



The predictive performance of commodity futures risk factors



Shamim Ahmed^{a,*}, Daniel Tsvetanov^b

^a Nottingham University Business School, University of Nottingham, Nottingham NG8 1BB, United Kingdom

^b Essex Business School, University of Essex, Colchester CO4 3SQ, United Kingdom

ARTICLE INFO

Article history:

Received 12 July 2015

Accepted 21 June 2016

Available online 9 July 2016

JEL Classification:

F37

G11

G13

G17

Q02

Keywords:

Commodity markets

Futures pricing

Out-of-sample predictability

Economic value

Time series

Econometric models

ABSTRACT

This paper investigates the time-series predictability of commodity futures excess returns from factor models that exploit two risk factors – the equally weighted average excess return on long positions in a universe of futures contracts and the return difference between the high- and low-basis portfolios. Adopting a standard set of statistical evaluation metrics, we find weak evidence that the factor models provide out-of-sample forecasts of monthly excess returns significantly better than the benchmark of random walk with drift model. We also show, in a dynamic asset allocation environment, that the information contained in the commodity-based risk factors does not generate systematic economic value to risk-averse investors pursuing a commodity stand-alone strategy or a diversification strategy.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Systematic risk factors, motivated by conventional asset pricing models, are insignificantly correlated with commodity futures excess returns (see, among others, Bessembinder, 1992; Daskalaki et al., 2014; Dusak, 1973; Erb and Harvey, 2006; Jagannathan, 1985).¹ On the other hand, several studies show that the difference between the current spot price and the contemporaneous futures price, which is commonly known as the basis, contains information about expected futures excess returns (see Fama and French, 1987; Gorton and Rouwenhorst, 2006; de Roon et al., 1998; Gorton et al., 2013). Assuming that the spot price is well approximated by a martingale, a contract whose price is lower (higher) than the current spot price is then expected to yield a positive (negative) excess return over the term to maturity. Such an intuition has led researchers like Yang (2013) and Szymanowska et al. (2014) to

identify commodity-based risk factors by analyzing the excess returns on a trading strategy that speculates on the slope of the term structure of futures contracts. In particular, Yang (2013) sorts a universe of commodity futures contracts into portfolios by basis and shows that two factors account for most of the in-sample time-series and cross-sectional variation in excess returns on these portfolios. The first factor is a commodity market factor, which corresponds to the equally weighted average excess return on a long position in all futures contracts (henceforth EWA). The second factor is the return difference between the high- and low-basis portfolios (henceforth HML).

The EWA and HML factors are far from being perfectly correlated and therefore capture the different aspects of systematic variation in commodity futures excess returns. Consistent with this risk-based interpretation, Bakshi et al. (2014) also show empirically that these commodity-based risk factors forecast changes in the investment opportunity set. Intuitively, when the EWA and HML factors are indeed proxies for systematic risk that are priced in the cross-section of commodity futures excess returns, combining expectations of these factors and knowledge of their betas should translate into expectations of future excess returns. There is extensive evidence, both empirical and anecdotal, that similar asset pricing intuition is often exploited in practice. More precisely, financial managers almost always use the traditional capital asset

* Corresponding author. Tel.: +44 (0) 115 823 2359; fax: +44 (0) 115 846 6667.

E-mail addresses: shamim.ahmed@nottingham.ac.uk (S. Ahmed), dtsvet@essex.ac.uk (D. Tsvetanov).

¹ Carter et al. (1983) provide evidence in support of systematic risk for commodity futures when the market portfolio includes both stock and commodity futures indexes. However, Marcus (1984) argues that the amended market portfolio constructed by Carter et al. (1983) is inappropriate and therefore questions the validity of the corresponding empirical results.

pricing model (CAPM) and the multifactor models motivated by the arbitrage pricing theory (APT) as a primary tool to compute expectations of returns, especially in the context of equity markets (see, for example, Fama and French, 1997; Gitman and Mercurio, 1982; Graham and Harvey, 2001; Jagannathan and Meier, 2002; Simin, 2008; Ferson et al., 2013). It is, therefore, appealing for practitioners, especially active portfolio managers who trade commodity futures as part of their broader diversification strategy, to know whether the EWA and HML factors contain information that adds to the out-of-sample predictability of futures excess returns. Said differently, whether these commodity-based risk factors, utilized in a contemporaneous regression model setup based on asset pricing theory, can ultimately help minimize real time uncertainty faced by market participants when forecasting future excess returns, making investment decisions, and/or modeling future risk exposure.² Knowing this is timely and important considering the fact that the use of commodity futures contracts as an alternative asset class has grown rapidly in recent years.³

The goal of our paper is thus threefold. First, we examine the ability of EWA and HML factors, identified by Yang (2013) and Szymanowska et al. (2014), to generate accurate expectations of monthly commodity futures excess returns in an out-of-sample setting. In other words, we investigate whether the asset pricing models that utilize these risk factors including their unconditional and conditional expectations provide accurate one-step ahead forecasts of excess returns in the time-series domain. Our test assets include individual commodity futures as well as basis-sorted portfolios. Furthermore, in the case of individual commodity futures, we also exploit the information content embedded in commodity-specific characteristics, such as the hedging pressure and the open interest, and perform the out-of-sample forecasting exercises using the Fama and MacBeth (1973) cross-sectional procedure.

Second, we examine whether the factor model forecasts translate into systematic economic value to risk-averse investors. In particular, we quantify the economic value due to a dynamic mean-variance efficient asset allocation strategy that exploits predictability in commodity futures excess returns. This empirical exercise is motivated by the evidence from a growing body of literature that statistical significance does not necessarily guarantee economic significance (see, among others, Della Corte et al., 2009; Leitch and Tanner, 1991; McCracken and Valente, 2014; Thornton and Valente, 2012, and references therein).

Finally, we aim to investigate the economic gains accruing to an investor who diversifies her exposure to conventional assets, such as stocks and bonds, by dynamically allocating commodity futures contracts in her portfolio conditioning on the predictive ability of the risk factors out of sample. The potential benefits of allocating commodity futures to conventional portfolios have long been a subject of academic research. In fact, several empirical studies show that investors can improve the risk-return profile of their portfolios by also investing in commodities (see, among others, Bodie and Rosansky, 1980; Erb and Harvey, 2006; Fortenberry and Hauser, 1990; Jensen et al., 2000). For ease of exposition throughout this paper, we refer to the optimal asset allocation strategies in our latter two objectives as the commodity stand-alone strategy and the diversification strategy, respectively.

Our paper makes several contributions to the extant literature on commodity futures. First, we evaluate the performance of a set of commodity-based risk factors in an out-of-sample setting, whereas the existing evidence is based solely on in-sample data fitting framework.⁴ Our analysis therefore shows the extent to which asset pricing models with commodity-based risk factors can be useful for practical applications including modeling future risk exposure. Second, an out-of-sample analysis enables us to assess the performance of the asset pricing models in terms of forecast errors and circumvents well-known issues, such as useless factor biases, errors-in-variables problem, the use of weak instruments, and data snooping biases.⁵ Third, an assessment of economic significance due to out-of-sample predictability allows us to investigate whether investors who trade commodity futures contracts as part of their commodity stand-alone strategy or diversification strategy can gain value by conditioning on expectations of returns provided by the factor models. To the best of our knowledge, this is the first study that examines the diversification benefits of commodity futures in the context of return predictability, more so using forecasts from asset pricing models with commodity-based risk factors.

We find a host of interesting results based on a cross-section of 15 commodity futures. First, the factor models hardly outperform a random walk with drift benchmark in the out-of-sample forecasting horse races. The poor statistical performance of the models remain invariant regardless of the way we form factor expectations, including forecasting methods, and whether we focus on predicting individual commodity futures excess returns or the excess returns on basis-sorted portfolios. We also find that the use of commodity-specific characteristics does not improve the ability of factor models to produce more accurate forecasts of excess returns.

Second, factor models using unconditional factor expectations generally produce step ahead forecasts more accurate than models using conditional expectations of risk factors. The implications of these findings, mentioned above, are largely consistent with those of Simin (2008), who also finds that neither the CAPM of Sharpe (1964) and Lintner (1965) nor the three-factor model of Fama and French (1993) is capable of producing more accurate expectations of future equity returns relative to the historical average of the market return.

Third, the out-of-sample economic value results for a commodity stand-alone investment strategy lead to a similar conclusion obtained for the statistical evaluation of the models. Put differently, the factor models with commodity-based risk factors fail to outperform the random walk with drift benchmark in economic terms. Expectations of commodity futures excess returns generated by asset pricing models do not offer systematic diversification benefits either. A risk-averse investor who is already exposed to conventional assets gains no tangible economic value from dynamically allocating commodity futures contracts to her portfolio by exploiting the next period return forecasts offered by the factor models. These results are fairly robust to the use of different performance evaluation metrics and echo those of Daskalaki and Skiadopoulos (2011). The authors show that a utility maximizing investor is better off with a portfolio of stocks and bonds only and there are no tangible economic gains from investing in a commodity index or in individual commodity futures contracts.

² Fama and MacBeth (1973, p. 618) also stress that “As a normative theory the model only has content if there is some relationship between future returns and estimates of risk that can be made on the basis of current information.”

³ According to a Staff Report prepared by the Commodity Futures Trading Commission (CFTC), the total value of commodity index-related instruments purchased by institutional investors increased markedly to US dollar (USD) 200 billion in 2008 from a modest figure of USD 15 billion in 2003. Recently, BarclayHedge reports that the commodity assets under management reached at about USD 320 billion in the last quarter of 2014.

⁴ Recently, Simin (2008) adopts a similar empirical approach but focusing only on conventional risk factors (i.e., Fama and French, 1993 factors) in the context of equity market alone. More importantly, we investigate not only the statistical significance but also the economic significance of the forecasting power of the factor models with commodity-based risk factors. Clearly, these dimensions of predictability assessment differentiates our paper from that investigated in Simin (2008).

⁵ Asset pricing models can also be evaluated based on out-of-sample pricing errors (see, for example, Ferson et al., 2013).

The remainder of this paper is organized as follows. Section 2 outlines the empirical factor models for commodity futures excess returns and the statistical criteria used to evaluate the predictive accuracy out of sample. Section 3 describes the data and the construction of the risk factors, and summarizes the empirical results. The dynamic asset allocation framework to build and assess optimal portfolios is outlined in Section 4 along with the findings on the economic value. Section 5 provides an analysis of the diversification benefits to investors. Finally, Section 6 concludes. A separate Internet Appendix contains additional robustness check results.

2. Asset pricing models and statistical evaluation of forecasts

2.1. Model specifications

We investigate the ability of EWA and HML factors to forecast excess returns in the out-of-sample using a contemporaneous regression framework combining factor model parameter estimates and factor forecasts. The rationale for adopting such a framework is that the successful pricing ability (i.e., explanatory power) of these non-conventionally measured risk factors in the cross-section of commodity futures excess returns has been recorded from a contemporaneous model setting (see Yang, 2013). In light of this empirical finding, and consistent with the literature (see, among others, Ferson and Harvey, 1997; 1999; Simin, 2008, and references therein), our starting point is a general representation that the conditional expected futures excess return on each commodity j is determined by the following asset pricing model:

$$\begin{aligned} E_t[\Delta f_{t+1}^j] &= \alpha_t^j + \beta_t^{j'} E_t[X_{t+1}], \\ \alpha_t^j &= a_0^j + a_1^j z_t^j, \\ \beta_t^j &= b_0^j + b_1^j z_t^j, \end{aligned} \quad (1)$$

where $E_t[\cdot]$ is the conditional expectation, given the information set available at time t , Δf_{t+1}^j is the logarithm of futures excess return (see Eq. (15)), X_{t+1} is a 2×1 vector of commodity-based risk factors, which comprises EWA and HML, and z_t^j is an instrumental variable. The parameters of the factor model: b_0^j is 2×1 , b_1^j is 1×2 , and a_0^j and a_1^j are scalars. The choice of the instrument z_t^j is central for the good performance of the model since it has to adequately summarize the different aspects of investors' information set. Several recent studies find that the futures basis is informative about the expected excess return on commodity futures (see, among others, Erb and Harvey, 2006; Fama and French, 1987; Gorton et al., 2013; Szymanowska et al., 2014, and references therein).⁶ For example, Yang (2013) sorts commodity futures contracts into portfolios by basis and shows that average portfolio excess return is increasing in basis. Building on this common finding, we choose commodity-specific futures basis (see Eq. (16)) as a natural candidate for the instrumental variable z_t^j capturing investor expectations of corresponding excess returns.⁷

⁶ Gorton et al. (2013) also provide a simple two-period mean-variance model linking basis to commodity futures excess returns.

⁷ We have also experimented with other potential instrumental variables, such as, liquidity (Amihud et al., 1997; Marshall et al., 2012), volatility (Dhume, 2011), inflation (Erb and Harvey, 2006; Gorton and Rouwenhorst, 2006), and momentum (Asness et al., 2013; Erb and Harvey, 2006; Gorton et al., 2013; Miffre and Rallis, 2007), and find results qualitatively similar to those based on the futures basis. To conserve space, these are omitted in this paper.

