PHBS WORKING PAPER SERIES

Gold is a Hedging Asset

Aoxiang Yang
Peking University

June 2024

Working Paper 20240605

Abstract

While industry investors commonly view gold as a hedging asset, academic studies often find the opposite. We show that gold is a prominent hedging asset via three different approaches: a state-space model, predictive regressions, and principal component analyses. We find that, ceteris paribus, gold prices increase with expected stock market return μ_t and expected dividend growth rate g_t. In bad times, μ_t rises while g_t declines. It thus may seem that gold prices fall in bad times and that gold prices insignificantly or even negatively predict stock returns. However, after addressing the omitted-variable-bias introduced by g_t, we find that gold prices significantly and positively predict stock returns.

JEL Classification: G0, G1, G2

Peking University HSBC Business School
University Town, Nanshan District
Shenzhen 518055, China

Gold is a Hedging Asset

Aoxiang Yang[†]

May 3, 2024

Abstract

While industry investors commonly view gold as a hedging asset, academic studies often find the opposite. We show that gold is a prominent hedging asset via three different approaches: a state-space model, predictive regressions, and principal component analyses. We find that, ceteris paribus, gold prices increase with expected stock market return μ_t and expected dividend growth rate g_t. In bad times, μ_t rises while g_t declines. It thus may seem that gold prices fall in bad times and that gold prices insignificantly or even negatively predict stock returns. However, after addressing the omitted-variable-bias introduced by g_t, we find that gold prices significantly and positively predict stock returns.

[†]Aoxiang Yang is with Peking University. Email: aoxiang.yang@phbs.pku.edu.cn. I thank Bjorn Eraker, Ivan Shaliastovich, Shen Zhao (discussant), and participants at the 5th PHBS-CUHKSZ Workshop in Economics and Finance for helpful feedback.

1 Introduction

Most industry investors believe that gold is a hedging asset, meaning two things. First, gold prices should rise in bad times (i.e., when expected stock market return goes up - when risk-aversion goes up as in habits model (Campbell and Cochrane (1999)) or when uncertainty goes up as in long-run risk (Bansal and Yaron (2004)) and time-varying disaster-risk (Wachter (2013)) models). Second, therefore, gold prices should positively predict stock market returns. However, academic studies often find the opposite. For instance, Huang and Kilic (2019) find that gold is a risky asset whose prices fall in higher-uncertainty times; Hou, Tang, and Zhang (2020) show that gold prices fall during political uncertainties; Erb and Harvey (2013) fail to find a relation between gold prices and future stock market returns.

We show that gold is indeed a prominent hedging asset, after accounting for an "omitted-variable-bias (OVB)"or "error-in-variable" associated with expected economic fundamental. We reach this conclusion via three different approaches: a state-space model, predictive regressions, and principal component analysis.

Consider first a parsimonious state-space model. To infer underlying states, we aggregate information in the world's two most important commodities - gold and oil.[1] Specifically, we estimate a model and find (roughly) the following coefficients:

$$\log G_t = const + \frac{1}{40}\mu_t + \frac{1}{4}g_t + \epsilon_t^G \tag{1.1}$$

$$\log O_t = const - \frac{1}{40}\mu_t + \frac{1}{4}g_t + \epsilon_t^O \tag{1.2}$$

$$r_{t+1} = \mu_t + \varepsilon_{r,t+1} \tag{1.3}$$

$$\Delta d_{t+1} = g_t + \varepsilon_{d,t+1} \tag{1.4}$$

[1]Gold and crude oil have a total market capitalization of respectively 14 and 3 trillion USD. CBOE computes and releases implied volatility index for gold and crude oil since around 2010.

Here $\log G_t$ and $\log O_t$ represent log real gold and oil prices detrended by log real dividends. Cointegration tests suggest that the detrended commodity prices are stationary processes. We model detrended commodity prices as linear functions in underlying states μ_t (expected stock market return) and g_t (expected stock market dividend growth) (both annualized in percentage), plus arbitrary pricing errors, which turn out to be small. We assume that both states follow an AR(1) process with correlated shocks. We estimate the model via maximum likelihood, using Kalman Filter to construct the likelihood, as in Van Binsbergen and Koijen (2010). Given our relatively short sample 1975-2022, we obtain deseasoned dividend data from Robert Shiller's website and estimate the model monthly. The objective is to leverage observed gold prices, oil prices, dividend growth rates, and stock returns collectively as observations to infer underlying states and explore the relation between commodity prices and these states.

Our estimation indicates that a 1% increase in expected stock return leads to a 2.5% increase (decrease) in gold (oil) price. Furthermore, a 1% rise in expected dividend growth rates is associated with a substantial 25% increase in both gold and oil prices. These patterns are statistically significant and align logically: gold and oil, serving as essential industrial production inputs, see their prices influenced positively by anticipated dividend growth. However, as investment assets, commodities are different in nature: gold is identified as a hedging asset, while oil a risky asset. Consequently, we observe distinct price behaviors: gold prices rise with expected stock market returns, whereas oil prices react inversely.

Our punchline is that gold prices indeed strongly rise with expected stock return, but only ceteris paribus, i.e., only after we correct the OVB introduced by expected economic fundamentals. The OVB helps explain many of the literature results. Suppose we want to study the relation between gold price and expected stock returns. To this end, we directly run future stock returns onto current gold price without controls (as in the literature):

$$r_{t+1} = const + \beta_1 \log G_t + \underbrace{\tilde{\varepsilon}_{t+1}} \tag{1.5}$$

$$= const + \beta_1(\frac{1}{40}\mu_t + \frac{1}{4}g_t) \underbrace{-\beta_1 \frac{1}{4} g_t + \varepsilon_{r,t+1}}. \tag{1.6}$$

Then a g_t loading would slip into OLS error $\tilde{\varepsilon}_{t+1}$, which would be correlated with the regressor $\log G_t$, leading to biased coefficients on the regressor $\log G_t$. To give a more intuitive explanation of the OVB, note that μ_t and g_t shocks are negatively correlated. In bad times, μ_t rises (i.e., future returns r_{t+1} tend to rise) and g_t falls. If the impact of g_t is strong enough, gold prices may fall rather than rise at such times (Huang and Kilic (2019)'s finding). It thus may appear that gold prices negatively or not predict future stock returns (Erb and Harvey (2013)'s finding), despite that the true predictive coefficient is positive $\beta_1 = 40$.

We show that the magnitude of the OVB is determined by a variance decomposition of gold prices into expected dividend growth effect, expected return effect, and their covariance. We find that the OVB is almost 100%, implying that observing any coefficient when regressing future returns on gold prices is not too surprising. When we control for several g_t proxies on the RHS, the unbiased coefficient $\beta_1 = 40$ is indeed recovered.

The OVB is a lesser concern for risky assets like crude oil. In this case, the influences of μ_t and g_t align in the same direction, leading to a clear drop in oil prices during economic downturns.[2] Consequently, oil prices significantly negatively forecast stock returns, consistent with established literature findings (Chen, Roll, and Ross (1986), Jones and Kaul (1996), Driesprong, Jacobsen, and Maat (2008), Huang, Masulis, and Stoll (1996), Sim and Zhou (2015), Ready (2018), Christoffersen and Pan (2018)).

To support our state-space model estimates, we further conduct OLS predictive regres-

[2]This is also consistent with the following fact: annualized realized volatility of returns on holding gold and oil are respectively 14% and 36%. The two forces μ_t, g_t tend to drive gold (oil) prices in the opposite (same) direction, making oil prices much more volatile. Out state-space model reproduces these statistics.

sions in the following way. We take the difference between the two log commodity prices ($logG/O$ hereafter) to cancel g_t loadings and reinforce μ_t loadings, thus obtaining a return predictor. We take the sum of the two log commodity prices ($logGO$ hereafter) to cancel μ_t loadings and reinforce g_t loadings, thus obtaining a dividend growth rate predictor. We confirm these predictive relations in the data, with OLS coefficients strongly supporting state-space model estimates.

Specifically, we show that in the 1975-2022 sample, a one standard deviation increase in $logG/O$ respectively significantly predicts a 6.6%, 7.3%, 7.2%, and 5.2% annualized increase in U.S. stock market excess returns over the following trading day, week, month, and year. Predictability is not consumed by any known predictors in the literature. On the contrary, in pairwise bivariate predictive regression tests, $logG/O$ consumes all the other known predictors including the gold-platinum price ratio of Huang and Kilic (2019), except Baker and Wurgler (2006) sentiment index. During the same period, a one standard deviation increase in $logGO$ respectively significantly predicts a 1.6%, 1.6%, and 1.5% annualized increase in U.S. dividend growth rates over the following month, quarter, and year.[3] The predictability holds in-sample and out-of-sample.

If, as our analyses suggest, $logG/O$ and $logGO$ are proxies respectively for expected return (countercyclical) and expected dividend growth (procyclical), then they should respectively carry negative and positive risk premiums in the cross-section of stock returns. We show this is indeed the case using Fama-French size/book-to-market 100 portfolios as test assets. Our results are robust under univariate and bivariate joint tests, under Fama-Macbeth and panel regressions.

Our state-space model studies the relation between commodity prices and stock market statistics μ_t, g_t. The model is not without restrictions. Recognizing this fact, we estimate an extended state-space model that incorporates constraints imposed by the price-

[3] If we only use the post-1986 sample, then the dividend growth rate predictability exists from a 1-month to a 5-year horizon. There seems to be a structural break around 1986.

dividend ratio (PD). Building on existing literature suggesting that PD is mainly driven by long-term discount rate news (Cochrane (2008), Cochrane (2011)), we introduce a slow-moving discount rate process, denoted θ_t, in addition to the already exisiting relatively fast-moving discount rate, μ_t, and expected dividend growth, g_t. Commodity prices are linear in the three states, along with pricing errors. Each of the processes follows an AR(1) with correlated shocks. We then use the Campbell-Shiller identity to derive a precise relation between the log PD ratio and the three underlying states, as in Van Binsbergen and Koijen (2010), and estimate model parameters under such restrictions. An additional benefit of this framework is that it allows us to study how commodity prices respond to discount rate news at different frequencies.

We filter a short-term discount rate, a long-term discount rate, and an expected dividend growth rate from the model. Consistent with our own and literature findings, we find that $logG/O$ is very informative about the short-term discount rate, log PD ratio is very informative about the long-term discount rate, and $logGO$ is most informative about the expected dividend growth. We show that the filtered processes serve as robust predictors for short-run returns, long-run returns, and dividend growth rates, respectively.

We confirm our main findings: ceteris paribus, gold prices strongly rise with short-term discount rate and expected dividend growth; ceteris paribus, oil prices strongly fall with short-term discount rate and rise with expected dividend growth. However, we do not find a clear relation between gold or oil prices and the long-term discount rate. Additionally, we find that there seems to be a structural break around 1986, with all observed relations strengthening statistically post-1986.

Finally, to generate richer evidence, we perform principal component analysis (PCA) on a large cross-section of commodity prices including gold, crude oil, silver, copper, and platinum. We find that the first PC, which accounts for 72% of all commodity price variations, can be largely interpreted as expected dividend growth. The second PC, which accounts for another 8%, is basically short-horizon expected stock return. Commodities' loadings on the PCs are consistent with our previous findings. The fifth PC, essentially a long-term expected return measure, only accounts for 2% of the variation, implying that commodity prices are not particularly responsive to long-term discount rates. The third and fourth PCs, which likely capture commodity-market-specific information, are elusive to interpret.

1.1 Literature

The literature on oil prices and stock markets is large. Chen, Roll, and Ross (1986) find that oil price risk is not separately rewarded in the stock market beyond traditional risk measures. Jones and Kaul (1996) find that the reaction of U.S. stock prices to oil shocks can be completely accounted for by expected cash-flow effects, not discount rates. Driesprong, Jacobsen, and Maat (2008) find that changes in oil prices strongly predict future stock market returns in many countries. Other studies include Huang, Masulis, and Stoll (1996), Sim and Zhou (2015), Ready (2018), Christoffersen and Pan (2018), and Gao, Hitzemann, Shaliastovich, and Xu (2022), etc. The typical literature finding is that oil price negatively predicts stock returns, and positively predicts cash flow growth, consistent with our findings. However, we are the first to show that oil prices predict stock returns at even a one-day horizon, after we use gold prices to cancel its cash flow growth loadings.

The literature on gold prices and stock markets is smaller. The nature of gold as an investment has been a debate. Jermann (2021) shows that gold price is negatively driven by real interest rate. Barro and Misra (2016) and Baur and Smales (2020) argue that gold is a hedge against uncertainty. On the contrary, Huang and Kilic (2019) argue that gold is a risky asset whose prices fall in uncertain times. Hou, Tang, and Zhang (2020) show that gold prices fall during political uncertainties. Erb and Harvey (2013) fail to find a relation between gold prices and stock market returns. Cheng, Tang, and Yan (2021) study hedging pressure in commodity option markets. Bakshi, Gao, and Zhang (2023) develop a model to price gold futures and options.

We reconcile divergent views on gold by emphasizing the necessity of considering economic fundamentals when evaluating the nature of gold and commodity assets more broadly. Failure to account for this factor can introduce significant error-in-variable bias. Our state-space model and PCA both indicate that expected economic fundamentals play a substantial role in influencing commodity price fluctuations.

Our paper is also related to the return and cash flow growth predictability literature, a key empirical issue and a ground to differentiate asset pricing models. Empirical studies include Cochrane (2008), Cochrane (2011), and Van Binsbergen and Koijen (2010), etc. Theories include Campbell and Cochrane (1999), Bansal and Yaron (2004), and Wachter (2013), etc. Our paper echoes Van Binsbergen and Koijen (2010), which find that both returns and cash flow growth are predictable, though not necessarily by the price-dividend ratio. They find that dividend growth rates are predictable by moving average terms of past dividend growth rates. We show that commodity prices also contain information that predicts dividend growth rates.

2 Cointegration Between log Commodity Prices and log Dividends

We obtain the longest monthly gold and WTI crude oil spot price data from Jan 1975 to Dec 2022 from Macrotrends.[4] We obtain monthly dividend and price-dividend ratio data from Robert Shiller's website,[5] and monthly stock market return data from Ken French's website. We aim to conduct return and dividend growth predictive regressions, and need

[4]The U.S. administration fully deregulated crude oil prices in 1981. The U.S. abandoned the gold standard in 1971.

[5]It is known that dividends have strong seasonality. As Professor Robert Shiller puts it, his monthly dividend data are computed from the S&P four-quarter totals for the quarter since 1926, with linear interpolation to monthly figures. I am not aware of any other acceptable way to deseason the original dividend data.

evidence that predictors are stationary. Neither gold nor oil prices are stationary processes.[6] We search for cointegration relations between log commodity prices and log dividends in three steps.

First, Table 1 performs Johansen (1988) rank tests. The main message is that we cannot reject the null that there is no cointegration relation between log commodity prices themselves. But we can reject the null that there is no cointegration relation between log commodity prices and log dividends. Second, we then search for cointegration relations as guided by the Johansen (1988) rank test using a Stock and Watson (1993) dynamic least square (DLS) estimation. Table 2 shows the estimated cointegration relations are[7]:

$$\log G_t \equiv \log G_t - 0.62 \log D_t$$

$$\log O_t \equiv \log O_t - 0.16 \log D_t$$

$$\log G_t/O_t \equiv \log G_t - \log O_t - 0.46 \log D_t$$

$$\log G_t \cdot O_t \equiv \log G_t + \log O_t - 0.78 \log D_t.$$

We will call these detrended commodity prices. In all that follows, when we use commodity prices, we always refer to detrended commodity prices. Third, we then confirm whether those coefficients truly imply cointegration using a one-sided augmented Dickey and Fuller (1979) unit root test. Table 3 shows that we can reject that $\log O_t$, $\log G_t/O_t$, and $\log G_t \cdot O_t$ contain a unit root at the 5% confidence level. We can reject that $\log G_t$ contains a unit root at the 10% confidence level.

[6] In fact, any asset that investors can hold for a long time cannot have stationary prices. Otherwise, one can immediately create an arbitrage strategy by buying low and selling high using limit orders. Gold and oil ETFs allow investors to hold for a long time.

[7] When constructing the regressors, we use real gold price, real oil price, and real dividends. Oil and gold prices are conceivably related to inflation. We want to get rid of the influence of inflation.

3 Gold-Oil Price Ratio as a Robust Short-Horizon Return Predictor

Table 3 summarizes our predictor $\log G/O$ and many other return predictors in the literature. Table 4 shows that $\log G/O$ significantly positively predicts returns at all horizons within five years. The R-squared peaks at roughly 1-year horizon. Tables 5, 6, and 7 report univariate overlapping return predictive regressions results for various predictors. Tables 8, 9, 10, and 11 report bivariate overlapping return predictive regression results where we pair $\log G/O$ separately with each known predictor in the literature.[8] The main finding is that $\log G/O$ can consume the predictive power of many other predictors at short horizons, such as the gold-platinum price ratio of Huang and Kilic (2019) and the implied cost of capital (ICC) of Li, Ng, and Swaminathan (2013). It also survives short-horizon predictors such as the VRP and other known medium- and long-horizon predictors.

We find that the only known predictors not consumed by $\log G/O$ at short horizons are the sentiment indexes of Baker and Wurgler (2006) and Huang, Jiang, Tu, and Zhou (2015). We find $\log G/O$ and the BW index are not significantly correlated. Rather, they have independent predictive power. We construct a three-factor model with $\log G/O$, BW sentiment index, and risk-neutral skewness. Table 13 shows that the 3-factor model has strong explanatory power for future returns. The R-squared reaches an amazing 50% at a 1-year horizon, and all three statistically significantly forecast returns at such horizons.

Table 12 shows that $\log G/O$ is strongly correlated with the gold-platinum price ratio of Huang and Kilic (2019), the sentiment index of Huang, Jiang, Tu, and Zhou (2015), and the implied cost of capital (ICC) of Li, Ng, and Swaminathan (2013). Moreover, $\log G/O$ also positively correlates with the VRP and risk-neutral higher-order moments.

Table 14 shows that, in the 1986-2022 sample[9], $\log G/O$ significantly predicts returns

[8] We add a time fixed effect on the RHS when regressing returns onto PD because PD has an upward trend during our sample 1975-2022.

[9] Daily oil price data was available only since 1986.

even at one-day to two-week horizons. We find $\log G/O$ together with the BW sentiment index stand out as the only two variables that can statistically significantly predict returns at a 1-day horizon. Overall, the evidence in this section suggests that $\log G/O$ is a very strong discount rate proxy primarily at relatively short horizons, e.g., less than one year, although we are uncertain whether this short-term discount rate is driven by investors' risk preferences (sentiments) or risk perceptions (tail risks). And it does not carry too much predictive power beyond that horizon. The next section analyzes why this is the case.

4 Regression Evidence: Expected Returns vs. Expected Dividend Growth

This section provides regression-based evidence of how $\log G/O$ is a discount rate proxy. Suppose we have an equilibrium model that prices gold and oil as investment assets. Under reasonable cointegration assumptions and log-linearization, the model will imply the following two pricing equations:

$$\log G_t - c^G \log D_t = a^G + \beta_\mu^G \mu_t + \beta_g^G g_t + \epsilon_t^G \tag{4.1}$$

$$\log O_t - c^O \log D_t = a^O + \beta_\mu^O \mu_t + \beta_g^O g_t + \epsilon_t^O, \tag{4.2}$$

where G_t and O_t are gold and oil prices, D_t is the dividend of the aggregate stock market, $\mu_t \equiv E_t[r_{t+1}^M]$ is the expected return of the aggregate stock market, $g_t \equiv E_t[\Delta d_{t+1}]$ is the expected dividend growth of the aggregate stock market, and ϵ_t^i is a stationary process that captures the portion of the commodity price left unexplained, where i=Gold or Oil.[10] In Section 2, we've estimated that $c^G = 0.62$ and $c^O = 0.16$. In all that follows, for simplicity, we always use $\log G_t$ and $\log O_t$ to denote log prices already detrended by log dividends. For now, assume the pricing errors on the RHS are relatively unimportant and have small variance.

