Diversification benefits of commodities: A stochastic dominance efficiency approach*

Charoula Daskalaki[1], George Skiadopoulos[2], Nikolas Topaloglou[3]

Abstract

We revisit the question whether commodities should be included in investors' portfolios. We employ for the first time a stochastic dominance efficiency (SDE) approach to construct optimal portfolios with and without commodities and we evaluate their comparative performance. SDE circumvents the necessity to posit a specific utility function to describe investor's preferences and it does not impose distributional assumptions on asset returns. We find that commodities provide diversification benefits both in- and out-of-sample. This evidence is stronger when commodity indices which mimic dynamic commodity trading strategies are used. We explain our results by documenting that commodity markets are segmented from the equity and bond markets.

JEL classification: C1, C4, C6, G10, G11

Keywords: Alternative investments, Commodity indices, Market integration, Portfolio choice, Stochastic dominance

---

* We would like to thank George Constantinides, Jens Jackwerth, Olga Kolokolova, Alexandros Kostakis, Kalle Rinne, Raman Uppal and participants at the 2015 International Conference on Computational and Financial Econometrics (London), 2016 Spring Conference of the Multinational Finance Society (Lemesos), Alternative Investments Conference (Monaco), 2016 Energy and Commodity Finance Conference and 2016 CRETE (Tinos) conference for useful discussions and comments. Financial support by the J.P. Morgan Center for Commodities at the University of Colorado Denver via the Commodities Research Fellowship Grant award is gratefully acknowledged. Any remaining errors are our responsibility alone.

[1] Department of Banking and Financial Management, University of Piraeus, GR, chdask@webmail.unipi.gr

[2] School of Economics and Finance, Queen Mary University of London, UK and Department of Banking and Financial Management, University of Piraeus, GR. Also Research Fellow with Cass Business School and Warwick Business School. g.skiadopoulos@qmul.ac.uk, gskiado@unipi.gr

[3] Department of International and European Economic Studies, Athens University of Economics and Business, GR, nikolas@aueb.gr

1. Introduction

Investments in commodities have grown significantly over the last years along with the revived academic interest in the properties of this alternative asset class (for reviews see e.g., Geman, 2005, Erb and Harvey, 2006, Gorton and Rouwenhorst, 2006, Skiadopoulos, 2013).[1] One of the explanations commonly invoked to explain investors' interest in commodities is their alleged diversification benefits. Commodity returns are expected to have low or negative correlation with the traditional asset classes like bonds and equities returns. Therefore, their inclusion in portfolios consisting of traditional asset classes is expected to increase the expected return per unit of risk.

Surprisingly, in the academic literature there is no consensus yet about whether investments in commodities do offer diversification benefits. In this paper, we revisit this open question by exploring whether investments in commodities make the investor better off once included in their portfolios. Our main contribution is that we deviate from the previous literature and we construct optimal portfolios and assess their performance in a non-parametric way. We do not postulate a *specific* utility function to describe investor's preferences and we do not make any assumption on the asset returns distributions. We manage to do so by employing a stochastic dominance efficiency (SDE) approach which extends the standard stochastic dominance (SD) concept.

SD is a natural setting to use for decision under uncertainty when partial information regarding the decision maker's risk preferences is available. SD offers criteria to rank two mutually exclusive investments when compared pairwise (Hadar and Russell, 1969, Hanoch and Levy, 1969, Levy and Hanoch, 1970). Commonly, first order and second order stochastic dominance criteria (FSD, SSD, respectively) are being used. Hadar and Russel (1969) and Bawa (1975) show that FSD (SSD) amounts

[1] The total value of commodity investments rose from $170 billion in July 2007 to $410 billion in February 2013 (Croft and Norrish, 2013) accompanied by the 2003-2008 remarkable increase in commodity prices (commodity boom, only the third since 1950).

to choosing the investment that maximizes investor's expected utility assuming that investors preferences are characterized by non-satiation (non-satiation and risk aversion). The theoretical attractiveness of SD lies in its nonparametric nature. SD criteria do not require any assumption on the distribution of returns of the two investments under consideration and they are consistent with a general class of preferences. Hence, SD criteria are a natural candidate to rank two investments because they do not impose strict assumptions on preferences and distribution of returns as the commonly used mean-variance portfolio construction setting.[2] The drawback of the FSD and SSD criteria though is that they can only compare pairwise any two *given* portfolios. Hence, they cannot be used to test whether a portfolio stochastically dominates *every* single portfolio because there is an *infinite* number of alternative portfolios. SD efficiency (SDE) introduced by Post (2003) and Kuosmanen (2004) circumvents this constraint. It is a direct extension of SD to the case where one can compare the return distribution of any portfolio constructed from a set of assets with another fixed portfolio by exhausting all possible combinations of portfolio weights.

We address our research question both in- and out-of-sample. First, we examine whether the introduction of commodities in the investor's asset universe yields diversification benefits in-sample. To this end, we employ a two-step procedure that uses Scaillet and Topaloglou (2010) SDE test to assess whether an asset universe which includes commodities yields a portfolio that stochastically dominates a portfolio orginated from the same asset universe without commodities. Scaillet and Topaloglou (2010) build on the general distribution definition of SD by relying on Kolmogorov-Smirnov type of tests and they develop consistent statistical tests for SDE at any order for time-dependent data. Next, we conduct

[2] The main disadvantage of the mean-variance approach is that it allows for violations of first-order stochastic dominance because it is not robust to outliers (extreme deviations are greatly overweighted and small deviations are relatively neglected). Mean variance optimal portfolio maximizes expected utility only in the case where investor preferences and return distributions obey highly restrictive conditions (i.e. quadratic utility function and/or normally distributed returns, see, e.g., Hanoch and Levy 1969, Levy 1992).

our analysis out-of-sample. At any point in time, we construct optimal portfolios based separately on an asset universe comprising traditional asset classes and on an asset universe augmented with commodities. We construct optimal portfolios by employing once again the Scaillet and Topaloglou (2010) SDE methodology in a rolling window fashion this time. We compare the optimal portfolios constructed from the two respective asset universes by evaluating their out-of-sample performance over January 2001-September 2013. First, we conduct the assessment non-parametrically by using the Scaillet and Topaloglou (2010) SDE test to check whether optimal portoflios based on the asset universe augmented with commodities dominate the ones based on the traditional asset universe. Next, in line with DeMiguel et al. (2009), we use a number of performance measures that take into account deviations from normality as well as transaction costs.

To ensure the robustness of our analysis, we employ the SDE approach under the first order and second order SDE criteria (FSDE, SSDE, respectively). We also use alternative sub-samples for the in-sample analysis. In addition, we employ ten first, second and third generation commodity indices as alternative vehicles to invest to the commodity asset class. First generation indices mimic passive commodity futures portfolio strategies where only long positions in the constituent commodity futures are allowed. Second and third generation commodity indices are gaining popularity because they mimic dynamic (long and long/short, respectively) commodity futures portfolio strategies which exploit popular commodity trading signals such as momentum and switches from backwardated to contangoed markets and vice versa (Miffre, 2012); Gorton et al. (2012) find that these dynamic strategies yield significant risk premia and Rallis et al. (2013) document that they outperform passive strategies similar in spirit to the first generation indices. Therefore, the use of second and third generation commodity indices allows us to address our research question by taking the characteristics of popular dynamic long/short commodity strategies into account. To the best of our knowledge, the previous literature has employed only first generation commodity indices to consider the diversification benefits of commodities. In addition, given that we consider dynamic commodity indices, we also consider investing in the Fama-French (1993) size and value factors on top of considering only a passive investment strategy in the equities asset class such as investing in the S&P 500. Fama-French factors mimic a dynamic long/short trading strategy. Hence, their use ensures that indices which represent dynamic trading strategies are used in both asset universes and hence it yields a fair comparison of the traditional asset classes universe with the augmented with commodities one.

We find that the inclusion of the commodity asset class in portfolios comprising traditional asset classes makes the investor better off both in-sample and out-of-sample. This holds regardless of the sub-sample period, portfolio performance evaluation measure and the SDE criterion under consideration. The documented diversification benefits of commodities are more pronounced in the case where the investor access commodities via the second and third order generation commodity indices. We explain our results on the outperformance of the augmented by commodities optimal portfolios by implementing Campbell and Hamao's (1992) approach which tests for market integration. We document that commodity markets are segmented from equity and bond markets. This is because we find that commodity portfolio returns cannot be forecasted by the instrumental variables that predict stock and bond market returns, thus extending the evidence in Gargano and Timmermann (2014) on the difficulty in predicting commodity *portfolio* returns.

Related literature. Our paper is related to three strands of literature: portfolio choice with commodities, the SDE literature and the evidence of whether commodity markets are integrated with other markets. Bodie and Rosansky (1980), Fortenbery and Hauser (1990), Ankrim and Hensel (1993), Abanomey and Mathur (1999), Anson (1999), Jensen et al. (2000), Conover et al. (2010), Belousova and Dorfleitner (2012) find that the investor is better off by including commodities in her portfolio whereas Cao et al.

(2010) find that commodities should not be included in investors' portfolios. These studies are conducted within an in-sample mean-variance Markowitz (1952) portfolio setting. In contrast, Daskalaki and Skiadopoulos (2011) account for the fact that real-time investors are interested in the out-of-sample performance and commodity futures returns deviate from normality (Gorton and Rouwenhorst, 2006, Kat and Oomen, 2007). They find that commodity investing could be beneficial in-sample, yet these benefits are not preserved out-of-sample.[3] Dai (2009) and Giamouridis et al. (2014) use a dynamic asset allocation setting and find that commodities offer diversification benefits in-sample and both in- and out-of-sample, respectively. However, all studies require the specification of investor's utility function and make specific modeling assumptions on the asset prices distributional characteristics. To the best of our knowledge, we are the first who explore whether commodities yield diversification benefits by employing an SDE approach; portfolio selection based on SDE is non-parametric.

Regarding the SDE literature, Post (2003) and Post and Versijp (2007), develop tests for SDE based on a duality representation of the investor's expected utility maximization problem which maps to the asset returns distributional characteristics. Their procedure relies on ranked observations in an identical independently distributed (i.i.d.) observations framework. Kuosmanen (2004) develops linear programming tests for SD efficiency that account for diversification possibilities. Even though these studies provide an important step in the evolution of the SDE literature, the proposed tests assume that asset returns are i.i.d.; the empirical evidence does not support this assumption though.

Post (2003), Kuosmanen (2004), Post and Versijp (2007), and Post and Kopa (2013) *test* whether the market portfolio is efficient from a stochastic dominance point of view; no portfolio construction and

[3] Buyuksahin et al. (2010), Chan et al. (2011), Tang and Xiong (2012), Delatte and Lopez (2013), Silvennoinen and Thorp (2013) and Buyuksahin and Robe (2014) also call the diversification benefits of commodities in question because they report that correlations between commodities and equities have increased over the last decade.

evaluation of its performance is undertaken though. Hodder et al. (2015) is the closest to our study from a SDE portfolio construction perspective. They construct stock portfolios that are SSDE over the CRSP all share index and they evaluate their out-of-sample performance. However, this study also relies on the assumption that asset returns are i.i.d. Scaillet and Topaloglou (2010) SDE methodology is more general than the previous SDE methodologies in that it does not assume that asset returns are i.i.d. Furthermore, it allows for a stronger definition of SDE allowing to test for global SDE. According to the previous studies, a portfolio is defined to be SD efficient if and only if it is not stochastically dominated by any other portfolio that can be constructed from a given asset universe for any given SDE criterion under consideration; this definition may give rise to multiple SD efficient portfolios though. Scaillet and Topaloglou (2010) use a stronger version of stochastic dominance efficiency where a portfolio is defined to be SD efficient when it stochastically dominates all other portfolios for any given SDE criterion under consideration. If a portfolio dominates all other portfolios then it is not dominated by any other portfolio, thus it is SD efficient.

Finally, there is mixed evidence on the integration of commodity markets with equity and bond markets. Bessembinder and Chan (1992) document that futures contracts (agricultural, metals and currency futures) are subject to different sources of priced risk than are equities. Bessembinder (1992) and de Roon et al. (2000) find that returns of commodity futures increase with net short positions of commodity hedgers after controlling for systematic risk. Erb and Harvey (2006) find that the Fama-French (1993) factors do not drive the returns of individual commodity futures. Gorton and Rouwenhorst (2006) suggest that the low correlations of commodities with other asset classes could be regarded as evidence for market segmentation. Daskalaki et al. (2014) test various asset pricing models which have been documented to explain equities markets' returns. They find that none of the models can explain the commodity futures returns per se which implies that commodity and equities market are segmented. On the other hand, Tang and Xiong (2012) argue that the financialization of commodity futures tends to integrate equity and commodity markets and Asness et al. (2013) and Giampietro et al. (2015) document that there are common factors which explain the pooled cross-section of various asset classes including commodities.

The remainder of the paper is structured as follows. Section 2 describes the tests for SDE and the construction of optimal portfolios under the SDE criteria, respectively. Section 3 describes the dataset. Sections 4 and 5 present and discuss results from the in-sample and out-of-sample analysis, respectively. Section 6 presents the evidence on the markets segmentation and Section 7 concludes and discusses the implications of our findings.