The empirical analysis proceeds by setting our benchmark consistent with a random walk with drift model specified as:⁸

$$\Delta f_t^j = a_0^j + e_t^j. \quad (2)$$

We then consider various model specifications of the general form in Eq. (1) to examine the predictive ability of commodity-based risk factors over time. These factor models are as follows:

$$\Delta f_t^j = a_0^j + b_{0,1}^j EWA_t + e_t^j, \quad (3)$$

$$\Delta f_t^j = a_0^j + b_{0,1}^j EWA_t + b_{0,2}^j HML_t + e_t^j, \quad (4)$$

$$\Delta f_t^j = a_0^j + b_{0,1}^j EWA_t + b_{1,1}^j z_{t-1}^j EWA_t + e_t^j, \quad (5)$$

$$\begin{aligned} \Delta f_t^j &= a_0^j + b_{0,1}^j EWA_t + b_{0,2}^j HML_t + b_{1,1}^j z_{t-1}^j EWA_t \\ &\quad + b_{1,2}^j z_{t-1}^j HML_t + e_t^j, \end{aligned} \quad (6)$$

$$\Delta f_t^j = a_0^j + a_1^j z_{t-1}^j + b_{0,1}^j EWA_t + e_t^j, \quad (7)$$

$$\Delta f_t^j = a_0^j + a_1^j z_{t-1}^j + b_{0,1}^j EWA_t + b_{0,2}^j HML_t + e_t^j, \quad (8)$$

$$\Delta f_t^j = a_0^j + a_1^j z_{t-1}^j + b_{0,1}^j EWA_t + b_{1,1}^j z_{t-1}^j EWA_t + e_t^j, \quad (9)$$

$$\begin{aligned} \Delta f_t^j &= a_0^j + a_1^j z_{t-1}^j + b_{0,1}^j EWA_t + b_{0,2}^j HML_t + b_{1,1}^j z_{t-1}^j EWA_t \\ &\quad + b_{1,2}^j z_{t-1}^j HML_t + e_t^j. \end{aligned} \quad (10)$$

It is important to note that the risk factor models in the form specified by Eqs. (3)–(10) are also adopted in the recent forecasting literature, especially in the context of equity and foreign exchange markets (see, among others, Ahmed et al., 2016; Fama and French, 1997; Malone et al., 2014; Simin, 2008; Verdelhan, 2016). Moreover, contemporaneous regressions of the kind specified above are utilized when the objective is to evaluate the predictive ability of a model, given a path for some unmodeled set of variables (see Ferraro et al., 2015; West, 1996, and references therein).

To generate one-month ahead time-series forecasts of excess returns on individual commodity futures from each of the models including the benchmark, we use rolling (ordinary least squares (OLS)) regressions with a 10-year window.^{9,10} More precisely, at the end of each forecast origin month t , we compute the in-sample parameter estimates of the models using the factor(s) ending at month t and if needed the lagged instrument z_{t-1}^j . We then substitute the expectations of the risk factor(s) for month $t+1$ (described below), the estimated parameters, and if needed the instrument z_t^j into Eqs. (3)–(10). This procedure gives us the factor

⁸ The random walk with drift benchmark is nested within each of the alternative factor models considered in this paper and is consistent with the hypothesis that commodity futures prices are unpredictable (e.g., Alquist and Kilian, 2010; Chinn and Coibion, 2014).

⁹ The choice of a rolling forecasting scheme is motivated by the fact that macroeconomic environment changes over time and therefore a very past data might be less informative for forecasting purposes (Daskalaki and Skiadopoulos, 2011; Thornton and Valente, 2012). But for robustness checks, we also employ a recursive forecasting scheme with an initial 10-year expanding window. The results of this exercise consistent with those from a rolling forecasting scheme are reported in the Internet Appendix.

¹⁰ Welch and Goyal (2008, p. 1464) note, "It is not clear how to choose the periods over which a regression model is estimated and subsequently evaluated." In one hand, a shorter estimation window length allows a model parameter to adapt to structural breaks relatively quickly though the parameter is estimated less efficiently. On the other hand, a large estimation window length generates a relatively efficient parameter estimate but leads to a lower power since fewer observations being available for out-of-sample forecast evaluation. Faced with this trade-off, our choice of a 10-year rolling window length seems reasonable. However, to address potential data snooping biases, we also carry out the forecasting exercises under three- and five-year rolling windows and find the out-of-sample results qualitatively identical to those reported in this paper. The results based on a three-year rolling window are provided in the Internet Appendix.

model time-series forecasts of excess returns on commodity futures j for month $t + 1$ in the out-of-sample period. We also carry out the forecasting exercises at the portfolio-level. In that case, Δf_t^j and z_{t-1}^j are replaced by $\overline{\Delta f}_t^k$ and \overline{z}_{t-1}^k , respectively, where $\overline{\Delta f}_t^k = \frac{1}{J_{t-1}^k} \sum_{j=1}^{J_{t-1}^k} \Delta f_t^j$, $\overline{z}_{t-1}^k = \frac{1}{J_{t-1}^k} \sum_{j=1}^{J_{t-1}^k} z_{t-1}^j$, and J_{t-1}^k denotes the number of commodity futures contracts available in portfolio k at the end of month $t - 1$.

We consider both unconditional and conditional expectations of our return-based risk factors, EWA and HML. In particular, we obtain the unconditional expectation of a risk factor at the end of month t by computing the time-series average of the factor from the 10-year rolling window ending at month t as the forecast for the following month.¹¹ To compute the conditional expectation of a risk factor, we also use the 10-year rolling window to perform the predictive regression, where the lagged instrument z_{t-1}^j (i.e., in our case, lagged futures basis) serves as a predictor for the factor in consideration. For the EWA factor, the regression specification is: $EWA_t = \phi_{0,1}^j + \phi_{1,1}^j z_{t-1}^j + \varepsilon_t^j$. We then use the parameter estimates from the factor predictive regression and the value of the instrument at time t to generate the one-step ahead conditional forecast $E_{t,j}[EWA_{t+1}|z_t^j]$. Similar empirical procedure is followed to compute the conditional expectation of the HML factor. For the basis-sorted commodity futures portfolios, \overline{z}^k and k replace z^j and j , respectively, in the above factor predictive regressions.

2.2. Statistical measures of forecast accuracy

We adopt three metrics for the statistical evaluation of the out-of-sample forecast accuracy of the factor models. The first one is the Campbell and Thompson (2008) measure of out-of-sample R^2 statistic denoted by R_{OOS}^2 . The statistic is defined as

$$R_{i,OOS}^{2,j} = 1 - \frac{\sum_{t=M+1}^{T-1} (\Delta f_{t+1}^j - \widetilde{\Delta f}_{i,t+1|t}^j)^2}{\sum_{t=M+1}^{T-1} (\Delta f_{t+1}^j - \widetilde{\Delta f}_{2,t+1|t}^j)^2}, \quad (11)$$

where Δf_{t+1}^j is the realized futures excess return at the end of month $t + 1$, $\widetilde{\Delta f}_{2,t+1|t}^j$ is the one-month ahead forecast of excess return from the benchmark random walk with drift model (Eq. (2)), $\widetilde{\Delta f}_{i,t+1|t}^j$ is the one-month ahead forecast from an alternative factor model specified in equation (i) ($i = 3, 4, \dots, 10$) that uses unconditional or conditional expectations of risk factor(s), T is the number of observations available in the full-sample period, and M is the number of observations in the estimation window. At the commodity portfolio-level forecasting, $\widetilde{\Delta f}_{2,t+1|t}^j$ and $\widetilde{\Delta f}_{i,t+1|t}^j$ are replaced by $\widetilde{\Delta f}_{2,t+1|t}^k$ and $\widetilde{\Delta f}_{i,t+1|t}^k$, respectively. If $R_{i,OOS}^{2,j}$ is positive, the alternative factor model in equation (i) ($i = 3, 4, \dots, 10$) has smaller mean squared forecast error (MSFE) than the random walk with drift benchmark. In other words, the alternative model generates more accurate step ahead forecasts. For statistical inference of the R_{OOS}^2 metric, we obtain one-sided (to the right) critical values using the fixed regressor bootstrap method proposed by Clark and McCracken (2012).

The second metric is due to Clark and West (2006, 2007), which tests the null hypothesis of equal predictive ability of two nested models. Importantly, the testing procedure takes into account the fact that under the null hypothesis, the MSFE from a larger alternative factor model in equation (i) ($i = 3, 4, \dots, 10$) is expected to be greater than that of the random walk with drift benchmark due

to the noise introduced into the alternative model forecasts. To circumvent this issue, Clark and West (2006, 2007) suggest an adjustment for the upward bias in the MSFE of a larger alternative model as follows:

$$MSFE_{i,adj}^j = \frac{1}{P} \sum_{t=M+1}^{T-1} (\Delta f_{t+1}^j - \widetilde{\Delta f}_{i,t+1|t}^j)^2 - \frac{1}{P} \sum_{t=M+1}^{T-1} (\widetilde{\Delta f}_{2,t+1|t}^j - \widetilde{\Delta f}_{i,t+1|t}^j)^2, \quad (12)$$

where $P = (T - 1 - M)$ is the number of one-month ahead forecasts. For computational convenience, we follow Clark and West (2006, 2007) to assess the null hypothesis by regressing $\hat{g}_{i,t+1|t}^j = (\Delta f_{t+1}^j - \widetilde{\Delta f}_{2,t+1|t}^j)^2 - [(\Delta f_{t+1}^j - \widetilde{\Delta f}_{i,t+1|t}^j)^2 - (\widetilde{\Delta f}_{2,t+1|t}^j - \widetilde{\Delta f}_{i,t+1|t}^j)^2]$ on a constant and obtain the $MSFE_{i,adj}^j$ - t statistic for a zero coefficient based on the Newey and West (1987) estimator. However, Clark and McCracken (2001, 2012) show that under the null hypothesis, the population forecast errors of a larger model (in our case, all of the alternative models in Eqs. (3)–(10)) are the same as those from a parsimonious baseline model (in our case, the random walk with drift model) nested by the larger model. This implies that the population analogue of $\hat{g}_{i,t+1|t}^j$ will have a degenerate distribution. Hence, we use the fixed regressor bootstrap method, as per Clark and McCracken (2012), to simulate the right-sided critical values of this test statistic for a pairwise model forecast comparison.

To account for the multiple-testing problem in the light of possible data mining (see, among others, Chatfield, 1995; Hoover and Perez, 1999; Lo and MacKinlay, 1990), we also compute the $\max_{i=3,4,\dots,10} MSFE_{i,adj}^j$ - t statistic for a reality check test advocated by Clark and McCracken (2012) including the fixed regressor bootstrap critical values. The null hypothesis for this composite test states that the baseline random walk with drift model is not inferior to any of the alternative factor models. Details of simulating the bootstrap critical values for the pairwise and reality check tests are provided in the Appendix A.

The third out-of-sample statistical evaluation metric that we consider is the equal conditional predictive ability test proposed by Giacomini and White (2006).¹² The test allows both nested and non-nested model comparisons and takes into account the possibility of model misspecification due to unmodeled dynamics, unmodeled heterogeneity, and incorrect functional form. However, its empirical application is restricted to forecasts generated only from a rolling scheme with a finite estimation window length $M \leq P < \infty$. For a given loss function and a given σ -field \mathcal{G}_t^j , the null hypothesis of equal conditional predictive ability can be stated as

$$H_0 : E[\Delta L_{i,M,t+1}^j | \mathcal{G}_t^j] = 0, \quad (13)$$

where $\Delta L_{i,M,t+1}^j$ is the difference between the squared forecast error loss functions of the benchmark and alternative models. Under the null hypothesis, when $\mathcal{G}_t^j = \mathcal{F}_t^j$ -measurable, the one-step ahead sequence $\{\Delta L_{i,M,t+1}^j, \mathcal{F}_t^j\}$ is a martingale difference sequence, so that for every \mathcal{F}_t -measurable function h_t we can write

$$H_{0,h_t} : E[h_{i,t} \Delta L_{i,M,t+1}^j] = 0. \quad (14)$$

The test statistic for one-step ahead forecasts is computed as $(P - 1)R_{uc}^{2,j}$, where $R_{uc}^{2,j}$ is the uncentered squared multiple correlation coefficient from the regression of $\Delta L_{i,M,t+1}^j$ on $h_{i,t}^j$. Under the

¹¹ Such an empirical approach is consistent with the common practice in industry for estimating the cost of equity within the context of CAPM or APT (e.g., Fama and French, 1997; Gitman and Mercurio, 1982; Graham and Harvey, 2001; Jagannathan and Meier, 2002).

¹² The Clark and West (2006, 2007) statistic tests the null at the population-level, whereas the Giacomini and White (2006) test is associated with finite sample. The former might be useful when evaluating models in population is of interest, and the latter when evaluating forecasts is of interest (Clark and McCracken, 2013).

Table 1
Description of data.

Commodity	Symbol	Delivery month											
		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Cocoa	CC			✓		✓		✓		✓			✓
Coffee	KC			✓		✓		✓		✓			✓
Copper	HG	✓		✓		✓		✓		✓			✓
Corn	C-			✓		✓		✓		✓			✓
Cotton	CT			✓		✓		✓		✓			✓
Crude oil	CL	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Gasoline	HU	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Gold	GC		✓		✓		✓		✓		✓		✓
Heating oil	HO	✓	✓	✓	✓	✓		✓		✓		✓	✓
Lean hogs	LH		✓		✓		✓	✓	✓		✓		✓
Live cattle	LC		✓		✓		✓		✓		✓		✓
Silver	SI			✓		✓		✓		✓			✓
Soybeans	S-	✓		✓		✓		✓		✓		✓	
Sugar	SB			✓		✓		✓			✓		
Wheat	W-			✓		✓		✓		✓			✓

The table describes the data used in the empirical analysis. Symbol is the mnemonic used by the Commodity Research Bureau for the corresponding futures contract. Delivery Month is the month used to construct excess returns for individual commodity futures. The sample period is from January 1986 to October 2013.

null hypothesis, the statistic is asymptotically χ^2_2 -distributed. For the choice of the test function h_t , we follow [Giacomini and White \(2006\)](#) in their empirical application and define $h_{i,t}^j = (1, \Delta L_{i,M,t}^j)'$. Whenever the null hypothesis is rejected in favor of the alternative (to the right-side), an alternative factor model in equation (i) ($i = 3, 4, \dots, 10$) is chosen over the random walk with drift benchmark if the fitted values from the regression of $\Delta L_{i,M,t+1}^j$ on $h_{i,t}^j$ are positive more than 50% of the time. Otherwise, the benchmark model is chosen in case of a rejection of the null hypothesis.