[10] One can think about μ_t and g_t as proxies for underlying economic states, which drive the pricing of commodity assets in equilibrium.

Our goal is to determine the values of the coefficients β_μ^i, β_g^i, keeping in mind that the coefficients reflect investors' preferences for the commodities, the commodities' roles in producing dividends, or the persistence of the processes μ_t and g_t themselves. In the remainder of this section, we will show regression-based evidence that supports the following parameterization (where μ_t and g_t are annualized in percentage):

$$\log G_t = const + \frac{1}{40}\mu_t + \frac{1}{4}g_t + \epsilon_t^G \tag{4.3}$$

$$\log O_t = const - \frac{1}{40}\mu_t + \frac{1}{4}g_t + \epsilon_t^O, \tag{4.4}$$

for ϵ_t^G and ϵ_t^O with small-variance. We now present evidence. First, we take the difference of the two equations to get:

$$\mu_t = const + 20\log G_t/O_t + \epsilon_t, \tag{4.5}$$

for some small-variance ϵ_t, which implies that if we run future returns (over a short period, say, less than 1-year) onto $\log G/O$, the slope coefficients should be roughly equal to 20 and should be statistically significant. Table 4 confirms this. On the contrary, if we run dividend growth onto the same regressor, the coefficient should be insignificant. And R-squared should be small. Table 16 confirms this. We then take the sum of the two equations to get:

$$g_t = const + 2\log G_t \cdot O_t + \epsilon_t, \tag{4.6}$$

for some small-variance ϵ_t, which implies that if we run future dividend growth (over a short period, say, less than 1-year) onto $\log G \cdot O$, the slope coefficients should be roughly equal to 2 and should be statistically significant. Table 16 confirms this. On the contrary, if we run returns onto the same regressor, then the coefficient should be small and insignificant. R-squared should also be small. Table 4 confirms this.

Second, if we run future returns (over a short period, say, less than 1-year) onto $\log G$

and $\log O$ simultaneously, then the least-square implies that slope coefficients on the two should be respectively equal to 20 and -20 and should be both statistically significant, because such coefficients not only cancel the g_t term but also match the μ_t term (see equation (4.5)). Table 15 confirms this. If we run future dividend growth (over a short period, say, less than 1-year) onto $\log G$ and $\log O$ simultaneously, the slope coefficients on the two should be respectively equal to 2 and 2 and should be both statistically significant (see equation (4.6)). Table 19 shows this is the case, though not precisely.

Third, using equations (4.3) and (4.4), we can derive other regression models. For instance, we can derive:

$$\mu_t = const + 40 \log G_t - 20 \log G_t \cdot O_t + \epsilon_t \tag{4.7}$$

$$\mu_t = const - 40 \log O_t + 20 \log G_t \cdot O_t + \epsilon_t, \tag{4.8}$$

for some small-variance ϵ_t. These are the correctly specified regression models to study how gold and oil price respectively predicts returns. That is, we need to control for expected cash flow growth g_t, as proxied by $\log G \cdot O$, in return predictive regressions. Table 20 shows the regression results, which again support our conjecture. Note that the above three conjectures should hold irrespective of how μ_t and g_t are correlated. And they did hold in the data.

Suppose we do not control for g_t and run future returns onto $\log G_t$ or $\log O_t$ directly. In that case, "omitted variable bias" or "attenuation bias" will typically bias slope coefficients toward zero, leading to insignificance. The bias depends on the correlation between μ_t and g_t. To formally study the bias, we estimate a state-space model in the next section.

Equations (4.3) and (4.4) are our main insights, which show that each of the log gold and oil prices contains a component that reflects cash flow expectation and a component that reflects discount rate. Gold and oil are valuable production inputs; thus, it is unsurprising that both prices rise with expected dividend growth. As investment assets, however, they are different. As conventional wisdom suggests, gold is a hedging asset (i.e., investors' preference for gold increases when discount rate goes up), while oil is a risky asset (i.e., investors' preference for oil decreases when discount rate goes up). As a result, log gold price rises while log oil price falls with expected return. When taking the difference, cash flow components cancel while discount rate components strengthen, leaving $\log G/O$ as an unambiguously positive return predictor (See Figure 2 bottom panel for a graphical illustration). When taking the sum, discount rate components cancel while cash flow components strengthen, leaving $\log GO$ as an unambiguously positive cash flow growth predictor (See Figure 2 top panel for a graphical illustration).

Our evidence so far is based on deseasoned monthly dividend data obtained from Robert Shiller's website. Table 18 performs dividend growth predictive regressions using quarterly dividend data directly obtained from CRSP. We see that $\log GO$ continues to predict dividend growth rates. Thus, it's unlikely that the deseasoning procedure introduces fake dividend growth predictability. The very low R-squared is not surprising, given LHS variables contain strong seasonality. Figure 3 repeats the exercise in Figure 2. We see that the patterns observed under monthly data continue to hold quarterly. We work with monthly data mainly because commodity prices contain substantial variation (information) at the monthly frequency that will get lost under lower frequency.

4.1 Out-of-Sample Predictability

We have been looking only at in-sample evidence. To assess out-of-sample predictability, we follow Welch and Goyal (2008) and compute OOS R-squared for return predictability by $\log G/O$ as

$$R^2_{OOS,R} = 1 - \frac{\sum_{t=0}^{T-1}(r_{t+1} - \hat{\mu}_t)^2}{\sum_{t=0}^{T-1}(r_{t+1} - \bar{r}_t)^2}, \tag{4.9}$$

where $\hat{\mu}_t$ is the filtered value of the expected return using data only up until time t to estimate OLS model parameters. The denominator $\bar{r}_t$ is the historical mean of returns up until time t. Similarly, we compute OOS R-squared for dividend growth predictability by $\log GO$ as

$$R^2_{OOS,Div} = 1 - \frac{\sum_{t=0}^{T-1}(\Delta d_{t+1} - \hat{g}_t)^2}{\sum_{t=0}^{T-1}(\Delta d_{t+1} - \bar{\Delta} d_t)^2}. \tag{4.10}$$

We start our OOS computations at the beginning of 2000. Using the data between Jan 1975 and Dec 1999 to compute the parameters, we compute the expected return (dividend growth) for Jan 2000. We compare this prediction with realized return (dividend growth). We then use the data between Jan 1975 and Jan 2000 to compute the parameters and then compute predictions for Feb 2000. We proceed in this way up until Dec 2022. For daily return predictability, we repeat this rolling-window process on a daily basis, with the first trading day of 2000 as the breakpoint.

We find that the OOS R-squared is 3.04% for monthly return predictability, 5.00% for monthly dividend growth rate predictability, and 0.025% for daily return predictability.

4.2 Cross-Sectional Evidence

If, as we analyze above, $\log G/O$ and $\log G \cdot O$ are proxies respectively for the discount rate and expected dividend growth, then they should be priced respectively negatively and positively in the cross-section of stock returns. Table 21 shows that this is the case using Fama-French size/book-to-market 100 portfolios as test assets. Panel A and Panel B respectively estimate the market price of risk for $\Delta \log G/O$ and $\Delta \log G \cdot O$ using Fama-Macbeth regressions and panel regressions, where we use AR(1) innovations as unexpected shocks to the factors. The results are highly consistent. Whether estimating $\lambda_{\Delta \log G/O}$ and $\lambda_{\Delta \log G \cdot O}$ separately or jointly, we always find that $\lambda_{\Delta \log G/O} < 0$ and $\lambda_{\Delta \log G \cdot O} > 0$.

Figure 4 left panel depicts the realized sample-time-series-average return to each portfolio against the portfolio's exposure to $\Delta \log G/O$ estimated using the full sample. As shown, portfolios' exposures to $\Delta \log G/O$ are mostly negative, as they should be if $\Delta \log G/O$ is negatively priced (or is countercyclical). Since portfolios largely have negative exposures to $\Delta \log G/O$ and positive average returns, a fitted line that is forced to pass through the origin must have a negative slope (i.e., $\lambda_{\Delta \log G/O} < 0$). The right panel depicts the realized average return to each portfolio against the expected return predicted by the one-factor ($\Delta \log G/O$) model without an intercept.

Figure 5 shows the same thing for $\Delta \log G \cdot O$. As shown, portfolios' exposures to $\Delta \log G \cdot O$ are mostly positive, as they should be if $\Delta \log G \cdot O$ is positively priced (or is procyclical). Since portfolios largely have positive exposures to $\Delta \log G/O$ and positive average returns, a fitted line that is forced to pass through the origin must have a positive slope (i.e., $\lambda_{\Delta \log G \cdot O} > 0$). The right panel depicts the realized average return to each portfolio against the expected return predicted by the one-factor ($\Delta \log G \cdot O$) model without an intercept. In both cases, we don't expect portfolios to line up along the 45-degree line extremely well because we're considering non-traded factors. However, our results do confirm that the factors are priced with an expected sign.

5 A State-Space Model

The analysis in the previous section is based on predictive regression coefficients at short horizons. However, we haven't been able to identify the dynamics that μ_t and g_t follow and their correlation. We accomplish this goal by estimating a state-space model.

5.1 The Model

Consider the model

$$
\begin{aligned}
\log G_t &= \alpha^G + \beta^G_\mu(\mu_t - \bar{\mu}) + \beta^G_g(g_t - \bar{g}) + \varepsilon_{G,t} \\
\log O_t &= \alpha^O + \beta^O_\mu(\mu_t - \bar{\mu}) + \beta^O_g(g_t - \bar{g}) + \varepsilon_{O,t} \\
r_t &= \mu_{t-1} + \varepsilon_{r,t} \\
\Delta d_t &= g_{t-1} + \varepsilon_{d,t} \\
\mu_t &= (1-\rho_\mu)\bar{\mu} + \rho_\mu \mu_{t-1} + \varepsilon_{\mu,t} \\
g_t &= (1-\rho_g)\bar{g} + \rho_g g_{t-1} + \varepsilon_{g,t}
\end{aligned}
$$

with

$$
Cov\left(\begin{bmatrix}\varepsilon_{G,t}\\ \varepsilon_{O,t}\\ \varepsilon_{r,t}\\ \varepsilon_{d,t}\end{bmatrix}\right) = \begin{bmatrix}\sigma^2_G & & & \\ & \sigma^2_O & & \\ & & \sigma^2_r & \\ & & & \sigma^2_d\end{bmatrix}
$$

$$
Cov\left(\begin{bmatrix}\varepsilon_{\mu,t}\\ \varepsilon_{g,t}\end{bmatrix}\right) = \begin{bmatrix}\sigma^2_\mu & \sigma_{\mu,g}\\ \sigma_{\mu,g} & \sigma^2_g\end{bmatrix}
$$

The system has four observed variables $(\log G_t, \log O_t, r_t, \Delta d_t)$ and two latent state variables (μ_t, g_t). The first four equations are measurement equations. The first two equations follow from equations (4.1) and (4.2) except that, for simplicity, we assume that the two shocks $\varepsilon_{G,t}$ and $\varepsilon_{O,t}$ are i.i.d. instead of being any persistent stationary processes. The third and fourth equations are the definitions of μ_t and g_t. For simplicity, we assume that the covariance matrix for the innovations to observations is diagonal. The last two equations are the transition equations of the latent state variables μ_t and g_t. We follow Van Binsbergen and Koijen (2010) and assume that each state variable follows an AR(1) process. We allow for an arbitrary covariance structure for the shocks to the AR(1) processes. All innovations follow normal distributions.

We estimate the model via maximum likelihood, using the Kalman Filter to construct the likelihood, like in Van Binsbergen and Koijen (2010).[11] We obtain deseasoned dividend data from Robert Shiller's website and estimate the model under a monthly frequency to sharpen parameter identification, given our relatively short sample. Because dividends contain strong seasonality, Van Binsbergen and Koijen (2010) estimate an annual model. However, commodity prices move substantially at a relatively high frequency. If we use annual dividend data instead, we will lose most of the information contained in commodity prices.

5.2 Estimation Results

The estimation results are shown below. The parentheses contain standard errors, that are derived by inverting the estimated information matrix.

$$\log G_t = const + \underset{(0.016)}{0.036}(\mu_t - \underset{(4.2)}{12.5}) + \underset{(0.068)}{0.228}(g_t - \underset{(1.5)}{6.3}) + \varepsilon_{G,t} \tag{5.1}$$

$$\log O_t = const - \underset{(0.015)}{0.022}(\mu_t - \underset{(4.2)}{12.5}) + \underset{(0.063)}{0.244}(g_t - \underset{(1.5)}{6.3}) + \varepsilon_{O,t} \tag{5.2}$$

$$r_t = \mu_{t-1} + \varepsilon_{r,t} \tag{5.3}$$

$$\Delta d_t = g_{t-1} + \varepsilon_{d,t} \tag{5.4}$$

$$\mu_t = (1 - \underset{(0.010)}{0.972})\underset{(4.2)}{12.5} + \underset{(0.010)}{0.972}\mu_{t-1} + \varepsilon_{\mu,t} \tag{5.5}$$

$$g_t = (1 - \underset{(0.005)}{0.991})\underset{(1.52)}{6.3} + \underset{(0.005)}{0.991}g_{t-1} + \varepsilon_{g,t} \tag{5.6}$$

[11]See Appendix A of Van Binsbergen and Koijen (2010) for how to use Kalman Filter to recursively update latent states and then write down the likelihood.

with

$$Cov\left(\begin{bmatrix}\varepsilon_{G,t}\\ \varepsilon_{O,t}\\ \varepsilon_{r,t}\\ \varepsilon_{d,t}\end{bmatrix}\right)=\begin{bmatrix}2\times10^{-31} & & & \\ & 7\times10^{-55} & & \\ & & \underset{(11.2)^2}{(54.5)^2} & \\ & & & \underset{(1.3)^2}{(6.9)^2}\end{bmatrix} \tag{5.7}$$

$$Cov\left(\begin{bmatrix}\varepsilon_{\mu,t}\\ \varepsilon_{g,t}\end{bmatrix}\right)=\begin{bmatrix}\underset{(2.2)}{3.0} & \underset{(0.22)}{-0.32}\\ \underset{(0.22)}{-0.32} & \underset{(0.05)}{0.07}\end{bmatrix}. \tag{5.8}$$

The above equations further imply the following two equations (standard errors are derived using the Delta method):

$$\log G_t/O_t = const + \underset{(0.021)}{0.058}(\mu_t - \underset{(4.2)}{12.5}) - \underset{(0.068)}{0.016}(g_t - \underset{(1.5)}{6.3}) + \varepsilon_{G/O,t} \tag{5.9}$$

$$\log G_t\cdot O_t = const - \underset{(0.024)}{0.014}(\mu_t - \underset{(4.2)}{12.5}) + \underset{(0.112)}{0.472}(g_t - \underset{(1.5)}{6.3}) + \varepsilon_{G\cdot O,t} \tag{5.10}$$

As seen, the pricing errors σ_G and σ_O are negligibly small, although we didn't assume it. Thus, given model parameters, the two observations $\log G_t$ and $\log O_t$ almost help pin down the two underlying states μ_t and g_t. Note, however, that this does not mean that return and dividend data are uninformative about the underlying states because they help determine model parameters. The essence of our state-space approach is indeed to search for the optimal 2×2 coefficient matrix that links $(\log G_t, \log O_t)$ and (μ_t, g_t) that makes the observed return and dividend growth data most likely.

When interpreting the estimation results, note that r_t, Δd_t, μ_t, and g_t are all annualized in percent. There are several key observations. First, the values of the estimated coefficients $(\beta^G_\mu, \beta^G_g, \beta^O_\mu, \beta^O_g)$ are similar to those in equations (4.3) and (4.4) both qualitatively and quantitatively. Our previous conjecture (1/40,1/4,-1/40,1/4) lies well within one standard error of the four parameter estimates. At the same time, all four parameter estimates are significantly different from zero (except β^O_μ, which is marginally so).

Second, the discount rate process has an AC(1) of 0.972 monthly. We can strongly reject the null that μ_t is nonstationary. The persistence is low relative to the typical findings in the literature. For example, Van Binsbergen and Koijen (2010) estimate this value at 0.994 monthly (or 0.932 annually) in a similar state-space model using price-dividend ratio and dividend growth as observations. Avdis and Wachter (2017) use the price-dividend ratio to proxy for the discount rate, thus estimating this value at 0.994 monthly (in our sample). Wachter (2013) uses a value of 0.993 monthly. Bansal and Yaron (2004) uses a value of 0.987 monthly. Figure 6 upper panel plots the filtered μ_t process, which moves closely with the log gold-oil price ratio.

Third, the expected dividend growth process has an AC(1) of 0.991 monthly. We can reject the null that g_t is nonstationary marginally. The persistence is very high relative to the typical findings in the literature, even higher than that in the long-run risks model of Bansal and Yaron (2004), 0.979 monthly. Figure 6 lower panel plots the filtered g_t process, moving closely with the sum of log gold and oil prices.

We then use corresponding filtered states to predict returns and dividend growth in the data. Table 23 shows that our filtered states μ_t and g_t are good predictors of returns and dividend growth, respectively. Moreover, we cannot reject that the slope coefficients are equal to one at various horizons, as expected.

5.3 Omitted Variable Bias

The state-space model allows us to highlight how the omitted variable bias (which in special cases reduces to the attenuation bias) affects slope coefficient estimates in return and cash flow growth predictability studies. We first present a proposition.

PROPOSITION 1 *Suppose the price of asset i follows:*

$$\log P_t^i = \alpha^i + \beta_\mu^i \mu_t + \beta_g^i g_t. \tag{5.11}$$

Then if we run $r_{t+1} = \mu_t + \varepsilon_{r,t+1}$ onto $\log P_t^i$ in a univariate OLS regression, the slope coefficient is

$$\hat{\beta}_{OLS} = \frac{1}{\beta_\mu^i} \frac{Var(\beta_\mu^i \mu_t) + Cov(\beta_\mu^i \mu_t, \beta_g^i g_t)}{Var(\log P_t^i)}. \tag{5.12}$$

Equivalently, the "omitted variable bias" (i.e., OVB) is $\frac{Var(\beta_g^i g_t) + Cov(\beta_\mu^i \mu_t, \beta_g^i g_t)}{Var(\log P_t^i)}$. When $Cov(\mu_t, g_t) = 0$, the OVB reduces to the "attenuation bias":

$$\hat{\beta}_{OLS} = \frac{1}{\beta_\mu^i} \frac{Var(\beta_\mu^i \mu_t)}{Var(\beta_\mu^i \mu_t) + Var(\beta_g^i g_t)}. \tag{5.13}$$

Similar expressions can be obtained for OLS slope coefficients from running $\Delta d_{t+1} = g_t + \varepsilon_{d,t+1}$ onto $\log P_t^i$.

