Stock return distribution and predictability:

Evidence from over a century of daily data on the DJIA Index

Abstract

This paper analyses the predictive power of the DJIA index returns, measured at different quantiles of its distribution, for future return distribution. The returns measured at quantile .75 have predictive power for most quantiles of future returns, except for their median. This result prevails after controlling for the predictive power of the lagged first four moments of returns and of other economic predictors used in the literature. Furthermore, this finding is stable over time. Forecasts of future mean returns based on predicted return quantiles have positive economic value, as do forecasts of future volatility, the latter especially for investors with low risk aversion. The predictive power of quantile .75 DJIA returns is shown to be the result of their ability to forecast shocks to future investment and consumption.

JEL classification: C21, C53, G11, G17

Keywords: Return predictability; DJIA index; distribution; forecasting.

1. Introduction

This paper analyses the predictive power of the DJIA stock index returns measured at different quantiles of their distribution for the future index return distribution. Traditionally, the extant literature has focused on directly predicting the center of the future return distribution, i.e., the future mean return. Different economic variables have been proposed (those reviewed by Welch and Goyal, 2008, and additional ones, e.g., Bollerslev, Tauchen, and Zhou, 2009; Kelly and Jiang, 2014) and have been found to possess predictive power for future returns in-sample, across markets, and up to four centuries of data (Golez and Koudijs, 2014). However, their predictive power out-of-sample has been questioned (e.g., Welch and Goyal, 2008), especially when choosing among only those predictive variables that would have been known to a hypothetical investor at the time (Turner, 2015). On the other hand, Cochrane (2008) points out that the out-of-sample tests can underestimate the predictive power of variables. Hence, the traditional variables as well as the new ones could generate valuable forecasts of future mean returns after all, especially when theory-derived restrictions on estimates and forecasts are imposed (Campbell and Thompson, 2008), market volatility is high (Marquering and Verbeek, 2004), nonlinearities in the predictive relationship are allowed for (Guidolin et al., 2014), commodities returns are utilised as predictors (Jordan et al., 2016) or innovative predictive approaches are used (e.g., Ferreira and Santa Clara, 2011, or others, as reviewed in Rapach and Zhou, 2013). Predictability is reported to have been stronger until the late 1970s to early 1980s (Marquering and Verbeek, 2004, Welch and Goyal, 2008, Ferreira and Santa Clara, 2011). Furthermore, whereas most of the traditionally employed predictors can be labelled as macroeconomic variables, Neely et al. (2014) point out that valuable information about the future index returns can be found in the historical behaviour of stock indices themselves, which they propose to extract using well-known technical trading rules such as the moving average rule.

In addition, a related branch of the literature focuses on forecasting the left tail of the return distribution, rather than the mean, as it is of relevance for the calculation of value at risk (VaR) measures. Engle and Manganelli (2004) propose direct dynamic quantile regression for calculating VaR, termed CAViaR, whereas Gerlach, Chen, and Chan (2011) propose a family of nonlinear CAViaR models (see, e.g., Chen et al., 2012, for a review of the literature, and Şener, Baronyan, and Mengütürk, 2012, on a comparative study of predictive performance of VaR estimators).

Yet another related branch of literature investigates the predictive power of observed economic variables for the entire future return distribution, as approximated by a set of different quantiles. Ma and Pohlman (2008) show that in-sample, different financial valuation factors can explain different quantiles of future return distribution. Pedersen (2015) also reports the in-sample and out-of sample predictive power of economic variables for stock and bond return distributions, finding that different variables predict different quantiles of future return distribution, most frequently in the tails and least strongly in the center. Zhu (2013) combines quantile regressions and the copula approach and also finds in-sample predictability of bond and stock returns using economic variables, with the predictive power of economic variables being heterogeneous across quantiles of the future return distribution. Cenesizoglou and Timmermann (2008) report predictive power for economic variables for future stock returns, too, especially in the right tail but not in the distribution's center. The same result of predictive power which is heterogeneous over future quantiles is obtained by Meligkotsidou et al. (2014), who also propose two approaches of forecasting future mean by combining individual variables' predictions—combining quantiles predicted by each variable first, and combining those predicted values across variables second, or combining predictions for each quantile separately (across all predictors) first, and combining predicted quantiles next, all with constant or time-varying endogenously optimised weights.

In this paper, we extend the earlier literature by adopting a complementary approach and investigate whether information valuable for the forecasting of the future distribution can be found in the whole distribution, rather than just the mean, of past DJIA index returns. In addition, we analyse the predictive power of past return distribution not only for the future mean return, but also for other quantiles of the future return distribution.[1] The ability to predict the distribution would be valuable for those interested in its specific parts, e.g., the center or the left-hand tail as used in VaR analysis. In addition, even if the center of the distribution (e.g., the mean or the median) is not directly predictable, if we could predict some points/quantiles of the future distribution, we could try to infer, e.g., other points of that distribution, some features of that distribution such as its variance, or the functional form of that distribution. This would yield potentially improved forecasts of the directly unpredictable points of the distribution of future returns, e.g., the mean or the median future return, even if those were not predictable directly.

Previous research finds that certain features of the lagged distribution, such as volatility, skewness, or kurtosis, can be used to predict certain features of the future return distribution.[2] We argue that measures of moments of the past return distribution can be too crude of estimates to capture its relevant characteristics, and changes therein. For example, an increase in volatility might suggest that the distribution is more spread out around the mean, but this might happen in a number of ways, e.g., in a specific sub-domain of the distribution (e.g., quantiles below .40), not necessarily across its entire domain. Similarly, a change in a skewness or a kurtosis measure can come across as a result of different changes in the shape

---

[1] Differentiated responses of returns across quantiles of its distribution to its own lagged values as well as to other variables have been documented in previous studies, e.g., Chuang et al. (2009), Baur et al. (2012), Baur (2013).

[2] For instance, Bollerslev et al. (2009) report that the variance risk premium predicts mean returns on international markets, and Amaya et al. (2015) demonstrate the predictive content of skewness and kurtosis for future mean returns. Paye (2012) shows that economic variables can generate superior volatility forecasts, especially prior to the 1980s.

of the distribution. Hence, it would be largely uncertain what an observed change in the past distribution's moment actually tells us about how the exact distribution changed. We argue that more information about those distributional changes, and potentially higher predictive power, can be extracted by looking at a wide range of specific quantiles of past distribution, even after controlling for the predictive power of the lagged first four moments. Our approach is more in spirit of the original proposal by Granger (1969), which defines causality in terms of conditional *distributions* of variables, than its subsequent operationalizations which focus almost exclusively on conditional means or variances.

We find that most DJIA returns measured at the 75th quantile have predictive power for most future quantile returns, with the exception of the median. This predictive power appears to prevail after controlling for the predictive power of other features of the lagged return distribution (i.e., its first four moments). It does not appear to be driven only by specific sub-periods (i.e., is stable over time). It also prevails when one controls for other potential predictors (as summarised in Welch and Goyal (2008), and Kelly and Jiang's (2014) tail risk). Predictions of future mean return based on predicted quantiles are found to have positive economic value, as are those utilising quantile predictions to forecast volatility, especially for investors with low aversion to risk. Lastly, the predictive power of quantile .75 DJIA returns is shown to be due to its ability to forecast future investment and consumption shocks.

2. Methodology

2.1. Quantile returns

For each year, using daily index log returns we obtain the estimated return at each quantile θ considered, i.e., $R_t(\theta = k)$, where k∈{.01, .02, .05, .10, .20, .25, .30, .40, .50, .60, .70, .75, .80, .90, .95, .98, .99}. This is accomplished by regressing the daily return on a constant in a quantile regression framework (Koenker and Bassett, 1978), with the resulting intercept value constituting the estimated daily return (in a given year) at a given quantile θ (but could also be done by ranking daily returns in each year and selecting the relevant quantile observation). This procedure is repeated for each quantile and year, and generates a time series of returns for each of the specified quantiles, at annual frequency.[3]

2.2. Predictive models using lagged quantile returns

The baseline (parsimonious) model employs the first to fourth moments of lagged daily returns as well as the lagged dependent variable as predictors and is of a form:

$$Y_t = \alpha_0 + \alpha_1 Mean_{t-1} + \alpha_2 Volatility_{t-1} + \alpha_3 Skewness_{t-1} + \alpha_4 Kurtosis_{t-1} + \alpha_5 Y_{t-1} + \varepsilon_t. \quad (1)$$

The dependent variable, Y_t, equals the returns calculated for a specific quantile, i.e., $Y_t = R_t(\theta = k)$, where k∈{.01, .02, .05, .10, .20, .25, .30, .40, .50, .60, .70, .75, .80, .90, .95, .98, .99}, calculated as explained in section 2.1. $Mean_{t-1}$ is the mean daily return in period t-1, with this period being one calendar year, hence periods are non-overlapping. The second, third and fourth moments are also calculated from daily returns in year t-1. The last component, $\alpha_5 Y_{t-1}$, accounts for potential autocorrelation in the dependent variable.

The larger (encompassing) model that nests the parsimonious model above (eq. (1)) and contains one more explanatory variable, X_{t-1} is:

$$Y_t = \alpha_0 + \alpha_1 Mean_{t-1} + \alpha_2 Volatility_{t-1} + \alpha_3 Skewness_{t-1} + \alpha_4 Kurtosis_{t-1} + \alpha_5 Y_{t-1} + \alpha_6 X_{t-1} + \varepsilon_t \quad (2)$$

---

[3] The estimated return quantiles are those for daily returns, each representing the average daily value (over each year). One could annualize those estimated values to obtain annual quantile return equivalents, but this would not affect their relevant statistical properties, only their magnitude.

This additional variable X_{t-1} is related to the quantile returns and can take on different forms. To conserve space, we concentrate on only one variant of it, i.e., X_{t-1} equals the returns calculated for a specific quantile, i.e., $X_{t-1} = R_{t-1}(\theta = k)$, where k∈{.01, .02, .05, .10, .20, .25, .30, .40, .50, .60, .70, .75, .80, .90, .95, .98, .99}, and $X_{t-1} \neq Y_{t-1}$. Therefore, in this variant model (2) includes the lagged value of quantile returns (for each quantile return on the LHS, we estimate models with each quantile return on the RHS).[4]

Predictive models such as (1) and (2) are estimated here using annual non-overlapping observations, as, e.g., in Ferreira and Santa-Clara (2011). Rather than using quantile regression technique here, we employ the OLS approach to estimate (1) and (2), which has the advantage of allowing for a straightforward calculation of standard errors corrected for autocorrelation and heteroscedasticity, where required. It also allows for direct comparability with studies on predictability of future stock returns. Further, as our variables (quantile returns and moments of return distribution) are estimated, a potential error-in-variable problem arises. However, these variables are not directly observable and have to be estimated. In addition, the potential error-in-variable problem in both dependent and independent variables will lead to t-statistics being biased downward, hence, any significant predictive power we observe will be even more meaningful than it would be otherwise, in absence of a (potential) bias.[5] Further, some of the predictability literature uses overlapping observations to increase the sample size, it is well known that this approach leads to

---

[4] We also considered two other options for X_{t-1} : firstly, where X_{t-1} equals to the differences between returns calculated for a specific quantile and the median return, i.e., $X_{t-1} = R_{t-1}(\theta = k) - R_{t-1}(\theta = .50)$, and secondly, there X_{t-1} equals to the differences between returns calculated for quantiles symmetrical around the median, i.e., $X_{t-1} = R_{t-1}(\theta = 1 - k) - R_{t-1}(\theta = k)$, e.g., $R_{t-1}(\theta = .99) - R_{t-1}(\theta = .01)$. Results for those alternative measures of return distribution were qualitatively similar to those obtained when using past quantile returns and are not reported to conserve space.

[5] In addition, unreported results indicate that model fits for (1) and (2) are significantly higher when using OLS rather than quantile regressions, especially for those quantiles in the tails for which the number of observations of extreme returns is limited. In the OLS framework, by contrast, the number of observations is identical across (estimated) quantile returns.

numerous problems and potentially incorrect inferences about the existence and magnitude of the predictive power of variables. Even though several measures have been proposed to account for potential biases in, for example, the estimated coefficients, their significance, and R-squares, no solution can be guaranteed to be perfect and we prefer to avoid the problem altogether.[6] Hence, non-overlapping annual variables are used.

2.3. Tests of out-of-sample performance

To assess the predictive power of X_{t-1} out-of-sample, over and above the lagged values of the first four moments of returns and the lagged dependent variable, two encompassing tests are used.

2.3.1. The Clark and McCracken (2001) test

Clark and McCracken (2001) propose a test for forecasting accuracy of one-step predictions derived from nested linear models, which is designed to perform well in small samples:

$$ENC-NEW = P\frac{P^{-1}\sum_t(\hat{u}_{1,t+1}^2 - \hat{u}_{1,t+1}\hat{u}_{2,t+1})}{P^{-1}\sum_t(\hat{u}_{2,t+1}^2)},$$

where $\hat{u}_{1,t+1}$ and $\hat{u}_{2,t+1}$ are estimated 1-step ahead prediction errors obtained recursively from the restricted and unrestricted (encompassing) model, respectively. P denotes the number of predictions (out-of-sample), whereas the in-sample period contains R observations. The distribution of the ENC-NEW statistic is shown to depend on $= {}^{P}/_{R}$, among other parameters, and is simulated and tabulated in Clark and McCracken (2001). These authors also demonstrate that the ENC-NEW test has superior power compared to those proposed by, e.g., Harvey, Leybourne, and Newbold (1998) and Ericsson (1992), especially in small samples (low values of P).

---

6 See, e.g., Richardson and Smith (1991), Boudoukh and Richardson (1993), Stambaugh (1999), Valkanov (2003), Ang and Bekaert (2007), Boudoukh, Richardson, and Whitelaw (2008).

2.3.2. The Clark and West (2007) test

Clark and West (2007) note that under the null that the data is generated by the parsimonious rather than the larger (encompassing) model, the latter introduces noise into its forecasts. Hence, under the null, when parameters are set to their population values on both models, the mean square prediction error (MSPE) of the parsimonious model will be smaller than that of the larger model, even though the models are essentially identical. Clark and West (2007) propose an approach to control for this noise by comparing the MSPE of the parsimonious model ($\hat{\sigma}_1^2$) with an adjusted MSPE of the larger model ($\hat{\sigma}_2^2 - adj$), where:

$$\hat{\sigma}_1^2 = P^{-1} \sum (y_{t+\tau} - \hat{y}_{1t,t+\tau})^2,$$

$$\hat{\sigma}_2^2 - adj = P^{-1} \sum \left(y_{t+\tau} - \hat{y}_{2t,t+\tau}\right)^2 - P^{-1} \sum (\hat{y}_{1t,t+\tau} - \hat{y}_{2t,t+\tau})^2,$$

$y_{t+\tau}$ denotes the observed value of the to-be-predicted variable at time $t + \tau$, τ denotes the prediction horizon, $\hat{y}_{1t,t+\tau}$ and $\hat{y}_{2t,t+\tau}$ stand for τ-period ahead predicted (at time t) values of y using the parsimonious and the larger models, respectively, and P is the number of predictions. They define $\hat{f}_{t+\tau} = (y_{t+\tau} - \hat{y}_{1t,t+\tau})^2 - (\left(y_{t+\tau} - \hat{y}_{2t,t+\tau}\right)^2 - (\hat{y}_{1t,t+\tau} - \hat{y}_{2t,t+\tau})^2)$ and notice that $\hat{\sigma}_1^2 - (\hat{\sigma}_2^2 - adj)$ is a sample average of $\hat{f}_{t+\tau}$, hence the test for equal MSPE (i.e., whether $\hat{\sigma}_1^2 - (\hat{\sigma}_2^2 - adj)$ equals zero) can be conducted by regressing $\hat{f}_{t+\tau}$ on a constant and using the resulting t-statistic (hereafter referred to as CW07) to test for significance of the resulting coefficient (one-sided test, $\alpha > 0$). The advantage of this regression-based test is that the standard errors can easily be adjusted for heteroskedasticity and autocorrelation, if necessary, using the usual approaches, hence the distribution of the test statistic does not have to be simulated.