3. Empirical results

3.1. Data and construction of risk factors

The data, used in this paper, consist of daily settlement prices of 15 most liquid commodity futures contracts obtained from the Commodity Research Bureau. The sample period spans from January 1986 to October 2013. [Table 1](#) lists all commodities including their delivery months within a calendar year that we use to construct the corresponding excess return series. It is important to emphasize that each of the 15 commodities considered in the empirical analysis is also included in either the Standard & Poor's Goldman Sachs Commodity Index (S&P GSCI) or the Dow Jones UBS Commodity Index.¹³ Therefore, our sample provides a diversified representation of commodity futures markets as an asset class.

From the daily settlement prices, we use the end-of-month value to compute the futures excess return on a fully collateralized long position in commodity j as

$$\Delta f_{t+1}^j = \log \left[\frac{F_{t+1}^{j,T_r}}{F_t^{j,T_r}} \right], \quad (15)$$

where F_t^{j,T_r} is the futures price at the end of month t for the contract with delivery month $t + T_r$. We consider returns on the next-to-nearest futures (i.e., $\tau = 2$) and exclude returns on contracts with less than one month to maturity since in this case the futures trader has to arrange a physical delivery of the underlying commodity. Therefore, the sequence $\{\Delta f_t^j\}_{t=1}^T$, used in our empirical

analysis, can be viewed as the realized excess return to an investor maintaining a long position in the futures contract on commodity j and maturity $t + T_2$, rolling over on the last trading day of the month prior to maturity.

Consistent with the standard practice in the literature (see, among others, [Gorton et al., 2013](#); [Yang, 2013](#), and references therein), we compute the futures basis of commodity j as the difference in log prices between the two nearest-to-maturity futures:

$$z_t^j = \text{Basis}_t^j = \frac{\log(F_t^{j,T_1}) - \log(F_t^{j,T_2})}{T_2 - T_1}, \quad (16)$$

where F_t^{j,T_1} and F_t^{j,T_2} are the futures prices of the nearby and next-to-nearest contracts, respectively.¹⁴ We then construct five commodity futures portfolios sorted by the basis. In particular, at the end of month $t - 1$, futures contracts of all commodities are assigned to five portfolios based on their corresponding basis observed at the end of month $t - 1$. The first portfolio comprises the 20% of sample commodity futures contracts with the lowest basis, while the last portfolio comprises the 20% of commodity futures contracts with the highest basis. We repeat this strategy at the end of every month so that the portfolios are rebalanced. The monthly excess return on a commodity futures portfolio constructed at the end of month $t - 1$, but realized at the end of month t is computed as the equally weighted average of excess returns for the constituent contracts. We construct the EWA factor as the equally weighted average excess return on a long position in all available futures contracts. On the other hand, the HML factor is created as the return difference between the last and first portfolios. The correlation between EWA and HML factors is 7.9% (p -value = 0.15) for the full-sample period.

3.2. Full-sample estimates

We first focus on the full-sample period (1986:01–2013:10) estimates of the factor models specified in [Eqs. \(3\)–\(10\)](#). For brevity, we report the OLS estimates with [Newey and West \(1987\)](#) standard errors for the models only at the portfolio-level in [Table 2](#).¹⁵ The

¹³ For robustness checks of the main results, we also carry out empirical analysis using samples of 21 and 30 commodity futures contracts as in [Szymanowska et al. \(2014\)](#) and [Hong and Yogo \(2012\)](#), respectively, for the period between January 1986 and October 2013. However, there is little evidence of out-of-sample predictability in excess returns from factor models using commodity-based risk factors. Hence, we abstain from reporting these results in the paper.

¹⁴ In theory, the basis of a commodity is defined as the difference between its current spot price and contemporaneous futures price with relevant maturity. However, in reality, commodity spot markets are quite illiquid and nearest-to-maturity futures price serves as a close approximation of the spot price.

¹⁵ The full-sample period OLS estimates of the factor models for individual commodity futures are reported in the Internet Appendix.

Table 2
Full-sample OLS estimates: portfolio-level.

Model	\hat{a}_0	\hat{a}_1	$\hat{b}_{0,1}$	$\hat{b}_{0,2}$	$\hat{b}_{1,1}$	$\hat{b}_{1,2}$	R^2_{adj}
Portfolio 1							
(3)	−0.755*** (0.256)		1.025*** (0.076)				42.13
(4)	−0.074 (0.161)		1.099*** (0.052)	−0.474*** (0.041)			77.12
(5)	−0.805*** (0.268)		0.766*** (0.084)		−0.098*** (0.025)		43.57
(6)	−0.086 (0.168)		1.080*** (0.077)	−0.373*** (0.051)	0.002 (0.025)	0.035 (0.022)	77.56
(7)	−0.660* (0.395)	0.037 (0.117)	1.026*** (0.076)				41.97
(8)	0.058 (0.248)	0.051 (0.075)	1.101*** (0.052)	−0.474*** (0.041)			77.08
(9)	−0.192 (0.435)	0.243 (0.183)	0.710*** (0.107)		−0.122*** (0.046)		43.90
(10)	0.163 (0.289)	0.100 (0.110)	1.056*** (0.079)	−0.371*** (0.052)	−0.009 (0.028)	0.035 (0.022)	77.58
Portfolio 2							
(3)	−0.389* (0.214)		0.893*** (0.061)				42.60
(4)	−0.348* (0.206)		0.897*** (0.064)	−0.028 (0.036)			42.60
(5)	−0.389* (0.215)		0.927*** (0.136)		0.035 (0.137)		42.44
(6)	−0.348* (0.207)		0.928*** (0.136)	−0.023 (0.061)	0.032 (0.137)	0.005 (0.059)	42.26
(7)	−0.247 (0.466)	0.146 (0.476)	0.893*** (0.061)				42.44
(8)	−0.240 (0.466)	0.112 (0.471)	0.897*** (0.063)	−0.028 (0.036)			42.43
(9)	−0.252 (0.466)	0.141 (0.475)	0.926*** (0.134)		0.034 (0.136)		42.28
(10)	−0.247 (0.461)	0.104 (0.465)	0.927*** (0.135)	−0.026 (0.061)	0.030 (0.135)	0.001 (0.058)	42.09
Portfolio 3							
(3)	0.118 (0.214)		0.887*** (0.081)				41.40
(4)	0.163 (0.206)		0.891*** (0.081)	−0.031 (0.039)			41.41
(5)	0.114 (0.211)		0.949*** (0.103)		0.132 (0.192)		41.33
(6)	0.166 (0.203)		0.949*** (0.104)	0.005 (0.058)	0.119 (0.198)	0.077 (0.080)	41.37
(7)	0.328 (0.280)	0.494 (0.522)	0.886*** (0.082)				41.36
(8)	0.356 (0.276)	0.461 (0.521)	0.890*** (0.082)	−0.029 (0.039)			41.36
(9)	0.326 (0.274)	0.501 (0.515)	0.949*** (0.105)		0.135 (0.198)		41.30
(10)	0.295 (0.282)	0.310 (0.546)	0.949*** (0.105)	0.001 (0.059)	0.122 (0.202)	0.064 (0.085)	41.24
Portfolio 4							
(3)	0.345 (0.223)		1.014*** (0.067)				43.52
(4)	0.333 (0.235)		1.013*** (0.069)	0.008 (0.041)			43.36
(5)	0.341 (0.225)		1.007*** (0.067)		0.082 (0.146)		43.45
(6)	0.313 (0.234)		1.008*** (0.068)	0.015 (0.040)	0.100 (0.157)	−0.046 (0.066)	43.28
(7)	0.355 (0.217)	−0.040 (0.472)	1.014*** (0.068)				43.35
(8)	0.342 (0.229)	−0.032 (0.471)	1.013*** (0.070)	0.007 (0.040)			43.19
(9)	0.349 (0.219)	−0.032 (0.469)	1.007*** (0.067)		0.082 (0.146)		43.27
(10)	0.312 (0.232)	0.003 (0.495)	1.008*** (0.068)	0.015 (0.041)	0.100 (0.157)	−0.046 (0.068)	43.10

(continued on next page)

Table 2 (continued)

Model	\hat{a}_0	\hat{a}_1	$\hat{b}_{0,1}$	$\hat{b}_{0,2}$	$\hat{b}_{1,1}$	$\hat{b}_{1,2}$	$R^2_{adj,i}$
Portfolio 5							
(3)	0.682** (0.267)		1.181*** (0.090)				45.92
(4)	−0.074 (0.161)		1.099*** (0.052)	0.526*** (0.041)			81.22
(5)	0.694*** (0.266)		0.949*** (0.113)		0.103*** (0.033)		46.93
(6)	−0.058 (0.160)		1.001*** (0.078)	0.466*** (0.063)	0.041 (0.026)	0.020 (0.014)	81.67
(7)	0.750 (0.470)	−0.025 (0.146)	1.181*** (0.090)				45.77
(8)	−0.097 (0.241)	0.008 (0.078)	1.099*** (0.052)	0.526*** (0.042)			81.16
(9)	0.722 (0.461)	−0.010 (0.141)	0.949*** (0.113)		0.103*** (0.033)		46.77
(10)	−0.037 (0.225)	−0.008 (0.064)	1.001*** (0.078)	0.465*** (0.063)	0.041 (0.026)	0.021 (0.014)	81.61

The table reports the ordinary least squares (OLS) estimates of the factor models in Eqs. (3)–(10) for basis-sorted commodity futures portfolios. All variables are in percentage points. Numbers in parentheses are the Newey and West (1987) standard errors. $R^2_{adj,i}$ denotes the adjusted R^2 (%). *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively. The sample period is from January 1986 to October 2013.

fairly flat coefficient estimates (i.e., very close to 1) of the EWA factor across all model and basis-sorted portfolio combinations turn out to be statistically significant at any conventional level. On the other hand, the coefficient estimates of the HML factor are statistically significant at the 1% level for the first and last portfolios. The observed pattern in the betas of the HML factor is very similar to the way the factor has been constructed. The adjusted R^2 values of the regressions range widely between 41.24% and 81.67%. These results in the time-series domain are largely consistent with those reported by Yang (2013).

3.3. Out-of-sample predictions

3.3.1. Commodity-level excess return forecasts: a time-series approach

Our out-of-sample empirical analysis begins by generating one-month ahead time-series forecasts of individual commodity futures excess returns based on unconditional expectations of risk factors. Table 3 summarizes the results for forecasting accuracy of the factor models relative to the random walk with drift benchmark. It is noticeable that the Campbell and Thompson (2008) R^2_{OOS} statistic is positive and statistically significant only for a handful alternative factor model and commodity futures combinations: one out of six models for coffee, heating oil, and lean hogs futures; three out of six models for crude oil futures; and four out of six models for gasoline and gold futures. The Clark and West (2006, 2007) test for pairwise model forecast comparison occasionally (i.e., only one out of six models for coffee, heating oil, and wheat futures; two out of six models for crude oil futures; four out of six models for gold futures; and five out of six models for gasoline futures) rejects the null hypothesis of equal predictive ability at the 10% significance level. In general, we observe similar level of predictive performance of the risk factors while focusing on the Giacomini and White (2006) test, which mostly fails to reject the null hypothesis of equal conditional predictive ability at the 10% significance level and to choose an alternative factor model over the benchmark random walk with drift model. Importantly, the reality check version of the Clark and West (2006, 2007) test, following Clark and McCracken (2012), suggests that with the exception of gasoline futures, none of the alternative models for commodity futures excess returns statistically outperforms the benchmark by producing more accurate forecasts in the out-of-sample period.

The empirical results obtained using time-series forecasts based on conditional expectations of commodity-based risk factors are reported in Table 4. We find the Campbell and Thompson

(2008) R^2_{OOS} statistic is significantly positive only for coffee (one out of eight models), copper, cotton (two out of eight models), gasoline (four out of eight models), heating oil (two out of eight models), and soybeans futures (three out of eight models). While the Clark and West (2006, 2007) pairwise test offers some evidence of predictability for copper and gasoline futures excess returns (i.e., factor models given by Eqs. (3)–(6), and by Eqs. (7)–(9), respectively), the results do not hold under the reality check version of the test. With few exceptions (i.e., only one out of eight models for coffee, crude oil, and soybeans futures; two out of eight models for cotton futures; and six out of eight models for copper futures), the equal conditional predictive ability test of Giacomini and White (2006) also fails to reject the null hypothesis at the 10% significance level and to choose an alternative model over the benchmark. The evidence of weak predictive power of commodity-based risk factors in the short-horizon is consistent with that based on the Clark and West (2006, 2007) test applied on a pairwise basis.