Proof. The true (unbiased) model is

$$\begin{aligned} r_{t+1} &= \mu_t + \varepsilon_{r,t+1} \\ &= -\frac{\alpha^i}{\beta_\mu^i} + \frac{1}{\beta_\mu^i} \log P_t^i - \frac{\beta_g^i}{\beta_\mu^i} g_t + \varepsilon_{r,t+1} \\ &= -\frac{\alpha^i}{\beta_\mu^i} + \beta_x x_t + \beta_y y_t + \varepsilon_{r,t+1} \end{aligned}$$

If we ignore the y_t term and run a univariate OLS regression of r_{t+1} on x_t (the biased model), standard formula for OVB implies that the OLS slope coefficient is

$$\hat{\beta}_{OLS} = \beta_x + \beta_y \delta, \tag{5.14}$$

where

$$\beta_x = \frac{1}{\beta_\mu^i}, \quad \beta_y = -\frac{\beta_g^i}{\beta_\mu^i}$$

and δ is the OLS slope from running y_t onto x_t:

$$\delta = \frac{Cov(x_t, y_t)}{Var(x_t)} = \frac{Cov(\log P_t^i, g_t)}{Var(\log P_t^i)} \tag{5.15}$$

Substituting (5.11) into (5.15), and then back to (5.14), one can easily obtain (5.12).[12] ■

Proposition 1 shows that OVB in OLS predictive regressions depends on asset i's loadings on μ_t and g_t and the relative variance of and the covariance between μ_t and g_t. We therefore first show a variance decomposition result. Table 22 Panel A decomposes the variance of log gold and oil prices respectively into three components: $Var(\beta^i_\mu \mu_t)$, $Var(\beta^i_g g_t)$, and $2Cov(\beta^i_\mu \mu_t, \beta^i_g g_t)$.

Our results show that more than 100% of the gold price variation is driven by expected dividend growth news. This implies if we run returns onto lagged gold price, the OVB will be huge. Oil price is also driven mostly by expected dividend growth, implying the OVB is also very large. Because log gold and oil prices load quite similarly on expected dividend growth (See Equations (5.1) and (5.2)), when we take the difference between them, cash flow components cancel (See Equation (5.9)). Thus, $\log G/O$ is a clean discount rate proxy, the variation of which is driven 91% by discount rate news. Expected dividend growth news only explains 1% of $\log G/O$. This is why it predicts returns robustly - it is clean and OVB is small. Also because log gold and oil prices load quite similarly on discount rate but with opposite signs (See Equations (5.1) and (5.2)), when we take the sum of them, discount rate components cancel (See Equation 5.10). Thus, $\log G \cdot O$ is a clean expected dividend growth proxy, the variation of which is driven 113% by cash flow news. Discount rate news only explains 1% of $\log G \cdot O$. This is why it predicts dividend growth so robustly.

[12] In this paper, we focus on the effect of OVB on the level of OLS slope estimates. The effect of OVB on the standard error of OLS slope estimator is known to be ambiguous. Specifically, standard error of $\hat{\beta}_x$ in the unbiased model is

$$s.e(\hat{\beta}_x) = \sqrt{\frac{\sigma_r^2}{\sum_{t=1}^{T}(x_t - \bar{x}_t)(1 - R^2_{x,y})}}, \tag{5.16}$$

where $\sigma_r^2 = Var(\varepsilon_{r,t+1})$ and $R^2_{x,y}$ is the R-squared if one regresses y_t onto x_t. There is a trade-off in comparing standard errors across unbiased and biased models. First, since the unbiased model is better able to explain the variation in r_{t+1}, σ_r^2 is smaller, which decreases $s.e.(\hat{\beta}_x)$ in the unbiased model relative to the biased model. Second, if x_t and y_t are correlated, then $R^2_{x,y} > 0$ in the unbiased model, which increases $s.e.(\hat{\beta}_x)$ in the unbiased model relative to the biased model. When we compare across tables, we didn't find that OVB significantly affects standard errors.

Table 22 Panel B then decomposes the variance of realized returns/dividend growth into the part due to expected returns/dividend growth news and unexpected shocks. At the monthly frequency, discount rate news only explains 1.9% of realized return variance. This number represents an upper bound on the R-squared from regressing one-month ahead returns onto discount rate proxies. That is, even if the regressor is an accurate measure of μ_t, the R-squared is 1.9% at most. At the monthly frequency, expected dividend growth news explains 8.4% of realized dividend growth variance. This number represents an upper bound on the R-squared from regressing one-month ahead dividend growth rates onto expected dividend growth proxies.

5.4 Predictive Regressions under Model-Simulated Data

To support the above arguments, we then simulate data from the model and perform similar OLS predictive regressions as in Section 4. Table 24 shows the results for return regressions. As seen, if we naively use gold or oil price as the regressor, then the slope coefficient will be biased toward 0. The bias is particularly severe in the case of gold. For gold, the OVB is $\frac{Var(\beta_g^G g_t)+Cov(\beta_\mu^G \mu_t, \beta_g^G g_t)}{Var(\log P_t^G)} = (1.54 - 1.01/2)/1 \approx 100\%$. The OVB is almost 100% - it's completely normal that one finds a weak or even negative relation between the gold price and future stock market returns, but, ceteris paribus, gold price indeed significantly rises with the expected return after one controls for expected cash flow growth. For oil, OVB becomes smaller ($(0.68 + 0.26/2)/1 \approx 81\%$) but is still substantial, which is again enough to lead to statistically insignificant estimates. Though not precisely, these are close to what we observe in the data (See Table 4).

Table 25 shows dividend growth predictive regression results. As seen, if we naively use log gold or oil price as the regressor, then there will again be OVB on slope coefficients. However, the biases are not large enough to lead to statistically insignificant estimates. OVB is $(0.48-1.01/2)/1 \approx 0\%$ for gold and $(0.07+0.26/2)/1 \approx 20\%$ for oil. The coefficients are again close to what we observe in the data (See Table 16). In a nutshell, if we compare across tables, all the data OLS slope coefficients are well within one standard error from their model counterparts at relatively short horizons. And vice versa.

5.5 Relation to the Literature

Huang and Kilic (2019) argue that gold is a risky investment asset whose price falls in bad times. Hou, Tang, and Zhang (2020) find that gold price falls during political uncertainty. Erb and Harvey (2013) fail to find a significant relation between gold price and future stock market returns. All these papers conclude that their findings cast doubt on conventional wisdom that gold is a hedging asset. We show that gold is indeed a hedging asset after controlling for the expected cash flow growth.

Huang and Kilic (2019) base the assumption that gold is a risky asset on an observation that gold price falls during recessions. This observation however is too casual. Figure 1 shows that gold price does not fall in every recession. For instance, it falls in prolonged recessions such as 2008, but not COVID-19. We argue that this contrast can be explained by cash flow and discount rate effects. Discount rate goes up a lot in both crises. Cash flow expectations fall a lot in 2008 but much less in COVID-19. Thus, cash flow effect dominates in 2008 while discount rate effect dominates in COVID-19. As a result, gold prices fall in 2008 but rise in COVID-19. In short, the conclusions in Huang and Kilic (2019), Hou, Tang, and Zhang (2020), and Erb and Harvey (2013) are one way or another affected by the omitted variable bias. Our marginal contribution is to formalize this omitted variable bias.

Our findings are consistent with Baur and Smales (2020), which find that gold price rises during geopolitical uncertainty, times during which discount rate goes up. But Baur and Smales (2020) didn't proceed to analyze then why gold price does not significantly positively predict stock returns in the data.

6 Disciplining the Model with the Price-Dividend Ratio

In Section 5, we didn't consider the restrictions on model parameters imposed by the Campbell and Shiller (1988) identity and the price-dividend ratio. As shown in Van Binsbergen and Koijen (2010), Campbell-Shiller present value calculation implies that, given the AR(1) dynamics for μ_t and g_t, the log price-dividend ratio is a linear function in μ_t and g_t. Thus, we do not have two free state variables once we observe the price-dividend ratio. In the next, we consider a model disciplined by the price-dividend ratio and show that not only our main findings hold well, but there are new insights.

6.1 The Model

Given the literature evidence that PD is primarily driven by long-term discount rate news (Cochrane (2011), Van Binsbergen and Koijen (2010)), we introduce another discount rate process component θ_t, on top of μ_t, g_t. The idea is that different linear combinations of μ_t and θ_t can help generate fast-moving and slow-moving discount rates that predict returns at corresponding horizons. Consider an extension of the model in Section 5.

$$\log G_t = \alpha^G + \beta_\mu^G(\mu_t - \bar{\mu}) + \beta_\theta^G \theta_t + \beta_g^G(g_t - \bar{g}) + \varepsilon_{G,t} \tag{6.1}$$

$$\log O_t = \alpha^O + \beta_\mu^O(\mu_t - \bar{\mu}) + \beta_\theta^O \theta_t + \beta_g^O(g_t - \bar{g}) + \varepsilon_{O,t} \tag{6.2}$$

$$r_t = \mu_{t-1} + \theta_{t-1} + \varepsilon_{r,t} \tag{6.3}$$

$$\Delta d_t = g_{t-1} + \varepsilon_{d,t} \tag{6.4}$$

$$\mu_t = (1-\rho_\mu)\bar{\mu} + \rho_\mu \mu_{t-1} + \varepsilon_{\mu,t} \tag{6.5}$$

$$\theta_t = \rho_\theta \theta_{t-1} + \varepsilon_{\theta,t} \tag{6.6}$$

$$g_t = (1-\rho_g)\bar{g} + \rho_g g_{t-1} + \varepsilon_{g,t} \tag{6.7}$$

with

$$Cov\left(\begin{bmatrix}\varepsilon_{G,t}\\ \varepsilon_{O,t}\\ \varepsilon_{d,t}\end{bmatrix}\right)=\begin{bmatrix}\sigma_G^2 & & \\ & \sigma_O^2 & \\ & & \sigma_d^2\end{bmatrix} \tag{6.8}$$

$$Cov\left(\begin{bmatrix}\varepsilon_{\mu,t}\\ \varepsilon_{\theta,t}\\ \varepsilon_{g,t}\end{bmatrix}\right)=\begin{bmatrix}\sigma_\mu^2 & \rho_{\mu,\theta}\sigma_\mu\sigma_\theta & \rho_{\mu,g}\sigma_\mu\sigma_g\\ \rho_{\mu,\theta}\sigma_\mu\sigma_\theta & \sigma_\theta^2 & \rho_{\theta,g}\sigma_\theta\sigma_g\\ \rho_{\mu,g}\sigma_\mu\sigma_g & \rho_{\theta,g}\sigma_\theta\sigma_g & \sigma_g^2\end{bmatrix}, \tag{6.9}$$

where note that we have normalized the unconditional mean of θ_t, $\bar{\theta}$, to zero since according to our specification $\bar{\mu}$ and $\bar{\theta}$ cannot be separately identified. In equation (6.8), we also have supressed parameter $\sigma_r^2 = Var(\varepsilon_{r,t})$, which will be automatically pinned down by the Campbell-Shiller identity once other parameters are estimated. We interpret μ_t and θ_t as two separate components of discount rates, and g_t is still expected dividend growth. Iterating on the Campbell-Shiller identity $r_{t+1} = \Delta d_{t+1} - pd_t + \kappa_0 + \kappa_1 pd_{t+1}$ allows us to obtain

$$pd_t = \frac{\kappa_0}{1-\kappa_1} + E_t\Big[\sum_{j=0}^{\infty}\kappa_1^j \Delta d_{t+1+j}\Big] - E_t\Big[\sum_{j=0}^{\infty}\kappa_1^j r_{t+1+j}\Big], \tag{6.10}$$

where κ_0 and κ_1 are log-linearization coefficients ($\bar{pd}$ is the sample average of log PD):

$$\kappa_0 = \log(1+e^{\bar{pd}}) - \kappa_1\bar{pd}$$
$$\kappa_1 = \frac{e^{\bar{pd}}}{1+e^{\bar{pd}}}.$$

We can then substitute Equations (6.3), (6.4), (6.5), (6.6), and (6.7) into Equation (6.10) to express pd_t as a linear function in states:

$$pd_t = A - B_1(\mu_t - \bar{\mu}) - B_2\theta_t + B_3(g_t - \bar{g}) \tag{6.11}$$

for

$$
\begin{aligned}
A &= \frac{\kappa_0 - \bar{\mu} + \bar{g}}{1-\kappa_1} \\
B_1 &= \frac{1}{1-\kappa_1 \rho_\mu} \\
B_2 &= \frac{1}{1-\kappa_1 \rho_\theta} \\
B_3 &= \frac{1}{1-\kappa_1 \rho_g}.
\end{aligned}
$$

It follows that we can reduce one state variable μ_t, and (by Campbell-Shiller identity) also reduce one observation r_t. Eventually, our system has two state variables (θ_t, g_t), whose dynamics are described by equations (6.6) and (6.7), and four observed variables $(\log G_t, \log O_t, pd_t, \Delta d_t)$, whose dynamics are described by equations (6.1), (6.2), (6.11), and (6.4), implicitly with equation (6.5). Assuming all the shocks are normally distributed, we can again use the Kalman Filter to construct the likelihood function, and estimate the model under monthly data.

6.2 Estimation Results

The estimation results are shown below. The parentheses contain standard errors.

$$\log G_t = const + \underset{(0.0200)}{0.0365}(\mu_t - \underset{(1.2)}{8.6}) + \underset{(0.0190)}{0.0354}\theta_t + \underset{(0.179)}{0.294}(g_t - \underset{(0.8)}{6.2}) + \varepsilon_{G,t} \tag{6.12}$$

$$\log O_t = const - \underset{(0.011)}{0.012}(\mu_t - \underset{(1.2)}{8.6}) - \underset{(0.011)}{0.010}\theta_t + \underset{(0.088)}{0.152}(g_t - \underset{(0.8)}{6.2}) + \varepsilon_{O,t} \tag{6.13}$$

$$r_t = \mu_{t-1} + \theta_{t-1} + \varepsilon_{r,t} \tag{6.14}$$

$$\Delta d_t = g_{t-1} + \varepsilon_{d,t} \tag{6.15}$$

$$\mu_t = (1 - \underset{(0.005)}{0.981})\underset{(1.2)}{8.6} + \underset{(0.005)}{0.981}\mu_{t-1} + \varepsilon_{\mu,t} \tag{6.16}$$

$$\theta_t = \underset{(0.005)}{0.982}\theta_{t-1} + \varepsilon_{\theta,t} \tag{6.17}$$

$$g_t = (1 - \underset{(0.007)}{0.981})\underset{(0.8)}{6.2} + \underset{(0.007)}{0.981}g_{t-1} + \varepsilon_{g,t} \tag{6.18}$$

with

$$Cov\left(\begin{bmatrix}\varepsilon_{G,t}\\ \varepsilon_{O,t}\\ \varepsilon_{d,t}\end{bmatrix}\right)=\begin{bmatrix}(6\times 10^{-133})^2 & & \\ & (3\times 10^{-122})^2 & \\ & & \underset{(0.1)^2}{6.8^2}\end{bmatrix}$$

and

$$\sigma_\mu=\underset{(4.5)}{22.5},\quad \sigma_\theta=\underset{(4.4)}{21.1},\quad \sigma_g=\underset{(0.09)}{0.30}$$

$$\rho_{\mu,\theta}=\underset{(0.0005)}{-0.998},\quad \rho_{\mu,g}=\underset{(0.26)}{-0.76},\quad \rho_{\theta,g}=\underset{(0.26)}{0.74}$$

The above equations further imply the following two equations:

$$\log G_t/O_t=const+\underset{(0.023)}{0.049}(\mu_t-\underset{(1.2)}{8.6})+\underset{(0.022)}{0.046}\theta_t+\underset{(0.236)}{0.142}(g_t-\underset{(0.8)}{6.2})+\varepsilon_{G/O,t}$$

$$\log G_t\cdot O_t=const-\underset{(0.023)}{0.024}(\mu_t-\underset{(1.2)}{8.6})-\underset{(0.022)}{0.025}\theta_t+\underset{(0.155)}{0.446}(g_t-\underset{(0.8)}{6.2})+\varepsilon_{G\cdot O,t}$$

together with the pricing equation for pd:

$$pd_t=\underset{(0.28)}{6.23}-\underset{(0.009)}{0.039}(\mu_t-\underset{(1.2)}{8.6})-\underset{(0.01)}{0.043}\theta_t+\underset{(0.012)}{0.039}(g_t-\underset{(0.8)}{6.2})$$

Again, the pricing errors σ_G and σ_O are negligibly small, although we didn't assume it. Together with the fact that pd_t is a linear function in the states without errors, the three observations $\log G_t$, $\log O_t$, and pd_t help almost pin down the three underlying states μ_t, θ_t, and g_t. The essence of our state-space approach in this case is thus to search for the optimal 3×3 coefficient matrix that makes the observed return and dividend growth data most likely, under the constraints on parameters imposed by the Campbell-Shiller identity. We're interested in seeing how commodity prices load on these states.

6.3 Short-Term vs. Long-Term Discount Rates

Let's make several definitions. We define $\mu_t + \theta_t$ as the "short-term discount rate" because it is a relatively fast-moving process (with monthly AC(1) of 0.975). Note from equation (6.14) that $\mu_t + \theta_t$ is also the conditional expected one-period return, which predicts immediate returns. Thus, it is appropriate to interpret it as a short-term discount rate. We then define $B_1\mu_t + B_2\theta_t$ as the "long-term discount rate" because it is a relatively slow-moving process (with monthly AC(1) of 0.994). The intuition for this definition can be seen from the Campbell-Shiller present value calculation:

$$pd_t = \frac{\kappa_0}{1-\kappa_1} + \underbrace{E_t\Big[\sum_{j=0}^{\infty}\kappa_1^j \Delta d_{t+1+j}\Big]}_{B_3 g_t} - \underbrace{E\Big[\sum_{j=0}^{\infty}\kappa_1^j r_{t+1+j}\Big]}_{B_1\mu_t + B_2\theta_t}, \tag{6.19}$$

where $B_1\mu_t + B_2\theta_t$ captures the expected weighted sum of returns at infinitely long terms. We find the correlation coefficient between short-term and long-term discount rates is 0.59. Third, note that g_t is still the expected dividend growth rate (with monthly AC(1) of 0.981). Moreover, we can now strongly reject the null that g_t is nonstationary, and its persistence looks more reasonable. For example, it is close to that in the long-run risks model of Bansal and Yaron (2004), 0.979 monthly.

Figure 7 then illustrates how the model extracts the three states. First, $\log G/O$ is very informative about the short-term discount rate. Second, the PD ratio is very informative about the long-term discount rate. Third, $\log G \cdot O$ is very informative about the expected dividend growth rate. These results are well consistent with our findings in previous sections.

Again, we find that both commodity prices load positively on g_t with similar coefficients, though not precisely the same (See Equations (6.12) and (6.13)). When we take the difference, the loadings get canceled, as we can see that $\log G/O$ loads insignificantly on g_t. When we take the sum, the loadings get strengthened, as we can see that $\log G \cdot O$ loads significantly positively on g_t and the coefficient is much larger relative to $\log G/O$'s and pd's. Overall, we find that the three price ratios, $(\log G, \log O, pd)$, contain independent information about the three states - none is redundant.

The above analysis is consistent with the variance decomposition results in Table 26 panel A, which shows that the variation in $\log G/O$ is 180% driven by (short-term) discount rate news. The variation in pd is 107% driven by long-term discount rate news (consistent with previous studies such as Cochrane (2011) and Van Binsbergen and Koijen (2010)). The variation in $\log G \cdot O$ is 102% driven by expected cash flow growth news. Variance decomposition results for realized returns and dividend growth do not change much, as shown in Table 26 Panel B.

6.4 How Do Gold/Oil Prices Respond to Long-/Short-Term Discount Rates?

Figures 8 and 9 plot our key results. Figure 8 (Figure 9) upper panel plots the portion of gold price explained by discount rates against short-term (long-term) discount rates, while the lower panel plots the portion of oil price explained by discount rates against short-term (long-term) discount rate. The two figures therefore illustrate how gold and oil prices respectively move with short-term/long-term discount rates after controlling for expected cash flow growth. As shown, gold price moves strongly positively with the short-term discount rate (correlation coefficient is 0.92); oil price moves strongly negatively with the short-term discount rate (correlation coefficient is -0.47). However, gold and oil's relation with long-term discount rates is less clear, though it seems that, in terms of long-run trends, both gold and oil prices rise with long-term discount rates (correlation coefficients are 0.26 and 0.42 respectively for gold and oil). Finally, it seems there was a structural break that occurred around 1986.

