Directional news impact curve

Stanislav Anatolyev*

CERGE-EI, Czech Republic and New Economic School, Russia

May 2020

Abstract

The directional news impact curve (DNIC) is a relationship between returns and a probability of next period's return to exceed a certain threshold, zero in particular. Using long series of S&P500 index returns and a number of parametric models suggested in the literature as well and flexible semiparametric models, we investigate the shape of DNIC, as well as forecasting abilities of these models. The semiparametric approach reveals that the DNIC has complicated shapes characterized by non-symmetry with respect to past returns and their signs, heterogeneity across the thresholds, and changes over time. Simple parametric models often miss some important features of the DNIC, but some nevertheless exhibit superior out-of-sample performance.

Keywords: stock returns, news impact curve, price directions, directional predictions

*Address: Stanislav Anatolyev, CERGE-EI, Politických vězňů 7, 11121 Prague 1, Czech Republic. E-mail: stanislav.anatolyev@cerge-ei.cz. I am grateful to the editor, associate editor and two anonymous referees whose comments and suggestions greatly improved the presentation. This research was supported by the grant 20-28055S from the Czech Science Foundation.

1 Introduction

The stock returns generally exhibit directional (sign) predictability (Gençay 1998, Hong and Chung 2003, Anatolyev and Gerko 2005, Christoffersen and Diebold 2006, Linton and Whang 2007). Directions of movements of asset prices, or more generally, exceedances by asset returns of certain thresholds, are important to model and predict. Observation-driven dynamic models for directional probability require specification of the probability equation, typically an input of the logistic link (i.e., the log odds ratio). This is useful, among others, in distributional regressions for returns (Foresi and Peracchi 1995, Anatolyev and Baruník 2019), in decomposition-based models (Rydberg and Shephard 2003, Anatolyev and Gospodinov 2010), and simply for predicting directions of financial market movements (Nyberg 2013, Taylor and Yu 2016). In practice, the probability equation is typically specified as a linear equation, possibly with a dynamic feedback, with (an) observable driving variable(s) from the past history.

There are no stylized facts behind the movement of directional probabilities, compared to, for example, volatility equations,[1] hence these driving variables are typically specified in an *ad hoc* manner. Most use past indicators of return movements and/or (simple functions of) past returns. With daily data, typically, all 'explanation power' comes from one or more such variables. Hong and Chung (2003) use a past indicator and past returns together with their powers of 2 to 4; Anatolyev (2009) uses past indicators only, Skabar (2013) uses up to 5 lags of returns; Liu and Luger (2015) and Frazier and Liu (2016) use a past indicator and past return squared; Taylor and Yu (2016) use, in a variety of specifications, a past indicator, a past signed indicator, a past absolute return or a past signed absolute return; Anatolyev and Baruník (2019) use a past indicator and a simple proxy for past daily volatility. This illustrates unawareness of researchers about what variables (or more precisely, what function(s) of past returns) one should use in probability equations as predictors. With ultra-high frequency (transaction) data, Rydberg and Shephard (2003) have to use a pretty long list of sophisticated past indicators and returns. With monthly data, most of the 'explanation power' is typically provided by various macroeconomic variables or financial ratios; nevertheless, researchers still include at least one term from the past related to the return. For example, Foresi and Peracchi (1995) and Nyberg (2013) use a past return; (in addition some extraneous predictors) Anatolyev and Gospodinov (2010, 2019) use only a

[1] It is this abundance of stylized facts about the temporal behavior of squared returns that spurred a great deal of research on volatility and a big volume of ARCH models (Bollerslev, 2010). The dynamics of other conditional features are much more problematic to specify; see, for example, Anatolyev and Petukhov (2016) for the dynamics of conditional skewness.

past indicator, while Anatolyev, Gospodinov, Jamali, and Liu (2017) use a past indicator together with a squared and cubed past return.

We do a careful investigation of what functions of past returns are most beneficial to include in the probability equation for daily stock return data. In other words, we try to uncover the *directional news impact curve* (directional NIC, DNIC), the effect of past returns ('news') on the current probability of positive/negative returns (or more generally, the probability of returns exceeding/not exceeding a threshold). This is a directional analog of volatility NIC for conditional variance (Engle and Ng 1993) and skewness NIC for conditional third moments (Anatolyev and Petukhov 2016). Using a long history of daily S&P500 returns, as well as DAX and Nikkei returns, we utilize a number of parametric and semiparametric specifications, and evaluate the models using both in-sample and out-of-sample criteria. We are interested in models for daily returns as it is most likely that for this frequency the purely autoregressive probabilistic forecasts are in a greatest need; however, the uncovered DNIC can also serve as a useful guide for models on intradaily and monthly data.

The semiparametric models we use reveal that the DNICs have complicated shapes characterized by non-symmetry with respect to past returns, their signs in particular. There is a lot of heterogeneity across the thresholds, as well as evidence that the DNICs have been subject to changes in shapes and in their parameters. Simple parametric models often miss these features, but in spite of that, some of them exhibit superior out-of-sample performance in terms of likelihood and/or Brier scores. While such evidence for the S&P500 index is moderately strong, that for other indices, such as DAX and Nikkei, is pretty weak.

The material is organized as follows. The models for DNIC are outlined in Section 2. Section 3 describes the data used on the experiments, and the criteria we use to discriminate the models. In Section 4, we report the results of uncovering the directional NIC and those of forecasting experiments. Finally, Section 5 concludes.

2 Models for directional NIC

Let r_t be a daily return on day t. We are interested in the event

$$r_t \leq c,$$

where c is a threshold that may be zero or non-zero, its indicator

$$I_t = \mathbb{I}\left\{r_t \leq c\right\},$$

and the corresponding conditional probability

$$p_t = \Pr_{t-1}\{r_t \leq c\} = E_{t-1}[I_t].$$

Here, $\Pr_{t-1}$ and E_{t-1} denote the probability and expectation, respectively, conditional on the history $\{r_\tau, \tau \leq t-1\}$.

The autologit model uses the traditional[2] logit link:

$$p_t = \frac{1}{1+\exp(-\theta_t)},$$

where the log-odds ratio $\theta_t = \ln(p_t/(1-p_t))$ follows the dynamics

$$\theta_t = \omega + \phi_{t-2} + \psi_{t-1},$$

where the term ψ_{t-1}, as a function of r_{t-1}, is the *directional news impact curve* (DNIC), and the term ϕ_{t-2} is a function of variables dated earlier.[3] The target is to investigate possible specifications for ψ_{t-1} from the point of view of in-sample fit and predictive ability.

The DNIC is a very important measure of the predicted probability of exceeding the threshold. In particular, if a realization of r_{t-1} leads to an extremely large positive (negative) realization of DNIC, then the probability that the next return does not exceed the threshold is very high (very low). On the other hand, if a realization of r_{t-1} leads to a very small in absolute value realization of DNIC, the predicted probabilities of exceeding and not exceeding the threshold are approximately equal to $\frac{1}{2}$.

We call *static* the model with no dynamics, i.e. when $\phi_{t-2} = \psi_{t-1} = 0$. We consider the following dynamic specifications for ψ_{t-1}.

Simple benchmarks The simplest choices are using the past return linearly and in the form of an indicator, resulting in the *only return* model

$$\psi_{t-1} = \alpha_r r_{t-1}$$

and the *only indicator* model

$$\psi_{t-1} = \alpha_I I_{t-1}.$$

[2]Most of the literature on financial directions-of-change of returns does use the logit link; in contrast, the literature on predicting business cycles uses the probit link (e.g., Saikkonen and Kauppi 2008, Nyberg 2014).

[3]Note that ψ_{t-1} may depend, beside r_{t-1}, also on earlier history, additively and/or multiplicatively. The term ϕ_{t-2} usually equals zero, or takes the same form as ψ_{t-1}, or has a feedback form $\varsigma\theta_{t-1}$; in our empirical experiments (see Section 4) we use the first two choices.

We also include both predictors together and call it the *return and indicator* model:

$$\psi_{t-1} = \alpha_r r_{t-1} + \alpha_I I_{t-1}.$$

The inclusion of past return makes the DNIC a linear form, while the inclusion of past indicator divides the DNIC at the threshold value into two separate pieces with a same slope.