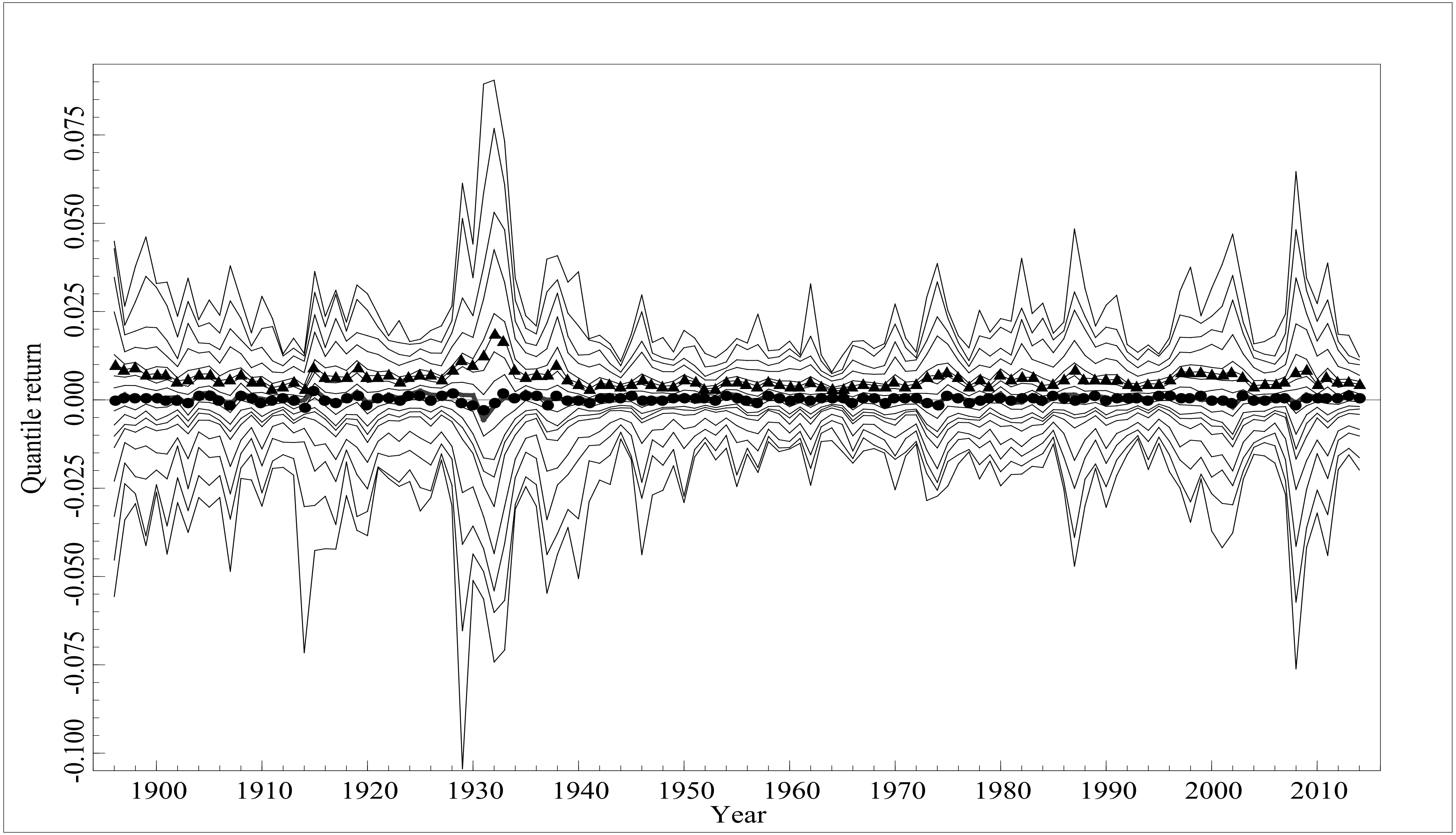
3. Data and estimated quantile returns

We employ daily values of the Dow Jones Industrial Average index, from May 26, 1896, to September 10, 2014. Observations up to and including March 2007 are available from WRDS and the remaining ones are obtained from Datastream.[7]

In every calendar year, the whole set of return quantiles for that year is estimated using daily index returns. Utilising one year of data at a time allows to mitigate potential biases due to existence of seasonalities in stock prices, e.g., the January effect, etc. The estimated quantile returns are presented in Figure 1 and their selected descriptive statistics in Table 1. It is apparent that the return distribution varies over time, in many respects. First, returns at each quantile are volatile, with those returns measured closer to the center of the distribution (median return is represented by a thick black line) showing lower levels of volatility than those situated in the tails. Second, distances among quantiles vary over time as well, implying that the shape of the distribution, and not only its location, is time varying. The mean and the median return are fairly close to each other, which suggest that the distribution is close to symmetry; this feature is also visible in Figure 1 to some extent. Those returns estimated for quantiles above (below) the mean have negative (positive) skewness, i.e., they "lean towards" the mean of the distribution. However, excess kurtosis is positive in all cases, indicating higher probability of extreme returns in any quantile, and for the distribution as a whole when mean return's kurtosis is considered.

[7] DJIA time series data is employed due to the high number of observations available, a feature which allows us to avoid incorrect inference due to subperiod-specific characteristics or inefficiency of estimates obtained from small samples. The latter issue is further analysed in section 4.3.

Figure 1: Estimated quantile returns



Note: Lines represent estimated returns at quantiles $\theta \in \{0.01, 0.02, 0.05, 0.10, 0.20, 0.25, 0.30, 0.40, 0.60, 0.70, 0.75, 0.80, 0.90, 0.95, 0.98, 0.99\}$ and the mean return, at annual frequency. The solid line represents the mean return, the dotted line in the middle represents the median return, ($R(\theta = .50)$), and triangles represent quantile .75 returns ($R(\theta = .75)$).

Table 1: Descriptive statistics for estimated quantile returns and the mean return.

Quantile	Mean	Standard Deviation	Skewness	Excess Kurtosis	Minimum	Maximum
1	-0.0282	0.0155	-1.9637	5.3883	-0.1044	-0.0110
2	-0.0226	0.0110	-1.5949	2.9996	-0.0654	-0.0079
5	-0.0165	0.0082	-1.9604	4.9035	-0.0541	-0.0059
10	-0.0117	0.0059	-2.3227	7.9554	-0.0435	-0.0041
20	-0.0070	0.0039	-2.8885	12.6666	-0.0302	-0.0023
25	-0.0053	0.0031	-2.6354	10.0304	-0.0219	-0.0016
30	-0.0040	0.0025	-2.6307	9.7613	-0.0169	-0.0012
40	-0.0017	0.0016	-2.3611	9.1484	-0.0103	0.0006
50	0.0004	0.0011	-1.6408	8.4259	-0.0057	0.0028
60	0.0025	0.0011	1.1384	1.7926	0.0006	0.0072
70	0.0047	0.0018	1.7668	5.5464	0.0023	0.0136
75	0.0061	0.0025	2.2772	8.6311	0.0031	0.0191
80	0.0077	0.0032	2.3229	8.6620	0.0035	0.0244
90	0.0120	0.0054	2.5686	10.2116	0.0053	0.0425
95	0.0160	0.0074	2.2821	7.5140	0.0066	0.0531
98	0.0210	0.0107	2.3895	8.0514	0.0074	0.0768
99	0.0261	0.0141	2.2148	6.7910	0.0079	0.0905
Mean	0.0002	0.0009	-0.8042	1.5863	-0.0030	0.0024

We also compare the predictive power of lagged quantile returns against that of a set of economic variables which has been widely used in the literature and are analysed in, e.g., Welch and Goyal (2008):[8]

- Default yield spread (dfy), calculated as the difference between BAA- and AAA-rated corporate bond yields.
- Inflation (infl), calculated utilising the data in the Consumer Price Index (All Urban Consumers) from the Bureau of Labor Statistics.
- Stock variance (svar), computed as sum of squared daily returns on S&P500.

[8] All data and its detailed descriptions can be obtained from Professor Goyal's webpage (http://www.hec.unil.ch/agoyal/). We thank those authors for making their updated data publicly available.

- Dividend payout ratio (de), calculated as the difference between log of dividends and log of earnings, where dividends (earnings) are twelve-month moving sums of dividends (earnings) paid on the S&P 500 index.
- Long-term government bond yield (lty).
- Term spread (tms), defined as the difference between the long-term yield on government bonds and the T-bill.
- Treasury bills rate (tbl).
- Default return spread (dfr), the difference between the return on long-term corporate bonds and return on the long-term government bonds.
- Dividend price ratio (dp), the difference between the log of dividends and the log of prices.
- Dividend yield (dy), difference between the log of dividends and the log of lagged prices.
- Long-term government bond returns (ltr).
- Earnings price ratio (ep), the difference between log of earnings and log of prices.
- Book to market ratio (bm) is the ratio of book value to market value for the Dow Jones Industrial Average.
- Investment to capital ratio (ik), the ratio of aggregate (private non-residential fixed) investment to aggregate capital for the whole economy, as proposed in Cochrane (1991).
- Net equity expansion (ntis) as a measure of corporate issuing activity, the ratio of twelve-month moving sums of net issues by NYSE listed stocks divided by the total market capitalization of NYSE stocks.
- Percent equity issuing (eqis), the ratio of equity issuing activity as a fraction of total issuing activity.

In addition, we use the Kelly and Jiang's (2014) tail risk measure (tail) as a proxy for extreme event risk.[9]

4. Results

4.1. Predictive power beyond the lagged first four moments and the lagged dependent variable.

Model (2) with lagged quantile returns employed as the RHS variable X_{t-1} is estimated and the encompassing tests are employed to test for the predictive power of lagged quantile returns in excess of that of lagged dependent variable and the lagged four moments of return distribution (i.e., model (1) is the parsimonious model and model (2) is the larger model). We conduct this analysis in two different ways with respect to how the whole sample is divided into the in-sample and out-of-sample periods. First, we consider a short in-sample period (R=30 years) and a long out-of-sample period (P=89 years), which corresponds to Clark and McCracken's (2007) π value of roughly 3, where π stands for the ratio of observations in out-of-sample (P) vs. in-sample (R) periods. Second, we also analyse an opposite case in which we allow for a long in-sample and short out-of-sample period, with R=85 and P=34 years in each, respectively. This case is denoted by π=0.4. However, in our discussions we tend to concentrate on cases where the out-of-sample period is long (π=3): Hansen and Timmermann (2012) show that out-of-sample tests of predictive ability have better size properties when the forecast evaluation period is a relatively large proportion of the available sample (Neely et al., 2014).

Table 2 (3) presents encompassing test statistics for the CW07 (ENC_NEW) tests when the out-of-sample period is long (π=3). The former test generates fewer significant

---

[9] We are grateful to Professor Kelly for providing us with his data on estimated tail risk.

results than the latter (i.e., the CW07 approach seems to be more conservative), but the common result is the superior forecasting accuracy of model (2) over model (1), mostly when it contains lagged returns from quantiles .70 and .75, and for the next-year returns measured at all quantiles except those in quantiles .70 and .75. There is also evidence of forecasting ability of returns in quantile .25 for the next-year returns in quantiles .80 and higher, and of those in quantiles .10-.25 for returns in quantile .60.

When we consider a shorter out-of-sample period (π=0.4), the results (presented in Tables 4 and 5) indicate fewer cases of significantly superior forecast accuracy of model (2) over model (1). This might be due to poor test size properties when forecast evaluation period is short (Hansen and Timmermann, 2012), or indicate a decline in forecasting ability of quantile returns in the later part of our sample. However, those cases which remain significant are largely in line with the results for a long out-of-sample period. Lagged returns measured at quantiles .70 and .75 show forecasting power for next-year returns across a wide range of quantiles of the latter. There is also evidence of predictive power of, for example, returns in quantiles .20-.60 and .99 for the future median return.

Taken together, the results in Tables 2-5 show evidence of predictive ability of lagged returns in quantiles .70 and .75 for a wide range of next-year's return quantiles. This predictive power is in addition to any information content captured by the lagged first four moments of the respective to-be-predicted returns, as well as their own lagged values. It does

Table 2: Encompassing tests statistics for CW07, variant 1 RHS variables, π=3.

LHS:	1	2	5	10	20	25	30	40	50	60	70	75	80	90	95	98	99
θ:									CW07								
0.99	-1.032	-1.624	0.171	0.461	0.480	0.158	0.107	0.972	1.3449	-0.714	-1.128	-1.360	-0.462	0.227	1.051	0.955	
0.98	0.733	-0.433	-0.393	-1.056	-1.072	-1.153	-1.223	-1.031	-0.906	-0.050	-0.225	-0.283	-1.086	-0.429	-0.535		0.643
0.95	-0.963	-0.601	0.562	1.110	-0.271	0.696	0.827	0.398	-1.005	0.851	-0.423	-0.151	-0.379	0.758		-0.482	0.206
0.90	-0.569	-0.925	-1.293	-1.064	-0.291	1.019	0.898	-1.256	-1.27	1.500	0.419	-1.462	-0.712		1.285	-0.647	-0.245
0.80	-0.343	-0.683	-0.088	-0.917	-1.138	-1.411	-1.186	-0.652	0.7656	0.114	-1.992	1.431		0.063	**1.747**	0.283	-0.124
0.75	**1.700**	**1.915**	**1.693**	**1.889**	**2.302**	**2.228**	**2.284**	**2.234**	1.4128	**2.651**	0.205		**2.268**	**2.300**	**2.738**	**2.867**	**2.754**
0.70	**2.496**	**2.155**	**1.995**	**1.979**	**2.094**	**2.350**	**2.276**	**2.626**	0.7284	**1.998**		-1.056	1.321	1.610	**2.448**	**2.253**	**2.465**
0.60	-0.868	0.320	0.706	-0.576	-0.844	-0.720	-0.233	-0.161	-0.455		0.998	-1.154	-1.208	-1.045	-0.913	-0.706	0.418
0.40	-0.963	-1.459	-1.270	-1.134	-1.145	-0.669	-0.067		1.3392	-0.386	0.327	-1.368	-1.309	-1.009	-0.585	-0.561	-1.272
0.30	-0.323	-0.827	0.629	0.032	0.873	-1.430		0.452	1.2093	-0.973	0.215	-0.501	0.958	0.959	1.141	0.963	0.792
0.25	0.358	-0.262	0.838	0.549	1.152		0.022	-0.215	0.882	**1.785**	1.165	0.542	1.435	1.417	1.486	1.564	1.406
0.20	-0.068	-0.719	-0.484	-1.025		-0.009	-0.015	-0.712	0.6902	1.316	-0.629	-1.114	0.978	0.827	1.426	0.747	0.320
0.10	-0.856	-0.803	-1.111		-1.032	-0.153	0.595	-0.825	-0.777	1.384	-0.533	-0.650	-1.059	-1.099	-0.711	-0.813	-0.515
0.05	1.098	0.208		-0.900	-1.026	-0.798	-0.369	-0.971	-1.127	-0.289	-0.993	-0.872	-1.255	-1.162	-1.139	-1.096	-1.025
0.02	0.760		-0.561	-0.581	-1.047	-1.010	-1.259	-1.093	-0.856	0.703	1.119	1.096	0.976	0.302	0.733	-0.738	0.731
0.01		0.420	-1.046	-1.014	-1.081	-1.121	-1.125	-1.020	0.6887	0.468	0.894	0.764	0.784	-0.161	0.717	-0.034	-0.049

Note: CW07 denotes the test statistic of Clark and West (2007). Shaded areas indicate significance at 10% level, bold numbers indicate significance at 5% level. π denotes the ratio of the number of observations in the out-of-sample (forecasting) period to the number of observations in the in-sample (estimation) period (Clark and McCracken, 2007). The critical values for CW07 and ENC_NEW at 10% and 5% level are: for π=3: 1.292/1.664 (one-sided test) and 1.442/2/374, and for π =0.4: 1.314/1.703 and 0.685/1.079.

Table 3: Encompassing tests statistics for ENC_NEW, variant 1 RHS variables, π=3.

LHS:	1	2	5	10	20	25	30	40	50	60	70	75	80	90	95	98	99
θ:									ENC_NEW								
0.99	-1.929	-0.963	0.234	0.954	1.269	0.402	0.401	**3.911**	**3.7794**	-0.771	-2.704	-0.926	-0.429	0.330	1.697	**12.647**	
0.98	1.651	-0.318	-0.377	-2.497	-4.803	-3.708	-4.030	-2.319	-6.045	-0.059	-0.399	-0.455	-1.491	-1.092	-0.716		**3.477**
0.95	-1.684	-0.522	0.772	1.923	-1.292	**2.691**	**5.548**	0.355	-4.769	**2.857**	-0.609	-0.150	-0.205	**3.008**		-0.211	0.215
0.90	-1.593	-3.142	-2.913	-3.253	-0.867	**3.453**	**8.065**	-1.505	-2.967	**5.989**	0.408	-0.865	-0.305		**6.734**	-0.637	-0.120
0.80	-0.375	-1.582	-0.128	-1.303	-2.978	-2.223	-2.373	-1.718	0.9834	0.117	-2.002	**4.137**		0.076	**3.479**	0.390	-0.173
0.75	**3.250**	**4.832**	**12.100**	**13.141**	**7.192**	**4.039**	**3.049**	**5.508**	**2.5513**	**5.101**	0.268		**9.540**	**6.028**	**11.727**	**9.865**	**9.946**
0.70	**7.468**	**13.528**	**19.149**	**18.751**	**9.951**	**7.888**	**8.296**	**6.947**	0.7085	**5.170**		-1.526	**3.518**	**4.281**	**7.786**	**11.718**	**15.892**
0.60	-1.932	0.493	1.630	-0.704	-2.078	-1.345	-0.348	-0.511	-1.076		1.560	-2.368	-3.933	-5.472	-4.340	-2.837	1.008
0.40	-1.080	-1.459	-2.291	-4.465	-2.493	-1.013	-0.092		4.1199	-0.764	0.918	-3.042	-3.280	-2.773	-0.847	-0.631	-1.039
0.30	-0.606	-2.423	**2.505**	0.112	**3.431**	-2.011		1.506	**10.763**	-0.770	0.451	-0.987	**7.348**	**10.650**	**15.121**	**4.422**	**1.723**
0.25	0.917	-0.802	**3.333**	1.852	**6.798**		0.017	-0.637	**4.163**	**3.140**	**2.908**	1.149	**8.968**	**9.698**	**19.414**	**7.033**	**3.508**
0.20	-0.149	-2.178	-1.441	-4.420		-0.017	-0.020	-3.461	**2.7479**	2.224	-0.793	-2.452	**2.978**	**3.303**	**13.406**	**2.378**	0.595
0.10	-3.404	-2.496	-3.130		-4.119	-0.290	1.677	-3.144	-1.381	**5.671**	-1.759	-2.602	-5.492	-7.272	-2.882	-2.775	-1.640
0.05	**3.640**	0.257		-4.055	-6.785	-3.014	-1.044	-5.358	-3.114	-0.238	-2.342	-2.176	-5.807	-9.362	-6.825	-5.543	-4.438
0.02	**3.043**		-1.639	-3.138	-7.310	-5.282	-5.761	-3.736	-1.847	2.260	**14.269**	**12.393**	**9.795**	1.433	**3.041**	-1.861	**3.343**
0.01		0.494	-2.927	-6.144	-7.186	-5.555	-3.850	-4.405	**4.0367**	1.065	**4.763**	**3.776**	**4.077**	-0.579	**3.193**	-0.100	-0.098

Note: ENC_NEW denotes the test statistic of Clark and McCracken (2001). Shaded areas indicate significance at 10% level, bold numbers indicate significance at 5% level. π denotes the ratio of the number of observations in the out-of-sample (forecasting) period to the number of observations in the in-sample (estimation) period (Clark and McCracken, 2007). The critical values for CW07 and ENC_NEW at 10% and 5% level are: for π=3: 1.292/1.664 (one-sided test) and 1.442/2/374, and for π =0.4: 1.314/1.703 and 0.685/1.079.