6.5 Predictive Regressions Under Model-Simulated Data

Table 27 reports results from using corresponding filtered states to predict returns and dividend growth rates in the data, close to the results in the last section.[13]

Tables 28 and 30 report univariate return and cash flow predictive regression results under model-simulated data. All these results are similar to those in Section 5 (Tables 24 and 25), further close to the data (Tables 4 and 16). One major miss, however, is that our model seems to generate too high R-squared in dividend growth predictive regressions. This is partially because our model fails to capture a structural break in the data which occurred around 1986. Table 17 shows that when we use data since 1986, the R-squared increases substantially.[14]

Table 29 further reports results from a bivariate return predictive regression under model-simulated data in which we use $\log G/O$ and log PD as the two predictors. As seen, the model largely reproduces the data (See Table 8) - at short horizons, $\log G/O$ dominates, while at long horizons, $\log PD$ dominates, consistent with our previous analysis that these two variables respectively proxy for short and long-term discount rates of the economy.

To sharpen our results even further, Table 31 reports bivariate predictive regressions in which we use $\log G$ and $\log O$ as the two regressors. For return regressions, the coefficients at short horizons are roughly 20 and -20. For cash flow regressions, the coefficients are close to 2 and 2, though not precisely. These patterns are similar to what we observe in the data (See Tables 15 and 19). Table 32 further reports results from bivariate return predictive regressions where we control for expected cash flow growth, as proxied by $\log G \cdot O$. As shown, the coefficients on $\log G$ and $\log G \cdot O$ are respectively 40 and -20, and the coefficients on $\log O$ and $\log G \cdot O$ are respectively -40 and 20. These results are again highly consistent with what we observe in the data (See Table 20). Overall, the message

[13] Again, we add a time fixed effect on the RHS when regressing returns onto long-term discount rate, which is closely correlated with PD, which has a trend in our sample.

[14] In fact, all this paper's found relations hold more significant if we only use post-1986 data.

in this section is highly consistent with that in the previous section.

6.6 How Accurate are the Returns Implied by the Campbell-Shiller Identity?

In Section 6, we didn't use returns data, which will be automatically pinned down from the Campbell-Shiller identity once we have supplied the price-dividend ratio and dividend growth data. But since the latter are both based on deseasoned monthly dividend data, a natural question is how far the recovered returns will deviate from their actual realized values without deseasoning. Figure 10 compares the actual realized monthly return r_{t+1} with the monthly return implied by Campbell-Shiller identity $r_{t+1}^{CS} = \Delta d_{t+1} - pd_t + \kappa_0 + \kappa_1 pd_{t+1}$ using deseasoned d_t and Δd_{t+1}.

As shown, the difference is negligible. This is because at monthly frequency, most of the variation in r_{t+1} comes from that in p_{t+1}, not d_{t+1} or Δd_{t+1}. The two return time series have a correlation coefficient of 0.98. Standard deviations are respectively $std(r_{t+1}) = 4.39\%$ and $std(r_{t+1}^{CS}) = 4.53\%$. Means are respectively $E(r_{t+1}) = 0.89\%$ and $E(r_{t+1}^{CS}) = 1.06\%$.

7 Principal Component Analysis

In this section, we perform PCA and relate to the previous sections. Specifically, we extract five principal components from prices of a cross-section of commodities including gold, crude oil, silver, copper, and platinum. We still use montly data from 1975-2022. We again first detrend real commodity prices using real dividends, which are shown in Figure 11.

We find that the five orthognal PCs respectively account for 72.3%, 8.3%, 5.0%, 4.3%, and 2.0% of commodity price fluctuations. These PCs are illustrated in Figure 12. As seen, the first PC is basically expected dividend growth g_t. The second PC is basically expected return μ_t.[15] These are confirmed by predictive regressions as shown in Table 33. These two PCs combined account for more than 80% of commodity price variations. We don't know how to interpret the third and fourth PCs, which likely capture commodity-market-specific information. The fifth PC can be interpreted as some long-term expected return, as suggested by predictive regressions in Table 33, and it is not subsumed by the PD ratio.

The loadings of commodity prices on the PCs are given by

$$\begin{array}{cc} & \begin{array}{ccccc} PC(1) & PC(2) & PC(3) & PC(4) & PC(5) \end{array} \\ \begin{array}{c} Gold \\ Oil \\ Silver \\ Copper \\ Platinum \end{array} & \begin{bmatrix} 0.41 & 0.44 & -0.05 & 0.28 & 0.75 \\ 0.45 & -0.62 & -0.48 & -0.36 & 0.21 \\ 0.61 & 0.49 & -0.19 & -0.16 & -0.57 \\ 0.37 & -0.16 & 0.84 & -0.35 & 0.08 \\ 0.35 & -0.40 & 0.13 & 0.80 & -0.24. \end{bmatrix} \end{array} \tag{7.1}$$

Consistent with our previous findings, gold and silver rise with short-term expected returns, PC(2). It seems gold and silver are short-term hedging assets. But their long-term performance are quite different: gold rises with long-term expected returns, PC(5); while silver behaves inversely. These loadings are also consistent with Huang and Kilic (2019), which take gold-platinum price ratio to cancel loadings on PC(1) and strengthen loadings on PC(2) and PC(5). This constructs them a return predictor at both short and long horizons. Our marginal contribution to the literature is to demonstrate that gold is a strong hedging asset using three different approaches: predictive regressions, state-space models, and PCA.

[15]We say basically, not essentially, because the definition of PCA implies that PCs are independent. But the state-space model suggests that g_t and μ_t are negative correlated (unconditional correlation is around -0.2). So there must be some small difference.

8 Conclusion

We show that gold serves as a prominent hedging asset, provided the "omitted-variable" or "error-in-variable" bias related to expected economic fundamentals is appropriately addressed. We reach this conclusion by employing three distinct methodologies: a state-space model, predictive regressions, and principal component analysis. Notably, each approach offers unique insights without redundancy.

The state-space model quantifies the impact of the omitted variable bias by decomposing the variance of gold prices into effects of expected dividend growth, expected returns, and their covariance. Remarkably, the model reveals a near 100% OVB when regressing future stock returns on current gold prices. This result explains the myriad contrasting findings in the literature concerning the relationship between gold prices and stock markets.

Through predictive regressions, we establish that the gold-oil price ratio emerges as a robust predictor of stock market returns, even at a one-day horizon, displaying predictive power independent of all existing literature predictors.

We perform PCA on a large cross-section of commodity prices. The first PC, which accounts for 72% of all commodity price variations, is basically expected dividend growth. The second PC, which accounts for another 8%, is basically short-horizon expected stock return. The fifth PC, basically some long-term expected return measure, only accounts for 2%, implying commodity prices are on average not reponsive to long-term discount rate news.

References

Avdis, Efstathios, and Jessica A Wachter, 2017, Maximum likelihood estimation of the equity premium, *Journal of Financial Economics* 125, 589–609.

Baker, Malcolm, and Jeffrey Wurgler, 2006, Investor sentiment and the cross-section of stock returns, *The journal of Finance* 61, 1645–1680.

Bakshi, Gurdip, Xiaohui Gao, and Zhaowei Zhang, 2023, Models of Gold Options Market and Evidence in Favor of Financialized Gold and Against Disasterization, *Available at SSRN 4652113*.

Bakshi, Gurdip, and Dilip Madan, 2000, Spanning and derivative-security valuation, *Journal of financial economics* 55, 205–238.

Bansal, Ravi, and Amir Yaron, 2004, Risks for the long run: A potential resolution of asset pricing puzzles, *The journal of Finance* 59, 1481–1509.

Barro, Robert J, and Sanjay Misra, 2016, Gold returns, *The Economic Journal* 126, 1293–1317.

Baur, Dirk G, and Lee A Smales, 2020, Hedging geopolitical risk with precious metals, *Journal of Banking & Finance* 117, 105823.

Bekaert, Geert, Eric C Engstrom, and Nancy R Xu, 2022, The time variation in risk appetite and uncertainty, *Management Science* 68, 3975–4004.

Bollerslev, Tim, George Tauchen, and Hao Zhou, 2009, Expected stock returns and variance risk premia, *The Review of Financial Studies* 22, 4463–4492.

Campbell, John Y, and John H Cochrane, 1999, By force of habit: A consumption-based explanation of aggregate stock market behavior, *Journal of political Economy* 107, 205–251.

Campbell, John Y, and Robert J Shiller, 1988, The dividend-price ratio and expectations of future dividends and discount factors, *The Review of Financial Studies* 1, 195–228.

Chabi-Yo, Fousseni, and Johnathan Loudis, 2020, The conditional expected market return, *Journal of Financial Economics* 137, 752–786.

Chen, Nai-Fu, Richard Roll, and Stephen A Ross, 1986, Economic forces and the stock market, *Journal of business* pp. 383–403.

Cheng, Ing-Haw, Ke Tang, and Lei Yan, 2021, Hedging Pressure and Commodity Option Prices, *Available at SSRN 3933070*.

Christoffersen, Peter, and Xuhui Nick Pan, 2018, Oil volatility risk and expected stock returns, *Journal of Banking & Finance* 95, 5–26.

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

Cochrane, John H, 2011, Presidential address: Discount rates, *The Journal of finance* 66, 1047–1108.

Dickey, David A, and Wayne A Fuller, 1979, Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American statistical association* 74, 427–431.

Driesprong, Gerben, Ben Jacobsen, and Benjamin Maat, 2008, Striking oil: another puzzle?, *Journal of financial economics* 89, 307–327.

Engle, Robert F, and Clive WJ Granger, 1987, Co-integration and error correction: representation, estimation, and testing, *Econometrica: journal of the Econometric Society* pp. 251–276.

Erb, Claude B, and Campbell R Harvey, 2013, The golden dilemma, *Financial Analysts Journal* 69, 10–42.

Gao, Lin, Steffen Hitzemann, Ivan Shaliastovich, and Lai Xu, 2022, Oil volatility risk, *Journal of Financial Economics* 144, 456–491.

Giglio, Stefano, and Dacheng Xiu, 2021, Asset pricing with omitted factors, *Journal of Political Economy* 129, 1947–1990.

Hou, Kewei, Ke Tang, and Bohui Zhang, 2020, Political uncertainty and commodity markets, *Fisher College of Business Working Paper* p. 025.

Huang, Dashan, Fuwei Jiang, Jun Tu, and Guofu Zhou, 2015, Investor sentiment aligned: A powerful predictor of stock returns, *The Review of Financial Studies* 28, 791–837.

Huang, Darien, and Mete Kilic, 2019, Gold, platinum, and expected stock returns, *Journal of Financial Economics* 132, 50–75.

Huang, Roger D, Ronald W Masulis, and Hans R Stoll, 1996, Energy shocks and financial markets, *Journal of Futures markets* 16, 1–27.

Jermann, Urban J, 2021, Gold's Value as an Investment, *Available at SSRN 3934052*.

Johansen, Søren, 1988, Statistical analysis of cointegration vectors, *Journal of economic dynamics and control* 12, 231–254.

Jones, Charles M, and Gautam Kaul, 1996, Oil and the stock markets, *The journal of Finance* 51, 463–491.

Li, Yan, David T Ng, and Bhaskaran Swaminathan, 2013, Predicting market returns using aggregate implied cost of capital, *Journal of Financial Economics* 110, 419–436.

Martin, Ian, 2017, What is the Expected Return on the Market?, *The Quarterly Journal of Economics* 132, 367–433.

Ready, Robert C, 2018, Oil prices and the stock market, *Review of Finance* 22, 155–176.

Sim, Nicholas, and Hongtao Zhou, 2015, Oil prices, US stock return, and the dependence between their quantiles, *Journal of Banking & Finance* 55, 1–8.

Stock, James H, and Mark W Watson, 1993, A simple estimator of cointegrating vectors in higher order integrated systems, *Econometrica: journal of the Econometric Society* pp. 783–820.

Van Binsbergen, Jules H, and Ralph SJ Koijen, 2010, Predictive regressions: A present-value approach, *The Journal of Finance* 65, 1439–1471.

Wachter, Jessica A, 2013, Can time-varying risk of rare disasters explain aggregate stock market volatility?, *The Journal of Finance* 68, 987–1035.

Welch, Ivo, and Amit Goyal, 2008, A comprehensive look at the empirical performance of equity premium prediction, *The Review of Financial Studies* 21, 1455–1508.

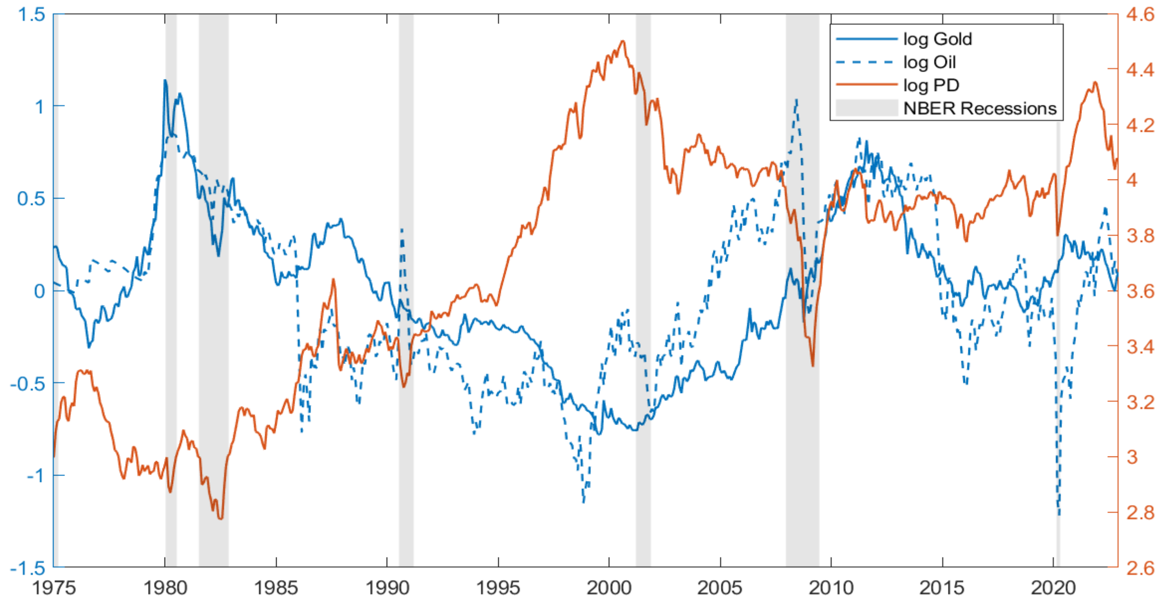


Figure 1. The figure plots detrended log real gold and oil prices, log PD ratio, and NBER recessions. Data are monthly from Jan 1975 to Dec 2022.

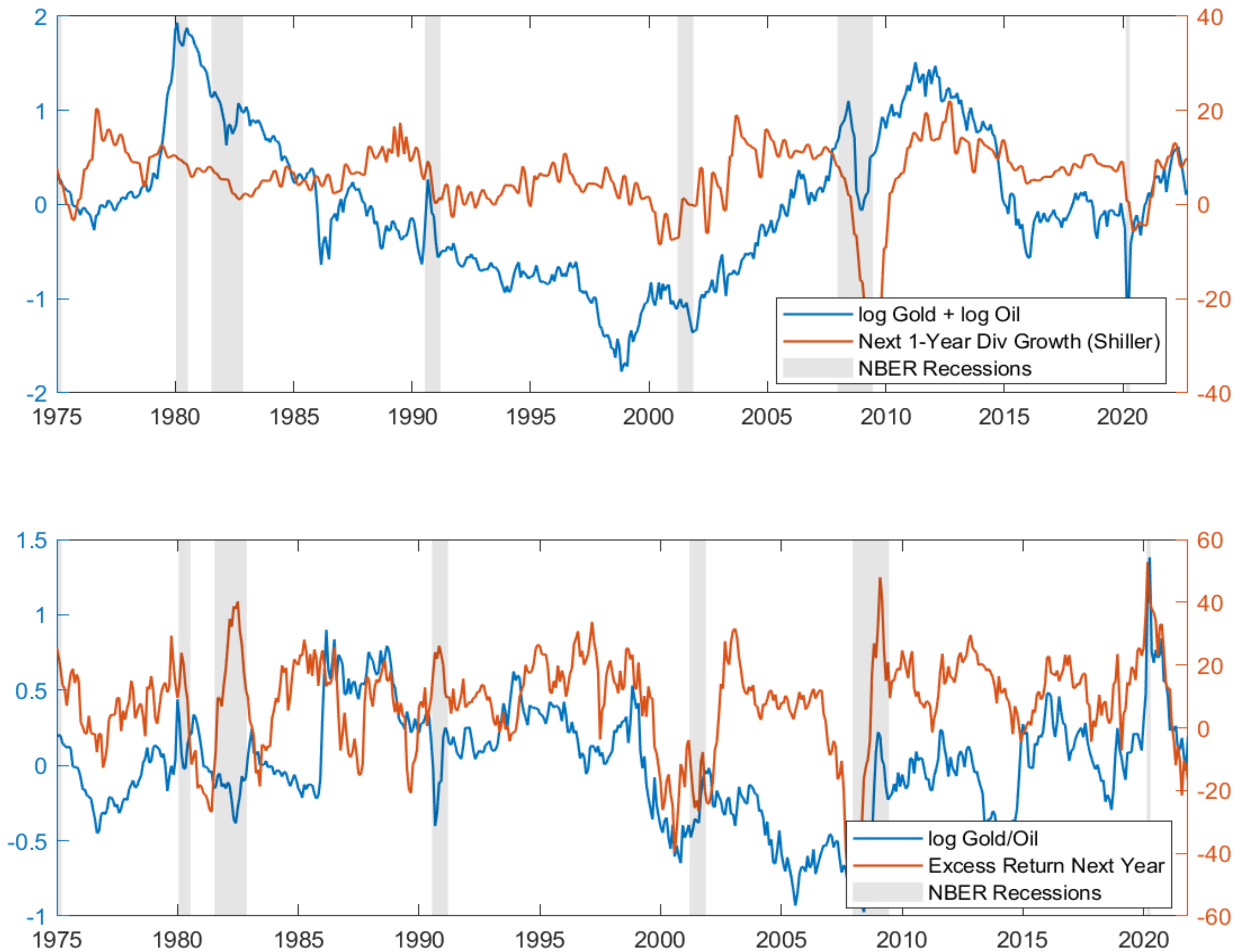
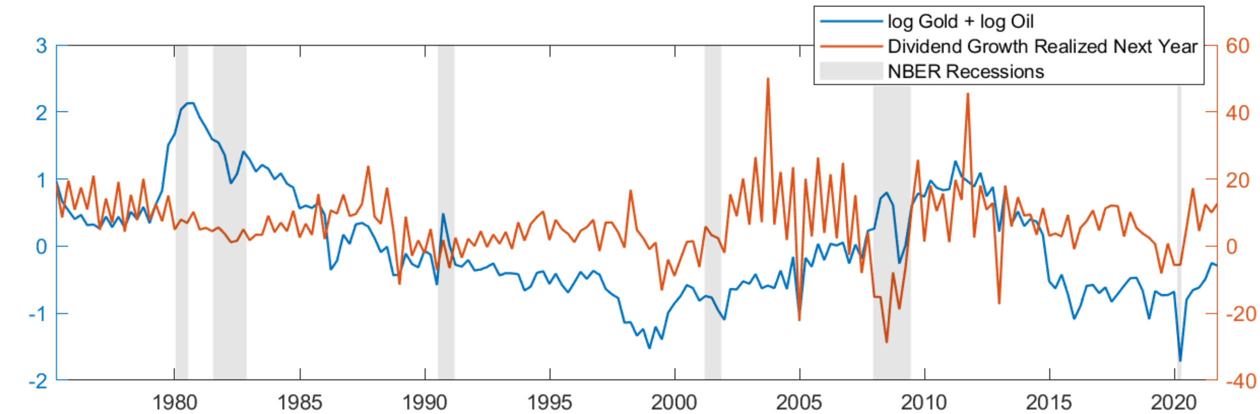


Figure 2. The top figure plots $\log G \cdot O$ and cumulative dividend growth rates realized over the next year, using dividend data from Robert Shiller's website. The bottom figure plots $\log G/O$ and cumulative excess returns realized over the next year. All series are annualized in percent. Data are monthly from Jan 1975 to Dec 2022.