2. Stochastic dominance efficiency: The test and portfolio construction

2.1. Description of the test

We describe the Scaillet and Topaloglou (2010) test for SDE. Let the asset returns be described by a strictly stationary process $\{\mathbf{Y}_t\}$ taking values in R^n. Observations consist of a realization of $\{\mathbf{Y}_t\ ;\ t = 1,...,T\}$. We denote by $F(\boldsymbol{y})$, the continuous cumulative distribution function of $\mathbf{Y}=(Y_1,\ldots Y_n)'$ at point $\mathbf{y}=(y_1,\ldots y_n)'$. Let a portfolio consisting of n assets and the vector $\boldsymbol{\lambda}$ of portfolio weights in L, where $L = \{\boldsymbol{\lambda} \in \mathrm{R}^n : \mathbf{e}'\,\boldsymbol{\lambda}=1\}$ with $\mathbf{e}$ being a vector of units. Let $G(z,\boldsymbol{\lambda};F)$ denote the cumulative density function of the portfolio return $\boldsymbol{\lambda}'\mathbf{Y}$ at portfolio return point z given by

$$G(z,\lambda;F) := \int_{R^n} I\{\lambda' u \leq z\} dF(u) \tag{1}$$

where $I(\ )$ denotes the indicator function taking the value of 1 if $\lambda' u \leq z$ and 0 otherwise. Further, define

$$J_1(z,\lambda;F) := G(z,\lambda;F),$$
$$J_2(z,\lambda;F) := \int_{-\infty}^{z} G(u,\lambda;F)du = \int_{-\infty}^{z} J_1(u,\lambda;F)du \tag{2}$$

The hypothesis for testing SDE of order *j* (*j*=1 for first SDE and *j*=2 for second SDE) can be written compactly as:

$$H_0^j : J_j(z,\tau;F) \le J_j(z,\lambda;F) \quad \text{for all } z \in R\text{, and for all } \lambda \in L,$$

$$H_1^j : J_j(z,\tau;F) > J_j(z,\lambda;F) \quad \text{for some } z \in R\text{, and for some } \lambda \in L.$$

Under the null hypothesis H_0^j there is no portfolio λ formed from the set of assets that dominates the benchmark τ at any order *j*, i.e. the benchmark portfolio τ is SDE. In this case, the function $J_j(z,\tau;F)$ is always lower than the function $J_j(z,\lambda;F)$ for any possible portfolio λ constructed from the set of alternative assets for any point *z*. Under the alternative hypothesis H_1^j, we can construct a portfolio λ that for some points *z*, the function $J_j(z,\tau;F)$ is greater than the function $J_j(z,\lambda;F)$, i.e. the benchmark portfolio τ is not SDE.

We test the null hypothesis by employing the Scaillet and Topaloglou (2010) test which uses a $\hat{S}_j$ Kolmogorov-Smirnov type test statistic of order *j*

$$\hat{S}_j := \sqrt{T} \sup_{z,\lambda} [J_j(z,\tau;\hat{F}) - J_j(z,\lambda;\hat{F})] \tag{3}$$

where $\hat{F}$ is the empirical distribution of *F*. We reject H_0^j if $\hat{S}_j > c_j$, where c_j is some critical value (for the test properties, see Scaillet and Topaloglou, 2010). Given that the distribution of $\hat{S}_j$ is not known, we calculate the *p*-value corresponding to c_j by bootstrap. We use Abadie's (2002) block bootstrap method. The method does not assume that asset returns are identically and independently distributed (i.i.d.); this is in contrast to SDE tests employed by the previous literature which rely on the i.i.d. assumption. It divides the original data into blocks. Blocks are resampled with replacement from the original data to mimic the time dependent structure of the original data. The block size equals the sample size once it is raised to the third power. We generate R=500 bootstrap samples. We define the p-value $p_j^* := P[S_j^* > \hat{S}_j]$, where S_j^* is the test statistic corresponding to each bootstrap sample. The p-value is approximated by

$$\tilde{p}_j = \frac{1}{R}\sum_{r=1}^{R} I\{\tilde{S}_j^r > \hat{S}_j\} \tag{4}$$

where the $\tilde{S}_j^r$ is the test statistic corresponding to the rth bootstrapped sample. Proposition 3.1 in Scaillet and Topaloglou (2010) ensures that the power of the test is preserved under the bootstrapped critical values.

Notice that rejection of the null hypothesis does not necessarily imply that there is a portfolio that stochastically dominates the benchmark portfolio. To examine whether such a portfolio exists in this case, we repeat the Scaillet and Topaloglou test to trace the SDE portfolio if any. We provide details on the implementation of the test under the first and second order SDE criteria in Appendix A.

2.2. Portfolio construction

We describe how we trace the SDE portfolio (if it exists) in the case where the null hypothesis is rejected. Portfolio λ is termed to dominate a benchmark portfolio $\boldsymbol{\tau}$ under the first order and second order stochastic dominance efficiency criteria (FSDE, SDEE) respectively, if it satisfies the following respective equations

$$\max_{z,\lambda}[G(z,\tau;F) - G(z,\lambda;F)] \tag{5}$$

$$\max_{z,\lambda}[\int_{-\infty}^{z} G(u,\tau;F)du - \int_{-\infty}^{z} G(u,\lambda;F)du] \tag{6}$$

The resulting portfolio is also termed efficient. Therefore, a portfolio is defined to be efficient when it stochastically dominates *all* other portfolios constructed from a given asset universe for any given SDE criterion under consideration.[4] Equations (5) and (6) highlight the difference between the standard SD and the SDE concepts. In the latter case, the dominant portfolio is derived by taking the maximum over *all* possible portfolios and returns whereas in the former case two *given* portfolios are being compared. Figures 1 and 2 display the FSDE and SSDE concepts. Note that FSDE is a sufficient but not a necessary condition for SSDE.

Hadar and Russel (1969) and Bawa (1975) show that FSD (SSD) amounts to choosing the investment that maximizes investor's expected utility assuming that investors preferences are characterized by non-satiation (non-satiation and risk aversion). Given Hadar and Russel (1969) and Bawa (1975) results, the optimal portfolio is the SD efficient portfolio derived from the solution of equations (5) and (6). Notice that the construction of optimal portfolios under the FSDE and SSDE criteria does not require an assumption on the specific form of a utility function. This is because both SDE criteria are consistent with a broad class of utility functions. FSDE is appropriate for both risk lovers and risk averters and permits a preliminary screening of investment alternatives eliminating those which no rational investor will ever choose. The SSDE criterion adds the assumption of global risk aversion.

From an implementation point of view, one needs to search for the portfolio weights so that the optimal portfolio will yield a cumulative distribution function that satisfies equations (5) and(6), i.e. a portfolio that maximizes the distance between the two cumulative distribution functions (FSDE) and the

[4] Notice that Markowitz (1952) concept of efficient portfolio differs from the efficient one under the stochastic dominance efficiency criterion. The latter is the optimal portfolio whereas the former may not be the optimal one even if Markowitz assumptions hold; from the set of efficient à la Markowitz portfolios, only one will be the optimal for a given investor if Markowitz assumptions hold.

distance between the integrals of the cumulative distribution functions (SSDE) for any given return. To this end, we need to choose a benchmark portfolio as well as the size of the historical sample of asset returns.

At any point in time (every month in our case), we use a two-step procedure that employs the Scaillet and Topaloglou (2010) SDE test to a portfolio construction setting in order to construct the optimal portfolio (if any) that is first- and second-order SDE with respect to a benchmark portfolio. We proceed in two steps. At any point in time, first, we choose the S&P 500 as a benchmark portfolio and compare it to portfolios derived from an asset universe which consists only of three asset classes: equities, bonds and cash. We apply the Scaillet and Topaloglou (2010) SDE methodology and we detect the optimal portfolio. If there are many portfolios consisting of these three asset classes that dominate the S&P 500, we take as optimal the one that maximizes the distance from the cumulative distribution function of the S&P 500 (according to the SD definition, dominant is the portfolio with the lowest cumulative distribution function). In case there is not a portfolio that dominates the S&P 500, we set the index as the optimal portfolio in that particular month. In the second step, we augment the asset universe by including the commodity asset class. We choose the optimal portfolio from the first step as the benchmark portfolio and we re-apply the Scaillet and Topaloglou (2010) SDE methodology. Every month, the described procedure delivers two optimal portfolios for the two respective asset universes. We repeat the two steps throughout our sample in a rolling window fashion. This delivers a time series of optimal portfolios for each one of the two asset universes. We construct optimal portfolios under both the FSDE and SSDE efficient criteria over January 2001-September 2013. We allow for short selling. Then, we evaluate their out-of-sample performance.

3. The data

We use data on monthly closing prices of a number of indices obtained from Bloomberg. We employ the S&P 500 Total Return Index, Barclays U.S. Aggregate Bond Index and the one-month Libor rate to proxy the traditional asset universe, i.e. the equity market, the bond market and the risk-free rate, respectively. To access the commodity asset class, we use various widely-followed commodity futures indices. We use first, second and third generation commodity indices. This is in contrast to the previous literature that uses only first generation commodity indices.

The first generation indices are long-only fully-collateralized investments. They reflect a strategy consisting of the shortest maturity commodity futures contracts which rolls to the subsequent month's contracts as the lead or front month expires. We consider the S&P Goldman Sachs Commodity Index (S&P GSCI), the Dow Jones-UBS Commodity Index (DJ UBSCI) and the Deutsche Bank Liquid Commodity Index (DBLCI) as representatives of the first generation indices. By construction, the first generation indices provide positive roll returns only in the case that the term structure of commodity futures prices is in backwardation. However, given that futures markets switch from backwardation to contango and vice versa, the first generation indices perform poorly in contangoed markets. In addition, the contracts close to expiration tend to be more contangoed and volatile and as a result they experience a negative roll yield.

The second generation indices take these facts into account and they attempt to minimize the harmful impact of contangoed markets by investing into sufficiently liquid contracts which are further out in the term structure of commodity futures prices. We consider the JP Morgan Commodity Curve Index (JPMCCI), the Deutsche Bank Liquid Commodity Index-Optimum Yield (DBLCI-OY), the Morningstar Long/Flat Commodity Index (MSDILF) and the Morningstar Long-Only Commodity Index (MSDIL) as representatives of the second generation indices.

Finally, the third generation indices allow taking both long and short positions. We consider the Morningstar Short/Flat Commodity Index (MSDISF), the Morningstar Short-Only Commodity Index (MSDIS) and the Morningstar Long/Short Commodity Index (MSDILS) as representatives of the third generation indices. The appendix provides details on the construction of these three generation indices. To ensure that both asset universes employ indices which represent dynamic trading strategies, we use also the Fama-French (1993) equity factors to proxy the equity market, i.e. the value and the size factors. We obtain these factors from Kenneth French's website. Table 1 describes the set of equity, bond and commodity indices employed in this study.

The dataset spans January 1990 to September 2013 with the exception of DJ-UBSCI that covers the period January 1991 to September 2013 due to data availability constraints. Table 2 reports summary statistics regarding the performance of the employed indices over this period. We can see that the monthly average return on commodity indices is higher than that of stocks and bonds and in most cases it exhibits greater standard deviation. With a few exceptions, the Sharpe ratio is considerably greater for bonds and stocks than commodity indices. The reported evidence is consistent with previous studies, which document that the stand-alone risk-adjusted performance of commodity indices is inferior to other asset classes (see e.g., Jensen et al., 2000, Daskalaki and Skiadopoulos, 2011).

Table 3 reports the pairwise correlation among the indices employed in the study. We can see that the pairwise correlations of commodity indices with the stock and bond ones are low or even negative. This indicates the potential diversification benefits of commodities. In most cases, there is a strong positive correlation among the commodity indices. However, the third generation indices constitute an exception; the indices with exclusively short positions exhibit negative correlation with the rest of the commodity indices.

Finally, we obtain data on a set of variables documented to forecast returns in equity and bond markets. We obtain data on the dividend yield on MSCI World, the junk bond premium (or default spread, defined as the excess of the yield on long-term BAA corporate bonds rated by Moody's over the yield on AAA-rated bonds), the term spread (defined as the difference between the Aaa yield and the one-month bill rate) and the Baltic Dry Index from Bloomberg. We obtain data on the 3-month Treasury Bill, the Industrial Production, and the money supply from the Board of Governors of the Federal Reserve System (U.S.).

4. In-sample Analysis: Results and discussion

We test whether the inclusion of commodities in the asset universe makes the investor better off compared to the case where the asset universe consists of only traditional asset classes (stocks, bonds and cash). In this section, we conduct an in-sample analysis by employing a SDE test. We proceed in two steps. In the first step, we define the benchmark portfolio τ to be the S&P 500 index. Then, we apply the Scaillet and Topaloglou (2010) test to assess the null hypothesis that the S&P 500 is SD efficient relative to any λ portfolio to be formed based on the constrained asset universe for every return level (described in Section 2). In the second step, we choose the benchmark portfolio to be the one dictated by the first step (the S&P 500 in the case the null hypothesis is not rejected or the portfolio based on the constrained asset universe if the null is rejected). Then, we re-apply the Scaillet and Topaloglou test to assess the null hypothesis that the redefined benchmark portfolio is efficient relative to any λ portfolio to be formed based on the augmented with commodities asset universe for every return level. We apply the two-step testing procedure for FSDE and SSDE criteria, separately.

We conduct the analysis using each one of the commodity indices under scrutiny separately. We access investment in commodities via the first generation indices (S&P GSCI, DJ-UBS CI, DBLCI), second generation indices (JPMCCI, DBLCI-OY, MSDILF, MSDIL) and the third generation indices (MSDIF, MSDILS, MSDIS), separately. The use of second and third generation commodity indices is novel in the literature on the diversification benefits of investments in commodities. It allows us to address our research question by taking into account the characteristics of dynamic long/short commodity strategies which exploit popular commodity trading signals such as momentum and switches from backwardation to contango and vice versa. Rallis et al. (2013) find that these dynamic strategies outperform strategies similar in spirit to the first generation indices. We consider two in-sample periods. The first is the full sample period, i.e. January 1990 to December 2013. The second is January 1990 – December 2000. The latter is the period we retain as a starting point to conduct the out-of-sample analysis over January 2001 – September 2013 in Section 6.

Table 4 reports the FSDE and SSDE test statistics and respective *p*-values for the null hypothesis that a set of benchmark assets consisting of stocks, bonds and the risk-free asset is efficient versus the benchmark asset universe augmented with the commodity asset class over the period from January 1990 to September 2013. We access investment in commodities via the first, second and third generation indices, separately. Panel A reports results for the case where the benchmark asset universe consists of the S&P 500 Total Return Index, Barclays Aggregate Bond Index and one-month LIBOR rate. Panels B and C report results when the benchmark set includes dynamic equity indices, i.e. the Fama-French (1993) size factor (Small minus Big, SMB) and the value factor (High minus Low, HML), respectively. We can see that the null hypothesis can be rejected in most cases at a 5% significance level and it can be rejected in almost all cases at a 10% significance level. Furthermore, we find that there is a portfolio that consists of stocks, bonds, cash and commodities which is SDE. This evidence suggests that the performance of traditional portfolios, consisting of stocks, bonds and cash, can be significantly improved by investing in commodities.[5] Qualitatively similar results are derived over the period from January 1990 to December 2000.

