b. Estimates of dependence ratios in the MGARL model for the Baltic stock markets.

TALSE-VILSE	TALSE-RIGSE	VILSE-RIGSE	TALSE-VILSE-RIGSE
1.048 0.020	1.032 0.021	0.983 0.021	1.065 0.039

Notes: Standard errors are below point estimates.

Table 3c. Estimates of dependence ratios and their dynamics in the MGARL model for the Chinese stock markets.

	$r = 0$	$r = 2$
κ_{12}	1.723 0.027	1.748 0.034
$\lambda_{1,1}$		0.036 0.031
$\lambda_{2,1}$		-0.048 0.031
$\lambda_{1,2}$		-0.038 0.025
$\lambda_{2,2}$		

Notes: Standard errors are below point estimates.

Table 4 reports results of testing hypotheses about dependence ratios. In European markets, the dependence ratios are so high that testing for dependence is superfluous. The two-market dependence ratios, however, are not far from each other so that the horizontal homogeneity may be tested. The test, however, rejects horizontal homogeneity even at the 1% significance level. In the Baltic markets, all three hypotheses of independence, horizontal homogeneity and vertical homogeneity are rejected at the 10% significance level, but are not at the 5% significance level. This in any case points at the symmetry and low, if at all, dependence among these emerging markets. If we believe in horizontal homogeneity and the latent binary factor, but not in vertical homogeneity in the case of Baltic markets, we can compute the implied values of $v = E(b)$ and $w = E(I_i|b=0)/E(I_i|b=1)$, the parameters of the corresponding association model. To this end, we apply minimum distance estimation minimizing the quadratic distance between the vector of estimated dependence ratios and the vector of those implied by the mapping from v and w . This yields point estimates $v = 0.40$ and $w = 0.74$, which imply that the latent fundamental factor is a bit favorable to downward movements (60/40 odds), which is somewhat surprising for bullish markets; given an upward fundamental movement, an individual market has 38% more chances to be positive than in case the fundamental movement is downward.

Table 4. Tests of dependence patterns in the MGARL model for the European and Baltic stock markets.

Hypotheses	European market	Baltic market
Independence		9.23 5.6%
Horizontal homogeneity	19.36 0.0%	5.81 5.5%
Vertical homogeneity		7.59 5.5%

Notes: The independence statistic is distributed as $\chi^2(4)$ under independence across the markets. The horizontal homogeneity statistic is distributed as $\chi^2(2)$ under horizontal homogeneity across the markets. The vertical homogeneity statistic is distributed as $\chi^2(3)$ under vertical homogeneity across the markets. P-values are below test statistics.

Our final exercise is on out-of-sample prediction, the point of which is to show a high ability of the model to perform well out-of-sample, even though such forecasting problem is barely feasible in real life. We use additional (“out-of-sample”) data for the European markets from the beginning to the end of 2001 (245 days in total). We compare several sets of directional predictions. One set represents two naïve forecasts: “straight up” when the prediction is

1 no matter what; “straight down” when the prediction is 0 no matter what. The second set of predictions, called “unconditional” and “conditional”, is generated by our “direct” model described and estimated before. To generate “unconditional” predictions, no information on directions-of-change in the two other markets is used. That is, such predictions use three individual separate univariate models, and we employ both static (corresponding to the absence of dynamics in the fitted MGARL model) and dynamic (using only lagged directions for the same markets) versions. To generate “conditional” predictions, information on directions-of-change in the other two markets is used as described at the end of section 3. The third type predictions used for comparison purposes are generated using the “indirect” approach to prediction of directions-of-change, i.e. via a fully specified multivariate dynamic model for the returns. In the “indirect” approach, we exploit a trivariate GARCH–DCC model of Engle (2002):

$$x_t|I_{t-1} \sim N(\mu, \Omega_t),$$

where $\Omega_t = D_t R_t D_t$,

$$\begin{aligned} D_t^2 &= \Psi + K \text{diag} \{ (x_{t-1} - \mu)(x_{t-1} - \mu)' \} + L D_{t-1}^2, \\ R_t &= \text{diag} \{ Q_t \}^{-1/2} Q_t \text{diag} \{ Q_t \}^{-1/2}, \\ Q_t &= \bar{Q} (1 - \alpha - \beta) + \alpha D_t^{-1} (x_{t-1} - \mu)(x_{t-1} - \mu)' D_t^{-1} + \beta Q_{t-1}, \end{aligned}$$