3.3.2. Commodity-level excess return forecasts: a cross-sectional approach

In this section, we aim at determining whether including commodity-specific characteristic(s) in a cross-sectional framework helps improve the forecasting performance of the factor models.¹⁶ For this purpose, we adopt the Fama and MacBeth (1973) cross-sectional method following the bias adjustment of Brennan et al. (1998) to generate one-month ahead forecasts of commodity futures excess returns in the out-of-sample period. At the end of each month t , the coefficients, α and β , of a given factor model in equation (i) ($i = 3, 4, \dots, 10$) for all commodity futures excess returns are estimated using time-series regressions based on a 10-year rolling window ending at month $t - 1$. Then the fitted values of the model for month t are obtained using the coefficient estimates (excluding the intercept a_0), the risk factor(s), and if needed the instrument, which in our case is the futures basis, to run the following cross-sectional regression:

$$\Delta f^j_t - \widehat{\Delta f^j_{i,t}} = \lambda_{i,t} C^j_{t-1} + u^j_{i,t}, \text{ for a given } t \text{ and } i, \forall j \quad (17)$$

where $\widehat{\Delta f^j_{i,t}}$ is the fitted value of the factor model in consideration and C^j_t denotes commodity-specific characteristic(s) other than

¹⁶ A cross-sectional approach employs information in the full cross-section of commodity futures.

Table 3
Forecasts based on unconditional expectations of factors.

		Model						RC
		(5)	(6)	(7)	(8)	(9)	(10)	
Cocoa	R^2_{005}	−0.14	0.03	−0.68	−0.80	−0.99	−1.25	0.25
	CW	−1.59	0.25	−0.16	−0.17	−0.32	−0.54	
	GW	0.02 ^b (0.11)	0.65 (0.69)	0.70 (0.07)	0.66 (0.07)	0.56 (0.05)	0.47 (0.02)	
Coffee	R^2_{005}	−0.41	0.36	−1.62	−2.16	−1.71	0.64**	1.12
	CW	−1.22	0.62	0.06	−0.13	0.11	1.12*	
	GW	0.00 ^b (0.26)	0.00 ^a (0.71)	0.00 ^b (0.18)	0.00 ^b (0.16)	0.00 ^b (0.19)	0.00 ^a (0.62)	
Copper	R^2_{005}	0.01	0.13	−0.55	−0.56	−1.07	−1.09	0.45
	CW	0.10	0.45	−0.03	−0.06	−0.20	−0.32	
	GW	0.51 (0.68)	0.55 (0.78)	0.45 (0.16)	0.41 (0.16)	0.18 (0.16)	0.14 (0.15)	
Corn	R^2_{005}	0.04	−1.13	−4.04	−3.99	−3.91	−1.87	0.25
	CW	0.25	−0.76	−1.07	−1.07	−1.12	−0.75	
	GW	0.99 (0.98)	0.11 (0.06)	0.11 (0.04)	0.07 ^b (0.04)	0.04 ^b (0.04)	0.00 ^b (0.15)	
Cotton	R^2_{005}	0.05	0.24	−3.06	−3.06	−2.92	−2.99	0.42
	CW	0.42	0.35	−1.45	−1.43	−1.21	−0.82	
	GW	0.65 (0.89)	0.06 ^a (0.76)	0.03 ^b (0.04)	0.04 ^b (0.02)	0.28 (0.00)	0.48 (0.00)	
Crude oil	R^2_{005}	−0.06	1.60***	0.31*	0.24	0.09	1.64***	1.39
	CW	−0.75	1.38*	0.97	0.92	0.78	1.39**	
	GW	0.40 (0.18)	0.00 ^a (0.78)	0.44 (0.76)	0.51 (0.72)	0.54 (0.60)	0.00 ^a (0.83)	
Gasoline	R^2_{005}	0.15	−1.62	1.52**	1.69**	1.57**	1.41***	2.14**
	CW	1.47*	−1.38	2.07***	2.14***	2.12***	2.08**	
	GW	0.40 (1.00)	0.17 (0.04)	0.35 (0.94)	0.31 (0.93)	0.39 (0.96)	0.27 (0.99)	
Gold	R^2_{005}	−0.08	−0.04	0.12*	0.13*	0.12*	0.26*	1.08
	CW	−1.47	−0.11	0.92*	0.92*	0.92*	1.08*	
	GW	0.32 (0.00)	0.74 (0.27)	0.06 ^a (0.59)	0.09 ^a (0.62)	0.03 ^a (0.58)	0.11 (0.66)	
Heating oil	R^2_{005}	−0.08	0.72**	−1.06	−1.04	−0.74	−0.37	1.36
	CW	−0.44	1.36*	−0.23	−0.30	0.06	−0.12	
	GW	0.82 (0.04)	0.00 ^a (0.73)	0.79 (0.01)	0.72 (0.02)	0.94 (0.00)	0.64 (0.07)	
Lean hogs	R^2_{005}	0.09	0.77*	−1.27	−1.14	−1.37	−0.95	1.12
	CW	0.84	1.12	−0.32	−0.22	−0.37	−0.03	
	GW	0.00 ^a (0.79)	0.78 (0.98)	0.75 (0.02)	0.85 (0.00)	0.76 (0.00)	0.33 (0.17)	
Live cattle	R^2_{005}	−0.14	−0.95	−3.49	−3.22	−3.66	−3.36	−0.32
	CW	−1.57	−1.52	−0.42	−0.32	−0.47	−0.68	
	GW	0.07 ^b (0.05)	0.06 ^b (0.09)	0.13 (0.06)	0.19 (0.06)	0.13 (0.05)	0.09 ^b (0.04)	
Silver	R^2_{005}	−0.10	0.09	−1.35	−1.31	−1.51	−1.29	0.55
	CW	−1.93	0.55	−2.56	−2.56	−2.34	−2.12	
	GW	0.02 ^b (0.12)	0.44 (0.78)	0.00 ^b (0.09)	0.00 ^b (0.08)	0.00 ^b (0.13)	0.00 ^b (0.15)	
Soybeans	R^2_{005}	−0.28	−0.99	−3.97	−4.02	−3.71	−5.06	0.09
	CW	−1.79	0.09	−1.81	−1.64	−1.49	−0.43	
	GW	0.00 ^b (0.20)	0.00 ^b (0.16)	0.02 ^b (0.01)	0.08 ^b (0.00)	0.01 ^b (0.03)	0.03 ^b (0.04)	
Sugar	R^2_{005}	−0.11	−2.32	−24.13	−23.61	−25.40	−15.67	0.04
	CW	−0.59	0.04	−1.26	−1.25	−1.58	−0.64	
	GW	0.02 ^b (0.25)	0.00 ^b (0.09)	0.51 (0.00)	0.47 (0.00)	0.43 (0.00)	0.00 ^b (0.05)	
Wheat	R^2_{005}	0.18	0.27	−1.40	−1.40	−1.20	−1.19	1.60
	CW	1.60*	1.29	−1.37	−1.40	−1.20	−1.20	
	GW	0.02 ^a (0.76)	0.14 (0.83)	0.25 (0.01)	0.24 (0.01)	0.34 (0.01)	0.28 (0.02)	

The table reports out-of-sample tests of forecast accuracy of factor models relative to a random walk with drift (RWD) benchmark. The one-month ahead forecasts of commodity futures excess returns are obtained from rolling regressions with a 10-year window using unconditional expectations of risk factors. For each alternative factor model specified in equation (i) ($i = 5, 6, \dots, 10$) and commodity j , R^2_{005} is the [Campbell and Thompson \(2008\)](#) out-of-sample $R^2_{i,j}(\%)$ statistic. CW is the [Clark and West \(2006, 2007\)](#) $MSFE^j_{i,adj} - t$ statistic based on the [Newey and West \(1987\)](#) estimator, which tests whether the adjusted mean squared forecast error (MSFE) difference between the RWD and the alternative model is zero. RC is the $\max_{i=5,6,\dots,10} MSFE^j_{i,adj} - t$ statistic for a reality check following [Clark and McCracken \(2012\)](#), which tests the composite null hypothesis that the RWD is not inferior to any of the alternative models. One-sided critical values for the R^2_{005} , CW, and RC statistics are obtained using 10,000 fixed regressor bootstrap replications, as per [Clark and McCracken \(2012\)](#). *, **, and *** denote statistical significance of the R^2_{005} , CW, and RC statistics at the 10%, 5%, and 1% levels, respectively. GW is the [Giacomini and White \(2006\)](#) p^j_i -value of the equal conditional predictive ability test. Numbers in parentheses are the proportion of times the alternative model outperforms the RWD in the out of sample. Superscript *a* (*b*) indicates that the GW test rejects the null hypothesis of equal conditional predictive ability at the 10% significance level and the alternative model outperforms (is outperformed by) the RWD more than 50% of the time. The out-of-sample data run from January 1996 through October 2013.

Table 4
Forecasts based on conditional expectations of factors.

		Model								RC
		(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	
Cocoa	R^2_{005}	-0.23	-0.17	-0.47	-0.34	-0.91	-0.91	-1.34	-1.62	-0.31
	CW	-1.42	-1.01	-1.99	-0.51	-0.31	-0.31	-0.50	-0.79	
	GW	0.05 ^b (0.11)	0.13 (0.12)	0.04 ^b (0.07)	0.24 (0.11)	0.65 (0.04)	0.65 (0.04)	0.49 (0.04)	0.36 (0.02)	
Coffee	R^2_{005}	0.17	0.69**	-1.19	-1.97	-1.70	-1.70	-2.78	-2.00	0.75
	CW	0.40	0.75	-0.32	-0.50	0.15	0.15	0.11	0.02	
	GW	0.00 ^a (0.72)	0.47 (0.94)	0.00 ^b (0.10)	0.10 (0.05)	0.00 ^b (0.17)	0.00 ^b (0.17)	0.00 ^b (0.17)	0.00 ^b (0.17)	
Copper	R^2_{005}	1.43***	1.46***	1.77*	1.81**	0.76*	0.76*	1.09*	1.13*	1.60
	CW	1.55*	1.57*	1.60**	1.59**	0.94	0.94	1.10	1.11	
	GW	0.00 ^a (0.86)	0.00 ^a (0.86)	0.00 ^a (0.86)	0.00 ^a (0.83)	0.10 (0.81)	0.10 (0.81)	0.09 ^a (0.81)	0.08 ^a (0.82)	
Corn	R^2_{005}	-1.59	-1.71	-4.66	-4.61	-5.51	-5.51	-6.32	-6.64	-0.23
	CW	-0.69	-0.70	-0.23	-0.46	-1.05	-1.05	-0.64	-0.37	
	GW	0.00 ^b (0.04)	0.00 ^b (0.04)	0.00 ^b (0.01)	0.01 ^b (0.01)	0.16 (0.02)	0.16 (0.02)	0.01 ^b (0.05)	0.06 ^b (0.04)	
Cotton	R^2_{005}	0.31**	0.31*	-0.58	-0.09	-2.77	-2.77	-4.27	-3.43	0.51
	CW	0.46	0.46	0.38	0.51	-0.77	-0.77	-0.35	-0.31	
	GW	0.00 ^a (0.73)	0.00 ^a (0.72)	0.00 ^b (0.20)	0.00 ^b (0.49)	0.17 (0.02)	0.17 (0.02)	0.20 (0.01)	0.13 (0.02)	
Crude oil	R^2_{005}	-0.90	-0.85	-0.81	-0.50	-0.17	-0.17	-0.13	-0.03	1.01
	CW	-0.08	-0.01	0.00	0.93	0.58	0.58	0.61	1.01	
	GW	0.72 (0.05)	0.74 (0.06)	0.81 (0.01)	0.38 (0.33)	0.20 (0.39)	0.20 (0.39)	0.32 (0.39)	0.06 ^a (0.52)	
Gasoline	R^2_{005}	-0.79	-1.13	-0.46	-2.61	0.48*	0.48*	0.76**	0.40**	1.56
	CW	-1.01	-1.28	-0.41	-0.77	1.32*	1.32*	1.56**	0.79	
	GW	0.34 (0.03)	0.22 (0.03)	0.65 (0.04)	0.42 (0.00)	0.75 (0.81)	0.75 (0.81)	0.80 (0.93)	0.82 (0.87)	
Gold	R^2_{005}	-0.19	-0.19	-0.24	-0.17	-0.32	-0.32	-0.36	-0.22	0.82
	CW	-0.31	-0.37	-0.56	-0.35	0.65	0.65	0.65	0.82	
	GW	0.04 ^b (0.28)	0.01 ^b (0.30)	0.33 (0.18)	0.00 ^b (0.31)	0.05 ^b (0.40)	0.05 ^b (0.40)	0.04 ^b (0.40)	0.08 ^b (0.42)	
Heating oil	R^2_{005}	0.44*	0.48*	0.59	0.32	-0.74	-0.74	-0.09	0.03	1.00
	CW	0.78	0.78	1.00	0.62	0.32	0.32	0.69	0.43	
	GW	0.88 (0.97)	0.81 (0.93)	0.69 (0.94)	0.97 (0.93)	0.58 (0.12)	0.58 (0.12)	0.65 (0.39)	0.89 (0.41)	
Lean hogs	R^2_{005}	-0.64	-0.79	-0.87	0.14	-1.71	-1.71	-2.02	-1.16	0.43
	CW	-0.63	-0.62	-0.71	0.43	-0.66	-0.66	-0.78	-0.20	
	GW	0.00 ^b (0.17)	0.00 ^b (0.15)	0.00 ^b (0.20)	0.99 (0.99)	0.63 (0.00)	0.63 (0.00)	0.54 (0.00)	0.82 (0.02)	
Live cattle	R^2_{005}	-0.10	-0.42	-0.03	-1.15	-3.50	-3.50	-3.45	-3.27	-0.22
	CW	-0.86	-1.29	-0.22	-1.10	-0.47	-0.47	-0.44	-0.73	
	GW	0.01 ^b (0.23)	0.00 ^b (0.15)	0.28 (0.26)	0.18 (0.04)	0.11 (0.06)	0.11 (0.06)	0.15 (0.05)	0.01 ^b (0.08)	
Silver	R^2_{005}	-0.01	-0.03	-0.02	0.09	-1.35	-1.35	-1.38	-1.31	0.34
	CW	0.08	0.02	0.07	0.34	-1.86	-1.86	-1.74	-1.41	
	GW	0.80 (0.41)	0.87 (0.28)	0.65 (0.37)	0.89 (0.83)	0.03 ^b (0.06)	0.03 ^b (0.06)	0.02 ^b (0.08)	0.06 ^b (0.08)	
Soybeans	R^2_{005}	0.36**	0.38	-0.07	1.65*	-3.52	-3.52	-3.32	1.29**	1.07
	CW	0.89	0.59	0.14	1.01	-1.95	-1.95	-1.76	1.07	
	GW	0.19 (0.95)	0.09 ^a (0.93)	0.99 (0.02)	0.58 (0.98)	0.03 ^b (0.00)	0.03 ^b (0.00)	0.05 ^b (0.00)	0.82 (0.98)	
Sugar	R^2_{005}	-1.04	-0.82	-52.31	-6.34	-18.21	-18.21	-33.28	-26.20	0.95
	CW	-0.68	-0.62	0.95	-0.67	-1.19	-1.19	0.57	-0.76	
	GW	0.02 ^b (0.03)	0.11 (0.02)	0.60 (0.00)	0.00 ^b (0.06)	0.16 (0.00)	0.16 (0.00)	0.59 (0.00)	0.00 ^b (0.04)	
Wheat	R^2_{005}	-0.31	-0.31	-0.24	-0.12	-1.62	-1.62	-1.51	-1.46	0.09
	CW	-0.67	-0.64	-0.31	0.09	-1.37	-1.37	-1.28	-1.34	
	GW	0.08 ^b (0.16)	0.06 ^b (0.17)	0.06 ^b (0.27)	0.04 ^b (0.36)	0.23 (0.02)	0.23 (0.02)	0.25 (0.02)	0.23 (0.02)	