Table 5 reports the results obtained for the period January 1990 to December 2000. Notice that the strong evidence on the diversification benefits from investing in commodities is not biased by the different nature of indices in the two asset universes under comparison. The use of the Fama-French (1993) factors ensures a fair comparison of the traditional asset classes universe with the augmented with commodities one. This is because in both asset universes, we use indices that represent dynamic trading strategies.

5. Diversification benefits: Out-of-sample analysis

We calculate optimal portfolios separately for two asset universes, one that includes “traditional” asset classes (i.e. equities, bonds, risk-free asset) and an “augmented” one that also includes commodity indices. We evaluate their relative performance in an out-of-sample setting which is the ultimate test given that at any given point in time, the investor estimates the optimal weights, sets the strategy in action, and then she reaps the realized returns over her investment horizon. We construct optimal portfolios by applying the two-step Scaillet and Topaloglou (2010) methodology described in Section 2 to a portfolio construction setting.

[5] Interestingly, we implemented Huberman and Kandel (1987) tests for mean-variance spanning and we find that that the augmented asset universe spans the constrained one, i.e. commodities do not provide diversification benefits. This highlights the well-known result that testing for SDE is a more general approach than testing for mean-variance spanning because the former does not impose any distributional assumptions on asset returns (for a review on mean-variance spanning, see also DeRoon and Nijman, 2001).

5.1. Performance evaluation

We evaluate the alternative investment opportunity sets (i.e. the traditional versus the augmented one) in terms of certain characteristics of the respective optimal portfolios that we construct in an out-of-sample setting. To this end, we employ a "rolling-window" approach. Assume that the dataset consists of T (in our case, $T=285$) monthly observations for each asset and K is the size of the employed rolling window used for the calculation of the portfolio weights. Standing at each month t, we use the previous K observations to estimate the SDE portfolio weights (i.e. weights that maximize expected utility). Next, we use the estimated weights to compute the out-of-sample realised return over the period $[t, t+1]$. We repeat this process by incorporating the return for the next period and ignoring the earliest one, until the end of the sample is reached. This approach allows deriving a series of monthly out-of-sample optimal portfolio returns. Then, we use the time series of realised portfolio returns to evaluate the out-of-sample performance of the derived optimal portfolios. We choose the size of the rolling window K=120. This delivers January 1990 –December 2000 as the starting time interval for the estimation of optimal portfolio weights and January 2001 – December 2013 as the out-of-sample period.

We compare the out-of-sample performance of the two optimal portfolios based on the respective asset universes by using non-parametric and parametric tests. First, we use the Scaillet and Topaloglou (2010) non-parametric test to test the null hypothesis that the optimal portfolio based on the augmented asset universe (stocks, bonds, cash and commodities) stochastically dominates the traditional asset universe. The alternative hypothesis is that the augmented asset universe based portfolio does not stochastically dominate the traditional asset universe based optimal portfolio. We test the null hypothesis by using first order and second order stochastic dominance criteria. The test does not require any assumptions either on investor's preferences (apart from non-satiation and risk aversion) or on the distribution of asset returns.

Next, in line with DeMiguel et al. (2009), Kostakis et al. (2010) and Daskalaki and Skiadopoulos (2011), we employ four commonly used parametric performance measures: the Sharpe ratio (SR), opportunity cost, portfolio turnover and a measure of the portfolio risk-adjusted returns net of transaction costs. These performance measures, yet parameteric, provide information useful to practitioners and hence they supplement the results from the previously discussed non-parametric SDE measure. To fix ideas, let a specific strategy denoted by c. The estimate of the strategy's SR_c is defined as the fraction of the sample mean of out-of-sample excess returns $\hat{\mu}_c$ divided by their sample standard deviation $\hat{\sigma}_c$, i.e.

$$\widehat{SR}_c = \frac{\hat{\mu}_c}{\hat{\sigma}_c} \tag{7}$$

To test whether the SRs of the two optimal portfolio strategies based on the traditional and augmented with commodities asset universes, are statistically different, we use the statistic proposed by Jobson and Korkie (1981) and corrected by Memmel (2003). The use of SR is in line with the finance industry practice however it is suitable to assess the performance of a strategy only in the case where the strategy's returns are normally distributed.

Next, we use the concept of opportunity cost (Simaan, 1993) to assess the economic significance of the difference in performance of the two optimal portfolios, respectively. Denote by r_{wc}, r_{nc} the optimal portfolio realized returns obtained by an investor with the augmented investment opportunity set that includes commodities and the investment opportunity set restricted to the traditional asset classes, respectively. The opportunity cost θ is defined to be the return that needs to be added (or subtracted) to the portfolio return r_{nc} so that the investor becomes indifferent (in utility terms) between the two strategies imposed by the different investment opportunity sets, i.e.

$$E\left[U\left(1 + r_{nc} + \theta\right)\right] = E\left[U\left(1 + r_{wc}\right)\right] \tag{8}$$

Therefore, a positive (negative) opportunity cost implies that the investor is better (worse) off in case of an investment opportunity set that allows commodity investing. Notice that the opportunity cost takes into account all the characteristics of the utility function and hence it is suitable to evaluate strategies even when the assets return distribution is not normal. To calculate the opportunity cost, we use an exponential and a power utility function alternatively.

The portfolio turnover (PT) is computed so as to get a feel of the degree of rebalancing required to implement each one of the two strategies. For any portfolio strategy *c*, the portfolio turnover, PT_c is defined as the average absolute change in the weights over the *T-K (T-120)* rebalancing points in time and across the *N* available assets, i.e.

$$PT_c = \frac{1}{T-K}\sum_{t=1}^{T-K}\sum_{j=1}^{N}\left(\left| w_{c,j,t+1} - w_{c,j,t+}\right|\right) \tag{9}$$

where $w_{c,j,t}, w_{c,j,t+1}$ are the derived optimal weights of asset *j* under strategy *c* at time *t* and *t+1*, respectively; $w_{c,j,t+}$ is the portfolio weight before the rebalancing at time *t+1*; the quantity $\left|w_{c,j,t+1} - w_{c,j,t+}\right|$ shows the magnitude of trade needed for asset *j* at the rebalancing point *t+1*. The *PT* quantity can be interpreted as the average fraction (in percentage terms) of the portfolio value that has to be reallocated over the whole period.

Finally, we also evaluate the two investment strategies under the risk-adjusted, net of transaction costs, returns measure proposed by DeMiguel et al. (2009). This metric provides an economic interpretation of the PT; it shows how the proportional transaction costs generated by the portfolio turnover affect the returns from any given strategy. To fix ideas, let *pc* be the proportional transaction cost and $r_{c,p,t+1}$ the realized portfolio return at *t+1* (before rebalancing). The evolution of the net of transaction costs wealth NW_c for strategy *c,* is given by:

$$NW_{c,t+1} = NW_{c,t}\left(1 + r_{c,p,t+1}\right)\left[1 - pc \times \sum_{j=1}^{N}\left(\left| w_{c,j,t+1} - w_{c,j,t+}\right|\right)\right] \quad (10)$$

Therefore, the return net of transaction costs is defined as

$$RNTC_{c,t+1} = \frac{NW_{c,t+1}}{NW_{c,t}} - 1 \quad (11)$$

The return-loss measure is calculated as the additional return needed for the strategy with the restricted opportunity set to perform as well as the strategy with the expanded opportunity set that includes commodity futures. Let μ_{wc}, μ_{nc} be the monthly out-of-sample mean of *RNTC* from the strategy with the expanded and the restricted opportunity set, respectively, and σ_{wc}, σ_{nc} be the corresponding standard deviations. Then, the return-loss measure is given by:

$$return - loss = \frac{\mu_{wc}}{\sigma_{wc}} \times \sigma_{nc} - \mu_{nc} \quad (12)$$

To calculate $NW_{c,t+1}$, we set the proportional transaction cost *pc* equal to 50 basis points per transaction for stocks and bonds (for a similar choice, see DeMiguel et al., 2009), 35 basis points for the commodity indices (based on discussion with practitioners in the commodity markets), and zero for the risk-free asset.

5.2 Results and discussion

This section discusses the results on the out-of-sample performance of the traditional asset classes portfolios and the augmented with commodities ones. Table 6 reports the Scaillet and Topaloglou (2010) test statistics and *p*-values (within parentheses). We employ first order and second order stochastic dominance criteria (FSD, SSD, respectively). The traditional asset universe set includes the S&P 500 Equity Index, the Barclays US Aggregate Bond Index and the 1-month LIBOR. Investors access investment in commodities via the first generation indices (S&P GSCI, DJ-UBS CI, DBLCI), second generation indices (JPMCCI, DBLCI-OY, MSDILF, MSDIL) and third generation indices (MSDIF, MSDILS, MSDIS), separately. We can see that we cannot reject the null hypothesis, i.e. we document that the augmented with commodities optimal portfolios stochastically dominate the optimal portfolios based on the traditional asset universe. This holds in all cases under both FSDE and SSDE criteria. Thus, we extend the in-sample evidence on the diversification benefits provided by commodities.

Next, we compare the out-of-sample performance of the two optimal portfolios using the four standard measures of performance. Table 7 reports results for each one of the four performance measures. Panels A and B report results for the cases where portfolio weights are calculated by FSDE and SSDE criteria, respectively. In the case of opportunity cost, we assume various levels of (absolute/relative) risk aversion (ARA, RRA=2, 4, 6) for the individual investor. To assess the statistical significance of the superiority in SRs, we also report the *p*-values of Memmel's (2003) test within parentheses. The null hypothesis is that the SRs obtained from the traditional investment opportunity set and the expanded with commodities investment opportunity set are equal.

In the case of the first generation indices, results are mixed. We can see that the optimal portfolios formed based on the augmented investment opportunity set yield greater SRs than the corresponding portfolio strategies based on the traditional investment opportunity set. However, the *p*-values of Memmel's (2003) test indicate that the differences in SRs are not statistically significant. Regarding the opportunity cost, we can see that this is positive in more than half of the cases. The positive sign indicates that the investor is demanding a premium in order to replace the optimal strategy that includes investment in commodities with the optimal one that invests only in the traditional asset classes. This implies that the investor is better off when the augmented investment opportunity set is considered. Interestingly, in most cases, the opportunity cost decreases (in absolute terms) as the risk aversion increases. This implies that the investor tends to becoming indifferent in utility terms between including and excluding commodities in her asset portfolio as she becomes more risk averse. Furthermore, with the exception of S&P GSCI, the portfolios that include only the traditional asset classes induce more portfolio turnover compared with the ones that include commodities. Finally, we can see that the return-loss measure that takes into account transaction costs is positive. The positive sign confirms the out-of-sample superiority of the portfolios that include commodity indices, even after deducting the incurred transaction costs. These findings hold regardless of the commodity index. Furthermore, they hold regardless of whether portfolio weights are calculated by the FSDE or SSDE criteria.

In the case of the second generation indices, there is strong evidence on the diversification benefits of commodities. We can see that the optimal portfolios formed based on the augmented investment opportunity set yield greater SRs than the corresponding portfolio strategies based on the traditional investment opportunity set. In contrast to the first generation indices, the *p*-values of Memmel's (2003) test show that the differences in SRs are statistically significant. Regarding the opportunity cost, we can see that this is positive in all cases which implies that the investor is better off when the augmented investment opportunity set is considered. Again, in most cases, the opportunity cost decreases as the risk aversion increases. Furthermore, with the exception of MSDIL, the portfolios that include only the traditional asset classes induce more portfolio turnover compared with the ones that also include commodities. Finally, we can see that the return-loss measure that takes into account transaction costs is positive which confirms the out-of-sample superiority of the portfolios that include commodity indices, even after deducting the incurred transaction costs. Again, these findings hold both for the cases where portfolios are constructed by FSDE and SSDE criteria just as was the case with the first generation commodity indices.

Regarding the third generation commodity indices, again, the diversification benefits of commodities are pronounced. We can see that the optimal portfolios formed based on the augmented investment opportunity set yield greater SRs than the corresponding portfolio strategies based on the traditional investment opportunity set. The *p*-values of Memmel's (2003) test indicate that the differences in SRs are statistically significant when investments in commodities are accessed by any commodity index but MSDIS. Regarding the opportunity cost, we can see that this is positive in all cases. Portfolios that include only the traditional asset classes induce more portfolio turnover compared with the ones that also include commodities; the only exception appears for MSDIS again. Finally, we can see that the return-loss measure is positive for all indices. These findings hold regardless of whether portfolio weights are constructed either by the FSDE or the SSDE criteria. Notice that the fact that MSDIS constitutes an exception for some performance measures highlights the superiority of commodity indices that mimic dynamic trading strategies; MSDIS is a passive short index.

Table 8 and Table 9 report results in the case where investors access investment in equities via dynamic indices, i.e. the Fama-French (1993) value factor (High minus Low, HML) and the size factor (Small minus Big, SMB), respectively. We can see that the optimal augmented portfolios that include commodity investing outperform the ones that do not just as it was the case where the passive equity index was considered. Interestingly, the diversification benefits of commodities appear unanimously even in the case of the first generation commodity indices in the case where the value factor is considered. Optimal portfolios formed based on the augmented investment opportunity set yield greater SRs than the corresponding portfolio strategies based on the traditional investment opportunity set. In most cases, the *p*-values of Memmel's (2003) test indicate that the differences in SRs are statistically significant. This evidence is stronger for the HML index and for the FSDE case. Regarding the opportunity cost, we can see that this is positive in almost all cases. Few exceptions occur in the case of the size factor for the SSDE cases. In most cases, the opportunity cost decreases (in absolute terms) as the risk aversion increases. Regarding the portfolio turnover, results differ when the dynamic equity indices are considered. In most cases, portfolios that include only the traditional asset classes induce less portfolio turnover compared with the ones that also include individual commodity investing. Finally, the return-loss is positive in all cases across the various commodity indices, both for the FSDE and SSDE cases. This implies that even though portfolios based on an investment opportunity set that includes dynamic indices have less turnover than the ones based on the expanded opportunity set, the investors can still earn positive risk-adjusted return by investing in commodities.

To sum up, the optimal portfolios formed based on the augmented investment opportunity set outperform the corresponding portfolios based on the traditional investment opportunity set. The diversification benefits of commodities are more pronounced when the investor accesses commodities via the second and third generation commodity indices. This implies that active commodity investing that exploits momentum and term structure signals and that allows for short positions as well increases diversification benefits to commodity investors compared to long only passive commodity strategies. This is in accordance with the evidence in Giamouridis et al. (2014) on the performance of dynamic commodity strategies in a portfolio setting; yet, their results depend on their modelling assumptions. Our findings hold regardless of the assumed equity index and regardless of whether portfolio weights are constructed by FSDE or SSDE criteria.