Complex benchmarks Based on *ad hoc* specifications in use in the previous literature, we also exploit the following, more complex, benchmarks. These benchmarks are related to attempts to account for volatility whose dynamics alone are able to generate directional predictability (Christoffersen and Diebold 2006). These specifications though are quite arbitrary choices given the approximate convoluted formula in Christoffersen, Diebold, Mariano, Tay, and Tse (2007), which is, however, also tied to arbitrary parameterizations of the conditional density. The specifications introduce various simple types of nonlinearity (beyond splitting it into two parts) into the directional NIC.

Based on Liu and Luger (2015) and Frazier and Liu (2016), we use the past indicator and return squared, and call it the *squared return and indicator* model:

$$\psi_{t-1} = \alpha_{r^2} r_{t-1}^2 + \alpha_I I_{t-1}.$$

The inclusion of a squared term introduces curviness (or non-constant returns) to the effect of past returns.

Next, we change the squared return to absolute return, and call it the *absolute return and indicator* model:

$$\psi_{t-1} = \alpha_{|r|} \left| r_{t-1} \right| + \alpha_I I_{t-1},$$

as well as change it to signed past absolute returns, and call it the *signed absolute return and indicator* model:

$$\psi_{t-1} = \alpha_{|r|-} \left| r_{t-1} \right| I_{t-1} + \alpha_{|r|+} \left| r_{t-1} \right| (1 - I_{t-1}) + \alpha_I I_{t-1},$$

as in Taylor and Yu (2016). The inclusion of an absolute value term makes the DNIC piecewise linear and not as curvy as the quadratic. Splitting the absolute value term in the latter specification is meant to introduce non-symmetry into the slope of DNIC with respect to the sign of past returns, a sort of a leverage effect.

Finally, following Anatolyev and Baruník (2019), we use a specification with a past indicator and a proxy for past daily volatility, and call it the *log-volatility and indicator* model:

$$\psi_{t-1} = \alpha_{\ln|r|} \ln\left(1 + |r_{t-1}|\right) + \alpha_I I_{t-1}.$$

The idea behind the first term is to involve an absolute value of the past return with 'penalization' of its high values.

Autoregressive score models Creal, Koopman and Lucas (2013) propose a family of models with dynamics based on the score function. In our case, the likelihood based on the Bernoulli distribution leads to the past indicator surprise as the predictor:

$$\psi_{t-1} = \alpha_{S0}\left(I_{t-1} - p_{t-1}\right).$$

Because p_{t-1} belongs to the past with respect to period $t-1$, the DNIC is essentially the same as in the simple 'only indicator' benchmark. Therefore, we use the following scaled version which we call the *autoregressive score* model:

$$\psi_{t-1} = \alpha_S \frac{I_{t-1} - p_{t-1}}{\sqrt{p_{t-1}\left(1 - p_{t-1}\right)}}.$$

The corresponding DNIC is also similar to that in the 'only indicator' benchmark, but with a time-varying coefficient, which depends on a more distant past history.

Semiparametric (piecewise linear) DNIC One semiparametric approach describes the directional NIC by a continuous piecewise linear function in a way similar to volatility NIC in Engle and Ng (1993) and skewness NIC in Anatolyev and Petukhov (2016). Let m_+, m_- be some nonnegative integers; $\{\tau_i\}_{i=-m_-}^{m_+}$ be a set of real numbers satisfying $\tau_{-m_-} < \tau_{(-m_-+1)} < \cdots < \tau_{(m_+-1)} < \tau_{m_+}$. Define the dynamics by

$$\psi_{t-1} = \sum_{j=0}^{m_+} \alpha_{j,+}(r_{t-1} - \tau_j)^+ + \sum_{j=0}^{m_-} \alpha_{j,-}(r_{t-1} - \tau_{-j})^-, \tag{1}$$

and α_0, α_{-1}, $\alpha_{j,+}$ $j = 0, 1, \ldots, m_+$, $\alpha_{j,-}$, $j = 0, 1, \ldots, m_-$ are parameters to be estimated; $x^+ = \max(0, x)$, $x^- = \min(0, x)$. For $r_{t-1} \in (\tau_0, \tau_1]$, this function has slope $\alpha_{0,+}$, for $r_{t-1} \in (\tau_1, \tau_2]$, this function has slope $(\alpha_{0,+} + \alpha_{1,+})$, ..., for $r_{t-1} \in (\tau_j, \tau_{j+1}]$, $j \geq 0$, this function has slope $\sum_{i=0}^{j} \alpha_{i,+}$, for $r_{t-1} \in (\tau_{-(j+1)}, \tau_{-j}]$, $j \geq 0$, this function has slope $\sum_{i=0}^{j} \alpha_{i,-}$. The parameters m_+, m_-, $\{\tau_j\}_{j=-m_-}^{m_+}$ are usually chosen by a researcher, although some automated algorithms can be used. Higher values of m_+ and m_- lead to higher flexibility of the model, but also to less precise parameter estimation. We describe in the empirical section how we choose the knots $\{\tau_j\}_{j=-m_-}^{m_+}$. We set $m_- = m_+ = m$, and call (1) the *piecewise M-knot linear* model, where $M = 2m + 1$.

Semiparametric (Fourier flexible form) DNIC Another semiparametric approach is based on approximating the DNIC by Fourier sine and cosine waves (Gallant 1981):[4]

$$\psi_{t-1} = \sum_{j=1}^{m} \alpha_{j\,\sin} \sin\left(2\pi j \frac{r_{t-1}}{R}\right) + \sum_{j=1}^{m} \alpha_{j\,\cos} \cos\left(2\pi j \frac{r_{t-1}}{R}\right), \tag{2}$$

where $R = \max_t |r_t|$. While the piecewise linear model provides a non-differentiable approximation to the DNIC, the Fourier flexible form model generates a smooth approximation. A shortcoming of this approach is spurious swings in the DNIC shape, especially towards the boundaries of the empirical distribution of returns, found even for small values of m such as $m = 2$. We call (2) the *Fourier of order m* model.

3 Data and performance criteria

The DNIC is hard to identify from the data, therefore we use a very long series in order to make reliable conclusions. Namely, we use daily S&P500 returns from 03.01.1950 to 08.08.2019, totalling to 17512 returns. The data are downloaded from `finance.yahoo.com`. We use the first $T = 5000$ returns for the in-sample analysis, and the rest $P = 12512$ for the out-of-sample analysis with a rolling window of $T = 5000$. The figure 5000 is pretty arbitrary, but large enough to make sure the dependencies are captured with sufficiently large precision, while the rolling window scheme leaves a possibility of long-run changes in parameters. Another measure taken because of weak identifiability is that we include, in the dynamic part of the models, either only the first lag(s) alone (i.e., $\phi_{t-2} = 0$) or together with the second lags of the indicator and return (i.e., ϕ_{t-2} has the same form as ψ_{t-1} in the 'indicator and return' benchmark but with an additional lag). In addition to S&P500 returns, we also experiment with DAX and Nikkei index return data, from 30.12.1987 to 08.08.2019 ($T + P = 7982$ returns, rolling window of $T = 3000$) for DAX, and from 05.01.1965 to 08.08.2019 ($T + P = 13436$ returns, rolling window of $T = 5000$) for Nikkei.