The setup is the same as in Table 3 except that conditional expectations of risk factors are used to generate one-month ahead forecasts of commodity futures excess returns from factor models in Eqs. (3)–(10). See also notes to Table 3.

Table 5

Forecasts based on unconditional expectations of factors: a cross-sectional approach with hedging pressure.

	Model					
	(5)	(6)	(7)	(8)	(9)	(10)
Cocoa	0.00 ^b (0.05)	0.00 ^b (0.02)	0.90 (0.78)	0.84 (0.64)	0.00 ^b (0.07)	0.00 ^b (0.02)
Coffee	0.00 ^b (0.03)	0.00 ^b (0.04)	0.11 (0.11)	0.16 (0.10)	0.00 ^b (0.04)	0.00 ^b (0.04)
Copper	0.00 ^b (0.05)	0.00 ^b (0.06)	0.21 (0.60)	0.04 ^a (0.61)	0.00 ^b (0.08)	0.00 ^b (0.07)
Corn	0.00 ^b (0.02)	0.00 ^b (0.02)	0.48 (0.00)	0.42 (0.00)	0.00 ^b (0.02)	0.00 ^b (0.02)
Cotton	0.00 ^b (0.13)	0.00 ^b (0.07)	0.28 (0.00)	0.24 (0.00)	0.00 ^b (0.13)	0.00 ^b (0.09)
Crude oil	0.00 ^a (0.61)	0.00 ^b (0.02)	0.06 ^b (0.37)	0.08 ^b (0.34)	0.00 ^a (0.57)	0.00 ^b (0.02)
Gasoline	0.01 ^b (0.00)	0.01 ^b (0.00)	0.67 (0.02)	0.41 (0.08)	0.07 ^b (0.00)	0.10 (0.00)
Gold	0.00 ^b (0.02)	0.00 ^b (0.01)	0.00 ^b (0.00)	0.00 ^b (0.00)	0.00 ^b (0.03)	0.00 ^b (0.01)
Heating oil	0.00 ^b (0.00)	0.00 ^b (0.05)	0.01 ^b (0.16)	0.02 ^b (0.13)	0.00 ^b (0.01)	0.00 ^b (0.05)
Lean hogs	0.00 ^b (0.13)	0.00 ^b (0.04)	0.03 ^b (0.01)	0.04 ^b (0.00)	0.01 ^b (0.06)	0.00 ^b (0.00)
Live cattle	0.00 ^b (0.00)	0.03 ^b (0.09)	0.13 (0.12)	0.08 ^b (0.13)	0.00 ^b (0.00)	0.01 ^b (0.05)
Silver	0.00 ^b (0.02)	0.00 ^b (0.00)	0.00 ^b (0.03)	0.00 ^b (0.03)	0.00 ^b (0.01)	0.00 ^b (0.00)
Soybeans	0.00 ^b (0.01)	0.01 ^b (0.00)	0.15 (0.01)	0.14 (0.00)	0.00 ^b (0.00)	0.00 ^b (0.00)
Sugar	1.00 (0.35)	0.00 ^b (0.08)	0.17 (0.06)	0.06 ^b (0.07)	0.00 ^b (0.04)	0.00 ^b (0.06)
Wheat	0.00 ^b (0.04)	0.00 ^b (0.03)	0.06 ^b (0.20)	0.06 ^b (0.22)	0.00 ^b (0.05)	0.00 ^b (0.03)

The table reports the [Giacomini and White \(2006\)](#) p^j -values of the equal conditional predictive ability tests for commodity j . The one-month ahead forecasts of commodity futures excess returns are obtained by the [Fama and MacBeth \(1973\)](#) cross-sectional procedure following [Brennan et al. \(1998\)](#). The forecasting scheme based on rolling regressions with a 10-year window utilizes unconditional expectations of risk factors and commodity-specific hedging pressure. Numbers in parentheses are the proportion of times the alternative factor model in equation (i) ($i = 5, 6, \dots, 10$) outperforms the random walk with drift (RWD) benchmark in the out of sample. Superscript a (b) indicates that the test rejects the null hypothesis of equal conditional predictive ability at the 10% significance level and the alternative model outperforms (is outperformed by) the RWD more than 50% of the time. The out-of-sample data run from February 1996 through October 2013.

the futures basis. Following [Simin \(2008\)](#), we compute one-month ahead forecast of excess return on commodity futures j at the end of month t as ¹⁷

$$\widehat{\Delta f}_{i,t+1|t}^j = \widehat{\Delta f}_{i,t+1|t}^j + \hat{\lambda}_{i,t} C_t^j. \quad (18)$$

The entire procedure, described above, is repeated for each of the factor models at a monthly frequency to generate the corresponding time-series of out-of-sample forecast errors for the commodity futures excess returns. As in the time-series approach, we continue to exploit both the unconditional and conditional expectations of the risk factors constructed in [Section 2.1](#).

To perform the cross-sectional exercises, we consider two variables as commodity-specific characteristics, namely, the hedging pressure and the open interest. There is an extensive literature that relates the futures risk premia to net positions of hedgers (see, among others, [Acharya et al., 2013](#); [Basu and Miffre, 2013](#); [Bessembinder, 1992](#); [Carter et al., 1983](#); [de Roon et al., 2000](#), and references therein).¹⁸ Consistent with [de Roon et al. \(2000\)](#), we compute the hedging pressure of a commodity futures contract as

the difference between the numbers of short and long hedge positions relative to the total number of hedge positions by commercial traders (i.e., hedgers).¹⁹ The choice of the open interest is motivated by the recent findings of [Hong and Yogo \(2012\)](#) that the amount of contracts outstanding in the commodity futures market predicts movements in commodity prices.

In [Table 5](#), we summarize the out-of-sample forecasting results based on the hedging pressure characteristic and unconditional expectations of the commodity-based risk factors. With the exception of copper and crude oil futures (i.e., factor models given by [Eq. \(8\)](#), and by [Eqs. \(5\) and \(9\)](#), respectively), the [Giacomini and White \(2006\)](#) test suggests that none of the alternative factor models for commodity futures excess returns statistically outperforms the benchmark random walk with drift model. The out-of-sample predictive performance of the factor models does not improve when the conditional expectations of the risk factors are utilized to generate the one-month step ahead forecasts. These results are summarized in [Table 6](#). The use of the open interest as a characteristic variable C_t^j offers qualitatively similar results to those in [Tables 5 and 6](#). To conserve space, we report these results in the Internet Appendix.

¹⁷ To avoid clustering of notations, we once again use $\widehat{\Delta f}_{i,t+1|t}^j$ to denote one-month ahead forecast from a factor model specified in equation (i) ($i = 3, 4, \dots, 10$).

¹⁸ The hypothesis of hedging pressure is underpinned by the theory of [Keynes \(1923,1930\)](#) and [Hicks \(1939\)](#), who argue that hedgers are ready to sell the futures at a discount from the expected future spot price (i.e., to pay a premium) in order to induce speculators to buy the contract. A formal equilibrium model of this hypothesis is provided by [Hirshleifer \(1988,1989\)](#).

¹⁹ The data on long and short hedge positions are obtained from the CFTC.

Table 6

Forecasts based on conditional expectations of factors: a cross-sectional approach with hedging pressure.

	Model							
	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Cocoa	0.48 (0.89)	0.50 (0.82)	0.65 (0.73)	0.65 (0.24)	0.88 (0.67)	0.76 (0.46)	0.96 (0.00)	0.60 (0.15)
Coffee	0.42 (0.09)	0.60 (0.05)	0.03 ^b (0.10)	0.23 (0.08)	0.05 ^b (0.13)	0.08 ^b (0.11)	0.02 ^b (0.11)	0.07 ^b (0.10)
Copper	0.04 ^a (0.77)	0.01 ^a (0.76)	0.06 ^a (0.78)	0.34 (0.84)	0.18 (0.77)	0.04 ^a (0.74)	0.11 (0.80)	0.88 (0.81)
Corn	0.30 (0.10)	0.27 (0.09)	0.00 ^b (0.09)	0.00 ^b (0.10)	0.07 ^b (0.13)	0.07 ^b (0.13)	0.01 ^b (0.12)	0.00 ^b (0.12)
Cotton	0.71 (0.06)	0.70 (0.05)	0.16 (0.11)	0.11 (0.08)	0.35 (0.00)	0.33 (0.00)	0.20 (0.05)	0.13 (0.04)
Crude oil	0.00 ^b (0.33)	0.05 ^b (0.26)	0.00 ^b (0.31)	0.02 ^b (0.29)	0.00 ^b (0.37)	0.07 ^b (0.28)	0.01 ^b (0.33)	0.02 ^b (0.30)
Gasoline	0.14 (0.00)	0.05 ^b (0.00)	0.17 (0.00)	0.06 ^b (0.00)	0.51 (0.00)	0.24 (0.03)	0.36 (0.03)	0.51 (0.01)
Gold	0.00 ^b (0.01)	0.00 ^b (0.00)	0.00 ^b (0.00)	0.00 ^b (0.00)	0.00 ^b (0.01)	0.00 ^b (0.00)	0.00 ^b (0.00)	0.00 ^b (0.00)
Heating oil	0.12 (0.27)	0.37 (0.12)	0.15 (0.29)	0.25 (0.17)	0.45 (0.05)	0.45 (0.00)	0.52 (0.12)	0.30 (0.10)
Lean hogs	0.01 ^b (0.04)	0.01 ^b (0.03)	0.01 ^b (0.06)	0.00 ^b (0.07)	0.03 ^b (0.00)	0.02 ^b (0.00)	0.03 ^b (0.02)	0.02 ^b (0.03)
Live cattle	0.00 ^b (0.20)	0.02 ^b (0.20)	0.00 ^b (0.21)	0.03 ^b (0.15)	0.13 (0.11)	0.20 (0.09)	0.13 (0.12)	0.08 ^b (0.12)
Silver	0.00 ^b (0.01)	0.00 ^b (0.01)	0.00 ^b (0.00)	0.00 ^b (0.03)	0.00 ^b (0.03)	0.00 ^b (0.03)	0.00 ^b (0.00)	0.00 ^b (0.05)
Soybeans	0.43 (0.00)	0.48 (0.00)	0.59 (0.00)	0.11 (0.24)	0.20 (0.01)	0.21 (0.02)	0.29 (0.02)	0.33 (0.06)
Sugar	0.53 (0.00)	0.26 (0.00)	0.28 (0.00)	0.00 ^b (0.07)	0.00 ^b (0.12)	0.00 ^b (0.10)	0.22 (0.00)	0.00 ^b (0.08)
Wheat	0.32 (0.17)	0.35 (0.24)	0.29 (0.13)	0.61 (0.10)	0.10 (0.13)	0.08 ^b (0.20)	0.05 ^b (0.15)	0.10 (0.15)

The setup is the same as in Table 5 except that conditional expectations of risk factors are used to generate one-month ahead forecasts of commodity futures excess returns from factor models in Eqs. (3)–(10). See also notes to Table 5.

3.3.3. Portfolio-level excess return forecasts

Table 7 presents the out-of-sample results for the basis-sorted commodity futures portfolios, where we use the unconditional expectations of the risk factors. We observe that none of the Campbell and Thompson (2008) R_{OOS}^2 statistic is significantly positive. The pairwise and reality check versions of the Clark and West (2006, 2007) test suggest that the factor model given by Eq. (5) for the futures portfolio 3 is statistically superior to that of the random walk with drift benchmark. Only for the commodity futures portfolios 4 and 5, the Giacomini and White (2006) test of equal conditional predictive ability rejects the null hypothesis at the 10% significance level and chooses one out of six alternative models over the benchmark. Moving to the out-of-sample forecasts based on the conditional expectations of the risk factors, we see that the forecasting performance of the alternative factor models deteriorates relative to the benchmark. For instance, none of the reality check test statistics in Table 8 turns out to be statistically significant at conventional levels.