6. Why do investments in commodities offer diversification benefits?

In this section, we investigate the reason of the outperformance of the augmented portfolios that include commodities versus the ones that include only the traditional asset classes documented in the previous sections. To this end, we test whether the commodity market is integrated/ segmented with the equity and bond market. Evidence of market segmentation would confirm that commodities form an alternative asset class and this would justify the documented diversification benefits. We follow Campbell and Hamao (1992) to develop the market integration test. Let a *K*-factor asset pricing model

$$R_{i,t+1} = E_t\left(R_{i,t+1}\right) + \sum_{k=1}^{K} \beta_{ik} f_{k,t+1} + e_{i,t+1} \tag{13}$$

where $R_{i,t+1}$ denotes the excess return on asset i held from time t to time t+1, $f_{k,t+1}$ the kth factor realization, β_{ik} the factor loading with respect to the kth factor and $e_{i,t+1}$ the error term. Equation (13) maps to an expected return-beta representation

$$E_t\left(R_{i,t+1}\right) = \sum_{i=1}^{K} \beta_{ik} \lambda_{kt} \tag{14}$$

where λ_{kt} is the market price of risk for the kth factor at time t. The notation in equation (14) indicates that the time variation in expected returns stems from time varying market prices of risk rather than time varying betas. The time variation in the kth factor market price of risk is modeled as

$$\lambda_{kt} = \sum_{n=1}^{N} \theta_{kn} X_{nt} \tag{15}$$

where there is a set of N forecasting (instrumental) variables, X_{nt}, n=1,2,…,N.

Market integration is defined to be the case where financial assets that trade in different markets yet they have identical risk characteristics, will have identical expected returns (Bessembinder, 1992, Campbell and Hamao, 1992). Equation (14) implies that in the case of market integration, the expected returns of different asset classes are driven by the same prices of risks.

The hypothesis of market integration can be tested by means of Fama-MacBeth (1973) two pass regressions (e.g., Bessembinder, 1992). This is not possible in our case though because at any point in time, we have only four observations for the respective four asset classes. We follow the approach of Bessembinder and Chan (1992) and Campbell and Hamao (1992) to bypass this constraint and test the market integration hypothesis. Assuming that (1) markets are integrated, i.e. they have the same prices of risk, and (2) the time variation in expected returns stems from time variation in the prices of risk, equations (14) and (15) imply that predictor variables that drive prices of risk should be the same across assets.[6] Therefore, given assumption (2), evidence that the predictors of the price of risk differ across asset classes would imply that the market price of risk is not the same across assets and hence markets are segmented.

To test whether the predictors of the prices of risk are the same across asset classes, we formulate the mapping between market price of risk and its predictors to an equivalent mapping between expected returns and the price of risk predictors. Equations (14) and (15) yield

$$\begin{aligned} E_t\left(R_{i,t+1}\right) &= \sum_{k=1}^{K}\beta_{ik}\sum_{n=1}^{N}\theta_{kn}X_{nt} = \\ &= \beta_{i1}(\theta_{11}X_{1t}+...+\theta_{1N}X_{Nt})+...+\beta_{iK}(\theta_{K1}X_{1t}+....+\theta_{KN}X_{Nt}) \\ &= (\beta_{i1}\theta_{11}+..+\beta_{iK}\theta_{K1})X_{1t}+...+(\beta_{i1}\theta_{K1}+...+\beta_{iK}\theta_{KN})X_{Nt} \\ &= \sum_{n=1}^{N}a_{in}X_{nt} \end{aligned} \tag{16}$$

Hence, to investigate whether price of risk predictors differ across asset classes, we run regressions of each asset class return on a set of predictors and we check whether predictors that predict the returns of one asset class also predict the returns of the other asset classes, too.

[6] Ferson and Harvey (1991) document that assumption (2) holds; expected returns are driven by time variation of prices of risk rather than time variation of betas.

We use instrumental variables that have been documented to predict equity and bond markets returns: the dividend yield, the yield on the three-month Treasury bills, the default spread (Bessembinder and Chan, 1992), the term spread (Fama and French, 1989), the industrial production growth (Fama, 1990), the money supply growth (Chen, 2007), and the growth in the Baltic Dry Index (Bakshi et al., 2011). Table 10 presents the evidence on the forecastability of the asset returns during the period 1999-2013; the choice of the period is consistent with the choice of the out-of-sample period in Section 5.1 where the out-of-sample performance of portfolios is evaluated. The coefficient estimates, the respective t-statistics, the adjusted R^2, the F-statistic and the respective p-values are reported. We can see that the S&P 500 equity index can be predicted by the T-bills, the term spread and the growth in the Baltic Dry Index. The Barclays Bond index can be predicted by the T-bills, the dividend yield and the term spread. On the other hand, these predictors do not forecast commodity indices returns. This evidence is also supported by the F-statistic and the respective p-values: the hypothesis that all coefficient estimates equal to zero can be rejected only for the traditional asset classes, i.e. the S&P 500 equity index and the Barclays Bond Index.

Our results show that commodity returns cannot be forecasted by variables that predict stock and bond market returns. This suggests that the price of risk is driven by different predictors across the various asset classes which in turn implies that the prices of risk differ. Hence, commodity markets are segmented from equity and bond markets. Evidence of market segmentation indicates that commodities form an alternative asset class. This explains the outperformance of the extended with commodities optimal portfolios documented in the previous sections.

A remark is in order at this point. Our results are not in contrast with the previous literature which documents that commodities can be predicted by various macroeconomic factors. Most studies employ *individual* commodity futures to examine the predictability of the chosen instruments (e.g., Bessembinder and Chan, 1992, Bjornson and Carter, 1997) or passive commodity indices (e.g., Bakshi et al., 2011, Hong and Yogo, 2012). We consider *dynamic portfolio* strategies via commodity indices rather than individual commodity futures which represent passive strategies. Interestingly, Gargano and Timmerman (2014) who examine the predictability of (spot) commodity indices also provide weak evidence that returns on commodity indexes can be predicted by a set of macroeconomic variables over the period 1947-2010.

7. Conclusions

One of the key yet still open questions is whether commodities offer diversification benefits once included in a portfolio that consists of traditional asset classes. The previous literature does not reach a unanimous agreement. Most importantly, to answer this question, it makes strong assumptions on investor's preferences and on the distributional properties of asset returns. We revisit this research question and we bypass these two obstacles by employing the non-parametric stochastic dominance efficiency (SDE) setting. The SDE setting is the natural choice for the purposes of our analysis because it accommodates any utility function to the extent that the investor is greedy or greedy and risk averse. SDE also accommodates any distribution for asset returns. To the best of our knowledge, the application of the SDE setting for the purposes of investigating the diversification benefits of commodities is novel.

We conduct our analysis both in- and out-of-sample by constructing and comparing optimal portfolios derived from two respective asset universes: one that includes only the traditional asset classes (equities, bonds and cash) and one that is augmented with commodities, too. To this end, we apply a two-step procedure that uses Scaillet and Topaloglou (2010) SDE test. We consider first and second order SDE criteria and we provide first time evidence on the performance of various first, second and third generation commodity indices in a portfolio setting.

We find that commodities provide diversification benefits. The results are robust both in- and out-of-sample for both SDE criteria, for a number of portfolio performance evaluation measures, stock indices and alternative periods. Interestingly, the evidence for diversification benefits is more pronounced in the case where the investor access commodities via second and third generation commodity indices. We explain the reported diversification benefits by documenting that commodity markets are segmented from equity and bond markets.

Our results have three implications. First, an investor should include commodities in her portfolio because commodity and traditional markets are segmented. Second, the new generation commodity indices should be preferred to the traditional first generation indices. Third, the investor should access commodities via dynamic commodity futures trading strategies such as the ones mimicked by the second and third generation indices. Future research should explore the use of the SDE setting in portfolio choice problems further because it does not need to make the strong assumptions typically made by studies on portfolio choice.

Appendix A: Mathematical formulations of Scaillet and Topaloglou (2010) test

In this section we present the mathematical formulation of the Scaillet and Topaloglou (2010) stochastic dominance efficiency (SDE) test under the first two SDE criteria and some details on its numerical implementation for the purposes of portfolio construction.

A.1 Mathematical formulation for FSDE

To test for first order stochastic dominance efficiency (FSDE), we optimize the test statistic

$$\hat{S}_1 := \sqrt{T} \sup_{z,\lambda} [J_1(z,\tau;\hat{F}) - J_1(z,\lambda;\hat{F})] \tag{17}$$

The above formulation permits testing the dominance of a given portfolio strategy τ over any potential linear combination λ of the *set* of the available assets. Hence, we implement a test of stochastic dominance efficiency and not a test of standard stochastic dominance. To formulate the mathematical model of the above optimization, we will use two auxiliary binary variables to express the $J_1(z,\tau;\hat{F})$ and $J_1(z,\lambda;\hat{F})$ cumulative distribution functions. These are step functions in our descrete case. The mathematical formulation of the problem is the following:

$$\begin{aligned}
&\max_{z,\lambda} S_1 = \sqrt{T}\frac{1}{T}\sum_{t=1}^{T}[L_t - Q_t] \\
&s.t. \\
&M(L_t - 1) \le z - \tau' Y_t \le M L_t, \forall t \\
&M(Q_t - 1) \le z - \lambda' Y_t \le M Q_t, \forall t \\
&e'\lambda = 1 \\
&Q_t \in \{0,1\},\ L_t \in \{0,1\}, \forall t
\end{aligned} \tag{18}$$

where *M* is the greatest portfolio return. The model is a mixed integer program maximizing the distance between the sum of two binary variables, $\frac{1}{T}\sum_{t=1}^{T} L_t, \frac{1}{T}\sum_{t=1}^{T} Q_t$, which represent $J_1(z,\tau;\hat{F})$ and $J_1(z,\lambda;\hat{F})$, respectively; sums are taken over all possible values of portfolio returns. We need to constrain the binary variable L_t to be equal to 1 if $z \geq \tau' Y_t$ and zero otherwise. This is achieved by using the first group of inequalities. If $z \geq \tau' Y_t$ these inequalities force L_t to equal 1, and 0 otherwise. Similarly, we need to constrain the binary variable Q_t to be equal to 1 if $z \geq \lambda' Y_t$ and zero otherwise. This is achieved using the second group of inequalities. Q_t equals 1 for each return t for which $z \geq \lambda' Y_t$. These two groups of inequalities ensure that the two binary variables are cumulative distribution functions. The third equation defines the sum of all weights to be unity.

To solve the problem described by equation (18), we discretize the variable z and we solve smaller problems *P(r)* in which z is fixed to a given return r (see Scaillet and Topaloglou 2010 for the proof). Then, we take the value for z that yields the maximum distance in equation (5). The advantage of doing so is that the optimal values of the L_t variables are known in $P(r)$ because L_t is not a function of the (unknown) optimal portfolio weights. Hence, problem *P(r)* boils down to the following minimization problem.

$$\begin{aligned}
&\min_{\lambda} \sum_{t=1}^{T} Q_t \\
&s.t. \\
&M(Q_t - 1) \leq r - \lambda' Y_t \leq M Q_t \\
&e'\lambda = 1, \\
&Q_t \in \{0,1\}, \forall t
\end{aligned} \tag{19}$$

In this case, we only have one auxiliary binary variable, Q_t, so the problem is solved much faster.

A.2 Mathematical formulation for SSDE

The model for second order stochastic dominance efficiency is formulated in terms of standard linear programming. Numerical implementation of first order stochastic dominance efficiency is much more computationally demanding because we need to develop mixed integer programming formulations. To test for second order stochastic dominance efficiency (SSDE) of portfolio τ over any potential linear combination λ, we optimize the test statistic

$$\hat{S}_2 := \sqrt{T} \sup_{z,\lambda} [J_2(z,\tau;\hat{F}) - J_2(z,\lambda;\hat{F})] \tag{20}$$

The mathematical formulation of the problem is the following

$$
\begin{aligned}
&\max_{z,\lambda} S_2 = \sqrt{T}\,\frac{1}{T}\sum_{t=1}^{T}[L_t - W_t] \\
&s.t. \\
&M(F_t - 1) \le z - \tau' Y_t \le M F_t \\
&-M(1 - F_t) \le L_t - (z - \tau' Y_t) \le M(1 - F_t) \\
&-M F_t \le L_t \le M F_t \\
&W_t \ge z - \lambda' Y_t \\
&e'\lambda = 1 \\
&W_t \ge 0,\ F_t \in \{0,1\}, \forall t
\end{aligned}
\tag{21}
$$

The model is a mixed integer program maximizing the distance between the sum of all returns over t of two variables, $\frac{1}{T}\sum_{t=1}^{T} L_t - \frac{1}{T}\sum_{t=1}^{T} W_t$ for each given value of z, which represent $J_2(z,\tau;\hat{F})$, $J_2(z,\lambda;\hat{F})$, respectively. Again, we will use one auxiliary binary variable F_t. According to the first group of inequalities, F_t equals 1 for each return t for which $z \ge \tau' Y_t$, and 0 otherwise. Analogously, the second and third groups of inequalities ensure that the variable L_t equals $z - \tau' Y_t$ for the scenarios for which the difference is positive, and 0 otherwise. The fourth and last inequalities ensure that W_t equals $z - \lambda' Y_t$ for the scenarios for which the difference is positive, and 0 otherwise. The fifth equation defines the sum of all weights to be unity.

For computational convenience, we reformulate the problem, following the same steps as for first-order stochastic dominance efficiency. Then, the model is transformed to the following linear program

$$
\begin{aligned}
&\min_{\lambda} \sum_{t=1}^{T} W_t \\
&s.t. \\
&W_t \ge r - \lambda' Y_t, \forall t \\
&e'\lambda = 1, \\
&W_t \ge 0, \forall t
\end{aligned}
\tag{22}
$$

A.3 Portfolio construction and numerical implementation

From a computational time cost perspective, it takes one hour to solve the FSDE model described in Section 2.2. This is a considerable amount of time given that to solve the model we discretize z and we solve the FSDE problem for each discrete value of z. We have 120 different z since we have 120 observations at each point in time (rolling window of 120 observations). Hence, we solve 120 mixed integer programming problems for each commodity index for each one of the 165 out-of-sample points in time. In contrast, the solution of the SSDE problem described in Section 2.3 is not computationally expensive since it takes less than one minute to solve it at each point in time. We solve the FSDE and SSDE problems with Gurobi solver on an iMac with 4*2.93 GHz Power, 16 GB of RAM. The Gurobi solver uses the branch and bound technique. We model the optimization problems by using GAMS (General Algebraic Modeling System).