Table 4: Encompassing tests statistics for CW07, variant 1 RHS variables, π=0.4

LHS:	1	2	5	10	20	25	30	40	50	60	70	75	80	90	95	98	99
θ:									CW07								
0.99	-1.896	-2.076	-0.840	-0.843	0.250	0.060	0.278	0.809	**2.572**	0.650	0.959	-0.254	-1.900	-0.817	0.709	-1.040	
0.98	-0.505	-2.495	-1.239	-1.754	-0.406	-1.226	-0.859	-0.283	-0.933	1.406	-1.469	-0.960	-1.972	-2.233	1.550		-2.062
0.95	-1.486	-0.532	0.976	0.770	0.807	1.631	1.179	1.287	-0.118	0.257	-1.199	0.230	0.610	2.414		**1.822**	-0.983
0.90	1.091	0.465	-0.728	-1.898	-1.944	-1.327	-1.459	-0.672	-1.331	1.250	0.615	-1.643	-0.003		**1.703**	0.839	**2.297**
0.80	-0.016	-0.372	-0.153	0.202	1.225	-1.506	-1.470	0.790	1.367	0.884	-1.361	0.095		-0.121	0.293	0.142	1.538
0.75	0.899	0.255	1.063	**1.848**	**1.989**	**1.707**	**1.858**	**3.024**	**1.830**	**1.987**	0.608		1.411	1.141	1.158	1.541	**2.708**
0.70	1.195	0.411	0.775	0.951	**1.931**	1.648	**1.821**	**2.654**	1.536	**2.580**		-0.524	1.216	0.661	0.809	1.086	**2.440**
0.60	-0.488	-0.904	-1.321	-1.228	-0.902	-0.156	-0.030	**2.217**	**1.789**		**1.944**	**1.802**	-2.386	-1.706	-1.333	-0.679	-0.028
0.40	-0.251	-0.584	-0.527	-0.149	-1.920	-1.256	-1.637		**2.023**	0.331	1.403	-2.524	0.057	0.613	0.726	0.308	0.302
0.30	-0.872	-0.852	-0.606	-0.443	-0.127	-1.343		-0.302	**1.906**	-0.255	0.219	-0.351	0.062	-0.584	0.146	-0.368	0.498
0.25	-0.395	-0.431	-0.119	-0.018	1.058		-0.454	-0.456	1.423	0.812	0.339	-0.100	0.059	-0.433	0.298	-0.322	0.623
0.20	-1.009	-1.165	-0.817	-0.354		-0.716	-1.650	0.156	2.425	-0.790	-1.604	-1.087	-1.251	-1.594	-0.057	-0.501	0.333
0.10	0.020	-0.653	-0.572		-1.357	-1.478	-2.093	**1.782**	-3.253	-0.254	**1.762**	**2.310**	-1.189	-1.016	0.244	-0.017	**1.954**
0.05	-0.107	-0.484		-1.297	-0.443	-0.368	-0.249	0.568	-1.717	0.238	1.323	1.378	0.675	-0.044	-1.677	-1.586	0.307
0.02	-0.489		-1.123	-1.084	-0.720	-0.771	-0.806	-0.260	0.333	-1.524	0.254	0.470	0.790	0.232	0.050	0.771	-0.605
0.01		-1.906	-1.411	-1.029	-0.753	-0.899	-0.995	-1.136	0.769	-1.294	-0.744	-0.590	-0.145	-0.576	-0.271	-0.042	-0.418

Note: CW07 denotes the test statistic of the test statistic of Clark and West (2007). Shaded areas indicate significance at 10% level, bold numbers indicate significance at 5% level. π denotes the ratio of the number of observations in the out-of-sample (forecasting) period to the number of observations in the in-sample (estimation) period (Clark and McCracken, 2007). The critical values for CW07 and ENC_NEW at 10% and 5% level are: for π=3: 1.292/1.664 and 1.442/2/374, and for π =0.4: 1.314/1.703 and 0.685/1.079.

Table 5: Encompassing tests statistics for ENC_NEW, variant 1 RHS variables, π=0.4

LHS:	1	2	5	10	20	25	30	40	50	60	70	75	80	90	95	98	99
θ:									ENC_NEW								
0.99	-0.547	-0.409	-0.374	-0.416	0.185	0.028	0.130	0.405	1.070	0.152	0.277	-0.055	-0.228	-0.431	0.312	-1.068	
0.98	-0.084	-0.056	-0.054	-0.096	-0.029	-0.046	-0.058	-0.051	-0.329	0.348	-0.318	-0.394	-0.219	-0.215	0.194		-0.474
0.95	-0.320	-0.025	0.215	0.623	0.460	1.073	0.899	0.523	-0.037	0.132	-0.093	0.146	0.186	**1.904**		0.381	-0.094
0.90	0.537	0.129	-0.091	-0.379	-1.055	-1.110	-0.934	-0.102	-0.454	1.023	0.167	-0.217	-0.001		**1.563**	0.182	0.481
0.80	-0.007	-0.213	-0.080	0.220	0.283	-0.180	-0.133	0.569	0.302	0.492	-0.099	0.144		-0.071	0.246	0.072	0.780
0.75	0.763	0.238	**1.318**	**3.272**	**5.002**	**4.140**	**3.314**	**7.043**	**2.771**	**1.852**	0.304		**3.276**	**1.524**	**1.896**	**2.075**	**3.705**
0.70	**1.093**	0.488	**1.174**	**1.849**	**2.532**	**2.406**	**2.619**	**4.701**	**1.766**	**3.460**		-0.240	0.849	0.531	0.842	**1.170**	**2.900**
0.60	-0.224	-0.465	-0.671	-0.806	-0.445	-0.081	-0.018	**1.791**	**1.360**		**2.193**	0.788	-0.184	-0.556	-0.504	-0.422	-0.022
0.40	-0.134	-0.183	-0.164	-0.050	-0.370	-1.196	-1.819		**2.764**	0.066	0.375	-0.162	0.008	0.284	0.487	0.166	0.117
0.30	-0.780	-0.676	-0.490	-0.491	-0.082	-0.520		-0.673	**2.850**	-0.061	0.043	-0.137	0.056	-0.811	0.239	-0.497	0.636
0.25	-0.423	-0.389	-0.107	-0.030	0.959		-0.254	-0.930	0.854	0.643	0.206	-0.133	0.080	-0.582	0.423	-0.441	0.779
0.20	-0.720	-0.727	-0.512	-0.418		-0.299	-0.174	0.258	0.996	-0.537	-0.474	-0.651	-1.026	-1.522	-0.068	-0.554	0.297
0.10	0.018	-0.454	-0.137		-0.194	-0.323	-0.250	0.815	-1.584	-0.303	0.396	**1.622**	-0.253	-0.147	0.219	-0.011	**1.238**
0.05	-0.051	-0.266		-0.187	-0.117	-0.093	-0.073	0.220	-0.516	0.124	0.429	0.771	0.140	-0.007	-0.096	-0.253	0.073
0.02	-0.089		-0.134	-0.104	-0.202	-0.277	-0.270	-0.089	0.155	-0.671	0.203	0.524	0.640	0.086	0.015	0.208	-0.113
0.01		-0.204	-0.324	-0.273	-0.215	-0.393	-0.406	-0.123	**1.147**	-1.576	-1.394	-1.098	-0.186	-0.532	-0.139	-0.030	-0.191

Note: ENC_NEW denotes the test statistic of Clark and McCracken (2001). Shaded areas indicate significance at 10% level, bold numbers indicate significance at 5% level. π denotes the ratio of the number of observations in the out-of-sample (forecasting) period to the number of observations in the in-sample (estimation) period (Clark and McCracken, 2007). The critical values for CW07 and ENC_NEW at 10% and 5% level are: for π=3: 1.292/1.664 and 1.442/2/374, and for π =0.4: 1.314/1.703 and 0.685/1.079

not seem to be restricted to a certain subperiod of the sample.[10]

We conduct further tests to analyse the robustness of this finding (results not reported but available on request). Firstly, we investigate all quantiles between .60 and .80 to establish whether quantile .75 return is indeed the one with superior predictive power, or whether it is just a proxy for other quantiles in its immediate proximity with superior forecasting performance. The results of the Clark and West (2007) and Clark and McCracken (2001) tests which compare the forecasting performance of model 2 vs. model 1 indicate that quantile .75 returns show significant forecasting power for the highest number of future return quantiles (jointly with quantile .74 returns). Hence, we conclude that quantile .75 returns are not dominated by any neighbouring quantile in terms of their forecasting power.

A second robustness test comprises of using a different stock market index to investigate broader validity of our findings. For reasons which will be demonstrated in section 4.3, this analysis requires daily data starting in 1920s or earlier: We have been able to obtain daily data on S&P500, starting in 1928, and conducted the forecasting tests in an equivalent way to those reported in Tables 2 and 3 (Clark and West (2007) and Clark and McCracken (2001), respectively, to compare the forecasting performance of model 2 vs. model 1). The results indicate that for both tests, predictions generated by quantile .75 returns are resulting in the highest number of significantly superior forecasts (i.e., for the highest number of future to-be-predicted quantiles), in one case jointly with quantile .70 and in

---

[10] A potential criticism could be that we are bound to find some significant results when considering outcomes from 256 independent tests (on average, in X% of cases when the significance level is X%), hence, these results could be due to data mining. However, as it is not known *a priori*, on theoretical level, which quantile(s) could have predictive power, considering a wide range of quantiles empirically is the only way to establish their individual forecasting ability. In other words, we do not search for the predictive quantile but rather analyse the predictive power of each separately in order to establish if, and which, quantiles possess predictive power. Further, even when we employ the false discovery rate approach, which accounts for multiple testing in a rather stringent way, and compute p-values following Benjamini and Hochberg (1995) and Benjamini and Yekutieli (2001) (see Harvey and Liu, 2014, for a discussion), quantiles 70 and 75 retain predictive power for a wide range of future quantiles (20-40, 60, 95-99), whereas other quantiles are overwhelmingly lacking predictive ability (results not reported to conserve space). Hence, our findings are not due to data mining.

another with .10. Hence, these results confirm that quantile .75 returns have a superior predictive power and this result is not restricted to one stock market index.

We also conducted our forecasting power analysis on stock level data. We identified stocks in the CRSP database with share codes 10 and 11 which traded over that entire period 1926-2014 by their PERMNO identifier (i.e., where the PERMNO identifier was reported for the start and the end of the entire sample). This procedure resulted in a sample of 31 stocks. We then analysed the forecasting power of quantile .75 returns using the same approach as above, with Clark and West (2007) test and $\pi=3$. The overall result was that quantile .75 returns did not outperform other quantile returns as a predictor. The possible reasons are that, firstly, stock-level returns are more noisy, due to idiosyncratic noise, than index-level returns, and hence more difficult to predict; secondly, index level predictability can be driven by other effects and does not require stock-level predictability to exist (Lo and MacKinlay, 1990); and, thirdly, any potential stock-level predictability is more likely to have been exploited due to historically lower transaction costs of trading in individual stocks rather than in an index. Virtually all papers in this branch of the literature focus on forecasting index-level rather than individual stock returns. Hence, when we construct an equally-weighted index out of these 31 stocks, this index's quantile .75 returns are the superior predictor of future market returns among all analysed quantiles, which is in line with our previous findings for other indices.

4.2. Predictive power of return quantiles beyond the historical mean

The results discussed above indicate that returns measured at quantiles .70 and .75 do possess superior predictive power for next-year returns at a wide range of quantiles (.20-.60 and .80 and above), which comes in excess of the potential predictive power of four moments of return distribution as well as the lagged to-be-predicted quantile return. However, it is not clear from those results whether the model with added lagged quantile returns (eq. (2)) performs any better than a model which uses only the historical mean of the LHS return variable. The historical mean is a standard benchmark in the literature. To address this question, we run encompassing tests comparing two models. The larger model is as shown in equation (2), and the nested, parsimonious one is a regression of the quantile return on a constant. The latter model generates estimates of historical means when it is estimated recursively for each period t, which are then used to predict one-year ahead ($t+1$) returns in the out-of-sample subperiod. We calculate the values of both statistics, however, Clark and McCracken (2001) do not report the critical values for the case of six excess parameters in the larger model (as is the case here), hence in our discussions we rely on the test statistics from the more conservative Clark and West (2007) approach.

To get an idea of the magnitude of the differences in predictability between the parsimonious (historical mean) and full models, we compute the out-of-sample goodness-of-fit measure as follows (Kelly and Jiang, 2014; Da, Jagannathan, and Shen, 2014):

$$OOS\_R^2 = 1 - \frac{\sum_{t=m}^{T}(r_{i+1} - \hat{r}_{i+1})^2}{\sum_{t=m}^{T}(r_{i+1} - \bar{r}_{i+1})^2},$$

where $r_{i+1}$is the observed return, $\hat{r}_{i+1}$ is the predicted return based on the larger model and $\bar{r}_{i+1}$ stands for the predicted return based on the parsimonious model, i.e., the historical mean return calculated for the sample up to period m. Positive values of this measure indicate better out-of-sample predictive accuracy of the larger model, negative values indicate that the historical mean generated less error-ridden forecasts.

It should be noted that the CW07 approach does not simply compare the MSPE of the parsimonious and larger models, but adjusts for the extra noise (induced if the null is correct) in the predictions made using the larger model. Hence, the test statistic is always larger than one which would be given by the difference in MSPE of the parsimonious vs. larger model.

An implication of this adjustment is that the OOS_R^2 measure, which does not adjust for the additional noise of the incorrectly specified larger model (under the null), can be negative and indicate the inferior performance of the larger model vis-á-vis the historical mean, but the CW07 statistic will be significant, implying the opposite (superior forecasting power of the larger model). We base our inference on the CW07 statistic, but report the values of the OOS_R^2 measure to give a flavour of the magnitude of the potential forecasting power of models including quantile returns (as well as first four moments and lagged dependent returns) as explanatory variables.

The results (Table 6) indicate that most models are better than the historical mean (parsimonious model), as the CW07 statistics are significant for both long (π=3) and short (π=.4) out-of-sample periods. A noticeable exception is when the median return is used as the dependent variable, as the CW07 statistics are insignificant (more so for π=3) and OOS_R^2 values negative.

Taken together and in conjunction with the previous findings, these results indicate that models using returns at quantiles .70 and .75 possess a superior forecasting power for future returns at various quantiles, over and above any forecasting ability of the lagged first moments of the return distribution as well as lagged values of to-be-predicted quantile return. However, the predictive content of those quantile returns is not exceeding that of the historical mean when it comes to forecasting the center of future return distribution. On the other hand, however, this predictive content of those quantile returns appears to be superior to both parsimonious models considered (historical mean and the one without lagged quantile returns only) when we forecast future returns at quantiles away from the center of the distribution.

Table 6: Values of the OOS_R^2 statistic.

LHS:	1	2	5	10	20	25	30	40	50	60	70	7	80	90	95	98	99
θ	Panel A: Short estimation/long forecast evaluation (OOS) period (π=3)																
99	0.10	0.23	0.21	0.23	0.11	0.16	0.08	0.18	-0.38	0.13	0.37	0.51	0.50	0.45	0.48	0.42	0.33
98	0.16	0.24	0.21	0.18	0.01	0.10	0.02	0.09	-0.73	0.13	0.38	0.50	0.48	0.42	0.45	0.36	0.33
95	0.04	0.23	0.22	0.25	0.04	0.16	0.09	0.15	-0.61	0.09	0.29	0.48	0.48	0.43	0.47	0.35	0.33
90	-0.14	0.09	0.14	0.16	0.08	0.20	0.14	0.12	-0.55	0.03	0.27	0.46	0.44	0.46	0.48	0.35	0.33
80	0.14	0.18	0.20	0.20	0.05	0.14	0.07	0.05	-0.40	0.18	0.40	0.54	0.51	0.44	0.48	0.34	0.29
75	0.13	0.22	0.27	0.33	0.18	0.21	0.15	0.16	-0.37	0.21	0.42	0.53	0.55	0.47	0.51	0.38	0.34
70	0.19	0.31	0.34	0.37	0.20	0.24	0.19	0.17	-0.41	0.21	0.43	0.50	0.51	0.45	0.47	0.37	0.38
60	0.10	0.23	0.22	0.21	0.06	0.14	0.09	0.09	-0.47	0.19	0.43	0.50	0.45	0.36	0.39	0.25	0.27
40	0.12	0.21	0.17	0.14	0.07	0.13	0.08	0.16	-0.34	0.15	0.41	0.49	0.46	0.41	0.44	0.35	0.31
30	0.08	0.14	0.15	0.17	0.13	0.14	0.13	0.06	-0.26	0.16	0.40	0.49	0.50	0.47	0.50	0.34	0.26
25	0.03	0.15	0.16	0.21	0.18	0.19	0.12	0.07	-0.35	0.18	0.41	0.51	0.54	0.50	0.55	0.36	0.25
20	0.10	0.16	0.10	0.10	0.13	0.14	0.10	-0.04	-0.41	0.19	0.40	0.47	0.48	0.41	0.46	0.30	0.20
10	-0.01	0.13	0.13	0.24	0.03	0.15	0.11	0.04	-0.81	0.08	0.33	0.47	0.41	0.33	0.40	0.25	0.23
5	0.13	0.23	0.22	0.12	-0.05	0.10	0.07	-0.05	-0.70	0.17	0.37	0.47	0.40	0.27	0.34	0.23	0.19
2	0.15	0.25	0.15	0.07	-0.10	0.03	-0.04	0.05	-0.63	-0.01	0.36	0.48	0.46	0.36	0.43	0.31	0.28
1	0.16	0.24	0.15	0.06	-0.06	0.05	0.02	-0.02	-0.55	0.15	0.40	0.48	0.46	0.36	0.43	0.29	0.23
θ	Panel B: Long estimation/shortforecast evaluation (OOS) period (π=0.4)																
99	0.13	0.15	0.11	0.07	0.06	0.06	0.00	0.01	-0.05	-0.05	-0.06	0.04	0.08	0.14	0.08	-0.05	0.01
98	0.15	0.17	0.13	0.10	0.08	0.07	0.01	0.01	-0.13	-0.04	-0.09	0.02	0.08	0.16	0.08	0.06	-0.02
95	0.14	0.17	0.14	0.11	0.09	0.09	0.02	0.03	-0.12	-0.08	-0.07	0.04	0.09	0.21	0.07	0.08	0.00
90	0.17	0.18	0.13	0.09	0.00	-0.03	-0.08	0.00	-0.14	-0.05	-0.06	0.03	0.09	0.17	0.09	0.07	0.03
80	0.15	0.15	0.12	0.09	0.10	0.06	0.00	0.02	-0.08	-0.06	-0.07	0.01	0.09	0.15	0.03	0.04	0.03
75	0.11	0.08	0.07	0.05	0.18	0.17	0.13	0.23	0.03	-0.07	-0.06	0.05	0.12	0.14	0.02	0.02	0.07
70	0.12	0.09	0.06	0.05	0.12	0.12	0.08	0.15	-0.01	0.00	-0.07	0.02	0.09	0.14	0.04	0.00	0.04
60	0.13	0.13	0.07	0.03	0.03	0.04	-0.01	0.05	-0.04	-0.06	0.01	0.08	0.08	0.13	0.03	0.00	-0.05
40	0.15	0.17	0.13	0.10	0.06	-0.03	-0.17	0.01	0.02	-0.06	-0.05	0.04	0.09	0.18	0.08	0.06	0.01
30	0.09	0.12	0.09	0.05	0.05	0.04	0.01	-0.21	-0.01	-0.07	-0.08	0.03	0.04	0.04	0.01	-0.02	-0.01
25	0.10	0.13	0.11	0.07	0.06	0.07	-0.01	-0.13	-0.07	-0.06	-0.08	0.00	0.03	0.06	0.02	-0.01	0.01
20	0.10	0.12	0.10	0.07	0.08	0.03	0.00	-0.02	-0.05	-0.14	-0.11	0.00	-0.01	0.03	0.00	-0.01	0.00
10	0.13	0.13	0.12	0.11	0.07	0.05	-0.01	0.05	-0.23	-0.17	-0.05	0.10	0.07	0.16	0.06	0.04	0.06
5	0.15	0.15	0.13	0.10	0.07	0.06	0.00	0.02	-0.15	-0.07	-0.05	0.07	0.09	0.17	0.06	0.05	0.01
2	0.15	0.18	0.13	0.10	0.07	0.05	-0.01	0.00	-0.10	-0.11	-0.10	0.02	0.10	0.17	0.06	0.07	0.00
1	0.16	0.17	0.12	0.09	0.07	0.04	-0.02	0.00	-0.06	-0.19	-0.22	-0.07	0.04	0.13	0.04	0.05	-0.01