Overall, the empirical results in Tables 3–8 offer feeble evidence that asset pricing models specified by Eqs. (3)–(10) generate step ahead forecasts of futures excess returns significantly better than the random walk with drift benchmark.²⁰ Said differently, commodity-based risk factors lack short-horizon predictive power

in the out of sample. Related to this observation, factor model forecasts using unconditional expectations of risk factors are generally more accurate than those using conditional factor expectations. These empirical findings are consistent with those of Simin (2008), who shows that asset pricing models produce poor out-of-sample expectations of equity returns relative to the historical average of the market return.

The failure of the factor models to provide accurate step ahead forecasts of commodity futures excess returns is possibly due to one or more of the following reasons: (i) futures basis as an instrumental variable is doing a poor job of capturing the relevant predictive information content; (ii) the linear specification of the factor predictive regressions is incorrect; (iii) the linear specification of the time-variation in beta is incorrect; and (iv) presence of structural breaks or sudden extreme events that introduce errors in the parameter estimation.

4. Economic evaluation

4.1. Asset allocation framework

The poor out-of-sample statistical performance of the asset pricing models using commodity-based risk factors does not necessarily indicate that they would not be valuable to utility maximizing investors. In this section, we examine the economic gains due to the factor model forecasts of excess returns. Standard in the recent literature (e.g., Della Corte et al., 2009; 2011; Thornton and Valente, 2012), we employ a mean-variance efficient framework to empirically assess the economic values of different port-

²⁰ In a separate empirical exercise, we also investigate predictability of excess returns on momentum and value portfolios for commodity futures contracts and find out-of-sample forecasting results qualitatively similar to those based on futures basis-sorted portfolios. The details on the construction of these portfolios and the one-month ahead forecasting results are provided in the Internet Appendix.

Table 7

Forecasts based on unconditional expectations of factors: portfolio-level.

		Model						RC
		(5)	(6)	(7)	(8)	(9)	(10)	
Portfolio 1	R^2_{OOS}	−0.13	−0.53	−0.84	−0.35	−1.12	−1.75	−0.07
	CW	−0.65	−0.81	−0.34	−0.07	−0.34	−0.39	
	GW	0.35	0.04 ^b	0.72	0.28	0.72	0.07 ^b	
		(0.16)	(0.09)	(0.01)	(0.05)	(0.00)	(0.07)	
Portfolio 2	R^2_{OOS}	−0.24	−0.36	−0.76	−0.79	−0.81	−0.88	−1.00
	CW	−1.00	−1.01	−1.79	−1.97	−1.72	−1.82	
	GW	0.14	0.06 ^b	0.09 ^b	0.06 ^b	0.08 ^b	0.06 ^b	
		(0.13)	(0.14)	(0.06)	(0.06)	(0.08)	(0.08)	
Portfolio 3	R^2_{OOS}	0.16	−0.07	−0.43	−0.35	−0.29	−0.19	1.73 [*]
	CW	1.73 ^{**}	−0.04	0.28	0.36	0.52	0.54	
	GW	0.12	0.51	0.17	0.16	0.24	0.18	
		(0.96)	(0.25)	(0.27)	(0.30)	(0.36)	(0.39)	
Portfolio 4	R^2_{OOS}	0.04	−0.05	−1.47	−1.44	−1.12	−1.11	0.39
	CW	0.39	0.18	−2.06	−2.07	−1.60	−1.64	
	GW	0.02 ^a	0.02 ^b	0.07 ^b	0.06 ^b	0.25	0.13	
		(0.69)	(0.40)	(0.00)	(0.01)	(0.00)	(0.08)	
Portfolio 5	R^2_{OOS}	0.09	−0.37	−2.14	−0.66	−2.12	−0.41	0.39
	CW	0.39	−1.02	−2.15	−1.66	−2.37	−1.39	
	GW	0.00 ^a	0.22	0.04 ^b	0.20	0.03 ^b	0.07 ^b	
		(0.60)	(0.13)	(0.00)	(0.00)	(0.01)	(0.14)	

The table reports out-of-sample tests of forecast accuracy of factor models relative to a random walk with drift (RWD) benchmark. The one-month ahead forecasts of excess returns for basis-sorted commodity futures portfolios are obtained from rolling regressions with a 10-year window using unconditional expectations of risk factors. For each alternative factor model in equation (i) ($i = 5, 6, \dots, 10$) and portfolio k , R^2_{OOS} is the Campbell and Thompson (2008) out-of-sample $R^2_{i,k}(\%)$ statistic. CW is the Clark and West (2006, 2007) $MSFE_{t,adj}^k$ - t statistic based on the Newey and West (1987) estimator, which tests whether the adjusted mean squared forecast error (MSFE) difference between the RWD and the alternative model is zero. RC is the $\max_{i=5,6,\dots,10} MSFE_{t,adj}^k$ - t statistic for a reality check following Clark and McCracken (2012), which tests the composite null hypothesis that the RWD is not inferior to any of the alternative models. One-sided critical values for the R^2_{OOS} , CW, and RC statistics are obtained using 10,000 fixed regressor bootstrap replications, as per Clark and McCracken (2012). *, **, and *** denote statistical significance of the R^2_{OOS} , CW, and RC statistics at the 10%, 5%, and 1% levels, respectively. GW is the Giacomini and White (2006) p^k -value of the equal conditional predictive ability test. Numbers in parentheses are the proportion of times the alternative model outperforms the RWD in the out of sample. Superscript *a* (*b*) indicates that the GW test rejects the null hypothesis of equal conditional predictive ability at the 10% significance level and the alternative model outperforms (is outperformed by) the RWD more than 50% of the time. The out-of-sample data run from January 1996 through October 2013.

Table 8

Forecasts based on conditional expectations of factors: portfolio-level.

		Model							RC
		(3)	(4)	(5)	(6)	(7)	(8)	(9)	
Portfolio 1	R^2_{OOS}	−1.97	−2.41	−6.29	−2.96	−3.09	−3.09	−7.98	−1.18
	CW	−1.75	−1.51	−1.59	−1.57	−1.18	−1.18	−1.33	
	GW	0.02 ^b	0.19	0.24	0.15	0.25	0.25	0.41	
		(0.06)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.00)	
Portfolio 2	R^2_{OOS}	0.53 ^{***}	0.51 ^{**}	0.55 ^{**}	0.35	−0.27	−0.27	−0.16	1.43
	CW	1.34	1.32	1.43 [*]	1.25	0.55	0.55	0.80	
	GW	0.45	0.50	0.58	0.74	0.71	0.71	0.90	
		(0.82)	(0.82)	(0.85)	(0.84)	(0.24)	(0.24)	(0.24)	
Portfolio 3	R^2_{OOS}	−0.17	−0.25	−0.16	−0.40	−1.05	−1.05	−1.10	0.35
	CW	0.13	0.00	0.16	0.01	0.25	0.25	0.30	
	GW	0.29	0.31	0.32	0.38	0.15	0.15	0.20	
		(0.19)	(0.17)	(0.20)	(0.20)	(0.23)	(0.23)	(0.22)	
Portfolio 4	R^2_{OOS}	−0.43	−0.48	−0.82	−1.13	−2.14	−2.14	−2.17	−0.57
	CW	−0.57	−0.64	−1.04	−1.43	−1.52	−1.52	−1.57	
	GW	0.24	0.33	0.09 ^b	0.17	0.07 ^b	0.07 ^b	0.05 ^b	
		(0.17)	(0.15)	(0.12)	(0.02)	(0.03)	(0.03)	(0.03)	
Portfolio 5	R^2_{OOS}	0.58 ^{***}	−0.78	0.88 ^{**}	−0.87	−1.53	−1.53	−1.35	1.21
	CW	1.09	−0.30	1.21	−0.34	−0.74	−0.74	−0.45	
	GW	0.75	0.68	0.60	0.64	0.38	0.38	0.45	
		(1.00)	(0.02)	(0.98)	(0.03)	(0.02)	(0.02)	(0.03)	

The setup is the same as in Table 7 except that conditional expectations of risk factors are used to generate one-month ahead forecasts of basis-sorted portfolio excess returns from factor models in Eqs. (3)–(10). See also notes to Table 7.

folios that condition on the forecasts generated from each of the asset pricing models including the benchmark random walk with drift model. More precisely, we consider an investment strategy of 100% in a risk-free asset and a self-financing dynamically re-balanced portfolio comprising J commodity futures contracts with similar maturity. In our analysis, we refer to this investment strat-

egy as the commodity stand-alone strategy. The investor's objective is to minimize the conditional portfolio variance subject to achieving a given target of expected excess return.

Let $\mu_{t+1|t} = E_t[rx_{t+1}]$ be the conditional expectation of the $J \times 1$ vector of commodity futures excess returns, $rx_{t+1} = (\Delta f_{t+1}^1, \Delta f_{t+1}^2, \dots, \Delta f_{t+1}^J)'$, and $\Sigma_{t+1|t} = E_t[(rx_{t+1} - \mu_{t+1|t})(rx_{t+1} - \mu_{t+1|t})']$.

$(rx_{t+1} - \mu_{t+1|t})'$ be the $J \times J$ conditional variance-covariance matrix of rx_{t+1} .²¹ At the end of each month t , the risk-averse investor solves the following problem:

$$\min_{w_t} \{w_t' \Sigma_{t+1|t} w_t\} \quad (19)$$

s.t. $w_t' \mu_{t+1|t} = \mu_p^*$ and $w_t' \iota = 0$,

where μ_p^* is the target of conditional expected excess return on the portfolio, w_t is the $J \times 1$ vector of time-varying portfolio weights on the commodity futures contracts, and ι is a conformable vector of ones. The solution to the constrained optimization problem yields the following weights on the commodity futures:

$$w_t = \Sigma_{t+1|t}^{-1} (\zeta_t' \Sigma_{t+1|t}^{-1} \zeta_t)^{-1} \zeta_t^*, \quad (20)$$

where $\zeta_t^* \equiv (\mu_p^*, 0)'$ and ζ_t is a $J \times 2$ matrix whose first column is the vector $\mu_{t+1|t}$, whereas the second column consists entirely of ones. Note that we assume that $\Sigma_{t+1|t} = \bar{\Sigma}$, where $\bar{\Sigma}$ is the unconditional variance-covariance matrix estimated using a 10-year rolling window.

4.2. Economic performance measures

We assess the economic value of predicting commodity futures excess returns by comparing the out-of-sample performance of optimal portfolios built using forecasts from factor models in Eqs. (3)–(10) to an optimal portfolio built using forecasts from a random walk with drift model (Eq. (2)), which we refer to as the benchmark portfolio.²² In line with the existing literature on portfolio performance evaluation, we adopt the most commonly used measure of economic gain – the Sharpe ratio, which is computed as the ratio of the realized average excess portfolio return to the standard deviation of excess portfolio returns. Several recent papers, however, point out that the Sharpe ratio can be misleading since it tends to underestimate the performance of dynamically rebalanced portfolios (see, among others, Han, 2006; Marquering and Verbeek, 2004, and references therein).

To provide further robustness checks of our empirical results on economic value, we consider two additional performance evaluation criteria. The first one quantifies the maximum performance fee that a risk-averse investor is willing to pay to switch from the benchmark portfolio to an alternative portfolio that exploits the corresponding factor model forecasts. Following Fleming et al. (2001), we compute the out-of-sample performance fee by equating the average realized utilities of selected pairs of portfolios. To illustrate, let $R_{p,t+1}$ be the gross return on the investor's optimal portfolio investment strategy such that

$$R_{p,t+1} = 1 + r_f + w_t' rx_{t+1}, \quad (21)$$

where r_f is the return on the risk-free asset. Under the assumption that the investor's preference is described by a quadratic utility function, the performance fee Φ is computed as:

$$\sum_{t=M+1}^{T-1} \left\{ (R_{p,t+1}^* - \Phi) - \frac{\gamma}{2(1+\gamma)} (R_{p,t+1}^* - \Phi)^2 \right\} = \sum_{t=M+1}^{T-1} \left\{ R_{p,t+1}^{ben.} - \frac{\gamma}{2(1+\gamma)} (R_{p,t+1}^{ben.})^2 \right\}, \quad (22)$$

where γ is the coefficient of relative risk aversion for the investor, $R_{p,t+1}^{ben.}$ is the gross return on the benchmark portfolio, and $R_{p,t+1}^*$ is the gross return on the portfolio that conditions on the forecasts

generated from the alternative factor model. If there is a systematic economic gain accruing to a mean-variance investor from the factor model using commodity-based risk factors, then $\Phi > 0$. On the other hand, if there is no predictive power embedded in the factors, then $\Phi \leq 0$. In this case, the investor is better off with the benchmark portfolio.

The next economic criterion that we consider is the manipulation-proof performance measure introduced by Goetzmann et al. (2007). This is defined as

$$\Theta = \frac{1}{1-\gamma} \log \left\{ \frac{1}{P} \sum_{t=M+1}^{T-1} \left(\frac{R_{p,t+1}^*}{R_{p,t+1}^{ben.}} \right)^{1-\gamma} \right\}. \quad (23)$$

In comparison to the performance fee Φ , the risk-adjusted measure Θ does not depend on the specific form of the underlying utility function to rank portfolios. Moreover, the manipulation-proof performance measure Θ is robust to the use of derivatives as alternative assets in the investor's portfolio as well as to the distribution of portfolio returns.