Appendix B: Description of the commodity indices

First generation commodity indices

S&P GSCI was launched in January 1991 with historical data available since January 1970. The index currently invests in twenty four commodities classified into five groups (energy, precious metals, industrial metals, agricultural and livestock) and is heavily concentrated on the energy sector (almost 70% of the total index value). The S&P GSCI is a world-production weighted index based on the average quantity of production of each commodity in the index over the last five years of available data. DJ-UBSCI was launched in July 1998 with historical data beginning on January 1991. The index invests in nineteen commodities from the energy, precious metals, industrial metals, agricultural and livestock sectors. In contrast to the S&P GSCI, the DJ-UBSCI relies on two important rules to ensure diversification: the minimum and maximum allowable weight for any single commodity is 2% and 15%, respectively, and the maximum allowable weight for any sector is 33%. The DJ-UBSCI construction algorithm constructs weights by taking liquidity and (to a smaller extent) production into account. DBLCI was launched in 2003 with available price history since 1 December 1988. It tracks the performance of six commodities in the energy, precious metals, industrial metals and grain sectors. The chosen commodity futures contracts represent the most liquid contracts in their respective sectors. DBLCI has a constant weights scheme which reflects world production and inventory, thus providing a diverse and balanced commodity exposure (for a detailed description of the first generation indices, see also for instance, Geman, 2005, Erb and Harvey, 2006).

Second generation commodity indices

JPMCCI was launched in November 2007 with historical data available since December 1989. JPMCCI includes thirty three commodities and it uses commodity futures open interest to determine the inclusion and relative weights of the individual commodity futures. DBLCI-OY was launched in May 2006 with historical data available since December 1988. The index components and weighting scheme are identical with these of DBLCI. DBLCI-OY is designed to select the futures contracts that either maximises the positive roll yield in backwardated term structures or minimises the negative roll yield in contangoed markets from the list of tradable futures that expire in the next 13 months. The Morningstar Commodity indices were launched in 2007 with historical data beginning on 1980.

To be considered for inclusion in the Morningstar Commodity Index family, a commodity should have futures contracts traded in one of the U.S. exchanges and being ranked in the top 95% by the 12-month average of total dollar value of open interest. Morningstar indices are built based on a momentum strategy. The weight of each individual commodity index in each of the composite indices is the product of two factors: magnitude and the direction of the momentum signal. In brief, for each commodity, they calculate a "linked" price series that converts the price of the contract in effect at each point in time to a value that accounts for contract rolls. The Morningstar Long-Only index (MSDIL) is a fully collateralized commodity futures index that is long all twenty eligible commodities. The Morningstar Long-Flat index (MSDILF) is a fully collateralized commodity futures index that is long the commodities whose linked price exceeds its 12-month moving average. If the linked price is lower than its 12-month moving average, the weight of that commodity is moved into cash, i.e. it is zeroed (flat position).

Third generation commodity indices

Morningstar Long/Short index is a fully collateralized commodity futures index that uses the momentum rule to determine if each commodity is held long, short, or flat. At each point in time, if the linked price exceeds its 12-month moving average, it takes the long side in the subsequent month. Conversely, if the linked price is below its 12-month moving average, it takes the short side. An exception is made for commodities in the energy sector. If the signal for a commodity in the energy sector translates to taking a short position, the weight of that commodity is moved into cash. The Morningstar Short/Flat commodity index is a fully collateralized commodity futures index that is derived from the positions of the Long/Short index. It takes the same short positions as the Long/Short index and replaces the long positions with flat positions. The Morningstar Short-Only commodity index is a fully collateralized commodity futures index that is short in all eligible commodities.

References

Abanomey, W.S., Mathur, I., 1999. The hedging benefits of commodity futures in international portfolio construction. Journal of Alternative Investments 2, 51-62.

Abadie, A., 2002. Bootstrap tests for distributional treatment effects in instrumental variable models, Journal of the American Statistical Association 97, 284–292.

Ankrim, E.M., Hensel, C.R., 1993. Commodities in asset allocation: A real-asset alternative to real estate? Financial Analysts Journal 49, 20-29.

Anson, M., 1999. Maximizing utility with commodity futures diversification. Journal of Portfolio Management 25, 86-94.

Asness, C.S., Moskowitz, T., Pedersen, L.H., 2013. Value and momentum everywhere. Journal of Finance 68, 929-985.

Bakshi, G., Panayotov, G., Skoulakis, G., 2011. The Baltic Dry Index as a predictor of global stock returns, commodity returns, and global economic activity. In AFA 2012 Chicago Meetings Paper.

Bawa, V. 1975. Optimal rules for ordering uncertain prospects. Journal of Financial Economics 2, 95-121.

Belousova, J., Dorfleitner, G., 2012. On the diversification benefits of commodities from the perspective of euro investors. Journal of Banking and Finance 36, 2455-2472.

Bessembinder, H. , 1992. Systematic risk, hedging pressure, and risk premiums in futures markets. Review of Financial Studies 5, 637-667.

Bessembinder, H., Chan, K., 1992. Time-varying risk premia and forecastable returns in futures markets. Journal of Financial Economics 32, 169-193.

Bjornson, B., Carter, C. A., 1997. New evidence on agricultural commodity return performance under time-varying risk. American Journal of Agricultural Economics 79, 918-930.

Bodie, Z., Rosansky, V.I., 1980. Risk and return in commodity futures. Financial Analysts Journal 36, 27-39.

Büyükşahin, B., Haigh, M.S., Robe, M.A., 2010. Commodities and equities: Ever a "Market of one"? Journal of Alternative Investments 12, 76-95.

Büyükşahin, B., Robe, M.A. 2014. Speculators, commodities and cross-market linkages. Journal of International Money and Finance 42, 38-70.

Campbell, J. Y., Hamao, Y., 1992. Predictable stock returns in the United States and Japan: A study of long-term capital market integration. Journal of Finance 47, 43-69.

Cao, B., Jayasuriya, S., Shambora, W., 2010. Holding a commodity futures index fund in a globally diversified portfolio: A placebo effect? Economics Bulletin 30, 1842-1851.

Chan, K. F., Treepongkaruna, S., Brooks, R., Gray, S. 2011. Asset market linkages: Evidence from financial, commodity and real estate assets. Journal of Banking and Finance 35, 1415-1426.

Chen, S. S., 2007. Does monetary policy have asymmetric effects on stock returns? Journal of Money, Credit and Banking 39, 667-688.

Conover, C.M., Jensen, G.R., Johnson, R.R., Mercer, J.M., 2010. Is now the time to add commodities to your portfolio? Journal of Investing 19, 10-19.

Croft, H., Norrish, K., 2013. The Commodity refiner: From an age of shortage to an era of enough. Barclays Commodities Research Report April.

Dai, R., 2009. Commodities in dynamic asset allocation: Implications of mean reverting commodity prices. Working paper, Tilburg University.

Daskalaki, C., Skiadopoulos G., 2011. Should investors include commodities in their portfolios after all? New evidence. Journal of Banking and Finance 25, 2606–2626.

Delatte, A. L., Lopez, C., 2013. Commodity and equity markets: Some stylized facts from a copula approach. Journal of Banking and Finance 37, 5346-5356.

DeMiguel, V., Garlappi, L., Uppal, R., 2009. Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? Review of Financial Studies 22, 1915-1953.

De Roon, F. A., Nijman, T. E., Veld, C., 2000. Hedging pressure effects in futures markets. Journal of Finance 55, 1437-1456.

DeRoon, F.A., Nijman, T.E., 2001. Testing for mean-variance spanning: A survey. Journal of Empirical Finance 8, 111-155.

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., 1989. Business conditions and expected returns on stocks and bonds. Journal of Financial Economics 25, 23-49

Fama, E. F., 1990. Stock returns, expected returns, and real activity. Journal of Finance 45, 1089-1108.

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

Ferson, W. E., Harvey, C. R., 1991. The variation of economic risk premiums. Journal of Political Economy 99, 385-415.

Fortenbery, T.R., Hauser, R.J., 1990. Investment potential of agricultural futures contracts. American Journal of Agricultural Economics 72, 721-726.

Gargano, A., Timmermann, A., 2014. Forecasting commodity price indexes using macroeconomic and financial predictors. International Journal of Forecasting 30, 825-843.

Geman, H., 2005. Commodities and Commodity Derivatives: Modeling and Pricing for Agriculturals, Metals and Energy. John Wiley & Sons Ltd, Chichester.

Giamouridis, D, Sakkas, A., Tessaromatis, N., 2014. The role of commodities in strategic asset allocation, Working paper, Athens University of Economics and Business.

Giampietro, M., Guidolin, M., Pedio, M., 2016. Can no-arbitrage SDF models with regime shifts explain the correlations between commodity, stock, and bond returns?. Working paper BAFFI CAREFIN Centre, Bocconi University.

Gorton, G.B., Rouwenhorst, G.K., 2006. Facts and fantasies about commodity futures. Financial Analysts Journal 62, 47-68.

Gorton, G.B., Hayashi, F., Rouwenhorst, K.G., 2012. The fundamentals of commodity futures returns. Review of Finance 17, 35–105.

Hadar, J., Russell, W.R., 1969. Rules for ordering uncertain prospects. American Economic Review 59, 25–34.

Hanoch, G., Levy, H., 1969. The efficiency analysis of choices involving risk. Review of Economic Studies 36, 335–346.

Hodder, J.E., Jackwerth, J.C., Kolokolova, O., 2015. Improved portfolio choice using second-order stochastic dominance. Review of Finance 19, 1623-1647.

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.

Huberman, G., Kandel, S., 1987. Mean-variance spanning. Journal of Finance 42, 873-888.

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.

Jobson, J.D., Korkie, B.M., 1981. Performance hypothesis testing with the Sharpe and Treynor measures. Journal of Finance 36, 889-908.

Kat, H.M., Oomen, R.C., 2007. What every investor should know about commodities Part I: Univariate return analysis. Journal of Investment Management 5, 4-28.

Kostakis, A., Panigirtzoglou, N., Skiadopoulos, G., 2011. Market timing with option-implied distributions: A forward-looking approach. Management Science 57, 1231-1249.

Kuosmanen, T., 2004. Efficient diversification according to stochastic dominance criteria. Management Science 50, 1390-1406.

Levy, H., Hanoch, G., 1970. Relative effectiveness of efficiency criteria for portfolio selection. Journal of Financial and Quantitative Analysis 5, 63–76.

Levy, H., 1992, Stochastic dominance and expected utility: Survey and Analysis. Management Science 38, 555-593.

Markowitz, H., 1952. Portfolio selection. Journal of Finance 7, 77-91.

Memmel, C., 2003. Performance hypothesis testing with the Sharpe ratio. Finance Letters 1, 21-23.

Miffre, J., 2012. Comparing first, second and third generation commodity indices, Working paper, EDHEC Business School.

Post, T., 2003. Empirical tests for stochastic dominance efficiency. Journal of Finance 58, 1905-1031.

Post T, Versijp, P. 2007. Multivariate tests for stochastic dominance efficiency of a given portfolio. Journal of Financial and Quantitative Analysis 42, 489–515.

Post, T., Kopa, M., 2013. General linear formulations of stochastic dominance criteria. European Journal of Operational Research 230, 321–332.

Rallis, G., Miffre, J., Fuertes, A. M., 2013. Strategic and tactical roles of enhanced commodity indices. Journal of Futures Markets, 33, 965-992.

Scaillet, O., Topaloglou, N. 2010. Testing for stochastic dominance efficiency, Journal of Business and Economic Statistics 28, 169-180.

Silvennoinen, A., Thorp, S., 2013. Financialization, crisis and commodity correlation dynamics. Journal of International Financial Markets, Institutions and Money 24, 42-65.

Simaan, Y., 1993. What is the opportunity cost of mean-variance investment strategies? Management Science 39, 578-587.

Skiadopoulos, G., 2013. Advances in the commodity futures literature: A review. Journal of Derivatives 20, 85-96.

Tang, K., Xiong, W., 2012. Index investing and the financialization of commodities. Financial Analysts Journal 68, 54-74.

Figure 1. First-degree stochastic dominance efficiency of portfolio λ over the market portfolio τ.

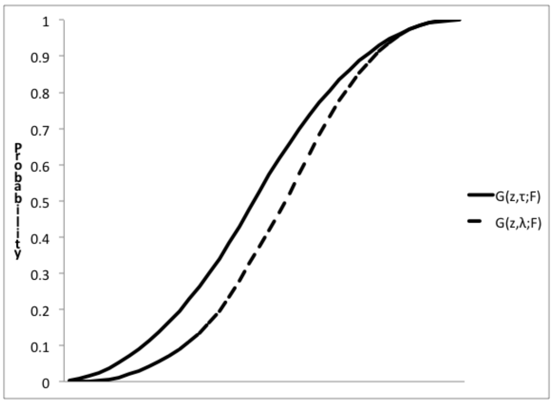

Figure 2. Second-degree stochastic dominance efficiency of portfolio λ over the market portfolio τ.

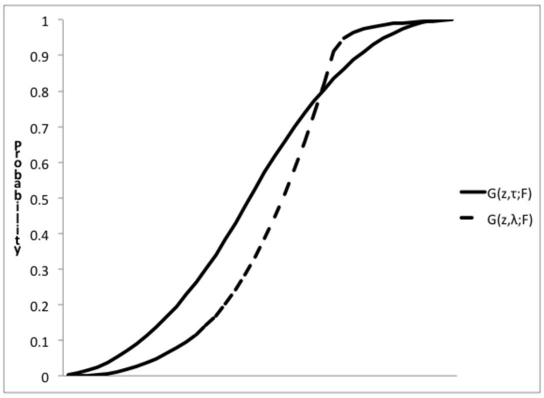

Table 1: Description of the Indices

Entries describe the set of stock, bond and commodity indices employed in this study.