Note: Positive values indicate that the MSPE is lower for model (2) as compared to a model using recursively estimated historical mean as a predictor. Shaded areas indicate cases with insignificant CW07 statistics.

4.3. *Time variations in the predictive power of quantile .75 returns*

To assess the time-varying nature of out-of-sample predictive power of returns at quantile .75, we apply the CW07 testing approach to model (2) vs. (1) recursively, in both directions. First, we fix the starting point and allow the sample window to increase by one observation, until it reaches the sample end; for each window, divided into estimation (in-sample) and prediction (OOS) subperiods, the CW07 test is performed, which results in a time series of test statistics and their corresponding p-values for the one-sided test. Similarly, we perform the recursive estimation with the fixed end point date while allowing the starting point to move over time (from the beginning of the sample). The smallest sample in each approach contains 40 observations, and we concentrate on cases where the out-of-sample period is long (π=3), due to superior size properties of the OOS tests (Hansen and Timmermann, 2012). The results from these forward and backwards recursive estimations are presented in Figures 2 and 3, respectively. More specifically, these figures show the estimated right-hand side, one-sided p-values of the CW07 tests. The last (first) period in Figure 2 (Figure 3) corresponds to the full-sample test. Figure 2 (fixed starting date, moving sample end date) shows most p-values to be higher than 10% for subsamples ending in the early part of the sample, which could suggest weak predictive power of quantile .75 returns until at least the mid-1940s. On the other hand, Figure 3 (fixed end date, moving sample start date) shows most p-values to be below 10% for subsamples starting in the 1930s, suggesting

Figure 2: p-values for CW07 statistics, recursive method (start-of-sample fixed)

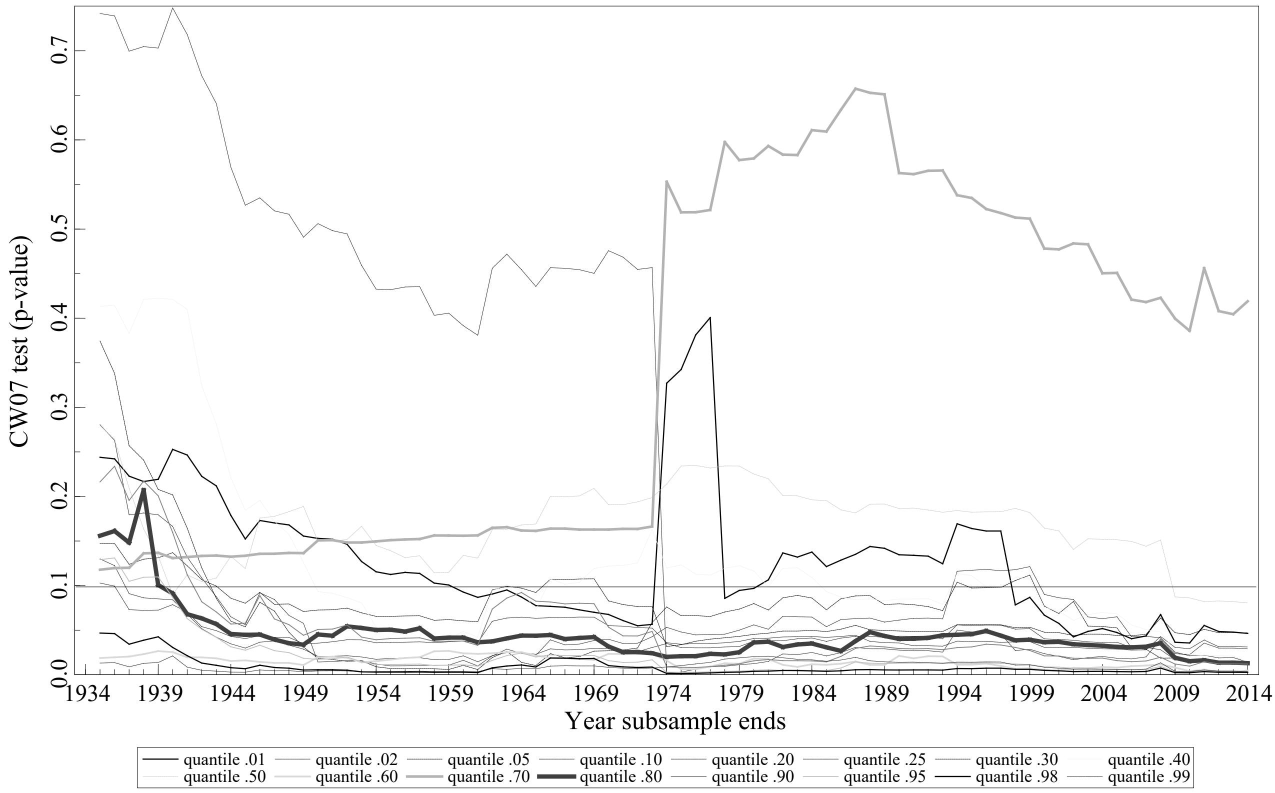
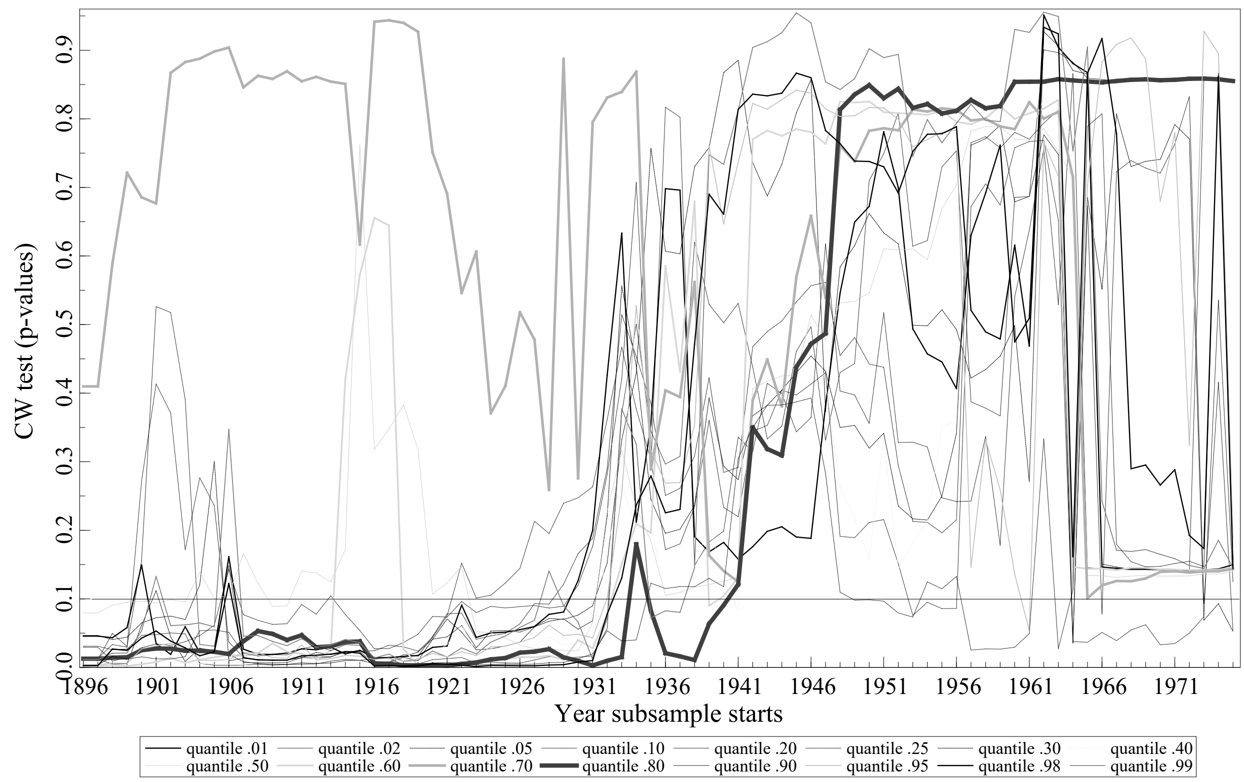


Figure 3: p-values for CW07 statistics, backwards recursive method (end-of-sample fixed)



better predictive power in the earlier part of the sample, but poorer in the later part of the sample. Considered separately, these two figures appear to provide contradictory evidence on the timing of predictive ability of quantile .75 returns. However, taken together, these non-rejections of the null of the predictive power in the early (Fig. 2) and late (Fig. 3) parts of the sample appear to be due to poor power of the CW07 tests when applied to short subsamples, as the underlying coefficients would have been imprecise (Campbell and Thompson, 2008, Cochrane, 2008). As the sample in the recursive estimations becomes larger, the coefficients estimated for forecasts are based on an increasing number of observations, hence becoming less prone to biases and generating a more reliable picture of the forecasting power of the model. Therefore, when we look at those results obtained from larger subsamples and hence the more precise and reliable (the RHS (LHS) side of Figure 2 (3)), most p-values are below the 10% (and even 5%) level, and we can conclude that the quantile 0.75 returns' predictive power was strong across the whole sample period and not confined to a narrow subperiod.

4.4. Predictive power of quantile .75 returns vs. other variables

We further analyse whether our predictive variable, the lagged index return measured at the 75th percentile of index return distribution, possesses predictive power for one-year ahead index returns which is not captured by other economic variables, as identified in the literature. To this end, we utilise the variables considered by Welch and Goyal (2008): Book-to-market ratio (bm), Default return spread (dfr), Default yield spread (dfy), Dividend payout ratio (de), Dividend price ratio (dp), Dividend yield (dy), Earnings price ratio (ep), Inflation (infl), Long-term return (ltr), Long-term yield (lty), Net equity expansion (ntis), Percentage equity issuing (eqis), Stock volatility (of S&P500, svar), Term spread (tms), Investment to capital ratio (ik), and Treasury-bill rate (tbl). In addition, we employ the tail risk measure from Kelly and Jiang (2014) (tail). Following the literature, we run single regressions, with each of the quantile returns as the dependent and one-year lag of each of the abovementioned variables, and our quantile .75 return, as explanatory ones, one variable at a time. We also estimate models with two independent variables at a time, one being always the lagged quantile .75 return, and the other one a variable from those listed above. All variables except the dependent ones and the quantile .75 return are standardised.

Table 7 presents the results from simple regressions for samples which vary across regressions, as different variables are available from different points in time (results for a unified sample starting in year 1963 are not reported to conserve space). The parameters are multiplied by 10,000 and R^2 statistics are expressed in percentage points, for ease of exposition. Looking at samples of varying lengths (Table 7), most parameters on lagged predictive variables appear to be insignificant. From 17 variables used here, based on prior research, only a few show a consistent pattern of significance, mostly the default yield spread and volatility (both for most quantiles of future returns) and, to a lesser extent, the investment to capital ratio (most cases) and the dividend payout ratio (only a few cases). Interestingly, when it comes to the predictive ability for the center of future return distribution, as measured by median returns here, variables such as net equity expansion and tail risk, in addition to investment to capital ratio, turn out to be significant, whereas default yield spread and volatility are not. Hence, the relevant set of significant predictors for the center of the future return distribution is different from that for predicting any other quantile of that distribution. When it comes to our variable, the lagged 75th quantile returns, it is a significant predictor for all studied quantiles of future return, with an exception of the future median. Its predictive power seems to be superior to that of the remaining variables, as the R^2 values from models with the 75th quantile return as an explanatory variable are in most cases higher than those from models using other predictors, and in cases where they are not, they came a close

Table 7: Simple predictive regressions, varying samples (full data utilisation).

Quantile	1			2			5			10			20			25			30			40			50		
	beta	t-stat	R^2	beta	t-stat	R^2	beta	t-stat	R^2	beta	t-stat	R^2	beta	t-stat	R^2	beta	t-stat	R^2	beta	t-stat	R^2	beta	t-stat	R^2	beta	t-stat	R^2
dfy	7.02	**3.94**	19.68	6.11	**4.85**	28.06	4.88	**5.24**	31.67	3.89	**4.02**	37.78	2.53	**3.57**	35.92	1.86	**3.47**	30.76	1.45	**3.19**	28.16	0.74	**4.17**	20.90	0.13	0.88	1.45
infl	3.85	1.16	5.62	3.29	1.24	8.36	2.44	1.12	8.25	1.72	0.99	7.73	0.92	0.81	5.02	0.63	0.71	3.73	0.41	0.57	2.41	0.20	0.49	1.50	0.02	0.20	0.04
svar	-6.16	**-4.22**	16.94	-5.26	**-5.25**	24.32	-4.06	**-5.58**	26.37	-3.19	**-5.87**	31.13	-1.91	**-4.69**	25.55	-1.42	**-4.28**	22.25	-1.12	**-4.04**	21.04	-0.58	**-4.10**	14.89	-0.07	-0.72	0.44
de	-4.64	**-2.74**	8.63	-3.67	**-2.96**	10.66	-2.73	-1.74	10.70	-1.89	-1.44	9.82	-1.17	-1.33	8.69	-0.82	-1.19	6.64	-0.58	-1.01	4.97	-0.31	-0.95	3.72	0.00	-0.01	0.00
lty	3.03	1.23	3.66	1.98	1.04	2.94	0.96	0.66	1.22	0.32	0.30	0.26	-0.12	-0.17	0.07	-0.08	-0.15	0.06	-0.17	-0.40	0.40	-0.16	-0.65	0.93	-0.10	-0.88	0.82
tms	-0.48	-0.23	0.09	-0.62	-0.40	0.29	-0.26	-0.21	0.09	-0.28	-0.32	0.19	-0.12	-0.20	0.08	-0.04	-0.09	0.02	-0.05	-0.12	0.03	-0.02	-0.10	0.02	-0.03	-0.25	0.07
tbl	2.88	1.23	3.31	2.03	1.12	3.08	0.96	0.68	1.21	0.41	0.39	0.42	-0.05	-0.08	0.01	-0.06	-0.11	0.03	-0.13	-0.32	0.24	-0.13	-0.55	0.64	-0.07	-0.66	0.47
dfr	2.34	1.62	2.05	1.47	1.46	1.53	0.91	1.23	1.04	0.30	0.57	0.22	0.14	0.38	0.10	0.04	0.12	0.01	-0.04	-0.15	0.02	-0.08	-0.57	0.25	-0.16	-1.42	2.28
dp	-1.55	-0.75	1.06	-1.17	-0.75	1.20	-0.74	-0.63	0.86	-0.57	-0.53	0.98	-0.40	-0.53	1.09	-0.27	-0.46	0.78	-0.23	-0.50	0.91	-0.12	-0.44	0.60	0.05	0.49	0.20
dy	-0.55	-0.26	0.14	-0.45	-0.28	0.19	0.06	0.05	0.01	0.23	0.26	0.16	0.24	0.44	0.42	0.23	0.53	0.59	0.18	0.52	0.54	0.10	0.55	0.48	0.14	1.50	1.88
ltr	-0.72	-0.41	0.19	-0.80	-0.64	0.45	-0.30	-0.32	0.11	-0.04	-0.06	0.00	0.06	0.13	0.02	0.12	0.33	0.12	0.11	0.39	0.16	0.14	0.86	0.76	0.20	1.85	3.83
ep	1.69	0.90	1.34	1.40	1.00	1.80	1.18	1.15	2.35	0.76	1.00	1.84	0.42	0.87	1.32	0.31	0.80	1.08	0.16	0.54	0.47	0.10	0.55	0.42	0.05	0.56	0.27
bm	0.49	0.20	0.10	0.12	0.06	0.01	-0.26	-0.18	0.09	-0.55	-0.40	0.74	-0.55	-0.83	1.71	-0.37	-0.71	1.24	-0.41	-1.01	2.31	-0.28	-1.23	3.00	-0.06	-0.52	0.29
ik	-1.65	-0.97	2.32	-1.09	-0.85	1.78	-1.47	-1.71	6.57	-1.21	**-2.02**	9.15	-1.04	**-2.75**	14.88	-0.90	**-2.93**	16.09	-0.81	**-3.43**	19.10	-0.54	**-4.27**	21.94	-0.23	**-2.58**	7.94
ntis	-0.52	-0.24	0.10	-0.57	-0.35	0.23	-0.98	-0.79	1.19	-0.87	-0.96	1.78	-0.68	-1.16	2.47	-0.70	-1.52	4.17	-0.61	-1.67	4.82	-0.38	-1.58	5.20	-0.23	**-2.05**	4.65
eqis	-2.36	-1.07	2.06	-1.77	-1.07	2.20	-1.64	-1.32	3.32	-1.28	-1.40	3.79	-0.91	-1.54	4.37	-0.76	-1.63	4.82	-0.65	-1.76	5.41	-0.40	-1.89	5.74	-0.13	-1.14	1.50
tail	0.05	0.02	0.00	-0.33	-0.20	0.14	0.17	0.15	0.07	0.32	0.44	0.60	0.33	0.71	1.42	0.41	1.07	3.08	0.37	1.34	3.70	0.30	1.80	6.45	0.26	**2.17**	9.14
q75 ret	-3.16	**-4.59**	25.62	-2.65	**-5.51**	35.86	-2.02	**-4.90**	38.10	-1.53	**-4.77**	41.38	-0.94	**-3.89**	35.44	-0.72	**-3.96**	32.75	-0.56	**-3.89**	29.40	-0.30	**-4.01**	22.50	-0.04	-0.91	0.95