4.3. Economic performance results

To implement the dynamic asset allocation framework and to compute the potential economic gains following the commodity stand-alone strategy, we require specifying the investor's degree of relative risk aversion, γ , and the target of the conditional expected excess return for the optimal portfolio, μ_p^* . Following Thornton and Valente (2012), we set $\gamma = 5$. As for a reasonable target of μ_p^* , we match the (annualized) excess return realized for the S&P GSCI over the full-sample period, which is 3%. The risk-free interest rate data are obtained from the Internet Data Library maintained by Kenneth French as a proxy for r_f .²³

Panel A in Table 9 summarizes the out-of-sample economic performance of eight alternative portfolios, based on the factor model forecasts, relative to a portfolio based on the benchmark random walk with drift model forecasts. In the left and right panels of the table, we differentiate, respectively, between the unconditional and conditional factor expectations based results. Our empirical analysis focuses on the Sharpe ratio, the Fleming et al. (2001) performance fee Φ , and the Goetzmann et al. (2007) manipulation-proof measure of performance Θ . In general, we observe empathetically feeble performance of the alternative portfolios relative to the benchmark. The annualized Sharpe ratios for the alternative portfolios are either economically small or negative (ranging between -0.262 and -0.019). The performance fees Φ and Θ expressed in decimals (i.e., $0.01=1$ annual percentage point) turn out to be negative. These empirical observations hold regardless of factor expectation methods.

Taken together, there is no evidence of tangible economic gains to an investor exploiting the information content embedded in the commodity-based risk factors. The economic value results for the commodity stand-alone strategy confirm and reinforce our preceding finding from the statistical tests that asset pricing models for commodity futures excess returns perform poorly in terms of out-of-sample expectations.

5. Diversification benefits

In this section, we extend our empirical investigation on the potential economic gains that factor models may offer to a risk-averse investor. We use the same mean-variance framework for efficient asset allocation as in Section 4.1, but rather than concentrating purely on commodity futures contracts we also consider

²¹ To simplify notations, we omit indexing of the models by i .

²² In the rest of this paper, each of the portfolios are a larger portfolio that always includes 100% investment in a risk-free asset together with the relevant dynamically rebalanced portfolio of risky assets.

²³ For details, see http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

Table 9
Economic value of commodity futures excess return predictability.

Model	Unconditional expectations of factors			Conditional expectations of factors		
	Sharpe ratio	Φ	Θ	Sharpe ratio	Φ	Θ
Panel A: Commodity stand-alone strategy						
RWD	0.261			0.261		
(3)				0.121	−0.004	−0.004
(4)				0.101	−0.005	−0.005
(5)	0.263	0.000	0.000	0.192	−0.002	−0.002
(6)	0.235	−0.001	−0.001	0.118	−0.004	−0.004
(7)	−0.262	−0.012	−0.012	−0.221	−0.011	−0.011
(8)	−0.233	−0.012	−0.012	−0.221	−0.011	−0.011
(9)	−0.185	−0.010	−0.010	−0.137	−0.009	−0.009
(10)	−0.019	−0.007	−0.007	0.119	−0.004	−0.004
Panel B: Diversification strategy						
S+B	0.579			0.579		
(3)				0.586	−0.005	−0.005
(4)				0.569	−0.005	−0.005
(5)	0.638	−0.003	−0.003	0.559	−0.005	−0.005
(6)	0.755	−0.002	−0.002	0.523	−0.007	−0.007
(7)	0.255	−0.012	−0.012	0.325	−0.011	−0.011
(8)	0.265	−0.012	−0.012	0.325	−0.011	−0.011
(9)	0.291	−0.012	−0.012	0.360	−0.011	−0.011
(10)	0.431	−0.009	−0.009	0.599	−0.007	−0.007

The table reports out-of-sample economic value of factor models. The one-month ahead forecasts of commodity futures excess returns are obtained from rolling regressions with a 10-year window. The left (right) panel contains results using forecasts based on unconditional (conditional) expectations of risk factors. Panel A reports results for a commodity stand-alone strategy, in which forecasts from each model including a random walk with drift (RWD) benchmark are utilized to build a corresponding minimum conditional portfolio variance strategy subject to a target annualized portfolio excess return of $\mu_p^* = 3\%$ for an investor who dynamically rebalances her portfolio comprising commodity futures on a monthly basis. Panel B reports results for a diversification strategy, which extends the minimum conditional portfolio variance strategy for an investor who diversifies her exposure to stocks and bonds by allocating commodity futures to her portfolio based on forecasts of excess returns from each factor model. The performance of a portfolio comprising commodity futures and conventional assets (i.e., stocks and bonds) in Panel B is compared against a benchmark portfolio comprising stocks and bonds (denoted by S+B) only. The Sharpe ratio is the ratio of the (annualized) mean excess portfolio return to the (annualized) standard deviation of excess portfolio returns. Φ is the Fleming et al. (2001) performance fee a risk-averse investor with quadratic utility function is willing to pay for switching from a benchmark strategy to an alternative factor model strategy. Θ is the Goetzmann et al. (2007) manipulation-proof measure of performance after adjusting for risk. Φ and Θ are computed with a relative risk aversion coefficient $\gamma = 5$ and are reported in decimals per annum (i.e., 0.01 = 1 annual percentage point). The out-of-sample data run from January 1996 through October 2013.

conventional assets such as stocks and bonds. We continue to assume that a risk-averse investor aims at achieving a given target of expected excess return by minimizing the conditional portfolio variance. To assess the potential diversification benefits of investing in commodity futures by conditioning on expectations from the risk factor models, we design two types of portfolio strategies. The first type involves building a dynamically rebalanced portfolio consisting of conventional assets only. For a set of N investable assets, the optimal weights, denoted by w_t^c , of the portfolio are given by

$$w_t^c = \frac{\mu_p^*}{V_t} \Sigma_{c,t+1|t}^{-1} \mu_{t+1|t}^c, \quad (24)$$

where $\mu_{t+1|t}^c$ and $\Sigma_{c,t+1|t}$ are, respectively, the conditional expectation and the conditional variance-covariance matrix of the $N \times 1$ vector of excess returns on the conventional assets and $V_t \equiv (\mu_{t+1|t}^c)' \Sigma_{c,t+1|t}^{-1} \mu_{t+1|t}^c$.²⁴

The second type of strategy involves building dynamically rebalanced portfolios consisting of both conventional assets and commodity futures contracts. To illustrate, let $\mu_{t+1|t}^d$ be the $(N+J)$ vector, whose first N elements are the conditional expectations of excess returns on the conventional assets and the remaining J elements are the conditional expected commodity futures excess returns from the factor model in consideration. In this case, the investor's problem is again to minimize the conditional portfolio variance but subject to $w_t^{d'} \mu_{t+1|t}^d = \mu_p^*$ and $w_t^{d'} D = 0$, where D is a $(N+J)$ vector, whose first N elements are zeros and the remaining J elements are all ones. The solution to the constrained opti-

mization problem with this type of portfolio investment strategy delivers the following weights:

$$w_t^d = \Gamma_{t+1|t}^{-1} \vartheta_t (\vartheta_t' \Gamma_{t+1|t}^{-1} \vartheta_t)^{-1} \zeta^*, \quad (25)$$

where $\vartheta_t \equiv (\mu_{t+1|t}^d, D)'$ and $\Gamma_{t+1|t}$ is the conditional variance-covariance matrix, whose first N elements on the main diagonal are the variances of excess returns on the conventional assets and the remaining J elements are the variances of commodity futures excess returns.

To prevent extreme investments, we follow Ferreira and Santa-Clara (2011), Welch and Goyal (2008), and Thornton and Valente (2012) in our empirical analysis and winsorize the portfolio weights on the conventional assets to $-1 \leq w_{n,t} \leq 2$, $n = 1, \dots, N$, for both types of portfolio strategies described above. Once we obtain the optimal weights as per Eqs. (24) and (25), we compute the returns and compare the economic performance of the portfolios built using the two types of strategies. For the clarity of discussion, we call the latter strategy comprising both conventional assets and commodity futures as the diversification strategy. As a dynamically rebalanced benchmark portfolio, we use a portfolio due to the first type strategy, which consists of stocks and bonds only. Moreover, we assume that the risk-averse investor gains exposure to conventional assets by investing in the S&P 500 and Barclays US Aggregate Bond indexes. We retrieve data on both of these indexes from Datastream.

The results for the diversification benefits (i.e., diversification strategy) of the commodity futures trading are presented in Panel B of Table 9. Once again, we use a 3% annual target of the conditional expected excess return for the optimal portfolio and employ performance measures similar to those in Section 4.2. We observe that only a few alternative portfolios (i.e., portfolios due to factor

²⁴ The conditional expectation and the conditional variance-covariance matrix are estimated, respectively, as the sample mean and the sample variance-covariance matrix using a 10-year rolling window.

Table 10
Economic value of commodity futures excess return predictability: a cross-sectional approach with hedging pressure.

Model	Unconditional expectations of factors			Conditional expectations of factors		
	Sharpe ratio	Φ	Θ	Sharpe ratio	Φ	Θ
Panel A: Commodity stand-alone strategy						
RWD	0.246			0.246		
(3)				−0.138	−0.016	−0.016
(4)				−0.081	−0.011	−0.011
(5)	−0.026	−0.005	−0.005	−0.074	−0.010	−0.010
(6)	−0.238	−0.007	−0.007	−0.164	−0.010	−0.010
(7)	−0.296	−0.010	−0.010	−0.192	−0.009	−0.009
(8)	−0.232	−0.010	−0.009	−0.112	−0.008	−0.008
(9)	−0.280	−0.007	−0.007	−0.233	−0.009	−0.009
(10)	−0.333	−0.007	−0.007	−0.080	−0.007	−0.007
Panel B: Diversification strategy						
S+B	0.554			0.554		
(3)				0.110	−0.013	−0.013
(4)				0.164	−0.012	−0.012
(5)	−0.157	−0.016	−0.016	0.003	−0.016	−0.016
(6)	−0.170	−0.016	−0.016	0.093	−0.014	−0.014
(7)	−0.117	−0.017	−0.017	−0.013	−0.015	−0.015
(8)	−0.117	−0.017	−0.017	0.006	−0.015	−0.015
(9)	−0.273	−0.016	−0.016	−0.070	−0.016	−0.016
(10)	−0.323	−0.016	−0.016	0.097	−0.014	−0.014

The table reports out-of-sample economic value of factor models. The one-month ahead forecasts of commodity futures excess returns are obtained by the [Fama and MacBeth \(1973\)](#) cross-sectional procedure following [Brennan et al. \(1998\)](#). The forecasting scheme based on rolling regressions with a 10-year window utilizes commodity-specific hedging pressure. The out-of-sample data run from February 1996 through October 2013. See also notes to [Table 9](#).

models in [Eqs. \(5\)–\(6\)](#) using unconditional factor expectations and due to factor models in [Eqs. \(3\) and \(10\)](#) using conditional factor expectations) generate annualized Sharpe ratios higher than the benchmark portfolio, which is 0.579. However, the negative values of the performance fee Φ and the manipulation-proof performance measure Θ corresponding to these portfolios suggest underperformance relative to the benchmark. More importantly, both Φ and Θ turn out to be negative for all alternative portfolios. They range between -1.2% and -0.2% per annum. These results clearly indicate that a risk-averse investor gains no diversification benefits from allocating commodity futures contracts to her portfolio by exploiting expectations of excess returns due to commodity-based risk factors that we consider in this paper.

Interestingly, our results are consistent with the recent findings of [Daskalaki and Skiadopoulos \(2011\)](#). The authors investigate the utility gains accruing to a risk-averse investor who trade an index of commodities as part of her broader diversification strategy. They conclude that apart from the early 2000s, when many commodities witnessed a synchronized boom in prices, investors are better off by holding an optimal portfolio that consists of conventional assets only. Therefore, apart from the weak predictive power of the risk factors, the empirical results emerged in this section could possibly be due to the fact that commodities, in general, offer little diversification benefits.

As for robustness checks of our preceding results on out-of-sample economic evaluation of factor models, we carry out similar dynamic asset allocation exercises for commodity stand-alone and diversification strategies using excess return forecasts obtained from the cross-sectional approach in [Section 3.3.2](#). The results of these exercises summarized in [Table 10](#) are qualitatively similar to those in [Table 9](#). To sum up, there is no systematic economic value associated with the empirical factor models for commodity futures excess returns.

6. Conclusion

Recent literature has identified some systematic risk factors that successfully price the cross-section of commodity futures excess returns. This paper extends the literature on commodity fu-

tures by examining the time-series predictability of commodity futures excess returns from asset pricing models. To investigate the predictive power of commodity-based risk factors, various factor model specifications are considered and their performance is compared against a random walk with drift benchmark. A battery of statistical tests performed under the null hypothesis of equal predictive ability suggest that the factor models hardly outperform the benchmark in the out-of-sample forecasting horse races. This finding is robust to the choice of underlying test assets (i.e., excess returns on individual commodity futures as well as basis-sorted futures portfolios), the computation of factor expectations, the forecasting methods including a cross-sectional approach, and the statistical evaluation metrics. Furthermore, there is no evidence that information contained in commodity-based risk factors generate systematic economic gains to risk-averse investors pursuing a commodity stand-alone strategy or a diversification strategy. These results provide valuable implications for practitioners including fund managers in the commodity futures markets who require accurate forecasts of returns to design active trading strategies. However, given the recent evidence offered by [Daskalaki and Skiadopoulos \(2011\)](#) on the poor diversification benefits of commodities, we remain agnostic as to whether the lack of systematic economic value to investors following a diversification strategy is due to the poor predictive performance of the asset pricing models. Finally, it is important to emphasize that the empirical exercises in this paper are meant for practical relevance of the asset pricing models with commodity-based risk factors to financial market participants facing real time uncertainty when forecasting returns rather than a formal test of the models.