We consider five thresholds c to investigate directional predictions with different amount

[4] An alternative set of basis functions may be (Hermite) polynomials. However, even inclusion of second powers of returns makes estimation less stable. An advantage of Fourier series is boundedness of all its basis functions.

of drift. We set these to be quantiles of the symmetrized distribution of returns, that is, of the empirical distribution of $|r_t|$, $t = 1, ..., T + P$. The leading case is the pure directional predictions, for which we set $c_{50\%}$ to be the leftmost point of this distribution. The other cases reflect (not) exceedances of returns of non-zero thresholds. We let $c_{75\%}$ be the 50%-empirical quantile and set $c_{25\%} = -c_{75\%}$; similarly, we let $c_{90\%}$ be the 80%-empirical quantile and set $c_{10\%} = -c_{90\%}$. Thus, $c_{50\%}$ is approximately the median, $c_{25\%}$ and $c_{75\%}$ are approximately the 25% and 75% unconditional quantiles of returns, and $c_{10\%}$ and $c_{90\%}$ are approximately their 10% and 90% unconditional quantiles. We call $c_{50\%}$ 'quantile 50%' or 'median,' call $c_{25\%}$ and $c_{75\%}$ 'quantile 25%' and 'quantile 75%,' respectively, and call $c_{10\%}$ and $c_{90\%}$ 'quantile 10%' and 'quantile 90%,' respectively. Of most interest is the DNIC for pure directional predictions, but DNICs for the other thresholds may also be important to, say, a trader who tries to time the market with low and high, respectively, transaction costs.

We judge the performance of different models by several criteria. For the in-sample analysis, we first look at t-ratios for the predictors in parametric models estimated using T observations, which serve as an approximate measure of importance/strength of particular predictors. Second, we use the information criteria (IC) for parametric models, AIC and BIC, based on T observations:

$$IC_T = -2T\bar{\ell}_T + w_T \|\alpha\|_0 ,$$

where $\bar{\ell}_T$ is in-sample average loglikelihood, α is a whole parameter vector, and where $w_T = 2$ for AIC and $w_T = \ln T$ for BIC. We use IC to select among the two specifications within one model (i.e. decide on whether only first lag(s) or also second lag(s) of basic predictors should be in) and to compare performance across different models. A better fit model according to a particular IC has a smaller value of IC_T.

The out-of-sample criteria based on P one-step-ahead forecasts and realizations, are the log-probability score

$$S_P = \frac{1}{P} \sum_{t=T+1}^{T+P} I_t \ln\left(\hat{p}_{t|t-1}\right) + (1 - I_t) \ln\left(1 - \hat{p}_{t|t-1}\right) ,$$

and the absolute Brier score

$$B_P = \frac{1}{P} \sum_{t=T+1}^{T+P} \left|I_t - \hat{p}_{t|t-1}\right| ,$$

where $\hat{p}_{t|t-1} = \left(1 + \exp(-\hat{\theta}_t)\right)^{-1}$ is the predicted probability of the event $r_t \leq c$ made at $t-1$. While the score S_P is larger for a better predictive model, the criterion B_P is smaller for a better predictive model. We apply the out-of-sample criteria to both parametric and semiparametric models.

4 Results

In this Section, we describe mostly the results obtained for the SP500 index. We also devote the last subsection to discussing their differences with the results for the DAX and Nikkei indices. All estimation and evaluation procedures were written in GAUSS (by Aptech Systems, Inc). For numerical optimization, the `cml` (constrained maximum likelihood) library and the BHHH algorithm (Berndt, Hall, Hall and Hausman 1974) were employed.

4.1 In-sample fit of parametric models

We compare the following parametric models: the 'only return' benchmark, the 'only indicator' benchmark, the 'return and indicator' benchmark, the 'squared return and indicator' model, the 'absolute return and indicator' model, the 'signed absolute return and indicator' model, the 'log-volatility and indicator' model, and the 'autoregressive score' model.

Tables 1 and 2 report median (across $P = 12512$ rolling samples of size $T = 5000$) t-ratios for predictor coefficient estimates. The comparison of top two panels of Table 1 eloquently leads to a conclusion that, except far in the tails, both past indicator and past return are important predictors when used solely on their own; however, the past indicator is much stronger. Far in the tails, it is the only strong predictor between the two. When they are used together, as the bottom panel shows, the past indicator dominates and encompasses the past return, the latter being practically useless. An attempt to make the DNIC time varying by using the autoregressive score model does not seem to produce a stronger predictor out of the past indicator, as the third panel indicates.

Table 2 contains median t-ratios in more complex parametric models. A general observation is that for some quantiles, though not all, a judiciously chosen second (+third) predictor, given the past indicator, is able to pull out statistical significance from the past indicator and become even stronger. Some of such predictors do this job better in the tails, some in the middle range. For example, the absolute return and signed absolute return are especially effective far in the tails, while in the center and in non-extreme tails, the log-volatility index used *ad hoc*-ly in Anatolyev and Baruník (2019) turns out to be very effective.

Figure 1 presents ranking of parametric models by median (across $P = 12512$ rolling samples of size $T = 5000$) values of information criteria AIC and BIC, for the five thresholds, which are measured relative to those for the static model $\theta_t = \omega$. As a reference point, the conventional α-level statistically significant difference between two nested models where the larger model has one more parameter occurs when the relative IC exceeds $|q_{1-\alpha}^{\chi^2(1)} - w_T|$, which for $\alpha = 5\%$ is approximately 2 for AIC and 4.5 for BIC. This implies that all deviations on the graph noticeable to a naked eye are substantial on the conventional significance scale.

One immediately notices high heterogeneity across the quantiles regarding which parametric model is optimal, as well as regarding the maximal deviation of the criterion from its trivial counterpart, which is relatively small for the median quantile but is relatively large for the rightmost quantile when AIC is considered; in terms of BIC, however, it is smallest for the leftmost quantile but large towards the center. There does not seem to be one model that is optimal for all parts of the distribution, be it in terms of AIC or BIC. At the same time, each parametric model is optimal or near optimal at least for one of quantiles considered, in terms of either AIC or BIC. The time-varying 'autoregressive score' model with the smart dynamics is near optimal only for the median quantile. Interestingly, for non-extreme quantiles, the 'only return' benchmark is preferred by IC to the 'only indicator' benchmark, despite the past indicator seems to be a stronger predictor than the past return if one makes judgements from median t-ratios (see the discussion above).

4.2 Out-of-sample predictability

Figure 2 depicts the two average out-of-sample criteria computed over $P = 12512$ forecast periods – the log-probability score and Brier score, for the same set of parametric models. These criteria are also measured relative to the static model $\theta_t = \omega$. In addition to the usual heterogeneity of model performances across the quantiles, one notices a much smaller dispersion for the extreme quantiles in the case of log-probability score and for the extreme and median quantiles in the case of Brier score. Note that for the 75% quantile, no parametric model among those under consideration is better than the static model in terms of either criterion. At the same time, for the whole left side of the return distribution, all parametric models (except the 'only return' benchmark) are no worse than the static model. For these quantiles, the 'only indicator' benchmark, the 'autoregressive score' model, and the 'squared return and indicator' model fair evenly and are best performing in terms of both the log-probability score and Brier score. For the rightmost quantile, however, the role of the best performing is taken by the models that use past indicator and the signed absolute value of past return or past volatility index, albeit these models fare better than others by a narrow margin. An interesting observation is that the 'only return' benchmark model is almost uniformly worse than the static model by both criteria.

Next, we have run the two semiparametric models – the piecewise linear model with $m = 0$ (i.e. $M = 1$ knot), with $m = 1$ (i.e. $M = 3$ knots) and with $m = 2$ (i.e. $M = 5$ knots), and the Fourier flexible form with $m = 1$ (i.e. of order 1) and $m = 2$ (i.e. of order 2). Figure 3 depicts the same out-of-sample criteria for these five semiparametric models relative to the static model, together with those for the quantile-specific and score-specific best faring parametric models revealed by Figure 2. Interestingly, in terms of both out-of-sample likelihood and Brier score, the best parametric model fares (weakly) better for the quantiles in the left side of the return distribution as well as for the extreme right quantile, while for the 75% quantile there are low-order (and thus not too parameterized) semiparametric models that outperform the best parametric model (which, for this particular quantile, is the static model).

4.3 Form of directional NIC

Now we investigate the form of the directional NIC. Figures 4–9 present the DNIC described by the most flexible semiparametric models – the 'piecewise 5-knot linear' model (in blue) and the 'Fourier of order 2' model (in orange). For Figures 4–6, the models are estimated within the first 5000-periods window, for Figures 7–9 – within the last 5000-periods window; Figures 4 and 7 present DNICs for purely directional predictions, Figures 5 and 8 – for intermediate quantiles, and Figures 6 and 9 – for extreme quantiles.