	Constituents	Strategy
First Generation Indices		
Deutshe Bank Liquid Commodity Index (DBLCI)	6 commodities from the energy, precious metals, industrial metals and grain sectors.	The relative weight of each commodity remains constrant and is determined by world production and inventory data.
S&P Goldman Sachs Commodity Index (S&P GSCI)	24 commodities from the energy, precious metals, industrial metals, agricultural and livestock sectors.	The relative weight of each commodity is determined by world production data over the last five years. The index is heavily concentrated on the energy sector (about 70%).
Dow-Jones-UBS Commodity Index (DJ UBS CI)	19 commodities from the energy, precious metals, industrial metals, agricultural and livestock sectors.	The relative weight of each commodity is determined primarily by liquidity data, and to a smaller extent by production. The index relies on two rules: the minimum and the maximum allowable weight for any single commodity is 2% and 15%, respectively, and the maximum allowable for any sector is 33%.
Second Generation Indices		
JP Morgan Commodity Curve Index (JPMCCI)	33 commodities from the energy, precious metals, industrial metals, agricultural and livestock sectors.	The relative weight of each commodity is determined by its open interest. The index adopts a simple, curve-neutral approach, holding exposure along the commodity futures curve according to the open interest of each tenor.
DBLCI-Optimum Yield (DBLCI-OY)	6 commodities, the same as DBLCI.	The relative weight of each commodity is identical with DBLCI. The index is designed to select the futures contracts that either maximises the positive roll yield in backwardated term structures or minimises the negative roll yield in contangoed markets from the list of tradeable futures that expire in the next 13 months.
Morningstar Long-Only Commodity Index (MSDIL)	20 most liquid commodity futures contracts, traded in U.S. exchanges and ranked in the top 95% by the 12-month average of total dollar value of open interest.	The index is long all eligible commodities.
Morningstar Long/Flat Commodity Index (MSDILF)	20 most liquid commodity futures contracts, traded in U.S. exchanges and ranked in the top 95% by the 12-month average of total dollar value of open interest.	The relative weight of each commodity is the product of two factors: magnitude and the direction of the momentum signal. The index is long the commodities whose product exceeds its 12-month moving average. If it is lower than its 12-month moving average, the weight of that commodity is moved into cash (flat position).
Third Generation Indices		
Morningstar Short-Only Commodity Index (MSDIS)	20 most liquid commodity futures contracts, traded in U.S. exchanges and ranked in the top 95% by the 12-month average of total dollar value of open interest.	The index is short all eligible commodities.
Morningstar Long/Short Commodity Index (MSDILS)	20 most liquid commodity futures contracts, traded in U.S. exchanges and ranked in the top 95% by the 12-month average of total dollar value of open interest.	The relative weight of each commodity is the product of two factors: magnitude and the direction of the momentum signal. If this product exceeds its 12-month moving average, it takes the long side in the subsequent month. If it is below its 12-month moving average, it takes the short side.
Morningstar Short/Flat Commodity Index (MSDISF)	20 most liquid commodity futures contracts, traded in U.S. exchanges and ranked in the top 95% by the 12-month average of total dollar value of open interest.	The relative weight of each commodity is the product of two factors: magnitude and the direction of the momentum signal. If the product is below its 12-month moving average, the index takes the short side. If it exceeds its 12-month average, the weight of that commodity is moved into cash (flat position).
Equity Indexes		
S&P 500 Total Return	500 large companies having common stock listed on the NYSE or NASDAQ.	The index has traditionally been capitalization-weighted; that is, movements in the prices of stocks with higher market capitalizations (the share price times the number of shares outstanding) have a greater impact on the value of the index than do companies with smaller market caps.
Small minus Big Factor (SMB)	6 value-weighted portfolios of stocks listed on the NYSE, AMEX, or NASDAQ formed based on market capitalization.	The difference in the return of a portfolio of small capitalization stocks and the return of a portfolio of large capitalization stocks.
High minus Low Factor (HML)	6 value-weighted portfolios of stocks listed on the NYSE, AMEX, or NASDAQ formed based on book-to-market figures.	The difference between the return of a portfolio of high book-to-market stocks and the return of a portfolio of low book-to-market stocks.
Bond Index		
Barclays US Aggregate Bond Index	Most U.S. traded investment grade bonds are represented. It is a long-term index that includes Treasury securities, Government agency bonds, Mortgage-backed bonds, Corporate bonds, and a small amount of foreign bonds traded in U.S.	It is a market capitalization-weighted index, meaning the securities in the index are weighted according to the market size of each bond type.

Table 2: Descriptive Statistics

Entries report the descriptive statistics for the alternative asset classes used in this study. The annualized mean returns (with the respective *t*-statistics), the standard deviations, the Sharpe Ratios as well as the skewness and the kurtosis figures are reported. The dataset spans the period from January 1990 to September 2013, with the exception of DJ-UBSCI that covers the period from January 1991 to September 2013.

	Annualized Average Return	tstat	Annualized Standard Deviation	Annualized Sharpe Ratio	Skewness	Kurtosis
First Generation Indexes						
Deutsche Bank Liquid Commodity Index	0.097	*(2.265)*	0.208	0.292	0.208	5.644
S&P Goldman Sachs Commodity Index	0.063	*(1.426)*	0.213	0.124	-0.148	5.078
Dow-Jones-UBS Commodity Index	0.055	*(1.777)*	0.149	0.142	-0.562	5.666
Second Generation Indexes						
JP Morgan Commodity Curve Index	0.078	*(2.421)*	0.157	0.265	-0.575	5.985
DBLCI-Optimum Yield	0.104	*(2.876)*	0.175	0.385	-0.190	5.455
Morningstar Long/Flat Commodity Index	0.087	*(4.026)*	0.105	0.482	0.117	5.654
Morningstar Long-Only Commodity Index	0.087	*(2.712)*	0.157	0.325	-0.361	5.607
Third Generation Indexes						
Morningstar Short/Flat Commodity Index	0.039	*(3.325)*	0.057	0.049	0.573	7.204
Morningstar Long/Short Commodity Index	0.093	*(4.193)*	0.107	0.529	0.225	4.992
Morningstar Short-Only Commodity Index	-0.007	*(-0.248)*	0.149	-0.293	0.499	5.558
Equity Indexes						
S&P 500 Total Return	0.099	*(3.228)*	0.149	0.419	-0.601	4.137
SMB	0.026	*(1.108)*	0.115	-0.088	0.823	11.382
HML	0.048	*(2.146)*	0.109	0.111	0.619	5.755
Bond Index						
Barclays US Aggregate Bond Index	0.065	*(8.455)*	0.037	0.769	-0.270	3.739

Table 3: Correlation matrix

Entries report the correlation matrix of the assets under consideration. One and two asterisks indicate that the correlation coefficient is statistically significant from zero at 10% and 5%, respectively.

.

	First Generation Indices			*Second Generation Indices*				*Third Generation Indices*						
	DBLCI	**S&P GSCI**	**DJ UBS CI**	**JPMCCI**	**DBLCI-OY**	**MSDILF**	**MSDIL**	**MSDISF**	**MSDILS**	**MSDIS**	**S&P 500**	**Barclays Index**	**SMB**	**HML**
DBLCI	1.00	0.94**	0.88**	0.87**	0.96**	0.73**	0.88**	-0.39**	0.51**	-0.81**	0.15**	-0.03	0.12**	0.04
S&P GSCI	0.94**	1.00	0.89**	0.92**	0.92**	0.75**	0.91**	-0.37**	0.54**	-0.85**	0.18**	-0.01	0.10	0.09
DJ UBS CI	0.88**	0.89**	1.00	0.96**	0.89**	0.78**	0.97**	-0.55**	0.48**	-0.93**	0.31**	0.05	0.10	0.07
JPMCCI	0.87**	0.92**	0.96**	1.00	0.91**	0.78**	0.94**	-0.49**	0.51**	-0.91**	0.26**	-0.02	0.07	0.10
DBLCI-OY	0.96**	0.92**	0.89**	0.91**	1.00	0.73**	0.90**	-0.44**	0.48**	-0.85**	0.21**	-0.03	0.12**	0.06
MSDILF	0.73**	0.75**	0.78**	0.78**	0.73**	1.00	0.82**	-0.21**	0.87**	-0.72**	0.04	-0.04	0.06	0.05
MSDIL	0.88**	0.91**	0.97**	0.94**	0.90**	0.82**	1.00	-0.57**	0.50**	-0.95**	0.23**	0.00	0.07	0.06
MSDISF	-0.39**	-0.37**	-0.55**	-0.49**	-0.44**	-0.21**	-0.57**	1.00	0.30**	0.64**	-0.32**	-0.04	-0.03	0.04
MSDILS	0.51**	0.54**	0.48**	0.51**	0.48**	0.87**	0.50**	0.30**	1.00	-0.38**	-0.12	-0.06	0.05	0.07
MSDIS	-0.81**	-0.85**	-0.93**	-0.91**	-0.85**	-0.72**	-0.95**	0.65**	-0.38**	1.00	-0.23**	0.01	-0.07	-0.04
S&P 500	0.15**	0.18**	0.31**	0.26	0.21**	0.04	0.23**	-0.32**	-0.12**	-0.23**	1.00	0.12**	0.10	-0.19**
Barclays Index	-0.03	-0.01	0.05	-0.02	-0.03	-0.04	0.00	-0.04	-0.06	0.01	0.12**	1.00	-0.15*	0.09
SMB	0.12**	0.10	0.10	0.07	0.12**	0.06	0.07	-0.03	0.05	-0.07	0.10*	-0.15**	1.00	-0.06
HML	0.04	0.09	0.07	0.10	0.06	0.05	0.06	0.04	0.07	-0.04	-0.19**	0.09	-0.06	1.00

Table 4: In-sample testing for Stochastic Dominance Efficiency: Jan 1990 – Sep. 2013

Entries report the Scaillet and Topaloglou (2010) FSDE and SSDE test statistics and respective *p*-values (within parentheses) for the null hypothesis that a set of benchmark assets consisting of stocks, bonds and the risk-free asset are efficient versus the benchmark assets plus the commodity asset. We access investment in commodities via first, second and third generation indices, separately. Panel A reports the results when the benchmark set of assets consists of the S&P 500 Total Return Index, Barclays Aggregate Bond Index and 1-month LIBOR. Panels B and C report the results when the benchmark set includes dynamic equity indices, i.e. the Fama-Frenech (1993) size factor (Small minus Big, SMB) and the value factor (High minus Low, HML), respectively. Results are based on monthly observations from Jan. 1990 –Sep. 2013. One and two asterisks indicate that the null hypothesis of spanning can be rejected at 10% and 5% significance level, respectively.

Test Asset	**Panel A: S&P 500**		**Panel B: SMB**		**Panel C:HML**	
	FSD	**SSD**	**FSD**	**SSD**	**FSD**	**SSD**
First generation indices						
Deutsche Bank Liquid Commodity Index	0.087**	0.005*	0.049*	0.048*	0.088**	0.047**
	(0.049)	*(0.053)*	*(0.068)*	*(0.055)*	*(0.037)*	*(0.014)*
S&P Goldman Sachs Commodity Index	0.087*	0.004*	0.088	0.019	0.087*	0.019*
	(0.093)	*(0.098)*	*(0.104)*	*(0.115)*	*(0.087)*	*(0.085)*
Dow-Jones-UBS Commodity Index	0.006**	0.005*	0.087*	0.048*	0.060**	0.048**
	(0.044)	*(0.051)*	*(0.053)*	*(0.052)*	*(0.039)*	*(0.041)*
Second generation indices						
JP Morgan Commodity Curve Index	0.087*	0.0031*	0.087**	0.029*	0.035**	0.029**
	(0.065)	*(0.077)*	*(0.048)*	*(0.056)*	*(0.039)*	*(0.044)*
Deutsche Bank Liquid Commodity Index-Optimum Yield	0.087*	0.0117*	0.013*	0.014*	0.011*	0.0065**
	(0.083)	*(0.073)*	*(0.063)*	*(0.065)*	*(0.054)*	*(0.047)*
Morningstar Long/Flat Commodity Index	0.014**	0.015**	0.014**	0.017**	0.005**	0.014**
	(0.029)	*(0.014)*	*(0.035)*	*(0.045)*	*(0.027)*	*(0.023)*
Morningstar Long-Only Commodity Index	0.087**	0.044*	0.086**	0.016*	0.070**	0.018**
	(0.038)	*(0.053)*	*(0.048)*	*(0.065)*	*(0.044)*	*(0.049)*
Third generation indices						
Morningstar Short/Flat Commodity Index	0.084**	0.016**	0.084**	0.047*	0.070**	0.048**
	(0.017)	*(0.014)*	*(0.037)*	*(0.053)*	*(0.039)*	*(0.034)*
Morningstar Long/Short Commodity Index	0.021**	0.023**	0.009*	0.009*	0.071**	0.009**
	(0.022)	*(0.049)*	*(0.063)*	*(0.075)*	*(0.034)*	*(0.039)*
Morningstar Short-Only Commodity Index	0.084**	0.023**	0.087*	0.048*	0.071**	0.049*
	(0.041)	*(0.049)*	*(0.057)*	*(0.066)*	*(0.044)*	*(0.051)*

Table 5: In-sample testing for Stochastic Dominance Efficiency: Jan 1990 – Dec. 2000

Entries report the Scaillet and Topaloglou (2010) FSDE and SSDE test statistics and respective *p*-values (within parentheses) for the null hypothesis that a set of benchmark assets consisting of stocks, bonds and the risk-free asset are efficient versus the benchmark assets plus the commodity asset. We access investment in commodities via first, second and third generation indices, separately. Panel A reports the results when the benchmark set of assets consists of the S&P 500 Total Return Index, Barclays Aggregate Bond Index and1-month LIBOR. Panels B and C report the results when the benchmark set includes dynamic equity indices, i.e. the Fama-Frenech (1993) size factor (Small minus Big, SMB) and the value factor (High minus Low, HML), respectively. Results are based on monthly observations from Jan. 1990 –Dec 2000. One and two asterisks indicate that the null hypothesis of spanning can be rejected at 10% and 5% significance level, respectively.