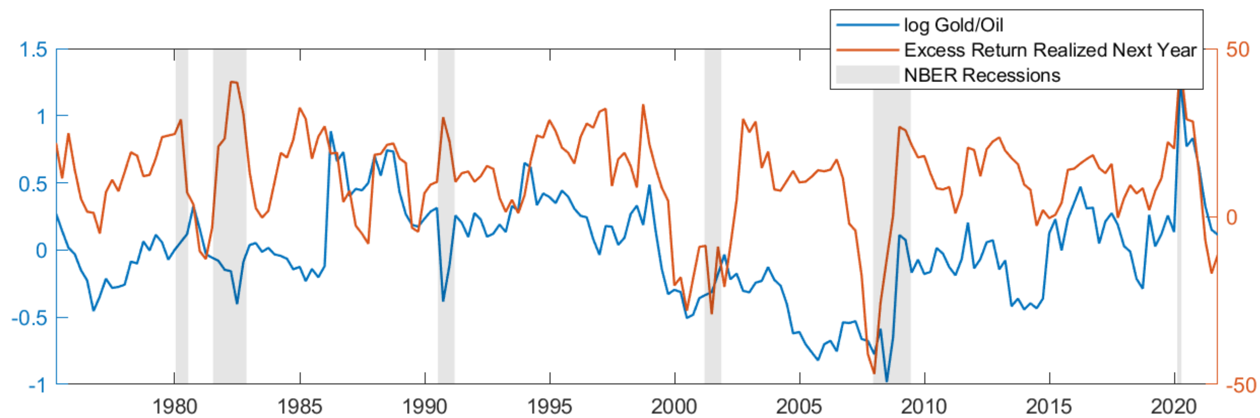


Figure 3. The top figure plots $\log G \cdot O$ and cumulative dividend growth rates realized over the next 4 quarters, using quarterly dividend data from CRSP. The bottom figure plots $\log G/O$ and cumulative excess returns realized over the next 4 quarters. All series are annualized in percent. Data are quarterly from 1975 Q1 to 2022 Q4.

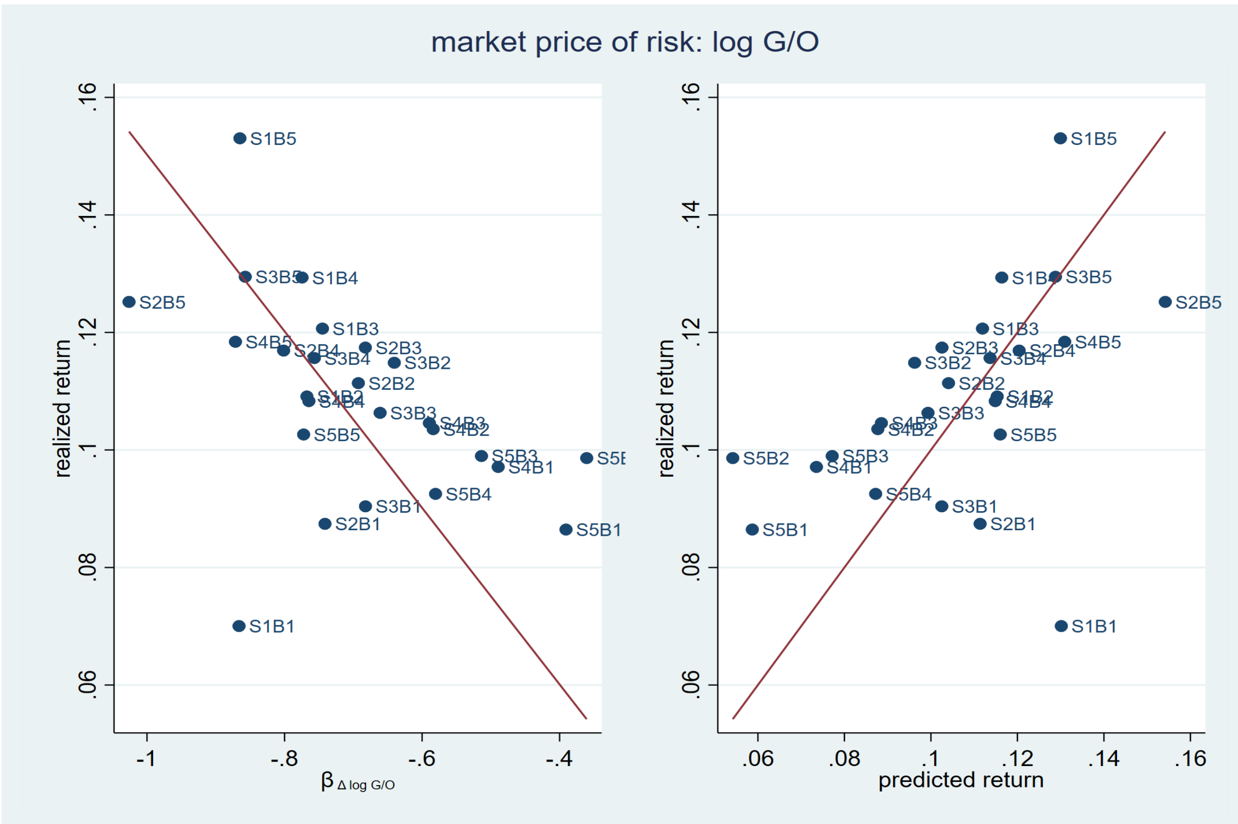


Figure 4. Left panel depicts realized sample-time-series-average return to each portfolio against the portfolio's exposure to $\Delta \log G/O$ estimated using the full sample. Right panel depicts realized average return to each portfolio against the expected return predicted by the one-factor ($\Delta \log G/O$) model without an intercept. Test portfolios are Fama-French size/book-to-market 25 portfolios. Data are monthly from Jan 1975 to Dec 2022.

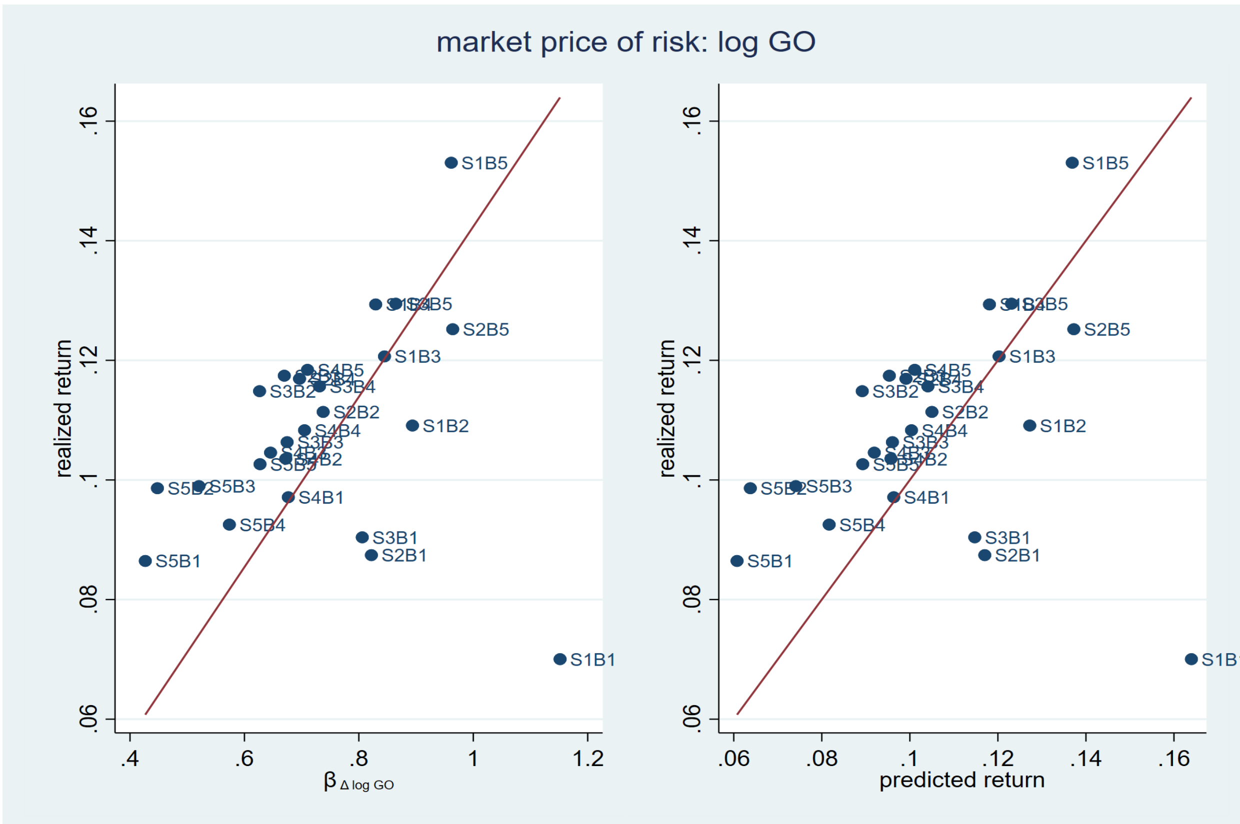


Figure 5. Left panel depicts realized sample-time-series-average return to each portfolio against the portfolio's exposure to $\Delta \log G \cdot O$ estimated using the full sample. Right panel depicts realized average return to each portfolio against the expected return predicted by the one-factor ($\Delta \log G \cdot O$) model without an intercept. Test portfolios are Fama-French size/book-to-market 25 portfolios. Data are monthly from Jan 1975 to Dec 2022.

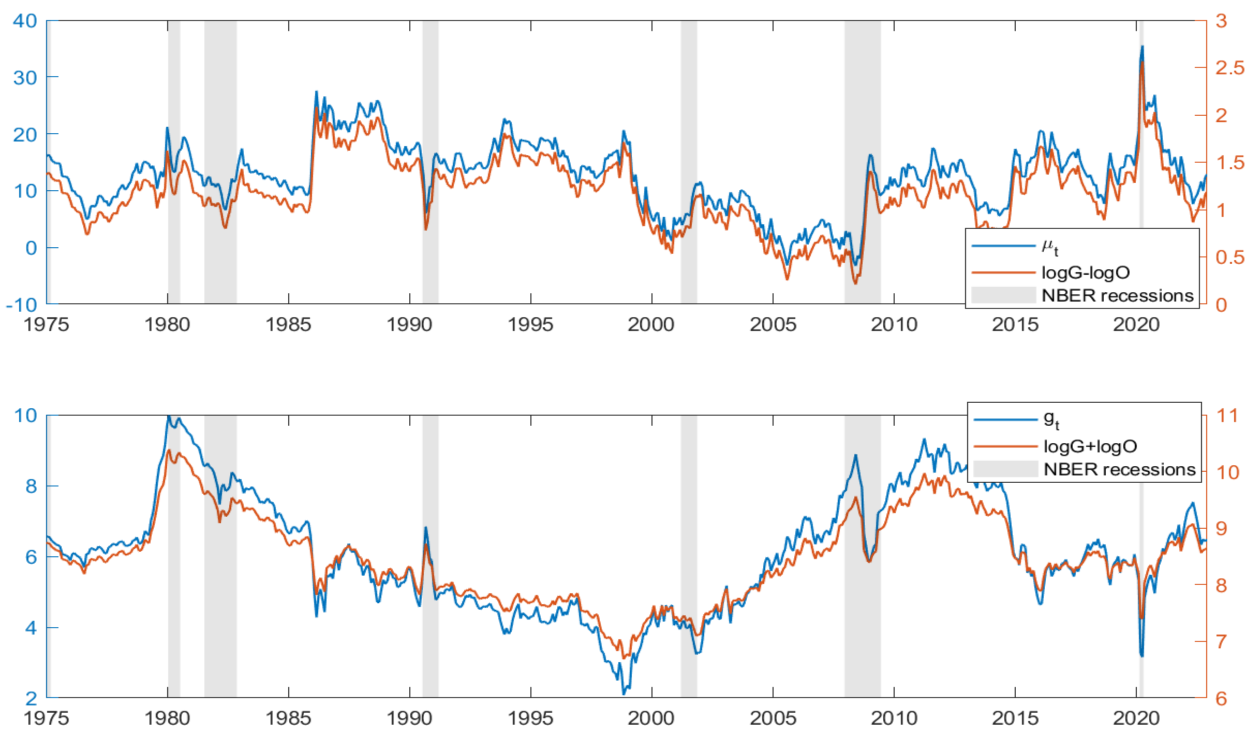


Figure 6. The upper figure plots filtered μ_t and $\log G/O$. The lower figure plots filtered g_t and $\log G \cdot O$. All numbers are annualized in percent. Data are monthly from Jan 1975 to Dec 2022.

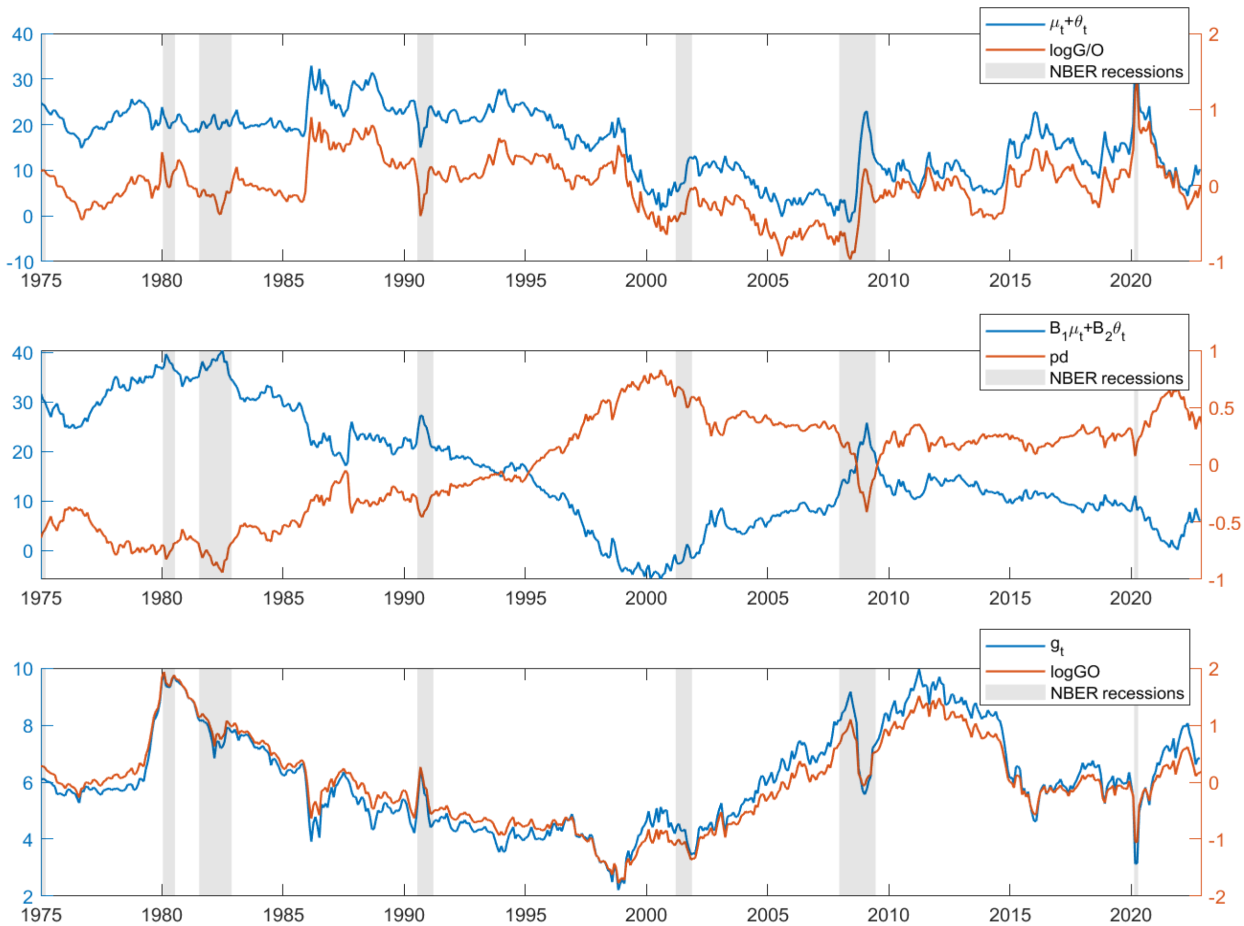


Figure 7. The top figure plots filtered expected return $\mu_t + \theta_t$ and $\log G/O$. The middle figure plots filtered long-term discount rate $B_1\mu_t + B_2\theta_t$ and $\log PD$. The bottom figure plots filtered expected dividend growth g_t and $\log G \cdot O$. All filtered series are annualized in percent. Data are monthly from Jan 1975 to Dec 2022.

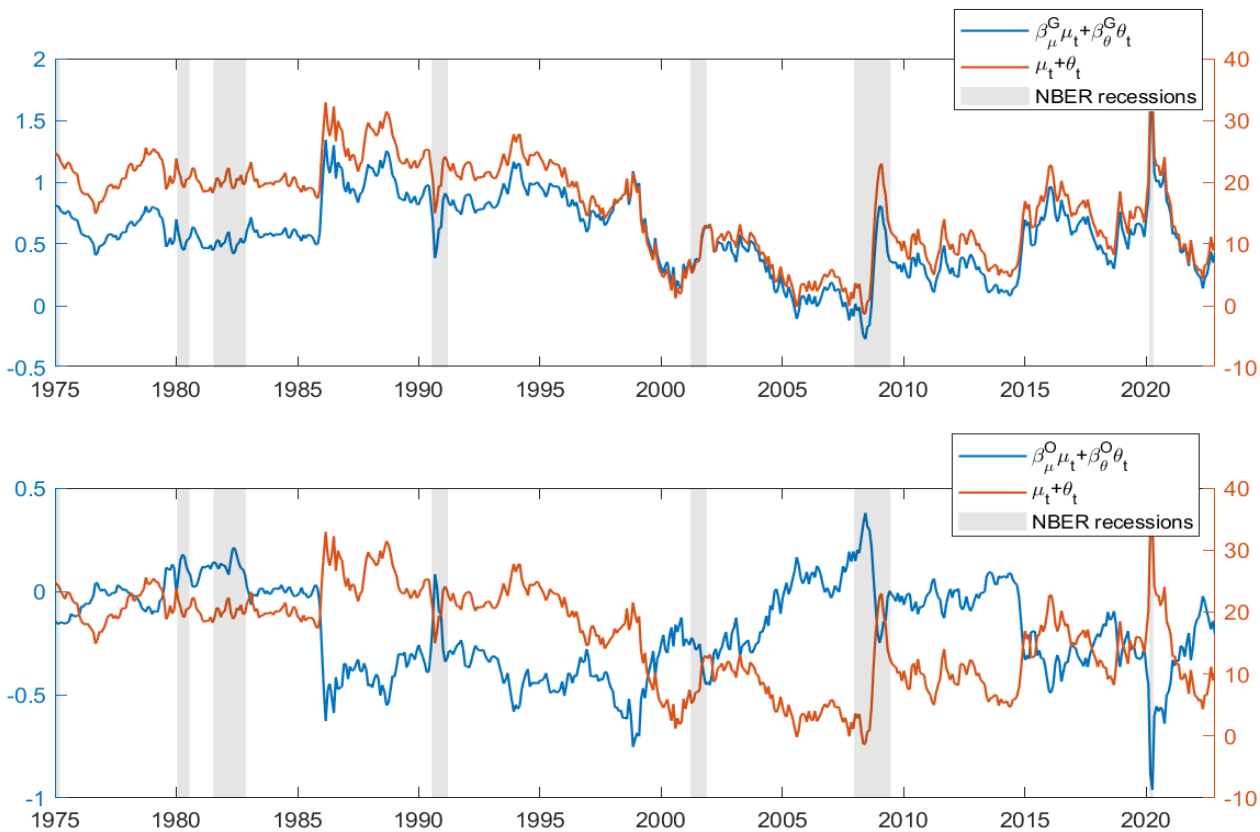


Figure 8. The upper figure plots filtered $\log G$'s loading on discount rates $\beta_\mu^G \mu_t + \beta_\theta^G \theta_t$ and filtered short-term discount rate $\mu_t + \theta_t$. The lower figure plots filtered $\log O$'s loading on discount rates $\beta_\mu^O \mu_t + \beta_\theta^O \theta_t$ and filtered short-term discount rate $\mu_t + \theta_t$. $\mu_t + \theta_t$ is annualized in percent. Data are monthly from Jan 1975 to Dec 2022.