Table 7 continued

Quantile	60			70			75			80			90			95			98			99		
	beta	t-stat	R^2	beta	t-stat	R^2	beta	t-stat	R^2	beta	t-stat	R^2	beta	t-stat	R^2	beta	t-stat	R^2	beta	t-stat	R^2	beta	t-stat	R^2
dfy	-0.50	**-4.30**	21.37	-1.19	**-4.39**	40.59	-1.78	**-4.26**	47.85	-2.40	**-4.55**	49.55	-4.06	**-4.57**	51.36	-5.28	**-4.59**	45.44	-7.07	**-3.85**	39.11	-8.50	**-3.75**	31.57
infl	-0.29	-1.34	6.60	-0.56	-1.08	8.78	-0.80	-1.05	9.66	-0.97	-0.99	8.36	-1.58	-0.97	8.02	-2.18	-1.02	8.02	-3.12	-0.98	7.89	-4.44	-1.12	9.00
svar	0.47	**3.04**	18.99	0.98	**6.50**	31.72	1.46	**6.09**	37.87	2.00	**7.36**	41.56	3.28	**7.72**	40.03	4.47	**7.65**	38.83	5.89	**6.32**	32.48	7.27	**6.03**	28.47
de	0.34	**2.49**	9.20	0.66	1.84	13.03	0.97	1.75	14.96	1.25	1.71	14.63	2.11	1.74	14.88	2.59	1.59	11.72	3.61	1.52	10.99	4.20	1.35	8.56
lty	-0.13	-0.82	1.51	-0.14	-0.43	0.57	-0.12	-0.25	0.20	-0.13	-0.21	0.14	-0.10	-0.11	0.03	0.22	0.17	0.08	-0.06	-0.03	0.00	0.05	0.02	0.00
tms	-0.03	-0.19	0.06	0.06	0.23	0.10	0.08	0.22	0.09	0.26	0.55	0.58	0.65	0.83	1.30	1.22	1.14	2.40	0.84	0.54	0.55	1.42	0.69	0.87
tbl	-0.11	-0.67	0.94	-0.15	-0.48	0.66	-0.14	-0.31	0.29	-0.23	-0.40	0.46	-0.39	-0.42	0.48	-0.37	-0.29	0.22	-0.44	-0.24	0.15	-0.61	-0.25	0.16
dfr	-0.22	**-2.11**	3.90	-0.20	-1.24	1.02	-0.24	-1.15	0.82	-0.35	-1.24	0.97	-0.48	-1.02	0.67	-0.72	-1.09	0.78	-0.73	-0.76	0.39	-0.84	-0.64	0.29
dp	0.20	1.45	3.32	0.32	0.98	3.34	0.46	0.96	3.73	0.56	0.90	3.27	0.88	0.83	2.84	0.84	0.60	1.35	1.28	0.66	1.53	0.86	0.34	0.39
dy	0.13	0.94	1.52	0.11	0.41	0.42	0.14	0.38	0.37	0.09	0.19	0.09	-0.04	-0.06	0.01	-0.35	-0.32	0.24	-0.23	-0.14	0.05	-1.03	-0.51	0.58
ltr	0.22	1.92	4.08	0.23	1.15	1.45	0.28	0.99	1.07	0.28	0.77	0.65	0.34	0.55	0.33	0.66	0.78	0.66	0.36	0.30	0.10	0.63	0.38	0.16
ep	-0.04	-0.29	0.12	-0.13	-0.57	0.62	-0.20	-0.62	0.77	-0.30	-0.70	0.96	-0.58	-0.85	1.31	-0.97	-1.04	1.93	-1.23	-0.91	1.50	-2.13	-1.22	2.57
bm	0.13	0.58	1.50	0.25	0.55	1.75	0.48	0.71	3.35	0.67	0.77	3.85	1.14	0.81	3.99	1.27	0.65	2.59	1.58	0.61	1.92	1.25	0.52	0.67
ik	0.02	0.23	0.08	0.27	1.55	5.41	0.47	**2.11**	10.15	0.65	**2.08**	10.02	1.03	1.81	8.40	1.40	1.61	6.72	2.20	**1.99**	9.35	2.74	1.78	6.88
ntis	0.00	0.02	0.00	0.18	0.67	0.87	0.23	0.61	0.76	0.13	0.25	0.13	0.07	0.08	0.01	-0.12	-0.11	0.02	0.57	0.34	0.23	1.30	0.61	0.69
eqis	0.05	0.34	0.18	0.26	0.96	1.82	0.33	0.86	1.52	0.39	0.77	1.19	0.64	0.77	1.19	0.81	0.70	0.98	1.78	1.09	2.30	2.31	1.07	2.16
tail	0.17	1.38	3.91	0.11	0.48	0.83	0.12	0.38	0.55	0.10	0.24	0.20	0.00	0.00	0.00	0.24	0.23	0.19	-0.17	-0.13	0.05	0.19	0.10	0.03
q75 ret	0.22	**3.41**	24.06	0.48	**5.03**	42.86	0.70	**5.19**	50.05	0.91	**5.37**	49.89	1.43	**4.79**	44.68	1.96	**4.84**	44.06	2.75	**4.64**	41.24	3.47	**4.69**	36.88

Note: We further analyse whether our predictive variable, the lagged index return measured at the 75th percentile of index return distribution, possesses predictive power for one-year ahead index returns which is not captured by other economic variables, as identified in the literature. We utilise the variables considered by Welch and Goyal (2008): Book-to-market ratio (bm), Default return spread (dfr), Default yield spread (dfy), Dividend payout ratio (de), Dividend price ratio (dp), Dividend yield (dy), Earnings price ratio (ep), Inflation (infl), Long-term return (ltr), Long-term yield (lty), Net equity expansion (ntis), Percentage equity issuing (eqis), Stock volatility (of S&P500, svar), Term spread (tms), Investment to capital ratio (ik), and Treasury-bill rate (tbl). In addition, we employ the tail risk measure from Kelly and Jiang (2014) (tail). We run single regressions, with each of the quantile returns as the dependent and one-year lag of each of the abovementioned variables, and our quantile .75 return, as explanatory ones, one variable at a time. We also estimate models with two independent variables at a time, one being always the lagged quantile .75 return, and the other one a variable from those listed above. All variables except the dependent ones and the quantile .75 return are standardised. Tables 7 and 8 present the results from simple regressions for samples which vary across regressions, as different variables are available from different points in time (Table 7) and for a unified sample period starting in year 1963 (Table 8). Shaded areas indicate significance at 10% level, bold values additionally at 5% level. Values under "beta" were multiplied by a factor of 10,000.

second (except for models predicting the median). Hence, the evidence suggests that our variable possesses predictive power for the future return distribution which is not contained in any of the alternative predictors, and is superior to those alternative predictors.

The unreported results for a unified sample starting in year 1963 show that only a small number of previously identified economic variables possess predictive power for future returns, slightly more so for those returns in the upper part of the distribution. Volatility appears to have maintained its predictive power as compared to the whole sample, but the default yield spread seems to be more predictive for returns in the upper part of the distribution. The dividend payout ratio appears to have lost its predictive power, and investment to capital ratio can only predict a few quantiles of future returns. On the other hand, net equity expansion and default return spread emerge as significant predictors for several quantile returns. As above, the set of predictors for median return is different from that for other quantiles: Net equity expansion, percentage equity issuing, investment to capital ratio, and tail risk each appear to significantly predict future median returns. Most importantly, however, the lagged 75^{th} quantile return is a significant predictor for all but the median future returns, and its predictive power, as measured by R^2 values, tends to be the best, especially for upper quantiles of future return distribution.

Overall, we can conclude that the predictive power of lagged returns measured at the 75^{th} quantile is not captured by any of the traditional predictors for future returns. It appears to be stable over time, unlike that of some other predictors. However, quantile .75 returns don't have predictive power for the center of the future return distribution.

In addition to those simple predictive regressions analysed above, we also estimated models with two independent variables at a time, one being always the lagged quantile 75^{th} return, and the other one a variable from Welch and Goyal (2008) or the tail risk of Kelly and Jiang (2014). The results in Table 8 derived from unequal samples and utilising all available data for each variable clearly show that the quantile .75 lagged return is a significant predictor of one-year ahead returns at all quantiles except for the median (the relevant estimated parameter values are reported under "*q beta*" and the corresponding t-statistics under "*q tstat*"). Other variables tend to be insignificant; for each dependent variable, only up to six out of seventeen predictors are significant. Compared to the simple regressions, the default yield spread and the investment to capital ratio remain significant across a wide range of quantiles, whereas dividend yield appears to have lost some of its predictive ability across the quantiles of future returns. Volatility maintained its predictive ability when combined with quantile .75 returns only for some high quantiles of future returns. On the other hand, default return spread emerges as a significant predictor of mostly high quantile returns. As before, predictive regressions for the median fare much worse than those for other quantiles, both in terms of the number of significant predictors and the R^2 values of those models. Hence, predicting the future center of return distribution appears to be much more difficult than other parts of it.

When we look at the homogenised sample starting in year 1963 (unreported), the predictive pattern is somewhat similar to those in simple models. Default return spread and, to a lesser degree, net equity expansion and investment to capital ratio predict future returns at different, if not all, quantiles of the distribution; however, volatility seems to have lost its predictive power when combined with quantile .75 returns in one model. As for the latter variable, it significantly predicts future returns in all but central (.50 and .60) quantiles in almost all models. Interestingly, it is insignificant when combined with mostly the volatility, although the latter does not show any significance either. As before, the model predicting the median suffers from the worst performance as compared to those predicting other quantiles of future return distribution.

Table 8: Predictive regressions with two predictors, varying samples (full data utilisation).

Quantile	1					2					5					10					20					25				
	q beta	q tstat	beta	tstat	R^2	q beta	q tstat	beta	tstat	R^2	q beta	q tstat	beta	tstat	R^2	q beta	q tstat	beta	tstat	R^2	q beta	q tstat	beta	tstat	R^2	q beta	q tstat	beta	tstat	R^2
dfy	-2.92	-3.84	17.75	0.90	31.70	-2.28	-4.44	20.22	1.52	41.79	-1.69	-3.13	18.53	1.67	44.99	-1.20	-3.20	17.39	1.85	50.44	-0.64	-2.47	13.92	1.79	43.90	-0.53	-2.70	9.17	1.68	39.43
infl	-3.07	-4.78	16.61	1.03	27.28	-2.56	-6.36	14.74	1.43	38.87	-2.00	-5.81	10.10	1.02	42.00	-1.55	-6.31	6.14	0.79	45.90	-0.97	-5.61	2.37	0.43	38.15	-0.75	-5.41	1.00	0.23	35.37
svar	-3.10	-3.29	-1.77	-0.08	25.62	-2.50	-3.26	-4.26	-0.29	35.91	-1.87	-3.62	-4.38	-0.52	38.21	-1.29	-3.59	-7.06	-1.37	41.90	-0.84	-4.07	-2.86	-0.86	35.64	-0.69	-4.38	-0.85	-0.34	32.78
de	-2.90	-4.88	-17.41	-1.20	26.69	-2.47	-6.48	-12.07	-1.29	36.88	-1.90	-6.14	-8.36	-1.31	38.99	-1.46	-6.59	-4.46	-0.80	41.86	-0.89	-6.00	-2.95	-0.67	35.92	-0.70	-5.94	-1.31	-0.38	32.90
lty	-3.34	-5.98	25.13	1.71	33.55	-2.78	-5.79	15.47	**2.44**	41.98	-2.17	-5.30	6.21	1.44	43.12	-1.66	-5.42	0.67	0.21	46.52	-1.01	-4.69	-2.72	-1.19	38.66	-0.78	-4.52	-2.04	-1.03	35.91
tms	-3.50	-5.78	13.23	0.87	31.29	-2.88	-5.50	8.62	0.79	40.24	-2.24	-5.15	9.00	1.33	43.00	-1.71	-5.28	6.03	1.37	46.83	-1.03	-4.66	4.14	1.64	38.36	-0.79	-4.53	3.64	1.76	35.76
tbl	-3.30	-5.70	16.47	1.09	31.68	-2.76	-5.46	9.92	1.77	40.42	-2.16	-5.09	1.45	0.40	41.99	-1.67	-5.31	-2.16	-0.82	46.08	-1.02	-4.69	-4.34	**-2.19**	38.47	-0.78	-4.52	-3.50	**-2.06**	35.70
dfr	-3.56	-6.61	33.66	**2.34**	35.33	-2.91	-7.31	23.07	**2.51**	43.52	-2.24	-5.61	15.58	**2.28**	45.10	-1.70	-5.61	7.93	1.78	47.79	-1.02	-4.75	4.30	1.27	38.92	-0.78	-4.53	2.59	0.99	35.74
dp	-3.13	-5.53	-6.97	-0.50	25.84	-2.63	-7.26	-4.57	-0.52	36.04	-2.01	-5.99	-1.89	-0.29	38.16	-1.52	-5.87	-1.59	-0.30	41.46	-0.93	-5.32	-1.46	-0.36	35.59	-0.72	-5.18	-0.75	-0.24	32.82
dy	-3.16	-5.56	1.24	0.09	25.63	-2.65	-6.73	1.09	0.15	35.87	-2.04	-5.86	4.91	1.04	38.51	-1.55	-5.65	5.52	1.97	42.35	-0.95	-5.15	4.42	**2.39**	36.86	-0.74	-5.05	3.84	**2.45**	34.43
ltr	-3.42	-5.63	-3.45	-0.24	31.19	-2.81	-5.58	-4.95	-0.58	39.99	-2.18	-5.04	-0.59	-0.09	42.11	-1.67	-5.16	1.43	0.30	46.39	-1.01	-4.44	1.68	0.47	38.10	-0.77	-4.28	2.02	0.79	35.52
ep	-3.13	-5.43	3.44	0.26	25.67	-2.63	-6.44	2.67	0.36	35.92	-2.00	-5.51	3.28	0.69	38.28	-1.52	-5.44	1.10	0.32	41.42	-0.94	-4.96	0.27	0.11	35.45	-0.72	-4.87	0.00	0.00	32.75
bm	-3.56	-4.22	21.45	1.34	32.66	-2.92	-5.88	14.77	1.59	41.44	-2.23	-5.30	7.79	1.28	42.83	-1.68	-5.34	2.32	0.57	46.13	-0.99	-4.54	-0.93	-0.31	37.47	-0.76	-4.34	-0.21	-0.09	34.61
ik	-2.61	-2.88	-13.25	-0.95	15.08	-2.09	-3.09	-8.30	-0.80	16.22	-1.52	-3.53	-12.87	**-2.02**	21.79	-1.17	-4.11	-10.62	**-2.50**	28.09	-0.74	-3.93	-9.49	**-3.39**	31.43	-0.55	-3.47	-8.28	**-3.52**	29.35
ntis	-3.43	-5.63	1.57	0.10	31.15	-2.82	-5.93	-0.18	-0.01	39.81	-2.16	-5.62	-5.50	-0.42	42.48	-1.66	-5.74	-5.46	-0.71	47.03	-0.99	-4.95	-4.88	-0.90	39.20	-0.75	-4.99	-5.55	-1.17	37.76
eqis	-3.39	-4.14	-12.28	-0.49	32.16	-2.80	-6.51	-8.41	-0.55	40.78	-2.15	-5.74	-9.23	-0.96	43.51	-1.64	-5.63	-7.31	-1.39	47.93	-0.98	-4.67	-5.84	**-2.23**	40.25	-0.75	-4.55	-5.09	**-2.25**	37.87
tail	-2.56	-2.46	4.60	0.28	11.63	-2.04	-2.46	-0.05	0.00	13.32	-1.44	-2.64	3.96	0.46	13.26	-1.11	c	5.02	0.89	17.76	-0.69	-2.89	4.44	1.17	16.57	-0.51	-2.56	4.90	1.54	15.18