Test Asset	**Panel A: S&P 500**		**Panel B: SMB**		**Panel C:HML**	
	FSD	**SSD**	**FSD**	**SSD**	**FSD**	**SSD**
First generation indices						
Deutsche Bank Liquid Commodity Index	0.043**	0.017*	0.094*	0.075*	0.051**	0.064**
	(0.044)	*(0.059)*	*(0.081)*	*(0.073)*	*(0.045)*	*(0.029)*
S&P Goldman Sachs Commodity Index	0.075*	0.033*	0.103*	0.084	0.093*	0.039*
	(0.086)	*(0.074)*	*(0.097)*	*(0.106)*	*(0.094)*	*(0.089)*
Dow-Jones-UBS Commodity Index	0.013*	0.009**	0.095**	0.082**	0.077**	0.055**
	(0.053)	*(0.047)*	*(0.045)*	*(0.039)*	*(0.046)*	*(0.023)*
Second generation indices						
JP Morgan Commodity Curve Index	0.075*	0.076*	0.074**	0.094**	0.123**	0.073**
	(0.086)	*(0.056)*	*(0.035)*	*(0.047)*	*(0.045)*	*(0.041)*
Deutsche Bank Liquid Commodity Index-Optimum Yield	0.093*	0.038*	0.046**	0.083*	0.034*	0.024**
	(0.057)	*(0.077)*	*(0.074)*	*(0.069)*	*(0.051)*	*(0.036)*
Morningstar Long/Flat Commodity Index	0.087**	0.027**	0.025**	0.035**	0.016**	0.040**
	(0.014)	*(0.026)*	*(0.044)*	*(0.046)*	*(0.034)*	*(0.030)*
Morningstar Long-Only Commodity Index	0.074**	0.088**	0.114**	0.074**	0.067**	0.035**
	(0.037)	*(0.049)*	*(0.027)*	*(0.040)*	*(0.030)*	*(0.044)*
Third generation indices						
Morningstar Short/Flat Commodity Index	0.118**	0.049**	0.074**	0.089**	0.087**	0.074**
	(0.038)	*(0.033)*	*(0.030)*	*(0.039)*	*(0.014)*	*(0.029)*
Morningstar Long/Short Commodity Index	0.084**	0.066**	0.016*	0.040*	0.103**	0.013*
	(0.041)	*(0.039)*	*(0.052)*	*(0.055)*	*(0.029)*	*(0.055)*
Morningstar Short-Only Commodity Index	0.012**	0.074**	0.073**	0.130*	0.106**	0.099**
	(0.029)	*(0.040)*	*(0.047)*	*(0.055)*	*(0.033)*	*(0.040)*

Table 6: Out-of-sample testing for Stochastic Dominance: January 2001 – September 2013

Entries report the Scaillet and Topaloglou (2010) first and second order stochastic dominance (FSD, SSD, respectively) test statistics and respective *p*-values (within parentheses) for the null hypothesis that the optimal portfolio based on the traditional asset universe (stocks, bonds and cash) augmented with commodities stochastically dominates the optimal portfolio based on the traditional asset universe. Optimal portfolios' performance is evaluated out-of-sample. We access investment in commodities via the first generation indices, the second generation indices and the third generation indices, separately. Panel A reports results when the traditional asset universe consists of the S&P 500 Total Return Index, Barclays Aggregate Bond Index and Libor 1-month. Panels B and C report results when the stock asset class in the traditional asset universe is proxied by dynamic equity indices, i.e. the Fama-French (1993) size factor (Small minus Big, SMB) and value factor (High minus Low, HML), respectively. Results are based on out-of-sample monthly portfolio returns spanning January 2001 –September 2013.

Test Asset	**Panel A: S&P 500**		**Panel B: SMB**		**Panel C:HML**	
	FSD	**SSD**	**FSD**	**SSD**	**FSD**	**SSD**
First generation indices						
Deutsche Bank Liquid Commodity Index	-0.005	-0.004	-0.008	-0.005	-0.010	-0.006
	(0.583)	*(0.530)*	*(0.483)*	*(0.560)*	*(0.563)*	*(0.573)*
S&P Goldman Sachs Commodity Index	-0.004	-0.003	-0.007	-0.002	-0.005	-0.005
	(0.553)	*(0.490)*	*(0.507)*	*(0.577)*	*(0.570)*	*(0.627)*
Dow-Jones-UBS Commodity Index	-0.0019	-0.002	-0.006	-0.003	-0.006	-0.004
	(0.530)	*(0.557)*	*(0.533)*	*(0.537)*	*(0.507)*	*(0.573)*
Second generation indices						
JP Morgan Commodity Curve Index	-0.007	-0.005	-0.009	-0.005	-0.009	-0.008
	(0.533)	*(0.517)*	*(0.560)*	*(0.527)*	*(0.543)*	*(0.567)*
Deutsche Bank Liquid Commodity Index-Optimum Yield	-0.009	-0.006	-0.009	-0.007	-0.010	-0.009
	(0.517)	*(0.557)*	*(0.543)*	*(0.550)*	*(0.590)*	*(0.563)*
Morningstar Long/Flat Commodity Index	-0.006	-0.004	-0.008	-0.005	-0.007	-0.006
	(0.533)	*(0.560)*	*(0.543)*	*(0.603)*	*(0.583)*	*(0.570)*
Morningstar Long-Only Commodity Index	-0.009	-0.006	-0.009	-0.006	-0.008	-0.008
	(0.487)	*(0.487)*	*(0.507)*	*(0.583)*	*(0.610)*	*(0.567)*
Third generation indices						
Morningstar Short/Flat Commodity Index	-0.003	0.001	-0.003	-0.002	-0.004	-0.002
	(0.527)	*(0.503)*	*(0.477)*	*(0.547)*	*(0.550)*	*(0.563)*
Morningstar Long/Short Commodity Index	-0.002	-0.002	-0.007	-0.004	-0.009	-0.005
	(0.510)	*(0.560)*	*(0.510)*	*(0.570)*	*(0.510)*	*(0.517)*
Morningstar Short-Only Commodity Index	-0.005	-0.002	-0.004	-0.003	-0.005	-0.004
	(0.513)	*(0.547)*	*(0.493)*	*(0.523)*	*(0.547)*	*(0.520)*

Table 7: Passive Equity Index and Commodity Investing

Entries report the performance measures (annualized Sharpe Ratio, annualized Opportunity Cost, Portfolio Turnover, annualized Return-Loss) when the benchmark set includes the S&P 500 Equity Index, the Barclays US Aggregate Bond Index and the 1-month LIBOR. The *p*-values of Memmel's (2003) test are also reported within parentheses; the null hypothesis is that the SR obtained from the traditional investment opportunity set is equal to that derived from the expanded set that includes commodities. The results for the opportunity cost are reported for different degrees of absolute risk aversion (*ARA*=2,4,6) and different degrees of relative risk aversion (RRA=2,4,6). Investors access investment in commodities separately via first generation indices, second generation indices and third generation indices, separately. The dataset for all assets spans the period from Jan. 1990-Sep. 2013 with the exception of DJ-UBSCI that covers the period from Jan. 1991 to Sep.2013 due to data availability constraints. The out-of-sample analysis is conducted over the period from Jan. 2001-Sep. 2013. Panels A and B report the results for the first and second order stochastic dominance efficiency, respectively.

Panel A: First Stochastic Dominance

		1st generation indexes			2nd generation indexes				3rd generation indexes		
	Benchmark set	**S&P GSCI**	**DJ-UBS CI**	**DBLCI**	**JPMCCI**	**DBLCI-OY**	**MSDILF**	**MSDIL**	**MSDISF**	**MSDILS**	**MSDIS**
Sharpe Ratio	0.21	0.35	0.34	0.42	0.74	0.79	0.77	0.81	0.60	0.43	0.90
Memmel p-value		*0.341*	*0.358*	*0.253*	*0.055*	*0.037*	*0.041*	*0.040*	*0.115*	*0.273*	*0.019*
PortfolioTurnover	8.20%	13.37%	3.84%	3.49%	4.95%	3.45%	2.98%	9.19%	6.41%	0.95%	3.15%
Return-Loss		0.71%	1.40%	1.96%	6.16%	6.63%	6.98%	6.83%	5.23%	2.95%	9.38%
Opportunity Cost											
Exponential Utility											
ARA=2		1.80%	1.68%	3.60%	7.92%	9.00%	7.08%	9.24%	4.08%	2.64%	6.48%
ARA=4		-0.60%	1.08%	1.44%	6.96%	7.80%	6.96%	8.16%	4.56%	2.64%	7.20%
ARA=6		-3.24%	0.36%	-0.84%	6.00%	6.60%	6.96%	7.08%	5.04%	2.76%	7.80%
Power Utility											
RRA=2		1.80%	1.68%	3.24%	7.92%	9.00%	7.08%	9.24%	4.08%	2.64%	6.48%
RRA=4		-0.72%	0.96%	1.08%	6.96%	7.80%	6.96%	8.16%	4.56%	2.76%	7.20%
RRA=6		-3.36%	0.24%	-1.32%	5.88%	6.48%	6.96%	7.08%	5.04%	2.76%	7.92%

Table 7: Passive Equity Index and Commodity Investing (cont'd)

Entries report the performance measures (annualized Sharpe Ratio, annualized Opportunity Cost, Portfolio Turnover, annualized Return-Loss) when the benchmark set includes the S&P 500 Equity Index, the Barclays US Aggregate Bond Index and the 1-month LIBOR. The *p*-values of Memmel's (2003) test are also reported within parentheses; the null hypothesis is that the SR obtained from the traditional investment opportunity set is equal to that derived from the expanded set that includes commodities. The results for the opportunity cost are reported for different degrees of absolute risk aversion (*ARA*=2,4,6) and different degrees of relative risk aversion (RRA=2,4,6). Investors access investment in commodities separately via first generation indices, second generation indices and third generation indices, separately. The dataset for all assets spans the period from Jan. 1990-Sep. 2013 with the exception of DJ-UBSCI that covers the period from Jan. 1991 to Sep.2013 due to data availability constraints. The out-of-sample analysis is conducted over the period from Jan. 2001-Sep. 2013. Panels A and B report the results for the first and second order stochastic dominance efficiency, respectively.

Panel B: Second Stochastic Dominance

		1st generation indexes			**2nd generation indexes**				**3rd generation indexes**		
	Benchmark set	**S&P GSCI**	**DJ-UBS CI**	**DBLCI**	**JPMCCI**	**DBLCI-OY**	**MSDILF**	**MSDIL**	**MSDISF**	**MSDILS**	**MSDIS**
Sharpe Ratio	0.21	0.36	0.32	0.39	0.49	0.53	0.58	0.58	0.27	0.43	0.63
Memmel p-value		*0.339*	*0.376*	*0.317*	*0.224*	*0.191*	*0.161*	*0.166*	*0.437*	*0.289*	*0.132*
PortfolioTurnover	7.32%	28.75%	13.82%	7.89%	5.56%	2.78%	4.03%	7.65%	15.32%	0.95%	5.42%
Return-Loss		0.40%	0.77%	1.19%	2.66%	2.99%	4.19%	3.54%	2.87%	2.67%	5.93%
Opportunity Cost											
Exponential Utility											
ARA=2		2.04%	1.44%	2.52%	4.32%	5.16%	4.44%	5.76%	-0.12%	2.64%	3.24%
ARA=4		0.24%	0.60%	-0.48%	2.40%	2.64%	4.20%	3.96%	1.08%	2.52%	3.96%
ARA=6		-1.68%	-0.24%	-3.72%	0.24%	-0.12%	4.08%	1.80%	2.16%	2.28%	4.68%
Power Utility											
RRA=2		1.92%	1.44%	2.40%	4.20%	5.04%	4.44%	5.76%	-0.12%	2.64%	3.24%
RRA=4		0.12%	0.60%	-0.72%	2.16%	2.52%	4.20%	3.84%	1.08%	2.52%	3.96%
RRA=6		-1.92%	-0.36%	-4.32%	-0.24%	-0.48%	4.08%	1.56%	2.16%	2.28%	4.68%

Table 8: The Value Dynamic Equity Factor and Commodity Investing

Entries report the performance measures (annualized Sharpe Ratio, annualized Opportunity Cost, Portfolio Turnover, annualized Return-Loss) when the benchmark set includes the Value Factor (HML), the Barclays US Aggregate Bond Index and the 1-month LIBOR. The *p*-values of Memmel's (2003) test are also reported within parentheses; the null hypothesis is that the SR obtained from the traditional investment opportunity set is equal to that derived from the expanded set that includes commodities. The results for the opportunity cost are reported for different degrees of absolute risk aversion (*ARA*=2,4,6) and different degrees of relative risk aversion (RRA=2,4,6). Investors access investment in commodities separately via first generation indices, second generation indices and third generation indices, separately. The dataset for all assets spans the period from Jan. 1990 to Sep. 2013 with the exception of DJ-UBSCI that covers the period from Jan. 1991 to Sep.2013 due to data availability constraints. The out-of-sample analysis is conducted over the period from Jan. 2001-Sep. 2013. Panels A and B report the results for the first and second order stochastic dominance efficiency, respectively.