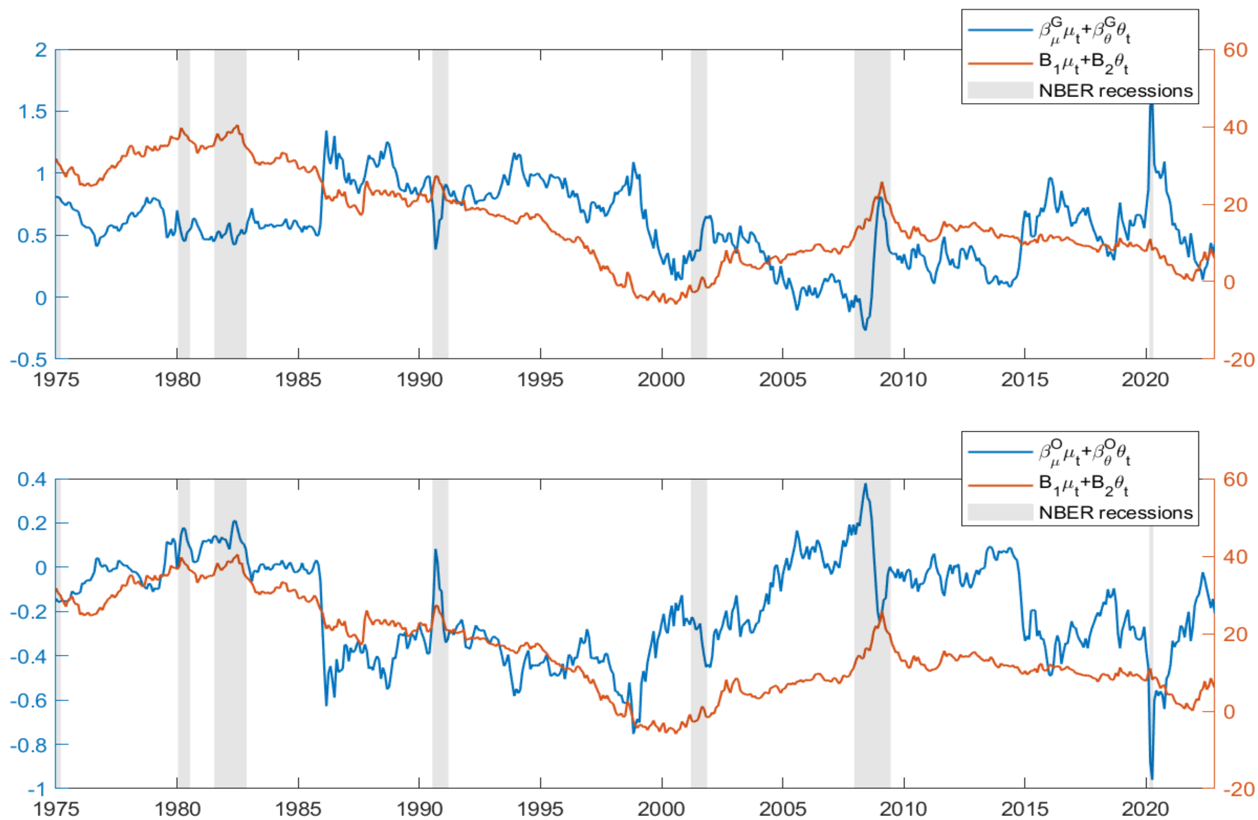


Figure 9. The upper figure plots filtered $\log G$'s loading on discount rates $\beta^G_\mu \mu_t + \beta^G_\theta \theta_t$ and long-term discount rate $B_1\mu_t + B_2\theta_t$. The lower figure plots filtered $\log O$'s loading on discount rates $\beta^O_\mu \mu_t + \beta^O_\theta \theta_t$ and long-term discount rate $B_1\mu_t + B_2\theta_t$. $B_1\mu_t + B_2\theta_t$ is annualized in percent. Data are monthly from Jan 1975 to Dec 2022.

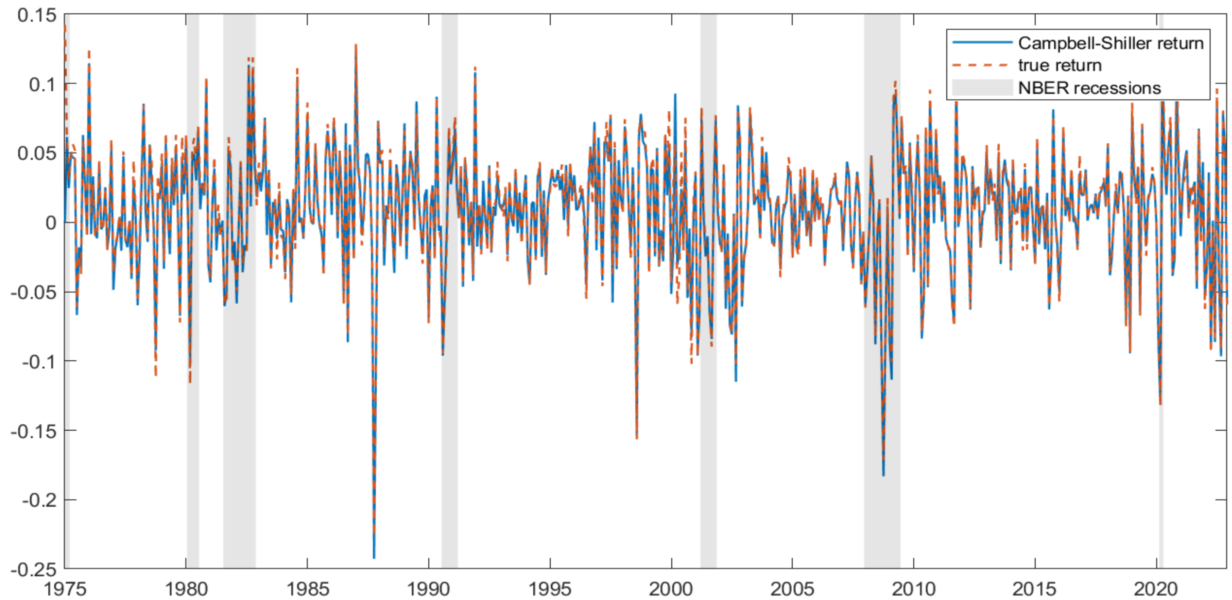


Figure 10. The figure plots monthly realized returns (true values vs. values implied by the Campbell-Shiller accounting identity and de-seasoned dividend data). The correlation coefficient is 0.98. Data are monthly from Jan 1975 to Dec 2022.

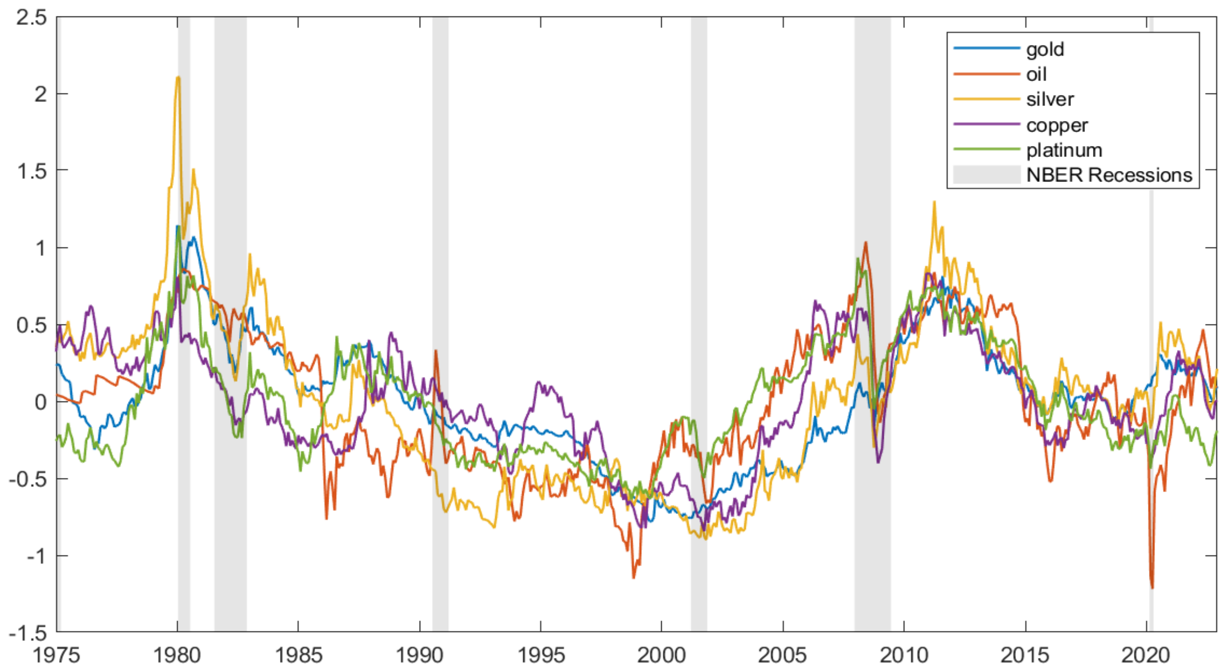


Figure 11. The figure plots real commodity (gold, oil, silver, copper, and platinum) prices detredned using real dividends (so that all processes are stationary). Data are monthly from Jan 1975 to Dec 2022.

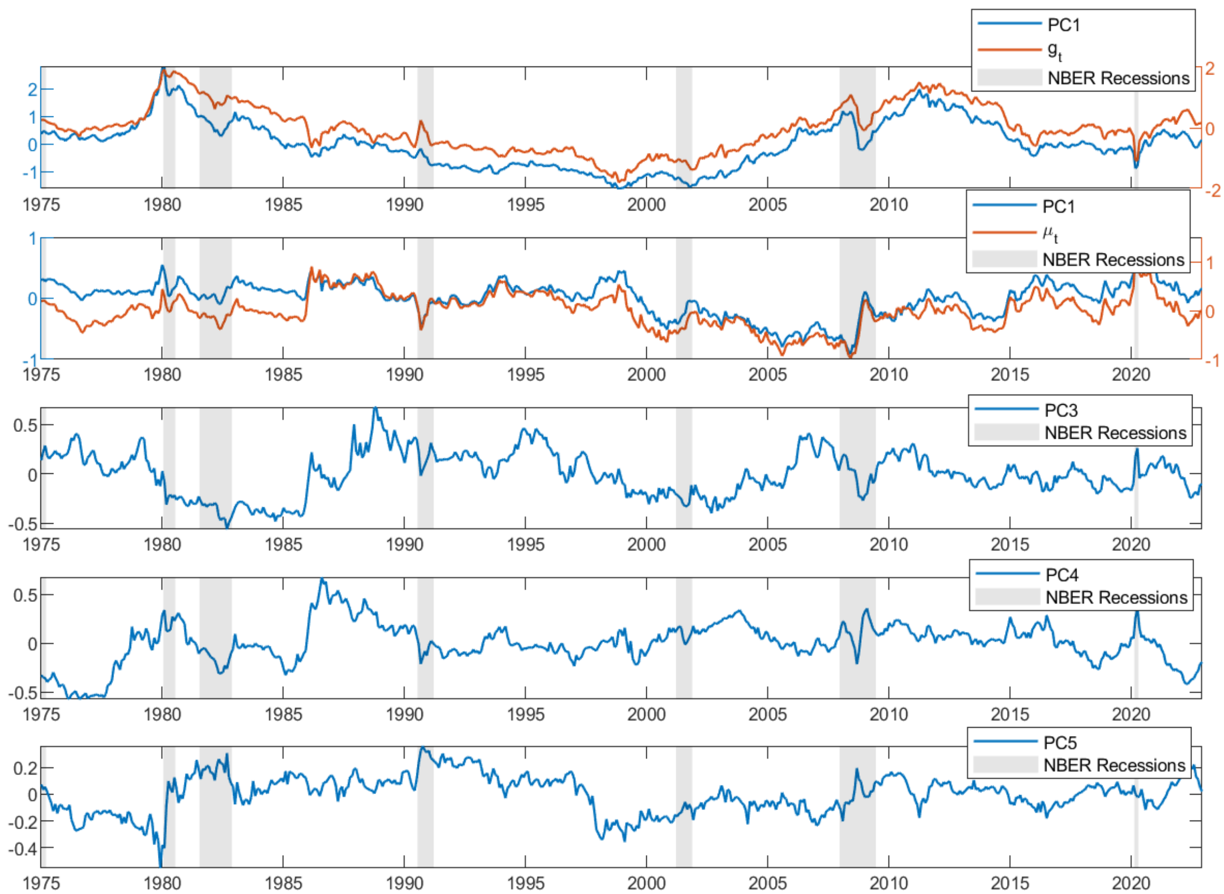


Figure 12. The figure plots the 5 principal components extracted from the cross-section of detrended real commodity prices. Data are monthly from Jan 1975 to Dec 2022. In the first panel, we compare PC(1) with g_t. In the second panel, we compare PC(2) with μ_t.

Table 1. Johansen (1988) Rank Test

We estimate an Engle and Granger (1987) vector error-correction model (VECM)

$$\Delta Y_t = \mu + \Pi Y_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta Y_{t-j} + \varepsilon_t$$

and conduct Johansen (1988) rank test for cointegration based on the rank of the matrix Π. The null hypothesis for the rank test is that there are no more than r cointegration relations.

	statistic	c-value	p-value
$Y_t = [\log G_t, \log O_t]'$			
$H_0 : r = 0$	12.6	15.5	0.129
$H_0 : r = 1$	1.0	3.8	0.454

$Y_t = [\log G_t, \log D_t]'$			
$H_0 : r = 0$	16.9	15.5	0.031
$H_0 : r = 1$	2.6	3.8	0.111
$Y_t = [\log O_t, \log D_t]'$			
$H_0 : r = 0$	23.9	15.5	0.003
$H_0 : r = 1$	5.1	3.8	0.024
$Y_t = [\log G_t/O_t, \log D_t]'$			
$H_0 : r = 0$	28.8	15.5	0.001
$H_0 : r = 1$	8.1	3.8	0.005
$Y_t = [\log G_t \cdot O_t, \log D_t]'$			
$H_0 : r = 0$	19.2	15.5	0.013
$H_0 : r = 1$	3.0	3.8	0.083
$Y_t = [\log G_t, \log O_t, \log D_t]'$			
$H_0 : r = 0$	33.4	29.8	0.018
$H_0 : r = 1$	10.5	15.5	0.269
$H_0 : r = 2$	2.3	3.8	0.129

Table 2. Stock and Watson (1993) Dynamic Least Square

We estimate cointegration vector between log commodity prices and log dividends using Stock and Watson (1993) dynamic least square with Newey-West robust standard errors.

$$\log X_t = \beta_0 + \beta_D \log D_t + \sum_{i=-k}^{i=k} \gamma_i \Delta \log D_{t-i} + \epsilon_t$$

	estimate	s.e.	t-stat	[95% CI]
$X = G$				
$\beta_D(k=1)$	0.62	0.06	10.54	[0.50, 0.73]
$\beta_D(k=2)$	0.62	0.07	8.58	[0.48, 0.76]
$\beta_D(k=3)$	0.62	0.08	7.46	[0.46, 0.78]
$X = O$				
	estimate	s.e.	t-stat	[95% CI]
$\beta_D(k=1)$	0.16	0.08	1.96	[0.00, 0.32]
$\beta_D(k=2)$	0.16	0.09	1.62	[-0.03, 0.35]
$\beta_D(k=3)$	0.16	0.11	1.44	[-0.06, 0.38]
$X = G/O$				
	estimate	s.e.	t-stat	[95% CI]
$\beta_D(k=1)$	0.46	0.08	5.64	[0.30, 0.62]
$\beta_D(k=2)$	0.46	0.10	4.76	[0.27, 0.65]
$\beta_D(k=3)$	0.46	0.11	4.25	[0.25, 0.67]
$X = G \cdot O$				
	estimate	s.e.	t-stat	[95% CI]
$\beta_D(k=1)$	0.78	0.12	6.66	[0.55, 1.01]
$\beta_D(k=2)$	0.78	0.14	5.47	[0.50, 1.05]
$\beta_D(k=3)$	0.78	0.16	4.80	[0.46, 1.10]

Table 3. Summary Statistics

All data are monthly. $\log G_t$ is log real gold price. $\log O_t$ is log real oil price. $\log D_t$ is log real dividend. $\log G_t/P_t$ is log gold to platinum price ratio from Huang and Kilic (2019). VRP=IV-RV from Bollerslev, Tauchen, and Zhou (2009). $\log PD_t$ is log price-dividend ratio. $Sentiment_t^{BW}$ is investor sentiment index from Baker and Wurgler (2006). $Sentiment_t^{PLS}$ is sentiment index estimated using partial least square from Huang, Jiang, Tu, and Zhou (2015). $Risk\_Aversion_t^{BEX}$ is risk aversion index from Bekaert, Engstrom, and Xu (2022). VIX is CBOE VIX index. $Interest\_Rate_t$ is one-month nominal interest rate from Ken French. $Inflation_t$ is inflation rate from FRED. DFSP is default spread from FRED. TMSP is term spread from FRED. ICC_t is implied cost of capital from Li, Ng, and Swaminathan (2013). $Skew_t^Q$ and $Kurt_t^Q$ are one-month risk-neutral skewness and kurtosis computed from S&P 500 option prices using spanning formula in Bakshi and Madan (2000). $Crash\_Prob_t$ is one-month physical probability of a 10% market crash backed out from SPX option prices as in Martin (2017) and Chabi-Yo and Loudis (2020). ADF p.val denotes p-value in the agumented Dickey and Fuller (1979) one-sided unit root test.