In most of the graphs, the two semiparametric methods agree on the DNIC shapes, though at times the Fourier flexible form shows spurious curls and loops. There is certain temporal non-stability of parameters, but the general lineaments are similar in the early and late periods: for the left quantiles, the DNICs are negative, asymmetric and possibly shifted V-shaped, while for the right quantiles, the DNICs are positive, asymmetric and possibly shifted inverted V-shaped. For the left quantiles, larger past returns, whether positive or negative, increase the probability of this period's return not to exceed a negative threshold, while for the right quantiles, they decrease this probability. These V-shapes are a reflection of (an extension of) the phenomenon documented by Christoffersen and Diebold (2006) who showed that volatility movements are able to generate directional predictability even in the absence of mean predictability. If one treats past absolute returns as a sort of volatility indicator, larger past absolute returns associated with higher volatility must bring higher probability of a return being smaller than a negative threshold and a smaller probability for a positive threshold. However, the slopes of DNIC for the left quantiles witness a sort of a leverage effect to directional predictability: the effect of past returns on the probability of this period's return to not exceed a negative threshold is larger for negative past returns than for positive past returns; in the right quantile, the evidence of the leverage effect is similar but more blurry and more prone to temporal changes.

For purely directional predictions (i.e., with the 50% threshold), the DNIC shape is most sophisticated and was changing over time most of all – from counterclockwise rotated inverted S-shaped to skewed and shifted W-shaped. On average, both DNICs are negative biased, which reflects a higher unconditional rate of positive returns, but this rate evidently has grown over the years. In earlier days, negative/positive past returns were leading to a probability of positive returns lower/higher than $\frac{1}{2}$, this relationship exhibiting decreasing returns to scale and not even being monotonic. In the latest decades, however, both negative and positive past returns have been leading to a probability of positive returns higher than $\frac{1}{2}$, just with a different intensity, negative past returns having a much higher impact.

Finally, it may be interesting to see if the parametric models capture the sophisticated shapes of the directional NIC that the semiparametric model uncover. We focus only on the early estimation period. Figures 10–12 show a parametric DNIC from the most curvy 'signed absolute return and indicator' model (in blue) and the 'log-volatility and indicator' model (in orange). The parametric DNICs do resemble the most adequate semiparametric DNIC – the one that comes out of the 'piecewise 5-knot linear' model – but miss some important features such as the leverage effect for the extreme quantiles and a smooth transition for small past returns for purely directional predictions. In spite of these facts, as we have seen before, these simple and tight parametric specifications often perform better in forecasting the directions than the 'more correct' flexible semiparametric specifications.

4.4 Other indices

In addition to S&P500 returns, we have also experimented with uncovering the DNIC for DAX and Nikkei index returns, for the median quantile. Perhaps surprisingly, the results for these two indices are strikingly more trivial. For DAX, the values of information criteria are practically indistinguishable from their values for the static model; for Nikkei, while the AIC prefers some models (with a past indicator coupled with a past squared return or past signed absolute return) to others, the BIC declines all parametric models in favor of the static model. The median t-ratios, although exhibiting similar patterns, are much smaller in value; none of them are statistically significant at conventional levels, the maximal (in absolute value) values barely reach 1.4 for DAX and 1.8 for Nikkei. Out-of-sample, for none of the parametric or semiparametric models the log-probability score or Brier score are better than those for the static model.

Figures 13–14 show the DNIC with the 'piecewise 5-knot linear' model (in blue) and the 'Fourier of order 2' model (in orange), for earlier estimation periods for DAX and Nikkei, respectively. If they are taken at face value despite very weak identifiability, the shapes are even more complex than in case of S&P500; however, the range of the impact is less than half of that for S&P500: approximately $[-0.3, 0.1]$ versus $[-0.7, 0.3]$. Such dissimilarity across different indices, as well as weaker identifiability for indices other than S&P500, reveals itself in the skewness NIC as well (see Anatolyev and Petukhov 2016).

5 Concluding remarks

We have analyzed the directional news impact curve – the relationship between returns and next period's probability of stock prices jumping by less or more than a certain threshold, zero in particular, – using a long (several decades) S&P500 and other index return data. We have used a number of simple parametric models suggested in the literature and that involve past indicators and/or various functions of past returns as driving processes in the specification for log-odds ratio, as well and more flexible semiparametric models such as the piecewise linear model and Fourier flexible form.

The semiparametric models reveal that the DNIC has complicated shapes characterized by asymmetry with respect to past returns and in particular their signs. There is a lot of heterogeneity across the quantiles considered, or, in other words, across the thresholds which the price needs to exceed in the next period. There is also evidence that the DNICs have been subject to changes in shapes and in their parameters during the decades. Simple parametric models, especially linear in past returns, often miss these features, but despite this fact, some of them, with the presence of past indicator being crucial, exhibit superior out-of-sample performance in terms of likelihood and/or Brier scores. These properties of DNIC for S&P500 returns are not shared by DNIC of other index returns, such as DAX and Nikkei, for which the dependence of conditional probabilities on the historical returns, at least for the directional predictions, turns out to be very weak.

References

Anatolyev, S. (2009) Multi-market direction-of-change modeling using dependence ratios. *Studies in Nonlinear Dynamics & Econometrics*, 13(1), article 5.

Anatolyev, S. and J. Baruník (2019) Forecasting dynamic return distributions based on ordered binary choice. *International Journal of Forecasting*, 35(3), 823–835.

Anatolyev, S. and A. Gerko (2005) A trading approach to testing for predictability. *Journal of Business & Economic Statistics*, 23(4), 455–461.

Anatolyev, S. and N. Gospodinov (2009) Modeling financial return dynamics via decomposition. *Journal of Business & Economic Statistics,* 28(2), 232–245.

Anatolyev, S. and N. Gospodinov (2019) Multivariate return decomposition: theory and implications. *Econometric Reviews*, 38(5), 487–508.

Anatolyev, S., N. Gospodinov, I. Jamali, and X. Liu (2017) Foreign exchange predictability and the carry trade: a decomposition approach. *Journal of Empirical Finance*, 42(C), 199–211.

Anatolyev, S. and A. Petukhov (2016) Uncovering the skewness news impact curve. *Journal of Financial Econometrics*, 14(4), 746–771.

Berndt, E., B. Hall, R. Hall, and J. Hausman (1974) Estimation and inference in nonlinear structural models. *Annals of Economic and Social Measurement*, 3(4), 653–665.

Bollerslev, T. (2010) Glossary to ARCH (GARCH*). Chapter in: *Volatility and Time Series Econometrics*, by T. Bollerslev, J, Russell and M. Watson (eds.), Oxford University Press.

Christoffersen, P.F. and F.X. Diebold (2006) Financial asset returns, direction-of-change forecasting, and volatility dynamics. *Management Science*, 52(8), 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(2), 3–24.

Creal, D. D., S. J. Koopman and A. Lucas (2013) Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, 28(5), 777–795.

Engle, R. F. and V. K. Ng (1993) Measuring and testing the impact of news on volatility. *The Journal of Finance*, 48(5), 1749–1778.

Foresi, S. and F. Peracchi (1995) The conditional distribution of excess returns: an empirical analysis. *Journal of the American Statistical Association*, 90(430), 451–466.

Frazier, D. T. and X. Liu (2016) A new approach to risk-return tradeoff dynamics via decomposition. *Journal of Economic Dynamics & Control*, 62(C), 43–55.

Gallant, A. R. (1981) On the bias in flexible functional forms and an essentially unbiased form. *Journal of Econometrics*, 15(2), 211–245.

Gençay, R. (1998) Optimization of technical trading strategies and the profitability in security markets. *Economics Letters*, 59(2), 249–254.

Hong, Y. and J. Chung (2003) Are the directions of stock price changes predictable? Statistical theory and evidence. Manuscript, Cornell University.

Linton, O. and Y. J. Whang (2007) The quantilogram: with an application to evaluating directional predictability. *Journal of Econometrics*, 141(1), 250–282.