Ψ , K , L are diagonal and positive, and \bar{Q} is symmetric positive definite, and $\text{diag} \{ \cdot \}$ replaces all non-diagonal elements with zeros. Because of conditional normality, conditional mean forecasts are linear in the other markets’ returns and can be computed in the following way: for market i , the conditional mean is $\mu_{i,t|-i} = S_i \mu + F' S_{-i} (x_t - \mu)$, the conditional variance is $\omega_{i,t|-i} = S_i \Omega_t S_i' + F_t' S_{-i} \Omega_t S_{-i}' F_t$, where $F_t = (S_{-i} \Omega_t S_{-i}')^{-1} S_{-i} \Omega_t S_i'$, and S_i , S_{-i} are selector matrices that select matrix rows related to market i and all markets but i , respectively. The conditional directional probability forecasts for market i are generated as $\phi_{i,t|-i} = \Phi(\mu_{i,t|-i} / \sqrt{\omega_{i,t|-i}})$. Of course, the feasible forecasts $\hat{\phi}_{i,t|-i}$ are constructed using estimates of the GARCH–DCC model.

We evaluate directional forecasts using the proportion of unsuccessful directional predictions U and the (absolute value) Brier score B popular in earth sciences and medicine, but sometimes used in economics and finance (see, e.g.,

Christoffersen, Diebold, Mariano, Tay, and Tse, 2007):

$$B_i = \frac{1}{P} \sum_{t=T+1}^{T+P} |\hat{\phi}_{i,t} - I_{i,t}|,$$

$$U_i = \frac{1}{P} \sum_{t=T+1}^{T+P} |\hat{I}_{i,t} - I_{i,t}|,$$

where $I_{i,t}$ is a realized direction indicator, $\hat{\phi}_{i,t}$ is a positive direction-of-change probability forecast, $\hat{I}_{i,t} = \mathbb{I}\{\hat{\phi}_{i,t} > \text{cut}_i\}$ is a direction forecast, cut_i is a specified cutoff level, and P is a number of predictions ($P = 244$ in our case). Both statistics are bounded between 0 and 1, the closer they are to zero the better the predictions are. In contrast to matching indicator predictions with realizations in U , directional probabilities and indicator realizations are compared in B . The cutoff level cut_i separates “up” forecasts from “down” forecasts. Reasonable values for cut_i are 0.5 and a proportion of positive returns in the sample. Note that by construction both measures for naïve forecasts in a given market are equal.

Table 5. Out-of-sample forecasting results for the European stock markets.

Market	DAX	CAC	FTSE	DAX	CAC	FTSE
Predictions	B			U		
Straight up	0.53	0.52	0.52	0.53	0.52	0.52
Straight down	0.47	0.48	0.48	0.47	0.48	0.48
Unconditional	0.50	0.50	0.50	0.53	0.52	0.52
Conditional	0.33	0.29	0.32	0.16	0.18	0.18
GARCH–DCC	0.36	0.41	0.41	0.18	0.25	0.24

Notes: Reported are statistics measuring the degree of out-of-sample fit of direction forecasts: B is the (absolute value) Brier score, and U denotes a proportion of unsuccessful directional predictions.

The results are presented in Table 5. We compute indicator forecasts for $\text{cut}_i = 0.5$; setting cut_i to a proportion of positive returns in the estimation sample when the European markets were slightly bullish delivers qualitatively similar results, but the difference between the “direct” and “indirect” approaches is more crispy; alternatively, setting it to a proportion of positive returns in the prediction sample when the European markets were slightly bearish diminishes that difference. The static unconditional forecasts in our

case coincide with “straight up” forecasts because the predicted probabilities for all markets are greater than 0.5. These forecasts are overall bad and their reversion to “straight down” improves the statistics, not appreciably though. Switching to dynamic unconditional forecasts leaves direction forecasts intact, while only slightly shifts the Brier scores.