Table 8 continued

Quantile	30					40					50					60					70					75				
	q beta	q tstat	beta	tstat	R^2	q beta	q tstat	beta	tstat	R^2	q beta	q tstat	beta	tstat	R^2	q beta	q tstat	beta	tstat	R^2	q beta	q tstat	beta	tstat	R^2	q beta	q tstat	beta	tstat	R^2
dfy	-0.42	-2.84	7.00	1.57	36.34	-0.24	-3.45	3.12	1.24	28.61	-0.04	-0.76	0.51	0.30	2.01	0.17	2.03	-1.93	-1.53	30.24	0.35	3.20	-5.58	**-2.01**	53.19	0.50	3.92	-8.80	**-2.41**	61.29
infl	-0.60	-5.47	-0.12	-0.03	32.90	-0.33	-5.75	-0.38	-0.18	26.87	-0.07	-1.45	-0.27	-0.17	2.46	0.19	2.80	-1.49	-1.06	24.53	0.45	4.65	-2.35	-1.09	43.90	0.68	5.03	-3.16	-1.15	51.65
svar	-0.51	-3.89	-1.38	-0.64	29.50	-0.30	-3.49	0.02	0.01	22.50	-0.06	-0.71	0.41	0.26	1.00	0.18	2.55	1.17	0.59	24.45	0.42	4.23	1.63	0.64	43.15	0.59	4.99	3.19	1.00	50.66
de	-0.55	-5.81	-0.40	-0.14	29.42	-0.30	-5.94	-0.19	-0.11	22.52	-0.05	-1.15	0.51	0.47	1.14	0.20	3.10	1.53	1.56	25.63	0.44	5.00	2.38	1.49	44.31	0.65	5.38	3.35	1.57	51.62
lty	-0.61	-4.40	-2.67	-1.69	33.84	-0.33	-5.41	-2.08	-1.34	28.32	-0.06	-1.39	-1.05	-0.96	2.87	0.22	3.09	-1.00	-1.08	29.43	0.50	4.69	-0.64	-0.52	48.68	0.74	4.88	-0.02	-0.01	55.20
tms	-0.61	-4.41	2.71	1.55	32.97	-0.33	-4.55	1.48	1.29	26.49	-0.05	-0.94	-0.02	-0.02	1.52	0.24	3.30	-1.51	-1.25	31.37	0.53	4.78	-2.11	-1.29	50.38	0.77	4.98	-3.19	-1.73	57.03
tbl	-0.61	-4.43	-3.63	**-2.59**	33.76	-0.33	-5.38	-2.54	-1.60	28.11	-0.06	-1.30	-0.93	-0.84	2.28	0.23	3.14	-0.20	-0.19	29.59	0.51	4.69	0.39	0.28	49.21	0.75	4.91	1.44	0.86	55.88
dfr	-0.61	-4.42	1.38	0.67	33.15	-0.32	-4.49	0.10	0.07	26.93	-0.05	-1.15	-1.44	-1.29	3.77	0.24	3.98	-2.87	**-2.80**	36.74	0.53	5.21	-3.46	**-2.55**	52.68	0.77	5.42	-4.61	**-2.38**	58.70
dp	-0.55	-4.86	-0.88	-0.33	29.53	-0.30	-4.87	-0.38	-0.24	22.57	-0.04	-1.11	0.59	0.62	1.28	0.22	3.43	1.40	1.70	25.73	0.47	5.17	1.97	1.42	44.12	0.69	5.40	2.77	1.46	51.42
dy	-0.57	-4.68	2.96	**2.17**	30.90	-0.31	-4.69	1.67	1.71	23.76	-0.05	-1.19	1.50	1.61	3.12	0.22	3.38	0.86	1.16	24.71	0.48	5.05	0.14	0.14	42.86	0.70	5.23	-0.02	-0.02	50.05
ltr	-0.60	-4.20	1.77	0.91	33.30	-0.32	-4.37	1.80	1.68	28.11	-0.06	-1.38	2.11	1.92	5.95	0.22	3.16	1.98	**2.20**	33.30	0.51	4.51	1.78	1.17	50.37	0.74	4.67	1.95	0.99	56.28
ep	-0.56	-4.62	-0.72	-0.39	29.49	-0.30	-4.62	-0.30	-0.27	22.55	-0.04	-0.97	0.34	0.36	1.06	0.23	3.50	0.57	0.73	24.35	0.48	5.06	0.66	0.65	43.01	0.71	5.22	0.94	0.72	50.21
bm	-0.58	-4.13	-1.43	-0.73	32.25	-0.31	-4.13	-1.39	-1.13	26.38	-0.05	-0.87	-0.35	-0.32	1.59	0.23	3.17	0.28	0.24	29.68	0.51	4.75	0.13	0.08	49.12	0.74	4.97	1.34	0.58	55.78
ik	-0.41	-3.11	-7.55	**-3.90**	29.75	-0.24	-3.10	-5.08	**-4.23**	31.41	-0.03	-0.47	-2.22	**-2.51**	8.25	0.14	2.12	0.05	0.06	6.61	0.37	4.40	2.27	1.81	27.38	0.52	5.27	4.09	**2.77**	37.33
ntis	-0.59	-4.90	-4.97	-1.46	36.06	-0.31	-5.07	-3.15	-1.62	30.55	-0.05	-0.96	-2.16	-0.65	6.14	0.23	3.02	-0.43	-0.28	30.23	0.51	4.45	0.81	0.49	49.70	0.75	4.60	0.85	0.31	55.85
eqis	-0.58	-4.39	-4.56	**-2.55**	36.18	-0.31	-4.37	-2.92	**-2.14**	30.83	-0.05	-0.96	-1.11	-0.68	3.14	0.23	3.03	-0.29	-0.21	30.04	0.51	4.43	0.93	0.52	49.79	0.74	4.60	0.82	0.38	55.91
tail	-0.39	-2.31	4.33	1.63	13.68	-0.23	-2.31	3.38	**2.11**	16.16	-0.03	-0.36	2.61	**2.18**	9.40	0.13	1.80	1.45	1.23	10.25	0.37	3.64	0.53	0.33	22.99	0.53	4.23	0.32	0.16	28.40

.

Table 8 continued

Quantile	80					90					95					98					99				
	q beta	q tstat	beta	tstat	R^2	q beta	q tstat	beta	tstat	R^2	q beta	q tstat	beta	tstat	R^2	q beta	q tstat	beta	tstat	R^2	q beta	q tstat	beta	tstat	R^2
dfy	0.64	4.22	-12.51	**-2.73**	61.94	0.92	3.53	-24.14	**-2.74**	60.56	1.39	4.20	-27.93	**-2.80**	56.48	2.04	3.54	-34.07	**-2.05**	50.60	2.75	3.52	-35.65	-1.76	43.20
infl	0.90	5.34	-3.32	-0.99	51.84	1.43	4.86	-5.66	-0.99	47.55	1.95	4.85	-7.93	-1.15	46.80	2.71	6.45	-11.97	-0.80	43.75	3.32	6.06	-20.77	-1.08	39.43
svar	0.67	4.46	6.92	1.71	51.57	0.88	3.50	15.80	**2.34**	47.91	1.23	3.50	20.90	**2.33**	47.02	2.16	3.23	17.16	1.76	42.18	2.88	3.40	16.95	0.93	37.40
de	0.85	5.64	4.23	1.49	51.36	1.31	5.12	7.96	1.47	46.59	1.85	5.12	7.29	1.07	44.90	2.60	6.69	10.35	0.97	42.04	3.32	6.62	9.84	0.66	37.28
lty	0.97	5.11	0.24	0.12	54.90	1.56	6.50	1.36	0.59	51.11	2.15	6.63	5.54	1.55	50.35	2.96	5.51	3.99	0.77	45.98	3.71	5.32	6.29	0.84	40.47
tms	1.00	5.12	-2.57	-1.08	55.81	1.59	4.54	-1.71	-0.47	51.44	2.15	6.32	1.12	0.23	50.20	3.03	5.33	-7.23	-0.89	46.34	3.78	5.10	-5.33	-0.51	40.66
tbl	0.99	5.12	1.39	0.61	55.43	1.59	4.62	2.01	0.54	51.48	2.18	4.75	4.46	0.83	50.49	3.02	5.52	6.89	1.57	46.31	3.79	5.36	8.08	1.35	40.82
dfr	1.01	5.74	-6.38	**-2.49**	58.77	1.63	4.97	-9.45	**-2.20**	54.61	2.23	5.29	-13.58	**-2.08**	53.78	3.08	4.66	-16.11	-1.66	48.91	3.87	4.57	-19.54	-1.38	43.40
dp	0.90	5.61	3.24	1.27	50.98	1.40	5.01	4.93	1.06	45.61	1.94	4.98	3.04	0.49	44.25	2.73	5.94	5.43	0.56	41.52	3.47	4.79	-0.64	-0.05	36.88
dy	0.92	5.43	-1.00	-0.56	49.99	1.44	4.88	-3.51	-1.15	45.17	1.98	4.96	-7.73	-1.62	45.30	2.78	5.63	-8.16	-1.40	41.89	3.53	4.87	-17.58	-1.93	38.59
ltr	0.98	4.90	1.76	0.67	55.80	1.59	4.42	1.63	0.37	52.12	2.17	4.56	4.22	0.73	51.29	3.01	5.40	0.35	0.04	47.04	3.79	4.16	2.10	0.19	41.88
ep	0.92	5.35	0.88	0.46	49.97	1.43	4.70	0.36	0.10	44.69	1.95	4.71	-1.30	-0.27	44.10	2.75	5.48	-0.61	-0.10	41.24	3.41	4.51	-7.04	-0.84	37.15
bm	0.97	5.24	2.24	0.75	55.68	1.55	6.79	4.22	0.98	51.86	2.14	6.93	2.79	0.45	50.28	2.97	5.51	2.03	0.25	46.02	3.78	5.33	-5.08	-0.52	40.74
ik	0.71	5.09	5.67	**2.73**	35.93	1.32	5.59	8.68	**2.48**	38.45	1.86	4.94	11.73	**2.10**	32.49	2.27	4.52	19.16	**2.56**	31.25	3.05	4.01	23.61	**2.09**	25.57
ntis	0.99	4.87	-0.68	-0.17	55.59	1.60	4.41	-2.45	-0.31	52.22	2.19	4.55	-5.56	-0.51	51.48	3.01	5.63	-0.27	-0.02	47.04	3.77	4.20	5.56	0.19	41.98
eqis	0.99	4.89	0.57	0.20	55.85	1.59	4.44	1.13	0.23	52.48	2.18	4.58	0.81	0.11	51.37	2.99	5.71	7.87	0.85	47.85	3.76	4.13	10.51	0.70	42.60
tail	0.71	3.99	-0.14	-0.05	25.86	1.29	4.39	-2.09	-0.45	29.55	1.71	3.69	-0.31	-0.04	22.96	2.05	3.24	-5.00	-0.50	18.61	2.89	3.07	-2.77	-0.19	17.02

Note: "q beta" ("q tstat") refer to estimated parameter values (t-statistics) for the lagged quantile .75 return as predictor, "beta" ("tstat") refer to those for the variable listed in column 1. For those latter variables only, shaded areas indicate significance at 10% level, bold values additionally at 5% level. Values under "beta" were multiplied by a factor of 10,000.

Overall, these results show the predictive power of lagged quantile .75 return which is consistent across the predicted return quantile as well as sample period, and is not captured by other variables.

5. Economic value of forecasts

Given the finding of the predictive power of quantile .75 returns, we investigate the economic implications of this finding.[11] Even if this variable does not predict the center of the distribution better than the competitors, and investors are interested in the center (e.g., the mean future return), there is evidence of superior predictive power for other quantiles of future return distribution. We propose to use these out-of-center superior predictions to estimate, rather than directly predict, the future distribution's center (as well as the volatility). To assess the economic values of those estimated predictions, we follow the framework employed in Marquering and Verbeek (2004), Ferreira and Santa-Clara (2011), and Neely et al. (2014), among others, and consider a risk-averse investor who re-allocates his wealth among stocks and bonds every year, based on his predictions of the next year's expected stock market return and volatility. It can be shown (see, Marquering and Verbeek, 2004) that for the expected utility function of a form $\hat{r}_{t+1} - \frac{1}{2}\gamma\hat{\sigma}^2_{t+1}$, the optimal weight for the fraction of wealth allocated to stocks at time *t* (for the holding period between *t* and *t+1*) is:

$$\omega_t = \left(\frac{1}{\gamma}\right)\left(\frac{\hat{r}_{t+1} - \widehat{rf}_{t+1}}{\hat{\sigma}^2_{t+1}}\right),$$

where $\hat{r}_{t+1}$ is the predicted stock market return between *t* and *t+1*, $\widehat{rf}_{t+1}$ is the predicted risk-free rate between *t* and *t+1*, $\hat{\sigma}^2_{t+1}$ is the predicted stock return volatility between *t* and *t+1,* and γ is the investor's risk aversion coefficient. The remaining fraction of investor's wealth (1-ω_t) is allocated to risk-free bonds. The resulting portfolio return at *t+1* is then given by:

---

[11] It is by no means obvious that a model with higher statistical forecasting performance will also generate forecasts with higher economic value, as Cenesizoglu and Timmermann (2012) demonstrate.

$$R_{P,t+1} = \omega_t r_{t+1} + (1 - \omega_t) rf_{t+1},$$

where r_{t+1} and rf_{t+1} are the realised values of the stock market and risk-free return, respectively. The certainty equivalent return (CER) for this investor can then be computed as:

$$CER_P = \hat{\mu}_P - \frac{1}{2}\gamma\hat{\sigma}_P^2,$$

where $\hat{\mu}_P$ and $\hat{\sigma}_P^2$ are the estimated mean and variance of the investor's portfolio returns in the evaluation period. The CER can be interpreted as a risk-free rate of return the investor would be willing to accept in exchange for his risky portfolio, or as a fee the investor would be willing to pay for access to the model generating forecasts $\hat{r}_{t+1}$ (Ferreira and Santa-Clara, 2011, Neely et al., 2014).

In our paper, the one-year ahead predicted mean stock return $\hat{r}_{t+1}$ and/or the accompanying predicted future volatility $\hat{\sigma}_{t+1}^2$ will be generated from the predicted quantile returns (as explained below), and the resulting CER values will be compared to those obtained from models using predictions based on other variables. Following Ferreira and Santa-Clara (2011), we assume the value of risk aversion coefficient γ to be two. To obtain more realistic values of weight, ω_t, we follow Campbell and Thompson (2008), Ferreira and Santa-Clara (2011) and Neely et al. (2014) and constrain ω_t to be between 0 and 1.5, these conditions exclude short selling and more than 50% leverage, respectively. Further, unless stated otherwise, the predicted variance equals the (moving) average of the last five period's variances (Campbell and Thompson, 2008, Neely et al., 2014).[12] Lastly, the next period's risk free rate (rf_{t+1}) is assumed to be known at period *t*, in line with the literature, and we employ data from Welch and Goyal (2008).

To obtain recursive forecasts of the next period's mean return, $\hat{r}_{t+1}$, using quantile .75 returns ($R_t(\theta = .75)$) in each period *t*, each of the to-be-predicted quantiles is regressed on a constant and the first lag of $R_t(\theta = .75)$. The resulting parameters as well as the current (period *t*) value of

[12] Kambouroudis and McMillan (2015) demonstrate that volatility forecasts based on short in-sample periods are most accurate.

$R_t(\theta = .75)$ are used to generate next year's (*t+1*) predicted value of returns at each considered quantile θ. This is done recursively; that is, in the sample beginning at a fixed date and running up to period *t,* with *t* expanding until the end of our data sample, and leaves us with a time series of annual predicted quantile returns for a set of quantiles, beginning 20 years after the sample's start (1963, to utilise data on all variables) to allow for the initial estimation window (the value of 20 is suggested by Ferreira and Santa-Clara, 2011). These predicted quantile returns (except for the median) are then employed to calculate the predicted mean return for each year in the prediction/OOS period.[13] In the next step, in the spirit of Rapach et al. (2010) and Meligkotsidou et al. (2014), we calculate the predicted mean at *t+1* as an equally weighted sum of *t+1* predicted quantile returns.[14] Two approaches are adopted here. First, all predicted quantile returns except for the median are averaged to obtain a prediction of next year's mean return, and, second, averages of pairs of symmetrical quantiles ($\theta =$.99 and .01, .98 and .02, etc.) are used for the same purpose.

For the sake of comparison, we use two alternative predictions for one-year ahead mean returns. First, the historical average of mean return, calculated recursively, is used as a prediction of the future return. Second, also calculated recursively, the mean index return is regressed on a lag of each economic variable, and the resulting coefficients and the current value of the relevant economic variable are used to predict the one-year ahead mean stock return.

To obtain forecasts of the next period's return volatility, $\hat{\sigma}_{t+1}^2$, using the quantile .75 return, we follow the approach proposed by Taylor (2005) using the estimated quantile return predictions. This approach utilises the feature that the distance between symmetric quantiles, θ and $1 - \theta$,

---

[13] We do not utilise the predicted median to infer about the same period's predicted mean as our previous results demonstrate that quantile .75 return has no predictive power for the next year's median. It does have predictive power for other quantiles of future return distribution, however.
[14] This is the simplest weighting scheme possible and more advanced techniques could be used, potentially to obtain forecasts that are more accurate; however, the literature demonstrated that a simple average often outperforms those more advanced approaches (see Meligkotsidou et al., 2014, or Timmermann (2006) and Aiolfi et al. (2011) who provide reviews of theoretical arguments and empirical evidence in support of simple forecast aggregation methods). However, our aim is to show that even such a naïve scheme can generate superior forecasts.

contains information about the variance of the distribution. Specifically, in the first step, our annual volatility measure, Vol_t, calculated as squared standard deviation of daily returns for each calendar year, is recursively regressed on squared symmetrical quantile returns:

$$Vol_t = \beta_0 + \beta_1 R_t^2(\theta) + \beta_2 R_t^2(1-\theta) + \beta_3 R_t^2(\theta) R_t^2(1-\theta) + u_t.$$

The resulting estimated parameter values $\hat{\beta}_0 - \hat{\beta}_3$ as well as quantile returns predicted using the 75$^{\text{th}}$ quantile return, $\hat{R}_{t+1}(\theta)$ and $\hat{R}_{t+1}(1-\theta)$, are substituted into the above model to obtain predictions of return volatility one period ahead. This procedure is repeated recursively, beginning 20 years after the sample's start to allow for the initial estimation window (the value corresponds to that of Ferreira and Santa-Clara, 2011), and results in a time series of predicted stock return volatilities which are based on predictions generated using $R_t(\theta = 0.75)$. This is done for all pairs of symmetrical quantiles (θ =.99 and 0.01, 0.98 and 0.02, etc.) considered in this paper. As noted above, the alternative volatility prediction follows the literature and equals the (moving) average of the last five period's variances (Campbell and Thompson (2008), Neely et al. (2014)).