Variable	Std	AR(1)	ADF p.val	Start	End
$\log G \equiv \log G_t - 0.62 \log D_t$	0.39	0.992	0.059	1975.1	2022.12
$\log O \equiv \log O_t - 0.16 \log D_t$	0.45	0.976	0.005	1975.1	2022.12
$\log G/O \equiv \log G_t/O_t - 0.46 \log D_t$	0.36	0.957	0.000	1975.1	2022.12
$\log G \cdot O \equiv \log G_t \cdot O_t - 0.78 \log D_t$	0.76	0.989	0.036	1975.1	2022.12
$\log G_t/P_t$	0.35	0.978	0.023	1975.1	2022.12
VRP_t	22.42	0.329	0.000	1990.1	2019.12
$\log PD_t$	0.44	0.994	0.073	1975.1	2022.12
$Sentiment_t^{BW}$	0.87	0.981	0.010	1975.1	2022.6
$Sentiment_t^{PLS}$	0.96	0.988	0.035	1975.1	2020.12
$Risk\_Aversion_t^{BEX}$	0.67	0.788	0.000	1986.7	2022.12
VIX_t	7.56	0.808	0.000	1990.1	2022.12
$Interest\_Rate_t$	0.29	0.978	0.005	1975.1	2022.12
$Inflation_t$	0.14	0.644	0.000	1975.1	2022.12
$DFSP_t$	0.45	0.959	0.000	1975.1	2022.12
$TMSP_t$	1.12	0.952	0.001	1975.1	2022.12
ICC_t	3.10	0.983	0.035	1977.1	2017.12
$Skew_t^Q$	0.41	0.847	0.000	1996.1	2021.12
$Kurt_t^Q$	1.11	0.863	0.000	1996.1	2021.12
$Crash\_Prob_t$	2.70	0.776	0.000	1996.1	2021.12

Table 4. Univariate Return Predictability: Gold and Oil

Univariate monthly return predictive regressions for the U.S. equity market, Jan 1975 to Dec 2022. S.E. are Newey-West HAC robust standard errors with 3 more lags than forecasting horizon. All LHS returns are annual in percent. Univariate regressions are run separately:

$$\frac{12}{h}\sum_{i=1}^{h}\left(r_{t+i}-r^f_{t+i}\right)=\beta_0+\beta_X X_t+\epsilon_{t+h}$$

	1m	2m	3m	6m	9m	1y	3y	5y
				$X=\log G/O$				
$\beta_{\log G/O}$	20.03	18.17	18.02	16.73	15.54	14.48	8.01	5.32
s.e.	[6.67]	[6.53]	[6.65]	[6.91]	[6.41]	[5.88]	[3.77]	[2.53]
R^2_{adj} (%)	1.58	2.63	4.08	7.31	9.46	11.04	11.74	9.78
				$X=\log G\cdot O$				
$\beta_{\log G\cdot O}$	-1.25	-0.86	-0.88	-0.91	-0.95	-0.71	1.88	2.29
s.e.	[3.17]	[3.04]	[3.00]	[3.03]	[2.98]	[2.83]	[2.10]	[1.49]
R^2_{adj} (%)	-0.14	-0.15	-0.13	-0.08	-0.01	-0.06	3.11	9.10
				$X=\log G$				
$\beta_{\log G}$	6.15	6.10	5.99	5.39	4.78	4.80	6.63	6.37
s.e.	[6.06]	[5.69]	[5.51]	[5.43]	[5.51]	[5.57]	[4.05]	[2.73]
R^2_{adj} (%)	0.02	0.20	0.38	0.74	0.90	1.27	10.57	18.72
				$X=\log O$				
$\beta_{\log O}$	-8.18	-7.03	-7.01	-6.64	-6.34	-5.66	0.42	1.83
s.e.	[5.42]	[5.31]	[5.35]	[5.54]	[5.28]	[4.76]	[3.62]	[2.69]
R^2_{adj} (%)	0.28	0.48	0.84	1.67	2.33	2.50	-0.13	1.81

Table 5. Univariate Return Predictability: I

Univariate monthly return predictive regressions for the U.S. equity market, Jan 1975 to Dec 2022. S.E. are Newey-West HAC robust standard errors with 3 more lags than forecasting horizon. All LHS returns are annual in percent. Univariate regressions are run separately:

$$\frac{12}{h}\sum_{i=1}^{h}\left(r_{t+i}-r^{f}_{t+i}\right)=\beta_0+\beta_X X_t+\epsilon_{t+h}$$

	1m	2m	3m	6m	9m	1y	3y	5y
				$X=\log G/P$				
$\beta_{\log G/P}$	14.54	14.43	14.66	15.54	16.07	16.53	14.01	12.07
s.e.	[6.55]	[6.35]	[6.49]	[7.15]	[7.20]	[6.52]	[3.20]	[1.46]
R^2_{adj} (%)	0.71	1.51	2.49	5.77	9.05	12.58	27.50	31.86
				$X=\log PD$				
$\beta_{\log PD}$	-19.14	-19.57	-19.13	-19.61	-19.93	-19.93	-15.56	-15.08
s.e.	[9.53]	[8.93]	[8.60]	[7.96]	[7.66]	[7.29]	[3.93]	[1.49]
R^2_{adj} (%)	0.57	1.53	2.43	5.68	8.96	12.28	30.85	51.27
				X=Interest Rate				
β_{IR}	-7.70	-7.42	-6.64	-6.31	-6.45	-6.49	-6.63	-3.93
s.e.	[7.60]	[7.48]	[7.58]	[7.44]	[6.97]	[5.88]	[3.19]	[3.83]
R^2_{adj} (%)	0.00	0.13	0.21	0.53	0.91	1.29	5.48	3.52
				X=DFSP				
β_{DFSP}	3.75	3.14	3.70	5.39	5.24	4.77	1.99	3.42
s.e.	[7.47]	[6.93]	[6.35]	[4.87]	[4.14]	[3.57]	[2.67]	[2.19]
R^2_{adj} (%)	-0.08	-0.04	0.11	1.04	1.55	1.74	1.09	6.95
				X=TMSP				
β_{TMSP}	0.34	0.33	0.40	0.46	1.22	1.69	3.05	2.53
s.e.	[2.36]	[2.19]	[2.14]	[1.89]	[1.73]	[1.64]	[1.32]	[0.81]
R^2_{adj} (%)	-0.20	-0.20	-0.19	-0.15	0.34	1.20	16.21	19.38

Table 6. Univariate Return Predictability: II

Univariate monthly return predictive regressions for the U.S. equity market, Jan 1975 to Dec 2022. S.E. are Newey-West HAC robust standard errors with 3 more lags than forecasting horizon. All LHS returns are annual in percent. Univariate regressions are run separately:

$$\frac{12}{h}\sum_{i=1}^{h}\left(r_{t+i}-r^{f}_{t+i}\right)=\beta_0+\beta_X X_t+\epsilon_{t+h}$$

	1m	2m	3m	6m	9m	1y	3y	5y
				X=VIX				
β_{VIX}	0.44	0.57	0.54	0.51	0.33	0.34	0.02	-0.01
s.e.	[0.60]	[0.51]	[0.47]	[0.28]	[0.25]	[0.22]	[0.26]	[0.27]
R^2_{adj} (%)	0.15	1.05	1.51	2.86	2.42	2.34	-0.26	-0.29
				X=VRP				
β_{VRP}	0.54	0.42	0.39	0.19	0.09	0.06	-0.00	-0.03
s.e.	[0.09]	[0.09]	[0.06]	[0.05]	[0.05]	[0.05]	[0.02]	[0.02]
R^2_{adj} (%)	6.09	7.08	8.87	3.81	1.19	0.55	-0.30	1.02
				X=Risk-Neutral Skewness				
β_{Skew^Q}	-6.30	-5.73	-4.61	-4.49	-4.24	-3.66	-5.15	-5.25
s.e.	[9.23]	[8.65]	[8.46]	[7.01]	[6.49]	[6.24]	[3.27]	[2.59]
R^2_{adj} (%)	-0.10	0.03	0.03	0.30	0.47	0.43	4.32	8.53
				X=Risk-Neutral Kurtosis				
β_{Kurt^Q}	0.30	-0.66	-0.30	0.40	0.95	0.73	1.93	1.93
s.e.	[2.90]	[2.91]	[2.84]	[2.33]	[2.16]	[2.12]	[1.12]	[1.06]
R^2_{adj} (%)	-0.32	-0.29	-0.31	-0.29	-0.03	-0.10	4.35	7.25
				X=Crash-Probability				
$\beta_{Crash-Prob}$	0.13	1.19	1.14	0.62	0.67	0.58	0.56	0.23
s.e.	[1.44]	[1.16]	[1.07]	[0.85]	[0.77]	[0.74]	[0.56]	[0.53]
R^2_{adj} (%)	-0.32	0.35	0.61	0.20	0.54	0.51	1.64	0.29

Table 7. Univariate Return Predictability: III

Univariate monthly return predictive regressions for the U.S. equity market, Jan 1975 to Dec 2022. S.E. are Newey-West HAC robust standard errors with 3 more lags than forecasting horizon. All LHS returns are annual in percent. Univariate regressions are run separately:

$$\frac{12}{h}\sum_{i=1}^{h}\left(r_{t+i}-r_{t+i}^{f}\right)=\beta_0+\beta_X X_t+\epsilon_{t+h}$$

	1m	2m	3m	6m	9m	1y	3y	5y
				X=Inflation				
$\beta_{Inflation}$	-24.72	14.45	-14.22	-17.84	-23.37	-21.87	-9.43	-5.95
s.e.	[18.27]	[17.71]	[16.47]	[10.07]	[10.45]	[9.64]	[3.38]	[3.95]
R^2_{adj} (%)	0.23	0.09	1.41	0.23	1.11	3.07	3.59	2.40
				X=Baker-Wurgler Sentiment				
$\beta_{Sentiment^{BW}}$	-4.87	-4.56	-4.53	-4.26	-4.10	-3.55	-0.02	0.15
s.e.	[2.45]	[2.29]	[2.21]	[2.10]	[2.13]	[2.24]	[1.32]	[1.04]
R^2_{adj} (%)	0.45	0.88	1.41	2.68	3.74	3.67	-0.19	-0.15
				X=PLS Sentiment				
$\beta_{Sentiment^{PLS}}$	-7.87	-7.41	-6.89	-5.69	-4.97	-4.45	-1.39	-0.17
s.e.	[1.97]	[1.79]	[1.70]	[1.81]	[1.96]	[2.06]	[1.42]	[0.95]
R^2_{adj} (%)	1.82	3.20	4.31	6.14	7.04	7.53	2.67	-0.12
				X=BEX Risk Aversion				
$\beta_{RA^{BEX}}$	5.34	6.53	5.71	6.17	5.29	4.82	1.37	1.61
s.e.	[7.79]	[6.44]	[5.89]	[2.95]	[2.14]	[1.74]	[1.81]	[1.66]
R^2_{adj} (%)	0.21	1.07	1.29	3.37	3.69	4.11	0.86	2.62
				X=ICC				
β_{ICC}	0.39	0.51	0.48	0.55	0.56	0.57	0.61	0.55
s.e.	[0.86]	[0.86]	[0.87]	[0.86]	[0.79]	[0.69]	[0.46]	[0.38]
R^2_{adj} (%)	-0.15	-0.03	0.02	0.39	0.72	1.10	5.16	7.79

Table 8. Bivariate Return Predictability: I

Bivariate monthly return predictive regressions for the U.S. equity market, Jan 1975 to Dec 2022. S.E. are Newey-West HAC robust standard errors with 3 more lags than forecasting horizon. All LHS returns are annual in percent. Bivariate regressions are run separately in which $\log G/O$ is always included:

$$\frac{12}{h}\sum_{i=1}^{h}\left(r_{t+i}-r_{t+i}^{f}\right)=\beta_0+\beta_{\log G/O}\log G_t/O_t+\beta_X X_t+\epsilon_{t+h}$$

	1m	2m	3m	6m	9m	1y	3y	5y
$\beta_{\log G/O}$	17.22	14.84	14.52	12.31	10.36	8.59	2.58	0.20
s.e.	[8.22]	[7.90]	[7.87]	[7.39]	[6.26]	[5.22]	[2.36]	[1.83]
$\beta_{\log G/P}$	5.70	6.77	7.09	8.87	10.18	11.48	12.54	11.93
s.e.	[7.73]	[7.34]	[7.25]	[7.25]	[7.13]	[6.26]	[3.00]	[1.74]
R^2_{adj} (%)	1.51	2.74	4.38	8.57	11.94	15.19	28.31	31.74
$\beta_{\log G/O}$	18.08	16.09	16.03	14.61	13.24	12.15	7.20	4.48
s.e.	[7.17]	[7.09]	[7.19]	[7.29]	[6.56]	[5.78]	[3.53]	[2.15]
$\beta_{\log PD}$	-13.84	-14.85	-14.42	-15.29	-15.98	-16.32	-13.87	-14.12
s.e.	[9.90]	[9.41]	[9.02]	[8.02]	[7.39]	[6.73]	[3.91]	[1.59]
R^2_{adj} (%)	1.74	3.43	5.43	10.88	15.39	19.56	39.57	57.63
$\beta_{\log G/O}$	20.54	18.65	18.46	17.14	15.92	14.85	8.93	5.89
s.e.	[6.53]	[6.39]	[6.53]	[6.85]	[6.32]	[5.75]	[3.75]	[2.60]
$\beta_{I.R.}$	-9.42	-8.97	-8.17	-7.71	-7.67	-7.61	-8.10	-4.94
s.e.	[7.27]	[7.12]	[7.24]	[7.26]	[6.96]	[6.06]	[3.74]	[4.05]
R^2_{adj} (%)	1.66	2.91	4.49	8.19	10.83	12.89	19.88	15.36
$\beta_{\log G/O}$	19.38	17.82	17.69	16.27	14.94	13.94	7.86	5.23
s.e.	[6.82]	[6.75]	[6.85]	[6.90]	[6.32]	[5.79]	[3.79]	[2.54]
$\beta_{Inflation}$	-20.72	-10.76	-10.58	-14.43	-20.44	-19.20	-8.57	-5.41
s.e.	[17.60]	[17.28]	[15.87]	[9.30]	[9.68]	[8.91]	[3.74]	[4.07]
R^2_{adj} (%)	1.69	2.61	4.14	7.99	11.77	13.77	13.72	11.21

Table 9. Bivariate Return Predictability: II

Bivariate monthly return predictive regressions for the U.S. equity market, Jan 1975 to Dec 2022. S.E. are Newey-West HAC robust standard errors with 3 more lags than forecasting horizon. All LHS returns are annual in percent. Bivariate regressions are run separately in which $\log G/O$ is always included:

$$\frac{12}{h}\sum_{i=1}^{h}\left(r_{t+i}-r_{t+i}^{f}\right)=\beta_0+\beta_{\log G/O}\log G_t/O_t+\beta_X X_t+\epsilon_{t+h}$$

	1m	2m	3m	6m	9m	1y	3y	5y
$\beta_{\log G/O}$	19.94	18.09	17.94	16.60	15.41	14.37	7.96	5.24
s.e.	[6.77]	[6.64]	[6.79]	[7.08]	[6.53]	[5.96]	[3.77]	[2.60]
β_{DFSP}	3.42	2.85	3.41	5.12	4.99	4.56	1.88	3.35
s.e.	[7.32]	[6.79]	[6.21]	[4.87]	[4.23]	[3.65]	[2.38]	[1.96]
R^2_{adj} (%)	1.49	2.57	4.15	8.25	10.87	12.63	12.72	16.47
$\beta_{\log G/O}$	19.63	17.88	17.61	16.08	15.10	14.33	7.80	4.78
s.e.	[6.92]	[6.80]	[7.00]	[7.35]	[6.86]	[6.24]	[3.34]	[2.40]
β_{TMSP}	0.28	0.27	0.35	0.43	1.19	1.67	2.88	2.40
s.e.	[2.33]	[2.13]	[2.06]	[1.79]	[1.62]	[1.48]	[1.09]	[0.85]
R^2_{adj} (%)	1.48	2.63	4.10	7.13	9.96	12.91	27.56	27.31
$\beta_{\log G/O}$	17.62	17.90	18.28	20.48	20.30	20.42	14.44	9.00
s.e.	[7.19]	[7.57]	[8.13]	[9.51]	[8.77]	[7.78]	[2.79]	[2.28]
β_{VRP}	0.51	0.39	0.35	0.15	0.05	0.02	-0.04	-0.05
s.e.	[0.09]	[0.08]	[0.06]	[0.05]	[0.05]	[0.05]	[0.02]	[0.02]
R^2_{adj} (%)	7.26	9.63	12.96	14.11	16.58	20.88	28.79	21.00

Table 10. Bivariate Return Predictability: III

Bivariate monthly return predictive regressions for the U.S. equity market, Jan 1975 to Dec 2022. S.E. are Newey-West HAC robust standard errors with 3 more lags than forecasting horizon. All LHS returns are annual in percent. Bivariate regressions are run separately in which $\log G/O$ is always included:

$$\frac{12}{h}\sum_{i=1}^{h}\left(r_{t+i}-r^{f}_{t+i}\right)=\beta_0+\beta_{\log G/O}\log G_t/O_t+\beta_X X_t+\epsilon_{t+h}$$

	1m	2m	3m	6m	9m	1y	3y	5y
$\beta_{\log G/O}$	25.77	24.29	24.29	23.05	21.95	21.15	14.02	9.10
s.e.	[7.85]	[7.80]	[8.17]	[8.59]	[7.67]	[6.84]	[2.91]	[2.22]
β_{VIX}	0.37	0.50	0.46	0.44	0.32	0.27	0.06	0.02
s.e.	[0.55]	[0.45]	[0.42]	[0.24]	[0.22]	[0.20]	[0.22]	[0.22]
R^2_{adj} (%)	3.03	6.25	9.61	17.16	21.67	25.83	26.98	21.02
$\beta_{\log G/O}$	32.26	31.96	33.14	31.24	28.72	26.87	14.13	5.59
s.e.	[7.66]	[6.85]	[7.16]	[8.30]	[7.61]	[6.69]	[4.21]	[4.28]
β_{Skew^Q}	2.42	2.90	4.34	3.95	3.51	3.59	-1.67	-3.82
s.e.	[8.49]	[7.65]	[7.24]	[5.76]	[5.52]	[5.29]	[3.62]	[3.18]
R^2_{adj} (%)	3.98	7.89	13.17	22.48	27.24	30.42	24.82	14.91
$\beta_{\log G/O}$	32.66	32.82	33.30	31.01	28.15	26.34	14.03	5.84
s.e.	[7.84]	[7.16]	[7.38]	[8.29]	[7.41]	[6.53]	[3.98]	[4.20]
β_{Kurt^Q}	-1.93	-2.91	-2.58	-1.73	-0.98	-1.07	0.91	1.41
s.e.	[2.54]	[2.38]	[2.21]	[1.58]	[1.51]	[1.46]	[1.32]	[1.40]
R^2_{adj} (%)	4.10	8.46	13.67	22.70	27.04	30.21	25.36	14.45
$\beta_{\log G/O}$	32.82	30.56	31.38	30.39	27.81	25.98	14.56	6.84
s.e.	[8.65]	[8.32]	[8.51]	[8.58]	[7.66]	[6.90]	[3.25]	[3.76]
$\beta_{Crash-Prob}$	-0.85	0.28	0.19	-0.29	-0.16	-0.19	0.45	0.16
s.e.	[1.35]	[1.09]	[0.97]	[0.69]	[0.65]	[0.65]	[0.39]	[0.45]
R^2_{adj} (%)	4.12	7.85	12.92	22.15	26.79	29.84	25.62	10.93

Table 11. Bivariate Return Predictability: IV

Bivariate monthly return predictive regressions for the U.S. equity market, Jan 1975 to Dec 2022. S.E. are Newey-West HAC robust standard errors with 3 more lags than forecasting horizon. All LHS returns are annual in percent. Bivariate regressions are run separately in which $\log G/O$ is always included:

$$\frac{12}{h}\sum_{i=1}^{h}\left(r_{t+i}-r_{t+i}^{f}\right)=\beta_0+\beta_{\log G/O}\log G_t/O_t+\beta_X X_t+\epsilon_{t+h}$$