Liu, X. and R. Luger (2015) Unfolded GARCH models. *Journal of Economic Dynamics and Control*, 58(C), 186–217.

Nyberg, H. (2013) Predicting bear and bull stock markets with dynamic binary time series models. *Journal of Banking & Finance*, 37(9), 3351–3363.

Nyberg, H. (2014) A bivariate autoregressive probit model: business cycle linkages and transmission of recession probabilities. *Macroeconomic Dynamics*, 18(4), 838–862.

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

Saikkonen, P. and H. Kauppi (2008) Predicting U.S. recessions with dynamic binary response models. *Review of Economics and Statistics*, 90(4), 777–791.

Skabar, A. (2013) Direction-of-change financial time series forecasting using a similarity-based classification model. *Journal of Forecasting*, 32(5), 409–422.

Taylor, J. W. and K. Yu (2016) Using auto-regressive logit models to forecast the exceedance probability for financial risk management. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 179(4), 1069–1092.

Tables and Figures

Table 1. The median t-statistics on return and indicator variables for the simple benchmarks and AS model.

		median t-statistics	
Model		on indicator	on return
	10%		−1.5
	25%		−3.1
only return	50%		−2.3
	75%		−2.2
	90%		1.0
	10%	4.3	
	25%	4.0	
only indicator	50%	3.4	
	75%	3.0	
	90%	2.3	
	10%	4.2	
	25%	4.0	
only scaled indicator	50%	3.4	
	75%	3.0	
	90%	0.9	
	10%	2.3	−1.0
	25%	3.1	−1.6
indicator and return	50%	2.2	−0.5
	75%	1.9	−0.4
	90%	3.4	1.7

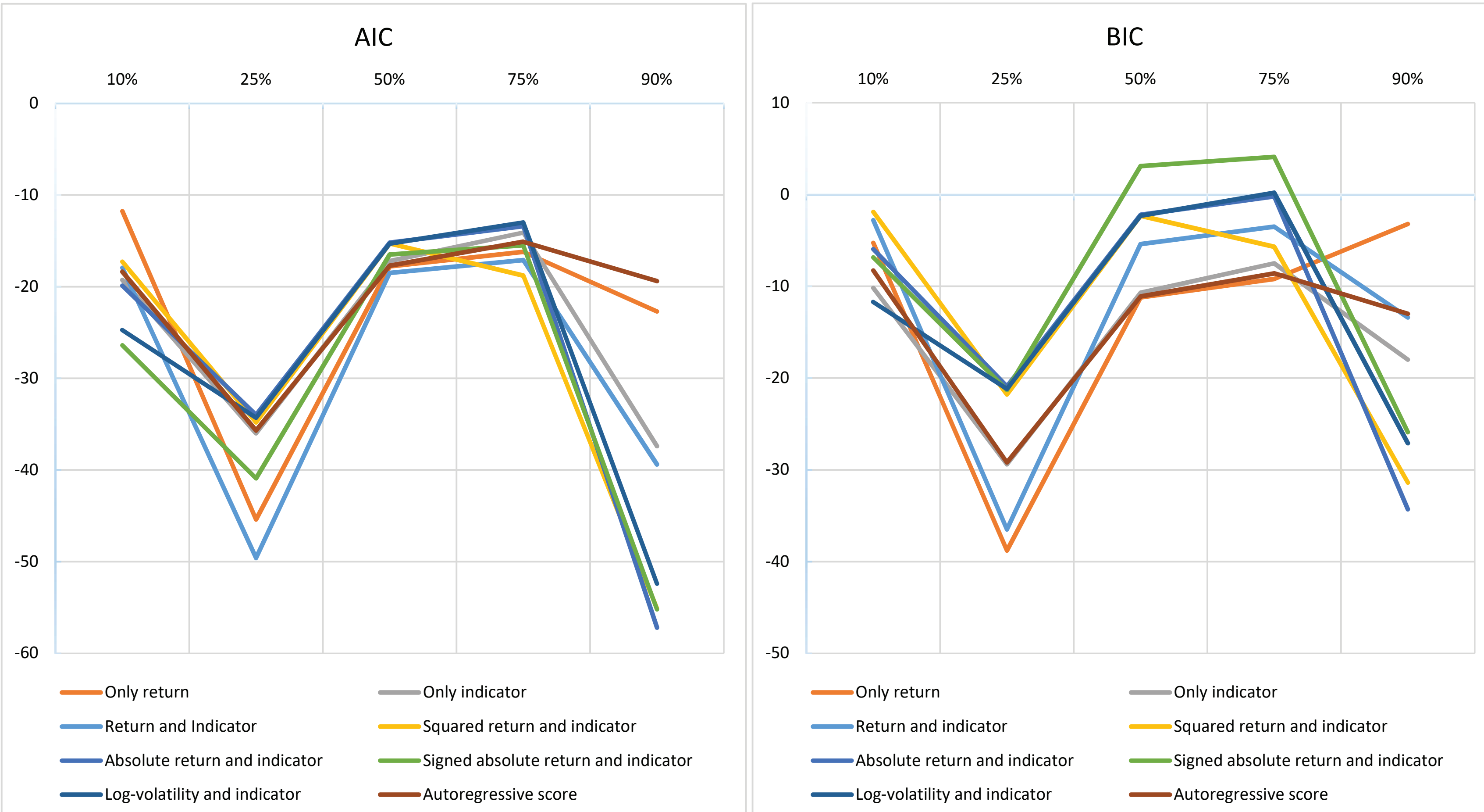
Notes: The table reports median t-statistics for select benchmarks computed over $P = 12512$ rolling samples of size $T = 5000$ of the SP500 index returns.

Table 2. The median t-statistics on return and indicator variables for the complex benchmarks.

		median t-statistics		
Model		on indicator	on other variables	
	10%	3.3	0.9	
	25%	4.1	−0.4	
indicator and return squared	50%	3.4	−0.4	
	75%	2.7	−2.6	
	90%	0.5	−4.0	
	10%	1.9	2.4	
	25%	4.0	−0.0	
indicator and absolute return	50%	3.4	−0.1	
	75%	2.5	−3.2	
	90%	−0.5	−6.4	
	10%	2.8	0.2	4.2
	25%	2.7	0.5	− 1.0
indicator and signed absolute return	50%	2.2	−0.0	− 1.1
	75%	0.6	−2.4	− 1.4
	90%	0.7	−6.9	− 1.1
	10%	−1.0	1.8	
	25%	−1.6	4.0	
indicator and log-volatility	50%	−0.5	3.4	
	75%	−0.4	2.9	
	90%	1.7	−0.4	