First, we evaluate the economic values of quantile return forecasts based on predictive power of quantile 0.75 returns by comparing them to those based on (moving) historical mean. Table 9, Panel A, presents the differences in annualised CER as well as those for Sharpe ratios (SR). Volatility is in both cases predicted using a five-year moving average, i.e., any differences in economic value stem from differences in forecasting performance for the mean. It is evident that, when either all predicted quantiles or those pairs in the tails are used (θ =.99 and .01, .98 and .02, .95 and .05), the predictions based on quantile .75 returns are superior to those based on historical mean, both in terms of the CER and the SR. However, for the remaining quantile pairs (.90-.10 and those closer to the distribution's center) the historical mean performs better. The predictive power of quantile .75 based predictions for the mean return versus those based on economic variables is further explored and the results presented in Table 9, Panel B. Here the quantile-based predictions perform better when utilising predicted quantile returns from the tails, outperforming each variable, except for the tail risk when all predicted quantiles or the pair $\theta = \{.95, .05\}$ are used. When mean returns are predicted using quantile pairs closer to the center of the distribution, they tend to perform worse than those based on macro variables, except for the pair $\{.75, .05\}$. Overall, combining quantile predictions based on lagged quantile .75 returns to forecast the mean return appears to generate economically successful predictions, as compared to the alternative predictors.

The advantage of being able to predict several points of the return distribution by means of return quantiles, as compared to just the mean of the future distribution when using economic variables, is that a prediction of volatility is also possible. We first compare the predictive power of quantile returns for future volatility against that of moving five-year average of observed volatility. Predicted quantile returns are utilised to obtain predictions of future volatility using the approach by Taylor (2005), as explained above. When models with historical mean as predictor for future mean returns are considered, those utilising quantile pairs perform consistently better than those using five-year volatility average. For all quantile pairs considered, the CER values are 0.05693 for quantile based and 0.056853 for historical volatility-based portfolios, and the values of SP are 0.29635 and 0.29592, respectively (not tabulated).[15]

Lastly, we compare models using quantile predictions for both the mean and the volatility against models using the moving average of volatility and lagged value of each economic variable to forecast the volatility and mean, respectively. While the results are too numerous to report (CER and SP values are calculated for combinations of each way to predict the mean using quantile returns, for each way to predict the volatility using quantile returns, and for each economic variable), some

---

[15] All pairs generate identical CER and SP values as their constrained weights ω_t are identical. When we lift the restriction on the weights to be within 0 and 1.5, different quantile pairs generate different values of weights and CER and SP.

Table 9: Economic value of predictive power of quantile .75 based returns for the mean

Variable	Measure	Predicted returns quantiles								
		average all	.99 and .01	.98 and .02	.95 and .05	.90 and .10	.80 and .20	.75 and .25	.70 and .30	.60 and .40
Panel A: Evaluated against the historical mean and 5 years moving average volatility										
Historical average	CER	0.002	0.007	0.007	0.002	-0.028	-0.035	-0.005	-0.036	-0.028
	SR	0.018	0.042	0.045	0.018	-0.155	-0.209	-0.029	-0.221	-0.173
Panel B: Evaluated against the macro variables' predictions and 5 years moving average volatility										
DFY	CER	0.016	0.021	0.021	0.016	-0.014	-0.021	0.009	-0.023	-0.014
	SR	0.102	0.126	0.129	0.102	-0.071	-0.125	0.055	-0.138	-0.090
INFL	CER	0.011	0.016	0.017	0.011	-0.018	-0.026	0.004	-0.027	-0.019
	SR	0.071	0.095	0.098	0.071	-0.102	-0.156	0.024	-0.168	-0.120
SVAR	CER	0.022	0.027	0.027	0.022	-0.008	-0.015	0.015	-0.017	-0.008
	SR	0.130	0.155	0.157	0.130	-0.042	-0.097	0.083	-0.109	-0.061
DE	CER	0.013	0.018	0.018	0.013	-0.017	-0.024	0.006	-0.026	-0.017
	SR	0.078	0.102	0.104	0.078	-0.095	-0.149	0.031	-0.162	-0.114
LTY	CER	0.038	0.043	0.044	0.038	0.009	0.001	0.031	0.000	0.008
	SR	0.242	0.266	0.268	0.242	0.069	0.014	0.195	0.002	0.050
TMS	CER	0.010	0.015	0.015	0.010	-0.020	-0.027	0.003	-0.029	-0.020
	SR	0.058	0.082	0.085	0.058	-0.115	-0.169	0.011	-0.181	-0.133
TBL	CER	0.013	0.018	0.018	0.013	-0.017	-0.024	0.006	-0.026	-0.017
	SR	0.081	0.106	0.108	0.081	-0.091	-0.146	0.034	-0.158	-0.110
DFR	CER	0.009	0.014	0.014	0.009	-0.021	-0.028	0.002	-0.030	-0.021
	SR	0.066	0.091	0.093	0.066	-0.106	-0.161	0.019	-0.173	-0.125
DP	CER	0.045	0.049	0.050	0.045	0.015	0.007	0.038	0.006	0.014
	SR	0.285	0.309	0.312	0.285	0.112	0.058	0.238	0.046	0.094
DY	CER	0.042	0.047	0.047	0.042	0.012	0.005	0.035	0.003	0.012
	SR	0.265	0.289	0.292	0.265	0.092	0.038	0.218	0.025	0.074
LPR	CER	0.021	0.026	0.027	0.021	-0.009	-0.016	0.014	-0.017	-0.009
	SR	0.124	0.149	0.151	0.124	-0.048	-0.103	0.078	-0.115	-0.067
EP	CER	0.056	0.060	0.061	0.056	0.026	0.019	0.049	0.017	0.025
	SR	0.377	0.401	0.404	0.377	0.204	0.150	0.330	0.137	0.185
BM	CER	0.027	0.032	0.033	0.027	-0.002	-0.010	0.020	-0.011	-0.003
	SR	0.165	0.189	0.192	0.165	-0.008	-0.062	0.118	-0.075	-0.027
IK	CER	0.013	0.017	0.018	0.013	-0.017	-0.025	0.006	-0.026	-0.018
	SR	0.072	0.097	0.099	0.072	-0.100	-0.155	0.025	-0.167	-0.119
NTIS	CER	0.007	0.012	0.013	0.007	-0.022	-0.030	0.000	-0.031	-0.023
	SR	0.041	0.065	0.068	0.041	-0.132	-0.186	-0.006	-0.199	-0.151
EQIS	CER	0.006	0.011	0.012	0.006	-0.023	-0.031	-0.001	-0.032	-0.024
	SR	0.034	0.059	0.061	0.034	-0.138	-0.193	-0.013	-0.205	-0.157
TAIL	CER	-0.003	0.002	0.003	-0.003	-0.033	-0.040	-0.009	-0.041	-0.033
	SR	-0.003	0.022	0.025	-0.003	-0.189	-0.250	-0.056	-0.265	-0.213

general patterns can be reported. First, portfolios formed using quantile-based predictions for both the mean and the volatility have higher CER and SR than those using historical volatility and economic variables when the former utilise quantiles .98 and .02 to predict the mean (regardless of which quantiles are used to predict future volatility).

When quantiles .99 and .01 are used to predict the mean (and any quantile pair for volatility prediction), the resulting portfolio has the highest SR and second-highest CER value (beaten only by the tail risk in the latter case). When all quantiles or the pair $\theta = \{.95, .5\}$ are used to predict the mean, the resulting portfolio still outperforms other approaches except for the one using tail risk as a mean return predictor. However, as before, using predicted returns from those quantiles closer to the median does not result in superior performance, except when the pair $\{.75, .05\}$ is used to forecast the mean. Hence, these results obtained using predicted tail quantiles to forecast the mean and predicted quantile returns to forecast volatility are in line with our previous findings when the volatility was predicted using its historical values.

However, the best models using quantile predictions for both future mean and volatility have, on average, lower CER and SP values than those using quantile predictions for the mean only. Apparently, better volatility forecasts can result in allocation of returns away from the stock market when volatility is predicted to be high, at a cost of foregoing higher returns. As both the CER and the SP impose a specific relationship between returns and risk, the realised gains from lower volatility might not be compensating for lower return, hence leading to lower CER and SP values when volatility forecasts are more accurate. Indeed, when we lower the risk aversion coefficient γ (from 2 to 0.5 in our example) to decrease the "penalty" for higher volatility, the CER values for portfolios relying on quantile estimates for both mean and volatility predictions become higher than those relying on quantile estimates for mean alone. Therefore, it appears that predicting mean using quantiles is superior, and additionally predicting volatility can also be superior if an investor is not too risk averse. For investors with sufficiently high-risk aversion, more accurate predictions of future volatility cause stronger allocation of wealth towards safer assets, resulting in lower risk but also lower return and lower economic value of those volatility forecasts.

Lastly, it is entirely possible that further economic value of the predictive ability of quantile .75 returns can be extracted if those are used to predict other features of the future return distribution, such as higher moments of the behaviour of tails for the VaR analysis.

6. Why are quantile .75 returns good at predicting future stock returns?

In this section, we empirically analyse potential reasons for quantile .75 returns to have superior forecasting performance for the next year's stock return distribution, as compared to other quantile returns. Specifically, we investigate if quantile .75 returns possess superior predictive power for future consumption and investment: if they do, and given that one would expect the stock market to be affected by (expected) changes in future realisations of those two variables, the rationale for the predictive power of quantile returns for next year's stock returns would be established. Our reasoning leading to selection of these two variables is as follows. Firstly, Vassalou and Liew (2000) demonstrate that FF factors predict future GDP growth, especially the news about it (Vassalou, 2003). Hence, variables which contain information about aspects of the future business cycle can explain stock returns, as the latter depend on future states of the economy, and we conjecture that this should be the case for any variable able to predict future economic growth. Henceforth, the ability of quantile .75 returns to predict GDP better than returns any other quantile do would help to explain why those quantile .75 returns also have the best predictive power for future returns.

In addition, Li et al. (2006) observe that investment component of the GDP has a stronger explanatory power for stock returns than the aggregated GDP figures do. Hence, we conjecture that a variable able to predict future investment should be able to predict future stock returns, too. In addition, it is well established analytically that economic agents find intertemporal smoothing of their consumption to be utility-maximising, and from the consumption-CAPM we know that they are willing to pay a premium for assets which allow them to do so. Hence, movements in stock prices will be related to movements in future changes to consumption opportunities, especially the unexpected ones, and we conjecture that a variable able to predict future shocks to consumption opportunities should also possess predictive ability for stock returns. In both cases, i.e., predicting shocks to future investment and consumption, we analyse whether the predictive ability of quantile .75 returns is superior to that of any other quantile returns, as this would explain why the former also perform best in predicting next year's stock returns.

We use data on gross private domestic investment and personal consumption expenditures, components of the GDP, from FRED. For each estimated annual return quantile $R_t(\theta = k)$, where k∈{.01, .02, .05, .10, .20, .25, .30, .40, .50, .60, .70, .75, .80, .90, .95, .98, .99}, we estimate its ability to explain shocks to the future growth of investment and consumption, calculated over a period of three years, three years ahead, i.e., years t+4 to t+6.[16] This is done by estimating the following model within the quantile regression framework:

$$SHOCK\_Z_{t+6}(\theta = j) = \beta_0 + \beta_1 R_t(\theta = k) + \varepsilon_t .$$

[16] It is not known theoretically what the relevant future horizon is which is the most relevant determinant for the current movements of the stock market. On the one hand, if we treat current prices as sums of future discounted cashflows, all future periods are important, but those cash flows laying further in the future are being discounted most. Hence, the nearest future may be argued to be more relevant. On the other hand, short term movements in economic variables may suffer from more noise than long-term trends, which would speak in favor of using longer term (expected) economic conditions as relevant variables driving stock market movements. This issue is addressed empirically here, by varying the length of time those future growth rates in Z are estimated, between 1 and 5 years, both starting in same year as the to-be-predicted returns are estimated and offset by one additional year. The results for 3-years growth rates measured over years t+4 to t+6 generate the most pronounced results, and are reported in this paper.

$SHOCK\_Z_{t+6}$ is a measure of shocks to the future growth of Z between t+4 and t+6, with Z being investment or consumption, and shocks are estimated as the observed minus the 10-periods moving average values of Z. By allowing for quantiles *j* of the conditional distribution of $SHOCK\_Z_{t+6}$ to range between 0 and 100, for each (potentially predictive) quantile k of the past stock returns we analyse its power to predict components of future distribution, and not only the expected value, of the dependent variable, $SHOCK\_Z_t$. As a measure of the predictive ability of returns at quantile *k* for future economic shocks at quantile *j*, we use the log-likelihood value of the estimated model. If stock returns at a particular quantile k^* possess superior predictive ability for future shocks to the economy ($SHOCK\_Z_{t+6}(\theta = j)$), then models using those returns ($R_t(\theta = k^*)$) as a RHS variable will have higher values of the log-likelihood function than models using returns measured at other quantiles $k \neq k^*$. A superior predictor would ideally generate higher log-likelihood values across a wide range of quantiles j of the dependent variable.

The results from models using investment to calculate the LHS variable Z are shown in Figure 4. Each line represents log-likelihood values for one particular explanatory variable ($R_t(\theta = k)$, with k fixed) across quantiles j of the dependent variable $SHOCK\_Z_{t+6}$. It can be seen that quantile .75 returns (the black thick line) do a rather good job in predicting future shocks to investment growth, across a wide range of quantiles. Specifically, they perform best in predicting the area around the center of the distribution (only outcompeted by closely related 70th quantile returns in some cases, denoted by a thick blue line), but also provide superior forecasts for areas around quantiles *j*=20, 30, 75, and 90. Although beaten slightly by the closely related 70th quantile returns when all predictions (for *j*=0 to 100) are considered jointly (e.g., by summing up all values of the log-likelihood function across quantiles *j*, or calculating the number of cases in which each predictive return quantile *k* ranks first), the evidence is rather strong that quantile .75 returns can generate superior predictions for future shocks to investment growth over a wide range of quantiles of the latter variable's distribution. If the stock market is concerned about the same points of the distribution of future investment shocks, this would explain why the quantile .75 returns and future stock market returns are related, i.e., why the former predicts the latter, as demonstrated in previous sections.

The results from models using consumption to calculate the LHS variable Z are shown in Figure 5. The quantile .75 return generates best predictions in almost all quantiles *j* between 30 and 60 of future consumption shocks, and it also performs very well in those quantiles above *j*=80 and 90. In addition, among all explanatory return quantiles *k* considered here, returns at *k*=.75 quantile have the highest average log-likelihood value when considered across all to-be-forecasted quantiles *j*, and the highest number of cases (quantiles *j*) where it provides the best model fit/prediction. Hence, the results support the notion that quantile .75 returns possess superior explanatory power for future shocks to private consumption.

To sum up, our findings imply that quantile 75 returns are a better predictor of next year's stock returns than those returns measured at any other quantile because quantile 75 returns are best, among returns from other quantiles, in predicting future shocks to investment and consumption, and these shocks appear to be the relevant determinants of next year stock returns.

We further analyse how the ability of quantile 75 returns to predict future shocks to investment and consumption compares to that of other economic variables. The discussion of the results in tables 7-10 concludes that, even when controlling for the impact of one economic variable at a time, quantile .75 returns are still a significant predictor for next year's stock returns across a wide range of its quantiles, with exception of the center of the distribution. We are interested to see if this in-sample predictive ability exists because of the

Figure 4: Log-likelihood values of model (3) for future investment shocks.