Acknowledgements

We thank Abhay Abhyankar, Amrit Judge, Giorgio Valente, Jerry Coakley, and Sanjay Banerji for helpful comments.

Appendix A

This appendix summarizes the fixed regressor bootstrap procedure to generate (one-sided to the right) critical values for

the pairwise and reality check test statistics following Clark and McCracken (2012).²⁵ The algorithm consists of the following steps:

1. Use OLS to estimate the parameters associated with the alternative factor model in equation (i) ($i = 3, 4, \dots, 10$). Store the residuals $\hat{\epsilon}_t$, $t = 1, 2, \dots, T - 1$. Only for the reality check test, estimate the parameters associated with the alternative model that includes all regressors considered across all models (i.e., the factor model given by Eq. (10)). Store the corresponding residuals and follow the steps below.
2. Use OLS to estimate the parameter of the benchmark random walk with drift model in Eq. (2). Store the fitted values Δf_t^{fit} , $t = 1, 2, \dots, T - 1$.
3. Let η_t , $t = 1, 2, \dots, T - 1$, denote an independently and identically distributed $N(0,1)$ sequence of simulated random variables. Generate time-series innovations $\hat{\epsilon}_t^* = \eta_t \hat{\epsilon}_t$, $t = 1, 2, \dots, T - 1$.
4. Construct artificial samples of Δf_t^* using the fixed regressor structure, $\Delta f_t^* = \Delta f_t^{fit} + \hat{\epsilon}_t^*$.
5. Use the artificial data to construct the out-of-sample forecasts and an estimate of the test statistics (e.g., $R_{i,005}^2$, $MSFE_{i,adj.-t}$, and $\max_{i=3,4,\dots,10} MSFE_{i,adj.-t}$) as if these were the original data.
6. Independently repeat steps 3–5 a large number of times, in our case, 10,000 replications.
7. Reject the null hypothesis at the $\alpha\%$ level if the test statistic is greater than the $(100 - \alpha)$ percentile of the empirical distribution of the simulated test statistics.

Appendix B. Supplementary data

Supplementary material associated with this article can be found, in the online version, at [10.1016/j.jbankfin.2016.06.011](https://doi.org/10.1016/j.jbankfin.2016.06.011)

References

- Acharya, V.V., Lochstoer, L.A., Ramadorai, T., 2013. Limits to arbitrage and hedging: evidence from commodity markets. *Journal of Financial Economics* 109, 441–465.
- Ahmed, S., Liu, X., Valente, G., 2016. Can currency-based risk factors help forecast exchange rates? *International Journal of Forecasting* 32, 75–97.
- Alquist, R., Kilian, L., 2010. What do we learn from the price of crude oil futures? *Journal of Applied Econometrics* 25, 539–573.
- Amihud, Y., Mendelson, H., Lauterbach, B., 1997. Market microstructure and securities values: evidence from the Tel Aviv stock exchange. *Journal of Financial Economics* 45, 365–390.
- Asness, C.S., Moskowitz, T.J., Pedersen, L.H., 2013. Value and momentum everywhere. *Journal of Finance* 68, 929–985.
- Bakshi, G., Gao, X., Rossi, A., 2014. Understanding the Sources of Risk Underlying the Cross-section of Commodity Returns. *Smith School of Business, University of Maryland. Working Paper*.
- Basu, D., Miffre, J., 2013. Capturing the risk premium of commodity futures: the role of hedging pressure. *Journal of Banking & Finance* 37, 2652–2664.
- Bessembinder, H., 1992. Systematic risk, hedging pressure, and risk premiums in futures markets. *Review of Financial Studies* 5, 637–667.
- Bodie, Z., Rosansky, V.I., 1980. Risk and return in commodity futures. *Financial Analysts Journal* 36, 27–39.
- Brennan, M.J., Chordia, T., Subrahmanyam, A., 1998. Alternative factor specifications, security characteristics, and the cross-section of expected stock returns. *Journal of Financial Economics* 49, 345–373.
- Campbell, J.Y., Thompson, S.B., 2008. Predicting excess stock returns out of sample: can anything beat the historical average? *Review of Financial Studies* 21, 1509–1528.
- Carter, C.A., Rausser, G.C., Schmitz, A., 1983. Efficient asset portfolios and the theory of normal backwardation. *Journal of Political Economy* 91, 319–331.
- Chatfield, C., 1995. Model uncertainty, data mining and statistical inference. *Journal of the Royal Statistical Society* 158, 419–466.
- Chinn, M.D., Coibion, O., 2014. The predictive content of commodity futures. *Journal of Futures Markets* 34, 607–636.
- Clark, T.E., McCracken, M.W., 2001. Tests of equal forecast accuracy and encompassing for nested models. *Journal of Econometrics* 105, 85–110.
- Clark, T.E., McCracken, M.W., 2012. Reality checks and comparisons of nested predictive models. *Journal of Business & Economic Statistics* 30, 53–66.
- Clark, T.E., McCracken, M.W., 2013. Advances in forecast evaluation. In: Elliott, G., Timmermann, A. (Eds.), *Handbook of Economic Forecasting*. North Holland, New York, pp. 1107–1201.
- Clark, T.E., West, K.D., 2006. Using out-of-sample mean squared prediction errors to test the martingale difference hypothesis. *Journal of Econometrics* 135, 155–186.
- Clark, T.E., West, K.D., 2007. Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics* 138, 291–311.
- Daskalaki, C., Kostakis, A., Skiadopoulos, G., 2014. Are there common factors in individual commodity futures returns? *Journal of Banking & Finance* 40, 346–363.
- Daskalaki, C., Skiadopoulos, G., 2011. Should investors include commodities in their portfolios after all? *Journal of Banking & Finance* 35, 2606–2626.
- Della Corte, P., Sarno, L., Tsiakas, I., 2009. An economic evaluation of empirical exchange rate models. *Review of Financial Studies* 22, 3491–3530.
- Della Corte, P., Sarno, L., Tsiakas, I., 2011. Spot and forward volatility in foreign exchange. *Journal of Financial Economics* 100, 496–513.
- Dhume, D., 2011. Using Durable Consumption Risk to Explain Commodities Returns. *Department of Economics, Harvard University. Working Paper*.
- Dusak, K., 1973. Futures trading and investor returns: an investigation of commodity market risk premiums. *Journal of Political Economy* 81, 1387–1406.
- Erb, C.B., Harvey, C.R., 2006. The strategic and tactical value of commodity futures. *Financial Analysts Journal* 62, 69–97.
- Fama, E.F., French, K.R., 1987. Commodity futures prices: some evidence on forecast power, premiums, and the theory of storage. *Journal of Business* 60, 55–73.
- Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33, 3–56.
- Fama, E.F., French, K.R., 1997. Industry costs of equity. *Journal of Financial Economics* 43, 153–193.
- Fama, E.F., MacBeth, J.D., 1973. Risk, return, and equilibrium: empirical tests. *Journal of Political Economy* 81, 607–636.
- Ferraro, D., Rogoff, K., Rossi, B., 2015. Can oil prices forecast exchange rates? An empirical analysis of the relationship between commodity prices and exchange rates. *Journal of International Money and Finance* 54, 116–141.
- Ferreira, M.A., 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.
- Ferson, W.E., Harvey, C.R., 1997. Fundamental determinants of national equity market returns: a perspective on conditional asset pricing. *Journal of Banking & Finance* 21, 1625–1665.
- Ferson, W.E., Harvey, C.R., 1999. Conditioning variables and the cross-section of stock returns. *Journal of Finance* 54, 1325–1360.
- Ferson, W.E., Nallareddy, S., Xie, B., 2013. The “out-of-sample” performance of long run risk models. *Journal of Financial Economics* 107, 537–556.
- Fleming, J., Kirby, C., Ostdieck, B., 2001. The economic value of volatility timing. *Journal of Finance* 56, 329–352.
- Fortenberry, R.T., Hauser, R.J., 1990. Investment potential of agricultural futures contracts. *American Journal of Agricultural Economics* 72, 721–726.
- Giacomini, R., White, H., 2006. Tests of conditional predictive ability. *Econometrica* 74, 1545–1578.
- Gitman, L.J., Mercurio, V.A., 1982. Cost of capital techniques used by major U.S. firms: survey and analysis of Fortune’s 1000. *Financial Management* 14, 21–29.
- Goetzmann, W., Ingersoll, J., Spiegel, M., Welch, I., 2007. Portfolio performance manipulation and manipulation-proof performance measures. *Review of Financial Studies* 20, 1503–1546.
- Gorton, G.B., Hayashi, F., Rouwenhorst, K.G., 2013. The fundamentals of commodity futures returns. *Review of Finance* 17, 35–105.
- Gorton, G.B., Rouwenhorst, K.G., 2006. Facts and fantasies about commodity futures. *Financial Analysts Journal* 62, 47–68.
- Graham, J.R., Harvey, C.R., 2001. The theory and practice of corporate finance: evidence from the field. *Journal of Financial Economics* 60, 187–243.
- Han, Y., 2006. Asset allocation with a high dimensional latent factor stochastic volatility model. *Review of Financial Studies* 19, 237–271.
- Hicks, J.R., 1939. *Value and Capital*. Oxford University Press.
- Hirshleifer, D., 1988. Residual risk, trading costs, and commodity futures risk premia. *Review of Financial Studies* 1, 173–193.
- Hirshleifer, D., 1989. Determinants of hedging and risk premia in commodity futures markets. *Journal of Financial and Quantitative Analysis* 24, 313–331.
- Hong, H., Yogo, M., 2012. What does futures market interest tell us about the macroeconomy and asset prices? *Journal of Financial Economics* 105, 473–490.
- Hoover, K.D., Perez, S.J., 1999. Data mining reconsidered: encompassing and the general-to-specific approach to specification search. *Econometrica* 67, 167–191.
- Jagannathan, R., 1985. An investigation of commodity futures prices using the consumption-based intertemporal capital asset pricing model. *Journal of Finance* 40, 175–191.
- Jagannathan, R., Meier, I., 2002. Do we need CAPM for capital budgeting? *Financial Management* Winter, 5–27.
- Jensen, G.R., Johnson, R.R., Mercer, J.M., 2000. Efficient use of commodity futures in diversified portfolios. *Journal of Futures Markets* 20, 489–506.
- Keynes, J.M., 1923. Some aspects of commodity markets. *Manchester Guardian Commercial* 13, 784–786.
- Keynes, J.M., 1930. *A Treatise on Money*, Vol. 2. Macmillan, London.
- Leitch, G.J., Tanner, E., 1991. Economic forecast evaluation: profits versus the conventional error measures. *American Economic Review* 81, 580–590.

²⁵ For simplicity, we omit indexing of the commodity futures and portfolios by j and k , respectively. Moreover, we use α to denote the level of significance in this appendix.

- Lintner, J., 1965. The valuation of risky assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics* 47, 13–37.
- Lo, A.W., MacKinlay, A.C., 1990. Data-snooping biases in tests of financial asset pricing models. *Review of Financial Studies* 3, 431–467.
- Malone, S. W., Gramacy, R. B., ter Horst, E., 2014. Timing Foreign Exchange Markets. SSRN Working Paper.
- Marcus, A.J., 1984. Efficient asset portfolios and the theory of normal backwardation: a comment. *Journal of Political Economy* 92, 162–164.
- Marquering, W., Verbeek, M., 2004. The economic value of predicting stock index returns and volatility. *Journal of Financial and Quantitative Analysis* 39, 407–429.
- Marshall, B.R., Nguyen, N.H., Visaltanachoti, N., 2012. Commodity liquidity measurement and transaction costs. *Review of Financial Studies* 25, 599–638.
- McCracken, M.W., Valente, G., 2014. Asymptotic Inference for Performance Fees and the Predictability of Asset Returns. Federal Reserve Bank of St. Louis. Working Paper 2012-049B.
- Miffre, J., Rallis, G., 2007. Momentum strategies in commodity futures markets. *Journal of Banking & Finance* 31, 1863–1886.
- Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55, 703–708.
- de Roon, F.A., Nijman, T.E., Veld, C.H., 1998. Pricing term structure risk in futures markets. *Journal of Financial and Quantitative Analysis* 33, 139–157.
- de Roon, F.A., Nijman, T.E., Veld, C.H., 2000. Hedging pressure effects in futures markets. *Journal of Finance* 55, 1437–1456.
- Sharpe, W.F., 1964. Capital asset prices: a theory of market equilibrium under conditions of risk. *Journal of Finance* 19, 425–442.
- Simin, T., 2008. The poor predictive performance of asset pricing models. *Journal of Financial and Quantitative Analysis* 43, 355–380.
- Szymanowska, M., de Roon, F.A., Nijman, T.E., Goorbergh, R., 2014. An anatomy of commodity futures risk premia. *Journal of Finance* 69, 453–482.
- Thornton, D.L., Valente, G., 2012. Out-of-sample predictions of bond excess returns and forward rates: an asset allocation perspective. *Review of Financial Studies* 25, 3141–3168.
- Verdelhan, A., 2016. The share of systematic variation in bilateral exchange rates. *Journal of Finance*, (forthcoming).
- Welch, I., Goyal, A., 2008. A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies* 21, 1455–1508.
- West, K.D., 1996. Asymptotic inference about predictive ability. *Econometrica* 64, 1067–1084.
- Yang, F., 2013. Investment shocks and the commodity basis spread. *Journal of Financial Economics* 110, 164–184.