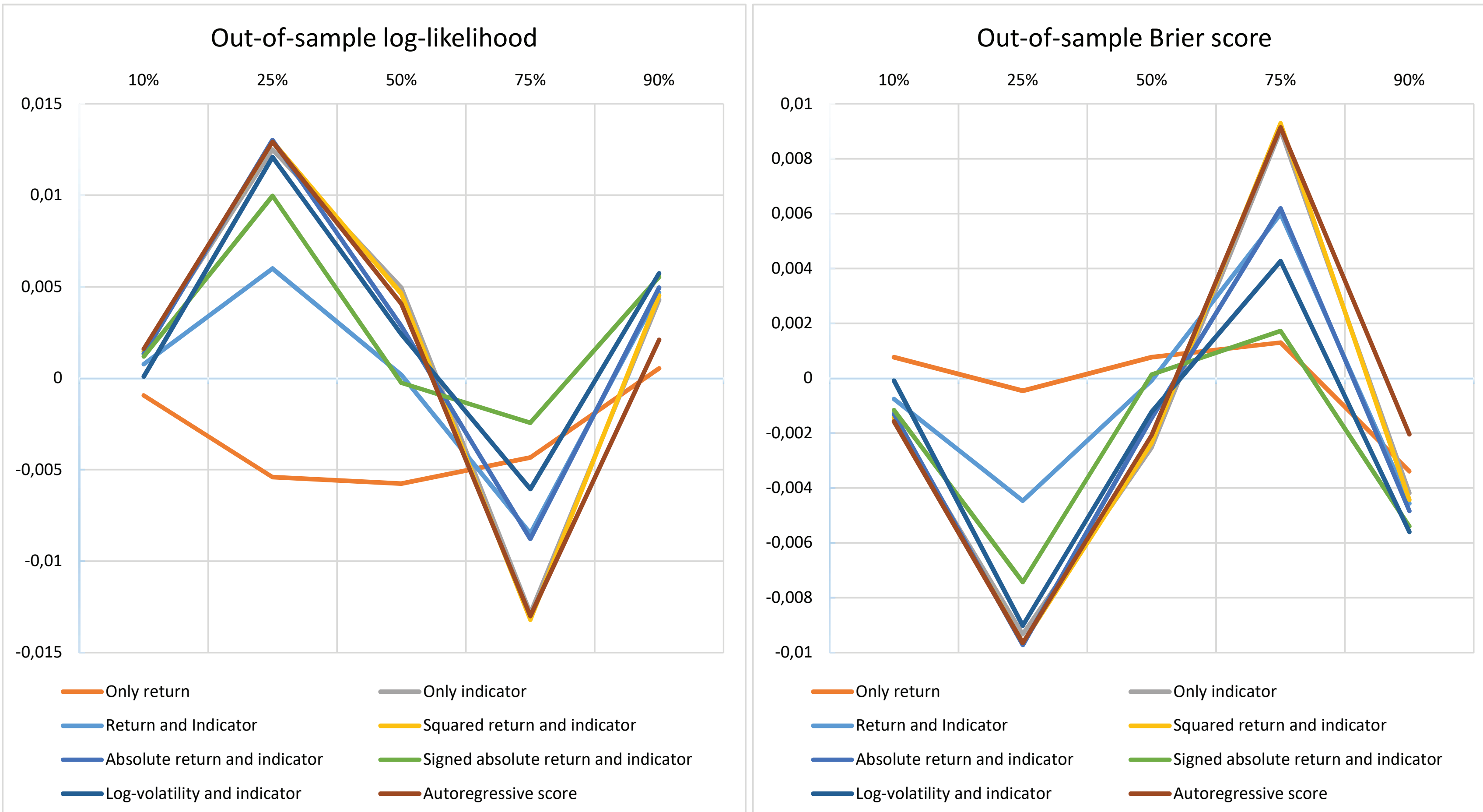
Notes: The table reports median t-statistics for select benchmarks computed over $P = 12512$ rolling samples of size $T = 5000$ of the SP500 index returns.

Figure 1: In-sample information criteria for parametric models



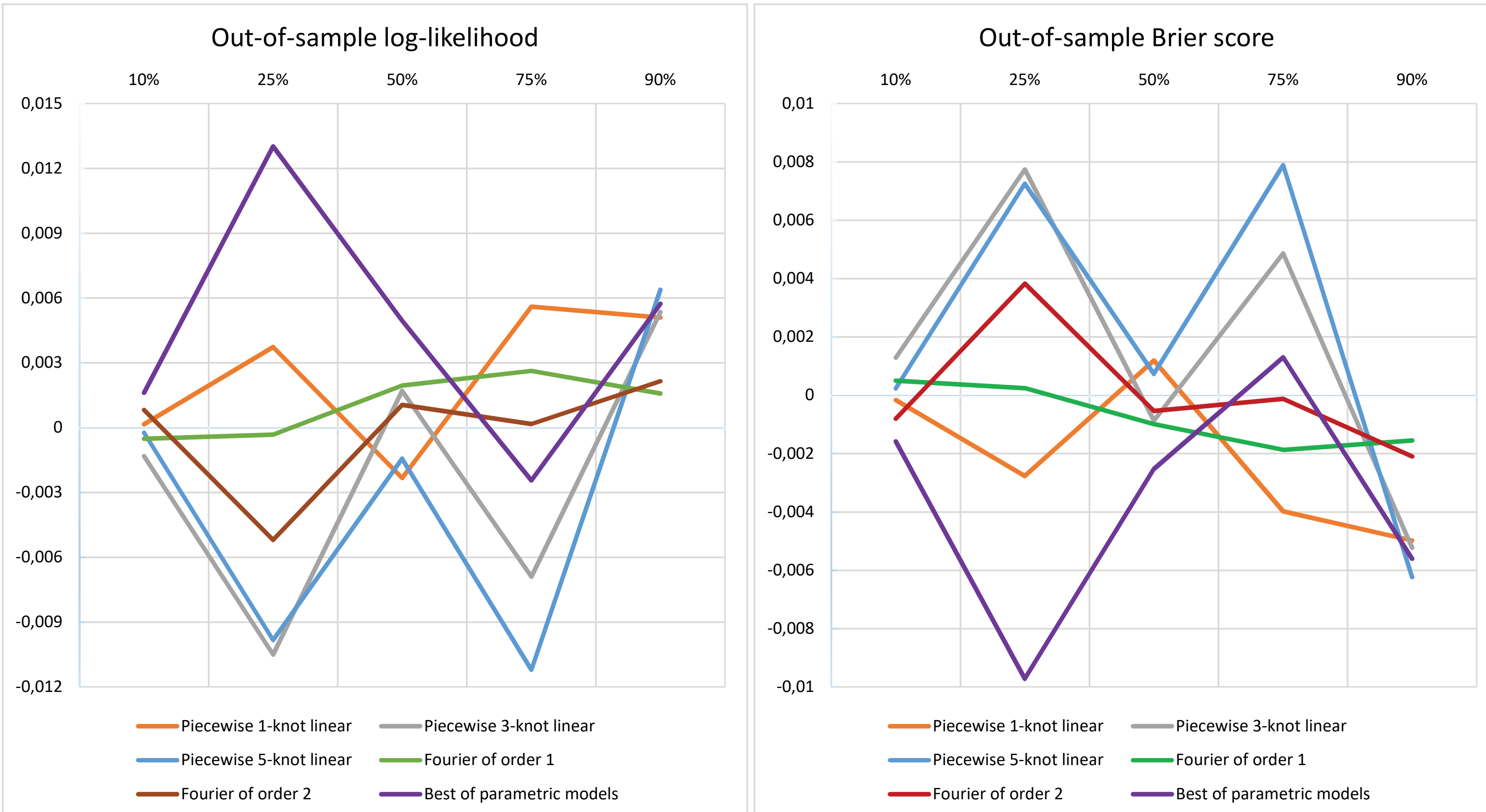
Notes to Figure 1: The figure depicts median AIC and BIC values for parametric DNIC models computed over $P = 12512$ rolling samples of size $T = 5000$ of the S&P500 index returns.

Figure 2: Out-of-sample probability score and Brier criteria for parametric models



Notes to Figure 2: The figure depicts average log-probability score and Brier scores for parametric DNIC models computed over $P = 12512$ forecast periods, for the S&P500 index.

Figure 3: Out-of-sample probability score and Brier criteria for semiparametric and best parametric models



Notes to Figure 3: The figure depicts average log-probability score and Brier scores for select DNIC models computed over $P = 12512$ forecast periods, for the S&P500 index.