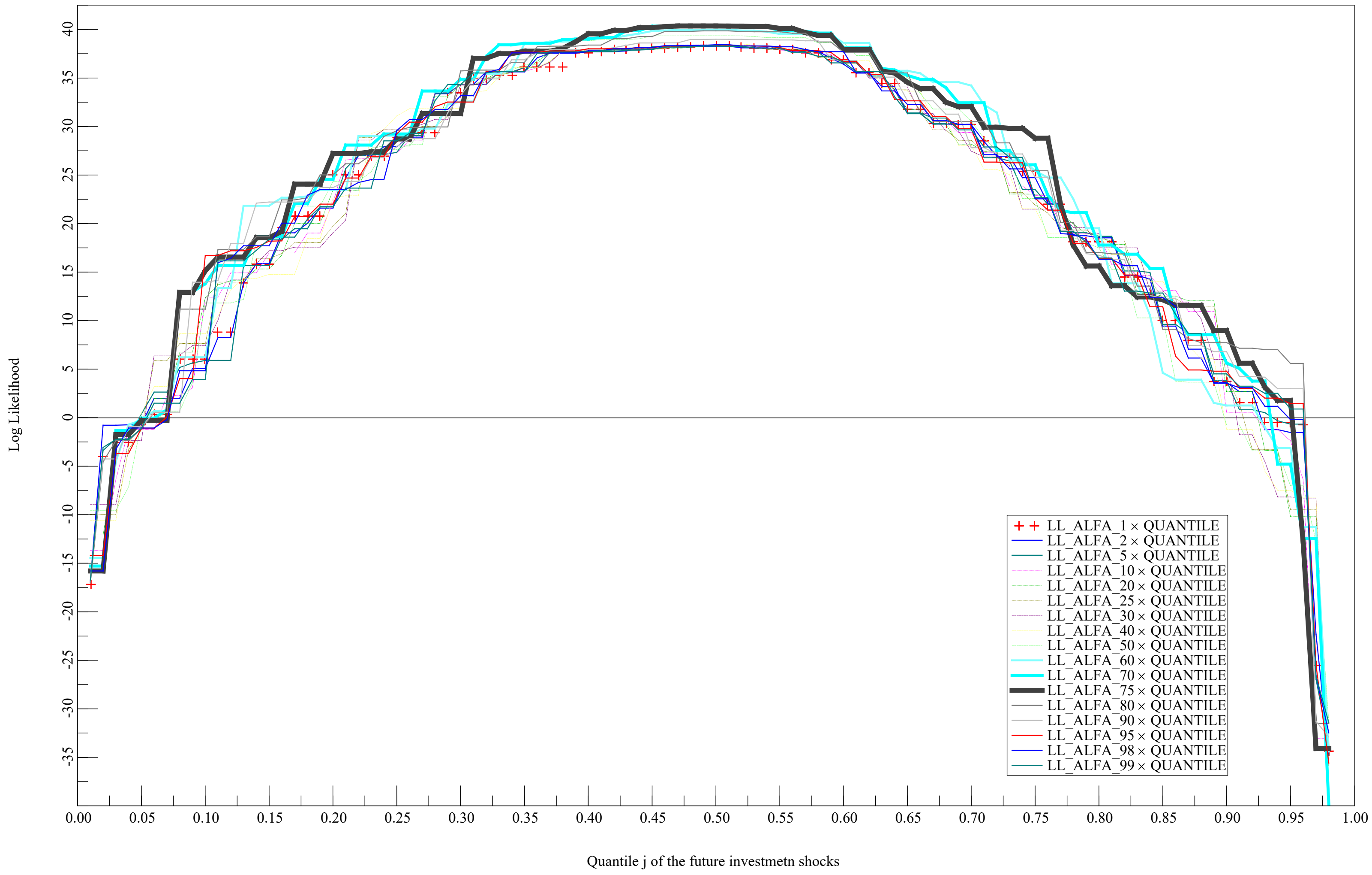
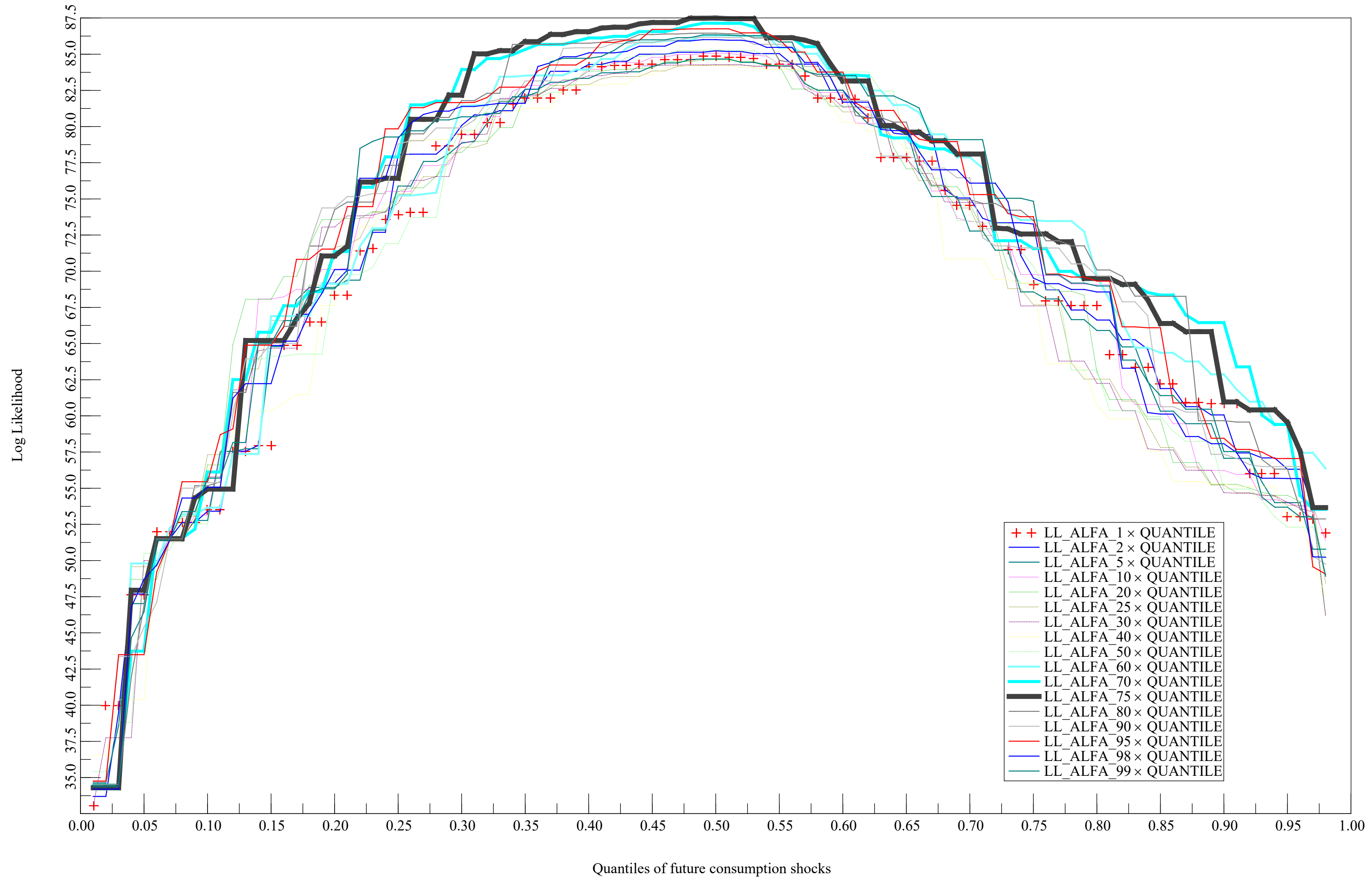


Figure5: Log-likelihood values of model (3) for future consumption shocks



ability to predict future consumption and investment shocks. To that end, we regress the future shocks, defined as above, on a constant, quantile .75 returns, and an economic variable, the latter one at a time. The results (not reported to conserve space) show that quantile 75 returns are a significant in-sample predictor for future economic shocks. For future consumption, the relevant coefficient is significant on average for 62% of all quantiles of the consumption shock distribution, and only when the volatility is included as another explanatory variable do the 75 quantile returns lose their predictive power for all quantiles of future shocks. For future investment shocks, a significant predictive power is observed in 20% of all cases, and quantile 75 returns have no predictive power (for any of the quantiles) only when the dividend payout ratio or the yield on long term government bond are included as another repressor. Overall, quantile .75 returns remain a significant in-sample predictor of either future investment or consumption shocks in over 67% of all quantiles of the shocks, on average, and significantly predict at least 39% quantiles of future shocks to investment or consumption, regardless of which economic variable is also used in the predictive regression. Hence, we conjecture that the ability of quantile returns to predict future mean stock returns, beyond what would be predicted by economic variables, stems to a large extent from its incremental ability to predict future investment or consumption shocks, which again goes beyond what other variables are able to predict.

7. Summary and conclusions

In this paper, we empirically demonstrate that DJIA returns measured at the 75th quantile possess predictive power for a wide spectrum of future quantile returns, with the exception of the median. This predictive power appears to prevail after controlling for the predictive power of other features of the lagged return distribution (i.e., its first four moments), is relatively stable over time, and prevails when one controls for other economic predictors (as summarised in Welch and Goyal, 2008), with Kelly and Jiang's (2014) tail risk being maybe the most difficult alternative to outcompete. Predictions of future mean return based on predicted quantiles possess positive economic value, as are those utilising quantile predictions to forecast volatility, especially for investors with low aversion to risk. Lastly, our results strongly indicate that this superior predictive power stems from quantile .75 returns' ability to predict future shocks to consumption and investment.

One potential issue with the result presented here is that it is not derived from any economic theory, hence it is unclear why the quantile .75 returns possess predictive power for the future return distribution (or that of future consumption and investment). Consequently, it could be argued that the uncovered causality is a statistical artefact. However, its robustness to empirical settings, as presented here, seems to suggest otherwise. More broadly, we would argue that many phenomena hotly debated in the academic literature, such as calendar anomalies, the causal impact of trading volume on stock return, or indeed one of the most prominent examples, the Fama and French (1993) three-factor model, started their life as empirical observations, by practitioners or academics, with little or no theoretical underpinnings. Attempts to find economic explanations for those phenomena only followed later. In this context, we observe that quantile .75 returns show similar predictive pattern to that of default yield spread and volatility in simple regressions, and that our predictive variable appears to be related to Kelly and Jiang's (2014) tail risk, which suggests that this variable might be capturing negative news about the state of the economy. However, our variable seems to be predicting more than the occurrence of negative shocks in the tail, as its increases also lead to subsequent increases in return quantiles in the right part of the distribution. In addition, it also predicts changes in both shoulders of that distribution (areas between the tails and the center), something which cannot be attributed to the extreme (and only negative) shocks. Therefore, the link between the quantile .75 DJIA returns and other variables, and economic reasons for its predictive ability could constitute an interesting avenue for further research. In addition, an investigation of whether our results also apply to other indices, possibly in other countries, would be another interesting avenue, assuming the availability of long-term daily data on other indices.

References

Aiolfi, M., Capistrán, C., and Timmermann, A. (2011) Forecast Combinations. In M. P. Clements and D. F. Hendry (Eds.), *The Oxford Handbook of Economic Forecasting* (pp. 355-390), New York: Oxford University Press.

Amaya, D., Christoffersen, P., Jacobs, K., and Vasquez, A. (2015) Does realized skewness predict the cross-section of equity returns? *Journal of Financial Economics* 118, 135–167.

Ang, A., and Bekaert, G. (2007) Stock return predictability: Is it there? *Review of Financial Studies* 20, 651–707.

Baur, D. G. (2013) The structure and degree of dependence: A quantile regression approach, *Journal of Banking & Finance* 37(3), 786-798.

Baur, D. G., Dimpfl, T., and Jung, R. C. (2012) Stock return autocorrelations revisited: A quantile regression approach, *Journal of Empirical Finance* 19(2), 254-265.

Benjamini, Y., and Hochberg, Y. (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing, *Journal of the Royal Statistical Society, Series B* 57, 289-300.

Benjamini, Y., and Yekutieli D. (2014) The Control of the False Discovery Rate in Multiple Testing Under Dependency, *Annals of Statistics* 29, 1165-1188.

Bollerslev, T., Tauchen, G., and Zhou, H. (2009) Expected stock returns and variance risk premia, *Review of Financial Studies* 22, 4463-4492.

Boudoukh, J., and Richardson, M. (1993) Stock returns and inflation: A long-horizon perspective, *American Economic Review* 83, 1346–55.

Boudoukh, J., Richardson, M., and Whitelaw, R. (2008) The myth of long-horizon predictability, *The Review of Financial Studies* 21, 1577-1605.

Campbell J.Y., and Thompson, S.B. (2008) Predicting the equity premium out of sample: Can anything beat the historical average? *Review of Financial Studies* 21, 1509–1531.

Cenesizoglu .T, and Timmermann, A. (2008) Is the distribution of stock returns predictable? available at: http, //ssrn.com/abstract=1107185.

Cenesizoglu, T., and Timmermann, A. (2012) Do return prediction models add economic value? *Journal of Banking & Finance* 36, 2974-2987.

Chen, C.W.S., Gerlach, R., Hwang, B.B.K., and McAleer, M. (2012) Forecasting Value-at-Risk using nonlinear regression quantiles and the intra-day range, *International Journal of Forecasting* 28, 557-574.

Chuang, C.-C., Kuan, C.-M., and Lin, H.-Y. (2009) Causality in quantiles and dynamic stock return–volume relations, *Journal of Banking & Finance* 33(7), 1351-1360.

Clark, T.E., and McCracken, M.W. (2001) Tests of equal forecast accuracy and encompassing for nested models, *Journal of Econometrics* 105, 85-110.

Clark, T.E., and West, K.D. (2007) Approximately normal tests for equal predictive accuracy in nested models, *Journal of Econometrics* 138, 291-311.

Cochrane, J. H. (1991) Production-based asset pricing and the link between stock returns and economic fluctuations, *Journal of Finance* 46, 209–37.

Cochrane, J. H. (2008) The dog that did not bark: A defense of return predictability. *Review of Financial Studies* 21, 1533-1575.

Da, Z., Jagannathan, R., and Shen, J. (2014) Growth expectations, dividend yields, and future stock returns, NBER working paper 20651, National Bureau of Economic Research, Inc.

Engle, R.F., and Manganelli, S. (2004) CAViaR, Conditional autoregressive value at risk by regression quantiles, *Journal of Business and Economic Statistics* 22, 367-382.

Ericsson, N.R. (1992) Parameter constancy, mean square forecast errors, and measuring forecast performance: An exposition, extensions, and illustration, *Journal of Policy Modeling* 14, 465-495.

Fama, E.F., and French, K.R. (1993) Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Ferreira, M.A., and Santa-Clara, P. (2011) Forecasting stock market returns: The sum of the parts is more than the whole, *Journal of Financial Economics* 100, 514–537.

Gerlach, R.H., Chen, C.W.S., and Chan, N.Y.C. (2011) Bayesian Time-Varying Quantile Forecasting for Value-at-Risk in Financial Markets, *Journal of Business & Economic Statistics* 29, 481-492.

Golez, B., and Koudijs, P. (2014) Four centuries of return predictability. NBER working paper 20814, National Bureau of Economic Research, Inc.

Granger, C.W.J. (1969) Investigating causal relations by econometric models and cross-spectral methods, *Econometrica* 37, 424–438.

Guidolin, M., Hyde, S., McMillan, D., and Ono, S. (2014) Does the macroeconomy predict UK asset returns in a nonlinear fashion? Comprehensive out-of-sample evidence, *Oxford Bulletin of Economics and Statistics* 76, 510-535

Hansen, P.R., and Timmermann, A. (2012) Choice of sample split in out-of-sample forecast evaluation, Economics Working Papers ECO2012/10, European University Institute.

Harvey, Campbell R., and Liu, Y. (2014) Evaluating trading strategies, *Journal of Portfolio Management* 40, 108-118.

Harvey, D.I., Leybourne, S.J., and Newbold, P. (1998) Tests for forecast encompassing, *Journal of Business & Economic Statistics* 16, 254-259.

Jordan, S.J., Vivian, A., and Wohar, M.E. (2016) Can commodity returns forecast Canadian sector stock returns? *International Review of Economics & Finance* 41, 172-188.

Kambouroudis, D.S., and McMillan, D.G. (2015) Is there an ideal in-sample length for forecasting volatility? *Journal of International Financial Markets, Institutions and Money* 37, 114-137.

Kelly, B., and Jiang, H. (2014) Tail risk and asset prices, *Review of Financial Studies* 27, 2841-2871

Koenker, R., and G. Bassett, Jr. (1978) Regression quantiles. *Econometrica* 46, 33-50.

Lo, A. W., and MacKinlay, A. C. (1990) When Are Contrarian Profits Due to Stock Market Overreaction? *The Review of Financial Studies* 3, 175–205.

Ma, L., and Pohlman, L. (2008) Return forecasts and optimal portfolio construction: A quantile regression approach, *The European Journal of Finance* 14, 409-425.

Marquering, W., and Verbeek, M. (2004) The economic value of predicting stock index returns and volatility, *The Journal of Financial and Quantitative Analysis* 39, 407-429.

Meligkotsidou, L., Panopoulou, E., Vrontos, I.D., and Vrontos, S.D. (2014) A quantile regression approach to equity premium prediction, *Journal of Forecasting* 33, 558-576.

Neely, C.J., Rapach, D.E. Tu, J., and Zhou, G. (2014) Forecasting the equity risk premium: The role of technical indicators, *Management Science* 60, 1772-1791.

Paye, B.S. (2012) 'Déjà vol': Predictive regressions for aggregate stock market volatility using macroeconomic variables, *Journal of Financial Economics* 106, 527-546.

Pedersen, T.Q. (2015) Predictable Return Distributions, *Journal of Forecasting* 34, 114-132.

Rapach, D.E., and Zhou, G. (2013) Forecasting stock returns. In: Elliott, G., and Timmermann, A., eds. Handbook of Economic Forecasting, Vol. 2A (Elsevier, Amsterdam), 328–383.

Rapach, D.E., Strauss, J.K., and Zhou, G. (2010) Out-of-sample equity premium prediction: Combination forecasts and links to the real economy, *Review of Financial Studies* 23, 821–862.

Richardson, M., and Smith, T. (1991) Tests of financial models in the presence of overlapping observations, *Review of Financial Studies* 4, 227–54.

Şener, E., Baronyan, S., and Mengütürk, L.A. (2012) Ranking the predictive performances of value-at-risk estimation methods, *International Journal of Forecasting* 28, 849-873.

Stambaugh, R. F. (1999) Predictive regressions, *Journal of Financial Economics* 54, 375–421.

Taylor, J.W. (2005) Generating volatility forecasts from value at risk estimates, *Management Science* 51, 712-725.

Timmermann, A. (2006), Chapter 4: Forecast Combinations, In G. Elliott, C.W.J. Granger, A. Timmermann (Eds.), *Handbook of Economic Forecasting, Volume 1* (pp. 135-196), Amsterdam: Elsevier.

Turner, J. A. (2015) Casting doubt on the predictability of stock returns in real time: Bayesian model averaging using realistic priors, *Review of Finance* 19, 785–821.

Valkanov, R. (2003) Long-horizon regressions: Theoretical results and applications, *Journal of Financial Economics* 68, 201-232.

Welch, I., and Goyal, A. (2008) A comprehensive look at the empirical performance of equity premium prediction, *Review of Financial Studies* 21, 1455-1508.

Zhu, M. (2013) Return distribution predictability and its implications for portfolio selection, *International Review of Economics & Finance* 27, 209-223.