Panel A: First Stochastic Dominance

		1st generation indexes			2nd generation indexes				3rd generation indexes		
	Benchmark set	**S&P GSCI**	**DJ-UBS CI**	**DBLCI**	**JPMCCI**	**DBLCI-OY**	**MSDILF**	**MSDIL**	**MSDISF**	**MSDILS**	**MSDIS**
Sharpe Ratio	0.05	0.76	0.49	0.77	0.76	0.87	0.75	0.67	0.56	0.87	0.53
Memmel p-value		*0.017*	*0.088*	*0.018*	*0.018*	*0.012*	*0.027*	*0.039*	*0.060*	*0.012*	*0.031*
PortfolioTurnover	3.88%	17.48%	10.57%	8.34%	16.17%	6.73%	8.29%	6.71%	14.46%	7.54%	11.54%
Return-Loss		8.04%	4.87%	8.32%	8.07%	9.73%	8.66%	7.19%	6.75%	9.95%	5.92%
Opportunity Cost											
Exponential Utility											
ARA=2		10.32%	6.24%	10.92%	10.20%	11.88%	8.64%	8.88%	5.40%	11.04%	5.88%
ARA=4		9.60%	5.52%	9.84%	9.60%	11.16%	8.76%	8.16%	6.12%	10.80%	6.12%
ARA=6		8.88%	4.80%	8.76%	8.88%	10.56%	9.00%	7.44%	6.96%	10.56%	6.48%
Power Utility											
RRA=2		10.32%	6.24%	11.04%	10.20%	11.88%	8.64%	8.88%	5.40%	11.04%	5.88%
RRA=4		9.60%	5.52%	9.84%	9.60%	11.16%	8.76%	8.16%	6.24%	10.80%	6.12%
RRA=6		8.88%	4.68%	8.76%	8.88%	10.56%	9.00%	7.44%	6.96%	10.56%	6.48%

Table 8: The Value Dynamic Equity Factor and Commodity Investing (cont'd)

Entries report the performance measures (annualized Sharpe Ratio, annualized Opportunity Cost, Portfolio Turnover, annualized Return-Loss) when the benchmark set includes the Value Factor (HML), the Barclays US Aggregate Bond Index and the 1-month LIBOR. The *p*-values of Memmel's (2003) test are also reported within parentheses; the null hypothesis is that the SR obtained from the traditional investment opportunity set is equal to that derived from the expanded set that includes commodities. The results for the opportunity cost are reported for different degrees of absolute risk aversion (*ARA*=2,4,6) and different degrees of relative risk aversion (RRA=2,4,6). Investors access investment in commodities separately via first generation indices, second generation indices and third generation indices, separately. The dataset for all assets spans the period from Jan. 1990 to Sep. 2013 with the exception of DJ-UBSCI that covers the period from Jan. 1991 to Sep.2013 due to data availability constraints. The out-of-sample analysis is conducted over the period from Jan. 2001-Sep. 2013. Panels A and B report the results for the first and second order stochastic dominance efficiency, respectively.

Panel B: Second Stochastic Dominance

		1st generation indexes			**2nd generation indexes**				**3rd generation indexes**		
	Benchmark set	**S&P GSCI**	**DJ-UBS CI**	**DBLCI**	**JPMCCI**	**DBLCI-OY**	**MSDILF**	**MSDIL**	**MSDISF**	**MSDILS**	**MSDIS**
Sharpe Ratio	0.04	0.51	0.42	0.41	0.65	0.65	0.62	0.61	0.42	0.57	0.56
Memmel p-value		*0.094*	*0.136*	*0.153*	*0.044*	*0.050*	*0.072*	*0.059*	*0.152*	*0.094*	*0.035*
PortfolioTurnover	4.88%	28.89%	14.10%	4.51%	18.12%	8.02%	1.97%	8.30%	23.63%	3.49%	61.64%
Return-Loss		4.52%	3.96%	3.67%	6.36%	6.38%	6.89%	6.16%	4.80%	6.23%	4.39%
Opportunity Cost											
Exponential Utility											
ARA=2		6.12%	4.92%	5.04%	8.76%	9.24%	7.08%	8.04%	3.72%	6.48%	5.64%
ARA=4		5.64%	4.32%	2.76%	7.56%	7.08%	6.96%	6.96%	4.56%	6.36%	6.00%
ARA=6		5.04%	3.84%	0.36%	6.48%	4.68%	6.72%	5.76%	5.40%	6.24%	6.36%
Power Utility											
RRA=2		6.12%	4.92%	4.92%	8.76%	9.24%	7.08%	8.04%	3.72%	6.48%	5.64%
RRA=4		5.64%	4.32%	2.64%	7.56%	6.96%	6.96%	6.84%	4.56%	6.36%	6.00%
RRA=6		5.04%	3.84%	0.00%	6.36%	4.20%	6.72%	5.76%	5.40%	6.24%	6.48%

Table 9: The Size Dynamic Equity Factor and Commodity Investing

Entries report the performance measures (annualized Sharpe Ratio, annualized Opportunity Cost, Portfolio Turnover, annualized Return-Loss) when the benchmark set includes the Size Factor (SMB), the Barclays US Aggregate Bond Index and the 1-month LIBOR. The *p*-values of Memmel's (2003) test are also reported within parentheses; the null hypothesis is that the SR obtained from the traditional investment opportunity set is equal to that derived from the expanded set that includes commodities. The results for the opportunity cost are reported for different degrees of absolute risk aversion (*ARA*=2,4,6) and different degrees of relative risk aversion (RRA=2,4,6). Investors access investment in commodities separately via first generation indices, second generation indices and third generation indices, separately. The dataset for all assets spans the period from Jan. 1990 to Sep. 2013 with the exception of DJ-UBSCI that covers the period from Jan. 1991 to Sep.2013 due to data availability constraints. The out-of-sample analysis is conducted over the period from Jan. 2001-Sep. 2013. Panels A and B report the results for the first and second order stochastic dominance efficiency, respectively.

Panel A: First Stochastic Dominance

		1st generation indexes			2nd generation indexes				3rd generation indexes		
	Benchmark set	**S&P GSCI**	**DJ-UBS CI**	**DBLCI**	**JPMCCI**	**DBLCI-OY**	**MSDILF**	**MSDIL**	**MSDISF**	**MSDILS**	**MSDIS**
Sharpe Ratio	0.16	0.61	0.61	0.65	0.81	0.88	0.87	0.84	0.56	0.77	0.79
Memmel p-value		*0.153*	*0.121*	*0.116*	*0.030*	*0.023*	*0.026*	*0.030*	*0.120*	*0.055*	*0.032*
PortfolioTurnover	4.66%	12.11%	15.87%	2.90%	10.97%	3.63%	4.27%	7.11%	9.43%	1.72%	2.23%
Return-Loss		3.88%	4.02%	4.63%	7.37%	8.34%	8.67%	7.84%	5.15%	7.47%	8.82%
Opportunity Cost											
Exponential Utility											
ARA=2		6.12%	5.52%	7.32%	9.60%	11.28%	9.00%	9.96%	4.32%	7.80%	5.76%
ARA=4		4.92%	5.04%	5.52%	8.76%	9.96%	8.88%	9.12%	4.80%	7.68%	6.48%
ARA=6		3.60%	4.44%	3.60%	7.92%	8.64%	8.88%	8.40%	5.28%	7.56%	7.32%
Power Utility											
RRA=2		6.12%	5.64%	7.32%	9.60%	11.28%	9.00%	9.96%	4.32%	7.92%	5.76%
RRA=4		4.92%	5.04%	5.52%	8.76%	9.96%	8.88%	9.12%	4.80%	7.68%	6.60%
RRA=6		3.60%	4.44%	3.60%	7.92%	8.64%	8.88%	8.40%	5.40%	7.56%	7.32%

Table 9: The Size Dynamic Equity Factor and Commodity Investing (cont'd)

Entries report the performance measures (annualized Sharpe Ratio, annualized Opportunity Cost, Portfolio Turnover, annualized Return-Loss) when the benchmark set includes the Size Factor (SMB), the Barclays US Aggregate Bond Index and the 1-month LIBOR. The *p*-values of Memmel's (2003) test are also reported within parentheses; the null hypothesis is that the SR obtained from the traditional investment opportunity set is equal to that derived from the expanded set that includes commodities. The results for the opportunity cost are reported for different degrees of absolute risk aversion (*ARA*=2,4,6) and different degrees of relative risk aversion (RRA=2,4,6). Investors access investment in commodities separately via first generation indices, second generation indices and third generation indices, separately. The dataset for all assets spans the period from Jan. 1990 to Sep. 2013 with the exception of DJ-UBSCI that covers the period from Jan. 1991 to Sep.2013 due to data availability constraints. The out-of-sample analysis is conducted over the period from Jan. 2001-Sep. 2013. Panels A and B report the results for the first and second order stochastic dominance efficiency, respectively.

Panel B: Second Stochastic Dominance

		1st generation indexes			**2nd generation indexes**				**3rd generation indexes**		
	Benchmark set	**S&P GSCI**	**DJ-UBS CI**	**DBLCI**	**JPMCCI**	**DBLCI-OY**	**MSDILF**	**MSDIL**	**MSDISF**	**MSDILS**	**MSDIS**
Sharpe Ratio	0.19	0.22	0.32	0.40	0.42	0.56	0.60	0.52	0.53	0.77	0.72
Memmel p-value		*0.465*	*0.363*	*0.289*	*0.269*	*0.166*	*0.166*	*0.194*	*0.194*	*0.055*	*0.083*
PortfolioTurnover	14.64%	30.79%	21.72%	4.39%	10.52%	0.22%	3.48%	5.15%	13.59%	1.72%	26.10%
Return-Loss		-0.45%	1.05%	2.25%	2.39%	4.26%	4.26%	3.82%	0.24%	7.47%	7.20%
Opportunity Cost											
Exponential Utility											
ARA=2		-0.48%	1.56%	3.12%	3.36%	6.00%	5.16%	5.04%	2.88%	7.80%	3.72%
ARA=4		-3.24%	0.24%	-0.12%	1.08%	3.36%	4.92%	3.12%	3.60%	7.68%	4.56%
ARA=6		-6.24%	-1.20%	-3.72%	-1.44%	0.48%	4.80%	1.08%	4.44%	7.56%	5.52%
Power Utility											
RRA=2		-0.48%	1.56%	3.00%	3.24%	6.00%	5.16%	5.04%	2.88%	7.92%	3.72%
RRA=4		-3.36%	0.12%	-0.36%	0.96%	3.24%	4.92%	3.00%	3.60%	7.68%	4.56%
RRA=6		-6.48%	-1.44%	-4.20%	-1.80%	0.12%	4.80%	0.84%	4.44%	7.56%	5.52%

Table 10: Forecasting assets' excess returns

Dependent variables for the regressions are the excess returns on commodity, equity and bond indices. The independent variables are the instrumental variables lagged one month: the dividend yield, yield on the three-month Treasury bills, the junk bond premium (or default spread) defined as the excess of the yield on long-term BAA corporate bonds rated by Moody's over the yield on AAA-rated bonds, the term spread defined as the difference between the Aaa yield and the one-month bill rate, the Industrial Production growth, the money supply growth, and the growth in the Baltic Dry Index. A constant term is included in all regressions. The coefficient estimates, the respective *t*-statistics, the Adjusted R^2, the *F*-statistic and the respective *p*-values are reported. The dataset spans 1999-2013. One and two asterisks indicate that the coefficient estimates are statistically significant different from zero at 10% and 5% significance level, respectively.

		Instrumental Variables										
Index		Intercept	T-bill	Dividend Yield	Junk bond yield	Term Spread	Money Growth	Industrial Production growth	Baltic Dry Index	Adj R squared	F-statistic	p-value
S&P GSCI		0.180	-0.008	-0.060	0.002	-0.008	-0.850	0.687	0.023	1.77%	1.449	0.189
	t-stat	*(1.574)*	*(-1.104)*	*(-1.495)*	*(0.128)*	*(-1.005)*	*(-1.559)*	*(0.815)*	*(0.962)*			
DJ -BS CI		0.057	-0.001	-0.022	0.005	-0.003	-0.266	0.710	0.012	-1.14%	0.719	0.656
	t-stat	*(0.680)*	*(-0.192)*	*(-0.762)*	*(0.429)*	*(-0.430)*	*(-0.663)*	*(1.143)*	*(0.696)*			
DBLCI		0.119	-0.003	-0.039	0.000	-0.006	-0.500	0.716	0.021	0.85%	1.213	0.298
	t-stat	*(1.112)*	*(-0.489)*	*(-1.032)*	*(0.022)*	*(-0.712)*	*(-0.979)*	*(0.906)*	*(0.955)*			
JPMCCI		0.088	-0.004	-0.025	0.005	-0.006	-0.555	0.668	0.019	0.72%	1.181	0.317
	t-stat	*(0.992)*	*(-0.700)*	*(-0.815)*	*(0.386)*	*(-0.874)*	*(-1.308)*	*(1.018)*	*(1.054)*			
DBLCI-OY		0.073	-0.002	-0.021	0.001	-0.004	-0.387	0.669	0.018	-0.28%	0.930	0.485
	t-stat	*(0.769)*	*(-0.297)*	*(-0.632)*	*(0.096)*	*(-0.559)*	*(-0.852)*	*(0.952)*	*(0.913)*			
MSDILF		0.026	0.001	-0.013	-0.001	0.001	-0.125	0.370	-0.017	-1.19%	0.707	0.666
	t-stat	*(0.424)*	*(0.338)*	*(-0.614)*	*(-0.145)*	*(0.248)*	*(-0.431)*	*(0.827)*	*(-1.371)*			
MSDILF		0.100	-0.003	-0.035	0.008	-0.005	-0.433	0.664	0.012	-0.59%	0.853	0.545
	t-stat	*(1.120)*	*(-0.594)*	*(-1.135)*	*(0.597)*	*(-0.737)*	*(-1.021)*	*(1.014)*	*(0.636)*			
MSDISF		0.008	0.001	-0.005	-0.007	0.001	0.088	-0.203	-0.002	-1.28%	0.687	0.683
	t-stat	*(0.259)*	*(0.285)*	*(-0.452)*	*(-1.523)*	*(0.637)*	*(0.634)*	*(-0.944)*	*(-0.308)*			
MSDILS		0.030	0.002	-0.016	-0.008	0.003	-0.034	0.163	-0.019	0.08%	1.020	0.419
	t-stat	*(0.493)*	*(0.510)*	*(-0.751)*	*(-0.898)*	*(0.566)*	*(-0.119)*	*(0.366)*	*(-1.546)*			
MSDIS		-0.061	0.004	0.018	-0.005	0.003	0.532	-0.647	-0.011	-0.98%	0.758	0.624
	t-stat	*(-0.725)*	*(0.649)*	*(0.599)*	*(-0.426)*	*(0.553)*	*(1.322)*	*(-1.040)*	*(-0.623)*			
S&P 500		0.137*	-0.012**	-0.032	0.008	-0.012**	-0.217	0.119	0.031**	3.24%	1.833	0.084
	t-stat	*(1.825)*	*(-2.542)*	*(-1.213)*	*(0.749)*	*(-2.139)*	*(-0.606)*	*(0.215)*	*(1.997)*			
Barclays Bond Index		-0.042**	0.003**	0.010*	-0.002	0.004**	0.123	0.103	-0.001	0.069	2.845	0.008
	t-stat	*(-2.471)*	*(2.773)*	*(1.692)*	*(-0.668)*	*(3.372)*	*(1.541)*	*(0.834)*	*(-0.321)*			