Figure 4: Directional news impact curves, early estimation period, 50% threshold

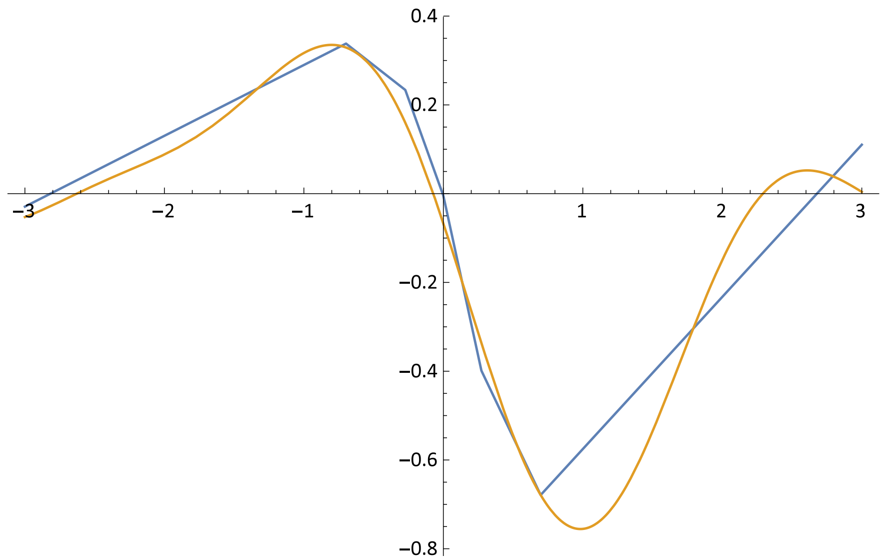


Figure 5: Directional impact news curve, early estimation period, 25% and 75% thresholds

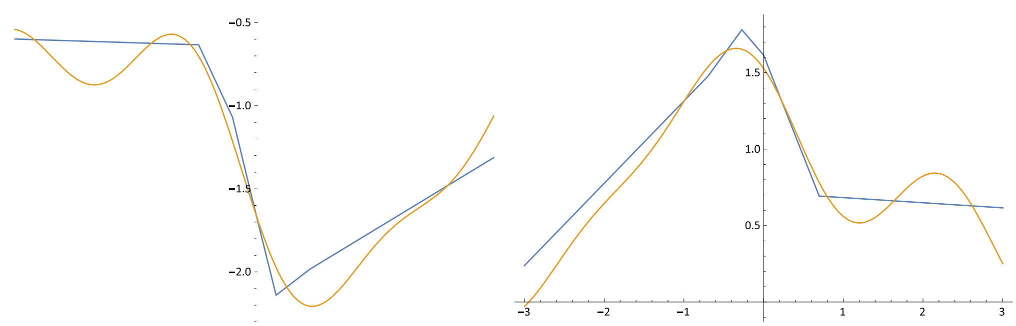
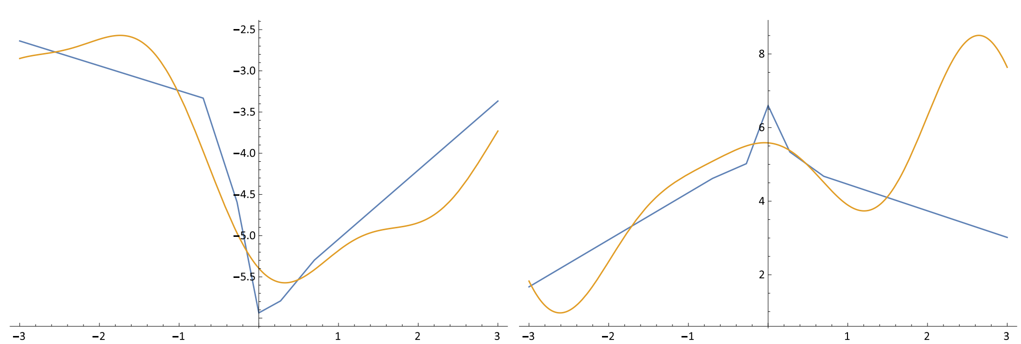


Figure 6: Directional impact news curve, early estimation period, 10% and 90% threshold



Notes to Figures 4–6: The figures depict DNIC for the SP500 index computed for $T = 5000$ returns in the period 03.01.1950–31.12.1969 with the five-knot piecewise linear model (in blue) and the order-two Fourier flexible form (in orange). Horizontal axis: lagged values of return r_{t-1}, vertical axis: values of DNIC.

Figure 7: Directional news impact curves, late estimation period, 50% threshold

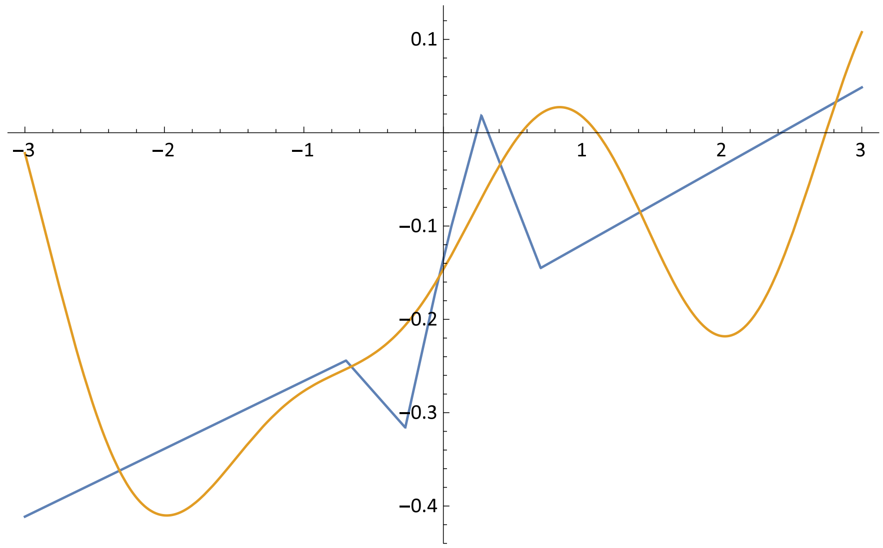


Figure 8: Directional impact news curve, late estimation period, 25% and 75% thresholds

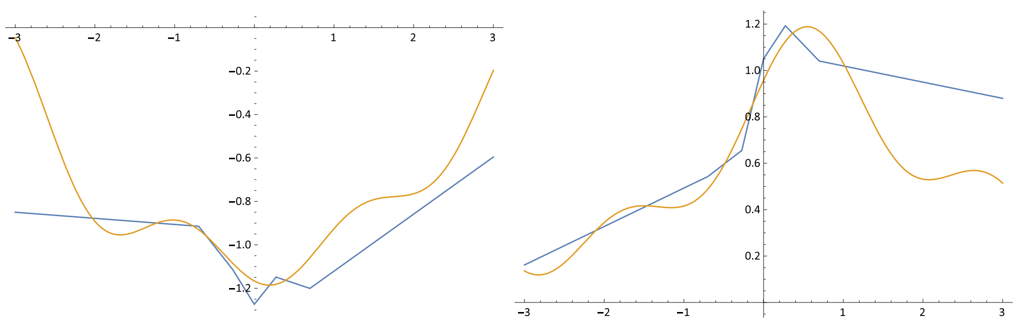
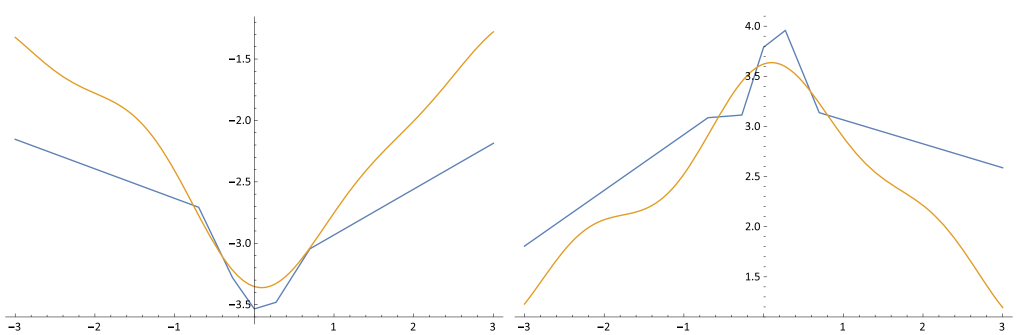


Figure 9: Directional impact news curve, late estimation period, 10% and 90% threshold



Notes to Figures 7–9: The figures depict DNIC for the SP500 index computed for $T = 5000$ returns in the period 23.09.1999–08.08.2019 with the five-knot piecewise linear model (in blue) and the order-two Fourier flexible form (in orange). Horizontal axis: lagged values of return r_{t-1}, vertical axis: values of DNIC.