The use of dependence ratios improves the quality criteria dramatically, from near non-predictability to very significant predictability, because of a high degree of dependence among the three markets. Importantly, the “indirect” approach using the GARCH–DCC model delivers worse directional predictions, both in terms of predicted directional probabilities and predicted directions themselves. Interestingly, the GARCH–DCC model has a tendency to be more “pessimistic” about “up” movements than the MGARL model: in total 160 (out of 375) “up” predictions generated by the MGARL model are “down” predictions when generated by the GARCH–DCC model, but only 4 (out of 357) predictions are switched in the opposite direction. On average, it is more (wrongfully) pessimistic about the CAC index, which is reflected in a larger increase in scores relative to scores provided by the MGARL model. We have also generated predictions from individual GARCH submodels implicit in the GARCH–DCC structure. These “indirect” univariate predictions have turned out to be very similar to the “straight down” forecasts.

The “indirect” approach shows worse performance presumably because of possibly incorrect shape restrictions (e.g., conditional normality) or/and incorrect dynamics (e.g., no conditional mean dynamics), and in addition because of relatively many parameters to estimate (the GARCH–DCC model contains 20 parameters, while only 7 parameters are used in the MGARL(0,0) model). Of course, it is possible that a more sophisticated “indirect” model with a skewed thick-tailed conditional density and complicated mean dynamics may eventually outperform the MGARL(0,0) model, but the computations may turn out to be enormous (in particular, the form of the conditional density may be very involved and not necessarily explicit). The proposed “direct” approach and the MGARL framework are less prone to these tendencies. Also, it is as well possible that including a nonlinear mean dynamics to the MGARL(0,0) equations would improve its forecasting abilities too.

5 Concluding remarks

We have developed a multivariate dynamic model for the joint distribution of binary outcomes associated with directions-of-change for several markets or assets, where the marginal distribution of each binary outcome follows a

dynamic binary choice model, while the association structure is parameterized via possibly time varying dependence ratios. We have illustrated the technique using daily stock index returns from three European markets, from three Baltic markets, and from two Chinese exchanges. While the European and Chinese data show little dynamics but high dependence, the Baltic data show a lot of dynamics but little dependence among the individual markets. Finally, modeling dependence ratios is able to produce out-of-sample conditional directional forecasts of high quality and dominate in this sense the “indirect” approach and univariate modeling.

References

Anatolyev, S. (2008) Nonparametric retrospection and monitoring of predictability of financial returns. *Journal of Business & Economic Statistics*, forthcoming.

Anatolyev, S. and A. Gerko (2005) A trading approach to testing for predictability. *Journal of Business & Economic Statistics* 23, 455–461.

Anatolyev, S. and N. Gospodinov (2009) Modeling financial return dynamics via decomposition. *Journal of Business & Economic Statistics*, forthcoming.

Bauwens, L. and S. Laurent (2005) A new class of multivariate skew densities, with application to generalized autoregressive conditional heteroscedasticity models. *Journal of Business & Economic Statistics* 23, 346–354.

Bekiros, S. and D.A. Georgoutsos (2008) Direction-of-change forecasting using a volatility based recurrent neural network. *Journal of Forecasting*, forthcoming.

Birchenhall, C.R., H. Jessen, D.R. Osborn, and P. Simpson (1999) Predicting U.S. business-cycle regimes. *Journal of Business & Economic Statistics* 17, 313–323.

Breen, W., L.R. Glosten, and R. Jagannathan (1989) Economic significance of predictable variations in stock index returns. *Journal of Finance* 44, 1177–1189.

Chauvet, M., and S. Potter (2005) Forecasting recessions using the yield curve. *Journal of Forecasting* 24, 77–103.

Christoffersen, P.F. and F.X. Diebold (2006) Financial asset returns, direction-of-change forecasting, and volatility dynamics. *Management Science* 52, 1273–1288.

Christoffersen, P.F., F.X. Diebold, R.S. Mariano, A.S. Tay, and Y.K. Tse (2007) Direction-of-change forecasts based on conditional variance, skewness and kurtosis dynamics: international evidence. *Journal of Financial Forecasting* 1, 3–24.