Figure 10: Parametric directional NIC, early estimation period, 50% threshold

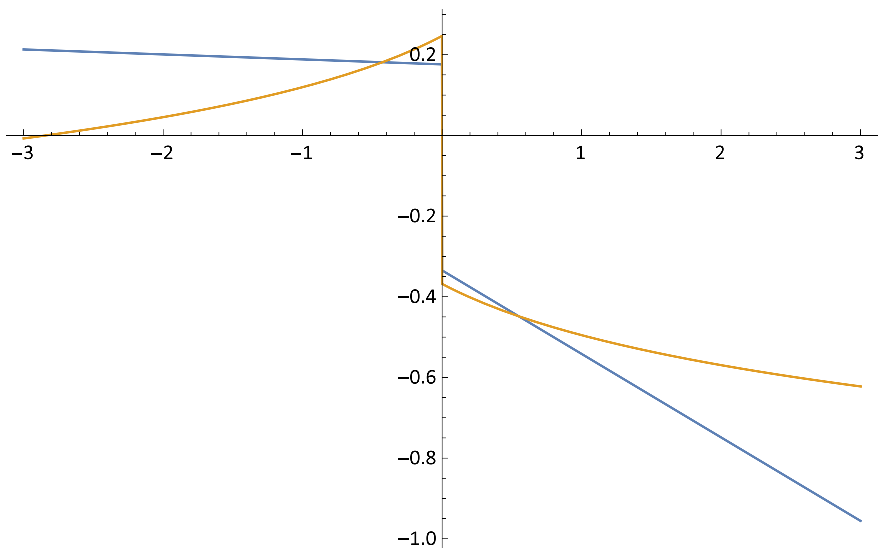


Figure 11: Parametric directional NIC, early estimation period, 25% and 75% thresholds

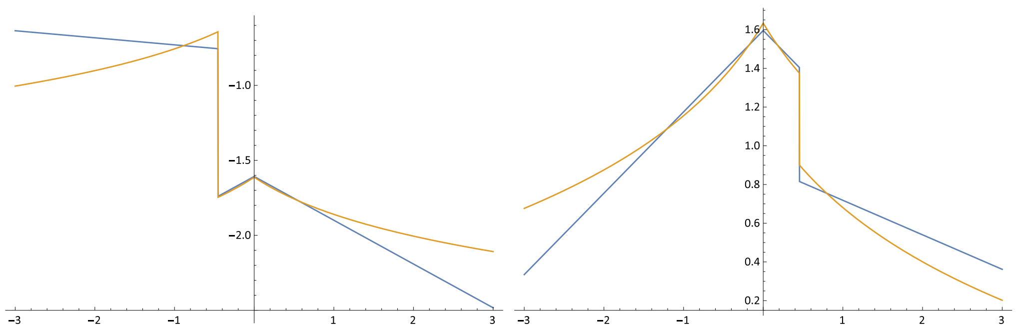
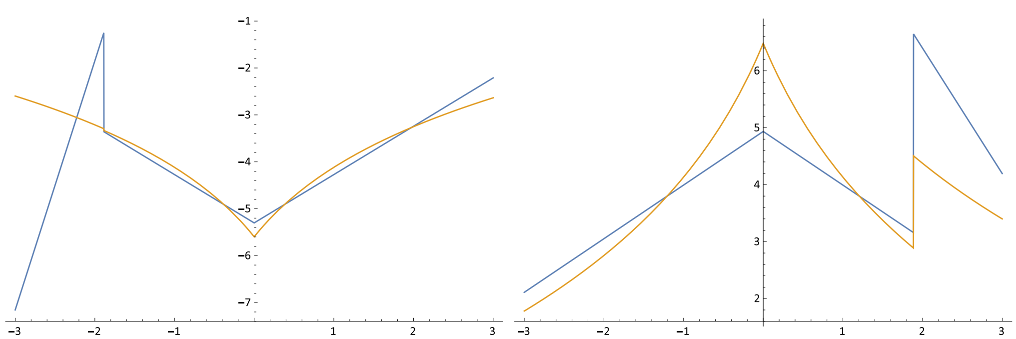
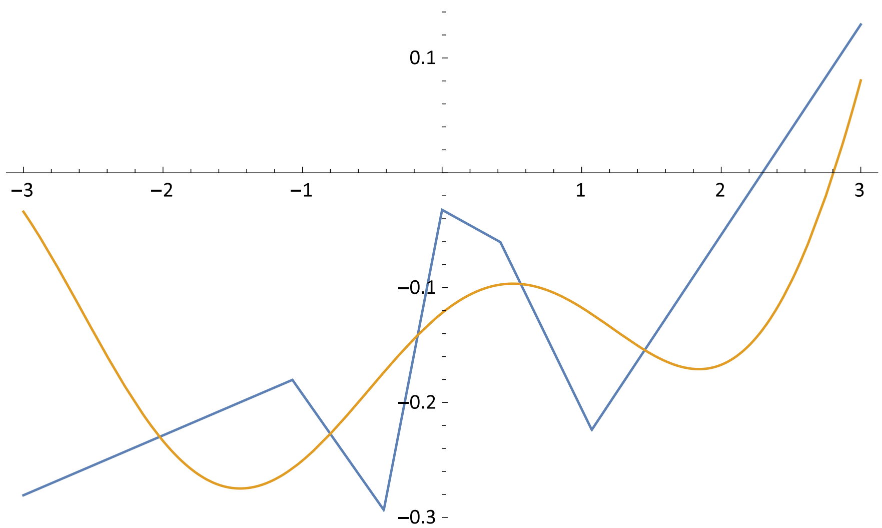


Figure 12: Parametric directional NIC, early estimation period, 10% and 90% threshold



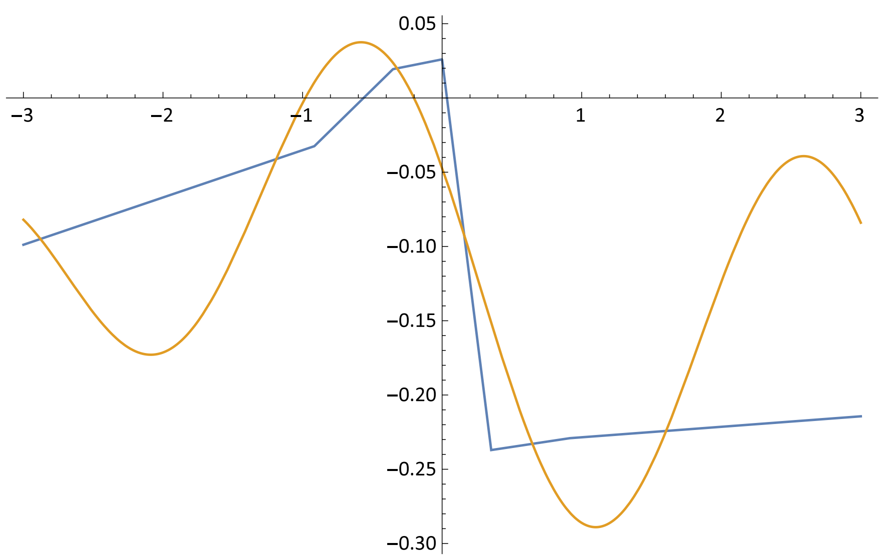
Notes to Figures 10–12: The figures depict DNIC for the SP500 index computed for $T = 5000$ returns in the period 03.01.1950–31.12.1969 with the past signed absolute return and past indicator model (in blue) and the model with a past log-volatility index and past indicator (in orange). Horizontal axis: lagged values of return r_{t-1}, vertical axis: values of DNIC.

Figure 13: Directional news impact curve for DAX index, early estimation period, 50% threshold



Notes to Figure 13: The figure depicts DNIC for the DAX index computed for $T = 3000$ returns in the period 30.12.1987–30.06.1999 with the five-knot piecewise linear model (in blue) and the order-two Fourier flexible form (in orange). Horizontal axis: lagged values of return r_{t-1}, vertical axis: values of DNIC.

Figure 14: Directional news impact curves for Nikkei index, early estimation period, 50% threshold



Notes to Figure 14: The figure depicts DNIC for the Nikkei index computed for $T = 5000$ returns in the period 05.01.1965–06.03.1984 with the five-knot piecewise linear model (in blue) and the order-two Fourier flexible form (in orange). Horizontal axis: lagged values of return r_{t-1}, vertical axis: values of DNIC.