Cox, D.R. (1981) Statistical analysis of time series: some recent developments. *Scandinavian Journal of Statistics* 8, 93–115.

de Jong, R. and T. Woutersen (2007) Dynamic time series binary choice. Manuscript, Ohio State University.

Dueker, M.J. (2005) Dynamic forecasts of qualitative variables: a qual VAR model of U.S. recessions. *Journal of Business & Economic Statistics* 23, 96–104.

Ekholm, A., P.W.F. Smith, and J.W. McDonald (1995) Marginal regression analysis of a multivariate binary response. *Biometrika* 82, 847–854.

Ekholm, A., J.W. McDonald, and P.W.F. Smith (2000) Association models for a multivariate binary response. *Biometrics* 56, 712–718.

Engle, R. (2002) Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics* 20, 339–350.

Fitzmaurice, G.M. and N. Laird (1993) A likelihood-based method for analysing longitudinal binary responses. *Biometrika* 80, 141–151.

Foresi, S. and F. Peracchi (1995) The conditional distribution of excess returns: an empirical analysis. *Journal of the American Statistical Association* 90, 451–466.

Glonek, G.F.V. (1996) A class of regression models for multivariate categorical responses. *Biometrika* 83, 15–28.

Glonek, G.F.V. and P. McCullagh (1995) Multivariate logistic models. *Journal of Royal Statistical Society B* 57, 533–546.

- Gouriéroux, C. (2000) *Econometrics of qualitative dependent variables*. Cambridge: Cambridge University Press.
- Gouriéroux, C. and A. Monfort (1981) Asymptotic properties of the maximum likelihood estimator in dichotomous logit models. *Journal of Econometrics* 17, 83–97.
- Greer, M. (2003) Directional accuracy tests of long-term interest rate forecasts. *International Journal of Forecasting* 19, 291–298.
- Hartzmark, M.L. (1991) Luck versus forecast ability: determinants of trader performance in futures markets. *Journal of Business* 64, 49–74.
- Henriksson, R.D. and R.C. Merton (1981) On market timing and investment performance II: statistical procedures for evaluating forecasting skills. *Journal of Business* 54, 513–533.
- Chung, J. and Y. Hong (2007) Model-free evaluation of directional predictability in foreign exchange markets. *Journal of Applied Econometrics* 22, 855–889.
- Kauppi, H. and P. Saikkonen (2008) Predicting U.S. recessions with dynamic binary response models. *Review of Economics and Statistics*, forthcoming.
- Leitch, G. and J.E. Tanner (1991) Economic forecast evaluation: profits versus the conventional error measures. *American Economic Review* 81, 580–590.
- Linton, O. and Y.J. Whang (2007) The quantilogram: with an application to evaluating directional predictability. *Journal of Econometrics* 141, 250–282.
- Mosconi, R. and R. Seri (2006) Non-causality in bivariate binary time series. *Journal of Econometrics* 132, 379–407.
- Patton, A.J. (2006) Modelling asymmetric exchange rate dependence. *International Economic Review* 47, 527–556.
- Pesaran, M.H. and A. Timmermann (1992) A simple nonparametric test of predictive performance. *Journal of Business & Economic Statistics* 10, 561–565.

Pesaran, M.H. and A. Timmermann (1995) Predictability of stock returns: robustness and economic significance. *Journal of Finance* 50, 1201–1228.

Pesaran, M.H. and A. Timmermann (2004) How costly is it to ignore breaks when forecasting the direction of a time series? *International Journal of Forecasting* 20, 411–425.

Rydberg, T.H. and N. Shephard (2003) Dynamics of trade-by-trade price movements: decomposition and models. *Journal of Financial Econometrics* 1, 2–25.

Startz, R. (2008) Binomial autoregressive moving average models with an application to U.S. recessions. *Journal of Business & Economic Statistics* 26, 1–8.

Tajar, A., M. Denuit, and P. Lambert (2001) Copula-type representation for random couples with Bernoulli margins. Discussion paper 0118, Université Catholique de Louvain.