	1m	2m	3m	6m	9m	1y	3y	5y
$\beta_{\log G/O}$	20.03	18.13	18.18	16.93	15.78	14.73	8.01	5.32
s.e.	[6.54]	[6.40]	[6.50]	[6.72]	[6.16]	[5.59]	[3.76]	[2.53]
$\beta_{Sentiment^{BW}}$	-5.03	-4.70	-4.67	-4.38	-4.25	-3.72	-0.07	0.11
s.e.	[2.26]	[2.09]	[1.99]	[1.88]	[1.90]	[1.99]	[1.14]	[0.94]
R^2_{adj} (%)	2.06	3.54	5.59	10.19	13.53	15.11	11.59	9.64
$\beta_{\log G/O}$	14.92	13.40	13.62	13.62	13.12	12.64	7.49	5.59
s.e.	[6.67]	[6.59]	[6.74]	[7.19]	[6.75]	[6.20]	[4.02]	[2.37]
$\beta_{Sentiment^{PLS}}$	-6.46	-6.14	-5.60	-4.40	-3.73	-3.26	-0.68	0.37
s.e.	[1.93]	[1.74]	[1.66]	[1.84]	[2.02]	[2.07]	[1.29]	[0.72]
R^2_{adj} (%)	2.60	4.49	6.47	10.78	13.55	15.60	12.22	9.96
$\beta_{\log G/O}$	19.09	18.13	18.51	18.05	16.96	15.73	9.31	6.15
s.e.	[7.51]	[7.43]	[7.60]	[7.81]	[7.13]	[6.51]	[3.99]	[2.64]
$\beta_{RA^{BEX}}$	4.89	6.61	5.27	5.73	4.85	4.40	1.53	1.72
s.e.	[7.34]	[5.99]	[5.42]	[2.56]	[1.92]	[1.60]	[1.63]	[1.42]
R^2_{adj} (%)	1.85	4.12	6.34	13.37	16.88	19.27	16.87	15.84
$\beta_{\log G/O}$	14.97	13.10	13.12	13.15	12.13	11.34	7.27	4.38
s.e.	[7.20]	[7.30]	[7.53]	[7.93]	[7.07]	[6.28]	[3.57]	[2.51]
β_{ICC}	-0.00	0.16	0.13	0.20	0.24	0.27	0.41	0.43
s.e.	[0.86]	[0.87]	[0.88]	[0.87]	[0.78]	[0.65]	[0.39]	[0.34]
R^2_{adj} (%)	0.63	1.17	1.96	4.50	6.02	7.38	14.67	14.01

Table 12. G/O and Other Variables

The table reports results from regressing variable X_t onto $\log G_t/O_t$, monthly from Jan 1975 to Dec 2022. S.E. are Newey-West HAC robust standard errors with 10 lags.

$$X_t = \beta_0 + \beta_{\log G/O} \log G_t/O_t + \epsilon_t$$

$X =$	$Sentiment^{PLS}$	ICC	$Skew^Q$	$Kurt^Q$	$\log G/P$	VRP
$\beta_{\log G/O}$	-0.67	1.99	-0.33	0.62	0.49	8.76
s.e.	[0.35]	[0.66]	[0.12]	[0.33]	[0.11]	[4.79]
R^2_{adj} (%)	6.12	5.08	8.67	3.96	24.90	1.40

Table 13. A Three-Factor Model

Multivariate monthly return predictive regressions for the U.S. equity market, Jan 1975 to Dec 2022. S.E. are Newey-West HAC robust standard errors with 3 more lags than forecasting horizon. All LHS returns are annual in percent. Regression:

$$\frac{12}{h}\sum_{i=1}^{h}\left(r_{t+i} - r^f_{t+i}\right) = \beta_0 + \beta_{\log G/O} \log G_t/O_t + \beta_{Sentiment^{BW}} Sentiment^{BW}_t + \beta_{Skew^Q} Skew^Q_t + \epsilon_{t+h}$$

	1m	2m	3m	6m	9m	1y	3y	5y
$\beta_{\log G/O}$	31.62	31.33	32.51	30.45	27.83	25.97	13.49	5.07
s.e.	[7.65]	[6.83]	[7.12]	[8.15]	[7.34]	[6.22]	[4.17]	[4.00]
$\beta_{Sentiment^{BW}}$	-8.40	-8.39	-8.38	-10.43	-11.61	-11.80	-3.61	-3.08
s.e.	[3.94]	[3.32]	[2.85]	[2.27]	[1.96]	[1.60]	[1.51]	[2.59]
β_{Skew^Q}	4.85	5.33	6.77	6.97	6.88	7.01	0.05	-2.26
s.e.	[8.68]	[7.63]	[6.92]	[4.49]	[3.12]	[2.31]	[3.23]	[3.12]
R^2_{adj} (%)	4.76	9.66	16.04	31.38	43.16	51.53	29.81	22.66

Table 14. Daily Regressions

Daily return predictive regressions for the U.S. equity market, Jan 1986 to Dec 2022. S.E. are Newey-West HAC robust standard errors with 3 more lags than forecasting horizon. All LHS returns are annual in percent.

$$\frac{260}{h}\sum_{i=1}^{h}\left(r_{t+i}-r_{t+i}^{f}\right)=\beta_0+\beta_X X_t+\epsilon_{t+h}$$

	1d	2d	3d	4d	1w	2w
			$X = \log G/O$			
$\beta_{\log G/O}$	18.26	18.77	19.33	19.95	20.38	21.01
s.e.	[8.30]	[7.79]	[7.47]	[7.28]	[7.15]	[7.09]
R^2_{adj} (%)	0.05	0.12	0.19	0.28	0.38	0.69
			$X = Sentiment^{BW}$			
β_{BW}	-10.74	-10.71	-10.65	-10.63	-10.64	-10.65
s.e.	[5.26]	[4.98]	[4.82]	[4.71]	[4.63]	[4.56]
R^2_{adj} (%)	0.04	0.10	0.16	0.23	0.29	0.50

$$\frac{260}{h}\sum_{i=1}^{h}\left(r_{t+i}-r_{t+i}^{f}\right)=\beta_0+\beta_{\log G/O}\log G_t/O_t+\beta_{BW}Sentiment_t^{BW}+\epsilon_{t+h}$$

	1d	2d	3d	4d	1w	2w
$\beta_{\log G/O}$	17.95	18.44	18.97	19.58	20.04	20.78
s.e.	[8.28]	[7.77]	[7.44]	[7.25]	[7.11]	[7.04]
β_{BW}	-10.54	-10.51	-10.45	-10.42	-10.42	-10.43
s.e.	[5.23]	[4.95]	[4.79]	[4.68]	[4.59]	[4.51]
R^2_{adj} (%)	0.09	0.22	0.36	0.51	0.67	1.19

Table 15. Bivariate Return Predictability: Gold and Oil

Univariate monthly return predictive regressions for the U.S. equity market, Jan 1975 to Dec 2022. S.E. are Newey-West HAC robust standard errors with 3 more lags than forecasting horizon. All LHS returns are annual in percent. Regression:

$$\frac{12}{h}\sum_{i=1}^{h}\left(r_{t+i}-r^f_{t+i}\right)=\beta_0+\beta_{\log G}\log G_t+\beta_{\log O}\log O_t+\epsilon_{t+h}$$

	1m	2m	3m	6m	9m	1y	3y	5y
$\beta_{\log G}$	20.77	19.23	19.03	17.54	16.09	15.24	11.53	9.12
s.e.	[7.97]	[7.57]	[7.51]	[7.56]	[7.30]	[7.20]	[5.08]	[3.19]
$\beta_{\log O}$	-19.71	-17.70	-17.58	-16.38	-15.29	-14.14	-6.42	-3.60
s.e.	[6.91]	[6.83]	[6.97]	[7.29]	[6.70]	[5.97]	[3.83]	[2.78]
R^2_{adj} (%)	1.41	2.48	3.94	7.19	9.32	10.95	17.51	22.79

Table 16. Univariate Dividend Growth Predictability: Gold and Oil

Univariate monthly dividend growth predictive regressions for the U.S. equity market, Jan 1975 to Dec 2022. S.E. are Newey-West HAC robust standard errors with 3 more lags than forecasting horizon. All LHS dividend growth rates are annual in percent. Univariate regressions are run separately:

$$\frac{12}{h}\sum_{i=1}^{h}\Delta d_{t+i} = \beta_0 + \beta_X X_t + \epsilon_{t+h}$$

	1m	2m	3m	6m	9m	1y	3y	5y
				$X = \log G/O$				
$\beta_{\log G/O}$	-1.91	-1.83	-1.77	-1.38	-0.85	-0.17	2.24	0.44
s.e.	[1.39]	[1.52]	[1.63]	[1.96]	[2.26]	[2.54]	[2.54]	[1.81]
R^2_{adj} (%)	0.77	0.73	0.69	0.39	0.05	-0.17	0.28	0.05
				$X = \log G \cdot O$				
$\beta_{\log G \cdot O}$	2.10	2.11	2.13	2.13	2.10	2.01	1.45	1.36
s.e.	[0.67]	[0.70]	[0.73]	[0.80]	[0.88]	[0.96]	[1.17]	[0.90]
R^2_{adj} (%)	4.93	5.20	5.43	5.87	6.01	5.89	6.07	11.57
				$X = \log G$				
$\beta_{\log G}$	3.20	3.26	3.31	3.48	3.63	3.74	3.62	2.76
s.e.	[1.37]	[1.45]	[1.52]	[1.68]	[1.81]	[1.91]	[2.39]	[1.90]
R^2_{adj} (%)	2.92	3.16	3.37	4.05	4.70	5.37	10.03	12.51
				$X = \log O$				
$\beta_{\log O}$	3.62	3.61	3.61	3.49	3.27	2.93	1.49	1.89
s.e.	[1.10]	[1.16]	[1.22]	[1.38]	[1.59]	[1.82]	[2.23]	[1.55]
R^2_{adj} (%)	5.10	5.30	5.46	5.48	5.09	4.34	2.03	7.45

Table 17. Univariate Dividend Growth Predictability: Gold and Oil, Post-1986

Univariate monthly dividend growth predictive regressions for the U.S. equity market, Jan 1986 to Dec 2022. S.E. are Newey-West HAC robust standard errors with 3 more lags than forecasting horizon. All LHS dividend growth rates are annual in percent. Univariate regressions are run separately:

$$\frac{12}{h}\sum_{i=1}^{h}\Delta d_{t+i} = \beta_0 + \beta_X X_t + \epsilon_{t+h}$$

	1m	2m	3m	6m	9m	1y	3y	5y
				$X = \log G/O$				
$\beta_{\log G/O}$	-1.32	-1.22	-1.16	-0.78	-0.26	0.39	2.80	0.87
s.e.	[1.50]	[1.63]	[1.75]	[2.09]	[2.41]	[2.72]	[2.64]	[1.65]
R^2_{adj} (%)	0.24	0.19	0.16	-0.04	-0.21	-0.18	4.47	0.75
				$X = \log G \cdot O$				
$\beta_{\log G \cdot O}$	2.78	2.88	2.97	3.17	3.26	3.25	2.92	2.38
s.e.	[1.20]	[1.24]	[1.27]	[1.31]	[1.37]	[1.45]	[1.40]	[0.97]
R^2_{adj} (%)	5.63	6.31	6.95	8.52	9.59	10.15	15.76	23.17
				$X = \log G$				
$\beta_{\log G}$	4.24	4.48	4.70	5.31	5.81	6.20	6.70	4.64
s.e.	[2.14]	[2.23]	[2.29]	[2.38]	[2.39]	[2.37]	[2.34]	[2.22]
R^2_{adj} (%)	3.48	4.09	4.65	6.45	8.27	10.10	23.78	25.28
				$X = \log O$				
$\beta_{\log O}$	4.01	4.10	4.19	4.26	4.15	3.85	2.69	2.79
s.e.	[1.77]	[1.85]	[1.91]	[2.10]	[2.37]	[2.69]	[3.26]	[2.10]
R^2_{adj} (%)	4.71	5.12	5.51	6.16	6.19	5.66	4.86	11.85

Table 18. Quarterly Dividend Growth Predictability: Gold and Oil

Univariate quarterly dividend growth predictive regressions for the U.S. equity market, 1975-Q1 to 2022-Q4. S.E. are Newey-West HAC robust standard errors with 3 more lags than forecasting horizon. Δd_{t+i} is quarterly dividend growth. All LHS variables are annual in percent. Univariate regressions are run separately:

$$\frac{4}{h}\sum_{i=1}^{h}\Delta d_{t+i} = \beta_0 + \beta_X X_t + \epsilon_{t+h}$$

	1Q	2Q	3Q	1y	3y	5y
			$X = \log G/O$			
$\beta_{\log G/O}$	2.50	0.12	1.55	1.47	-0.63	-2.45
s.e.	[5.84]	[3.42]	[3.32]	[3.46]	[2.76]	[2.13]
R^2 (%)	-0.47	-0.53	-0.37	-0.20	-0.39	5.30
			$X = \log G \cdot O$			
$\beta_{\log G \cdot O}$	5.98	2.97	2.89	1.96	0.99	0.98
s.e.	[2.29]	[1.48]	[1.48]	[1.47]	[1.30]	[0.90]
R^2 (%)	0.75	1.22	1.66	1.78	1.33	3.61
			$X = \log G$			
$\beta_{\log G}$	13.10	5.94	6.50	4.60	1.67	0.85
s.e.	[5.06]	[3.30]	[2.99]	[2.88]	[2.78]	[1.83]
R^2 (%)	1.03	1.24	2.27	2.71	0.80	0.21
			$X = \log O$			
$\beta_{\log O}$	7.12	3.92	3.34	2.11	1.56	2.13
s.e.	[3.79]	[2.39]	[2.62]	[2.72]	[2.12]	[1.31]
R^2 (%)	0.15	0.61	0.56	0.47	1.15	6.52

Table 19. Bivariate Dividend Growth Predictability: Gold and Oil

Bivariate monthly dividend growth predictive regressions for the U.S. equity market, Jan 1975 to Dec 2022. S.E. are Newey-West HAC robust standard errors with 3 more lags than forecasting horizon. All LHS dividend growth rates are annual in percent. Regression:

$$\frac{12}{h}\sum_{i=1}^{h}\Delta d_{t+i} = \beta_0 + \beta_{\log G}\log G_t + \beta_{\log O}\log O_t + \epsilon_{t+h}$$

	1m	2m	3m	6m	9m	1y	3y	5y
$\beta_{\log G}$	0.88	0.99	1.07	1.52	2.05	2.68	4.54	2.41
s.e.	[1.75]	[1.89]	[2.03]	[2.37]	[2.65]	[2.90]	[3.33]	[2.30]
$\beta_{\log O}$	3.13	3.06	3.01	2.64	2.13	1.44	-1.21	0.45
s.e.	[1.40]	[1.51]	[1.62]	[1.95]	[2.29]	[2.63]	[2.70]	[1.82]
R^2_{adj} (%)	5.08	5.31	5.51	5.79	5.85	5.84	10.66	12.58

Table 20. Return Predictability: Controlling For Expected Dividend Growth

Bivariate monthly dividend growth predictive regressions for the U.S. equity market, Jan 1975 to Dec 2022. S.E. are Newey-West HAC robust standard errors with 3 more lags than forecasting horizon. All LHS returns are annual in percent. Bivariate regression:

$$\frac{12}{h}\sum_{i=1}^{h}(r^M_{t+i} - r^f_{t+i}) = \beta_0 + \beta_X X_t + \beta_{\log G\cdot O}(\log G_t \cdot O_t) + \epsilon_{t+h}$$

	1m	2m	3m	6m	9m	1y	3y	5y
				$X = \log G$				
$\beta_{\log G}$	40.48	36.93	36.61	33.92	31.39	29.38	17.94	12.71
s.e.	[13.61]	[13.23]	[13.39]	[13.82]	[12.93]	[12.14]	[8.03]	[5.26]
$\beta_{\log G\cdot O}$	-19.71	-17.70	-17.58	-16.38	-15.29	-14.14	-6.42	-3.60
s.e.	[6.91]	[6.83]	[6.97]	[7.29]	[6.70]	[5.97]	[3.83]	[2.78]
R^2_{adj} (%)	1.41	2.48	3.94	7.19	9.32	10.95	17.51	22.79
				$X = \log O$				
$\beta_{\log O}$	-40.48	-36.93	-36.61	-33.92	-31.39	-29.38	-17.94	-12.71
s.e.	[13.61]	[13.23]	[13.40]	[13.82]	[12.93]	[12.14]	[8.03]	[5.26]
$\beta_{\log G\cdot O}$	20.77	19.23	19.03	17.54	16.09	15.24	11.53	9.12
s.e.	[7.97]	[7.57]	[7.51]	[7.56]	[7.30]	[7.20]	[5.08]	[3.19]
R^2_{adj} (%)	1.41	2.48	3.94	7.19	9.32	10.95	17.51	22.79

Table 21. Fama-Macbeth and Panel Regressions

Panel A shows Fama-Macbeth regression results. Test assets are Fama-French 100 size and book-to-market portfolios. Data are monthly from Jan 1975 to Dec 2022. We first use a rolling window of 180 months to update each portfolio's exposure to monthly AR(1) innovation in $\log G/O$ and $\log G \cdot O$. We then use exposure estimates as RHS variables in cross-sectional regressions. t-stat are based on Newey-West robust standard errors. Average R-squared across periods are reported. Panel B shows panel regression results in which we regress portfolio returns onto one-period lagged portfolio factor exposures. We include time and portfolio fixed effects and cluster standard errors by time.

Panel A: Fama-Macbeth Regressions			
$\lambda_{\Delta \log G/O}$	-0.14		-0.12
t-stat	[-3.48]		[-2.90]
$\lambda_{\Delta \log G \cdot O}$		0.10	0.07
t-stat		[2.88]	[2.28]
average R^2 (%)	8.23	6.74	12.56
Panel B: Panel Regressions			
$\lambda_{\Delta \log G/O}$	-0.09		-0.14
t-stat	[-4.91]		[-4.33]
$\lambda_{\Delta \log G \cdot O}$		0.16	0.14
t-stat		[8.83]	[7.21]
R^2 (%)	48.78	49.42	49.72

Table 22. Variance Decomposition: Model I

Panel A			
	Dis rate (μ_t)	Div growth (g_t)	Covariance
$Var(\log G)$	48%	154%	-101%
$Var(\log O)$	7%	68%	26%
$Var(\log G/O)$	91%	1%	9%
$Var(\log G \cdot O)$	1%	113%	-14%
Panel B			
$Var(r)$	Expected (μ_t)	Unexpected ($\varepsilon_{r,t}$)	
100%	1.9%	98.1%	
$Var(\Delta d)$	Expected (g_t)	Unexpected ($\varepsilon_{d,t}$)	
100%	8.4%	91.6%	

Table 23. Univariate Predictability Using Filtered States: Model I

Using filtered states to predict returns and dividend growth, monthly from Jan 1975 to Dec 2022. S.E. are Newey-West HAC robust standard errors with 3 more lags than forecasting horizon. All LHS variables are annual in percent. Univariate regressions are run separately:

$$\frac{12}{h}\sum_{i=1}^{h}\left(r_{t+i}-r^f_{t+i}\right)=\beta_0+\beta_\mu\mu_t+\epsilon_{t+h}$$

$$\frac{12}{h}\sum_{i=1}^{h}\Delta d_{t+i}=\beta_0+\beta_g g_t+\epsilon_{t+h}$$

	1m	2m	3m	6m	9m	1y	3y	5y
				Returns				
β_μ	1.21	1.10	1.09	1.01	0.94	0.88	0.51	0.36
s.e.	[0.40]	[0.40]	[0.40]	[0.42]	[0.39]	[0.36]	[0.23]	[0.15]
R^2 (%)	1.58	2.65	4.11	7.35	9.47	11.11	13.36	12.11
				Dividend growth				
β_g	1.00	1.00	1.01	0.99	0.95	0.88	0.56	0.59
s.e.	[0.31]	[0.32]	[0.34]	[0.37]	[0.42]	[0.47]	[0.57]	[0.41]
R^2 (%)	5.23	5.47	5.67	5.92	5.80	5.33	4.02	9.85

Table 24. Univariate Return Predictability: Model I

Univariate monthly return predictive regressions under data simulated from the model. The model is simulated 1,000 times at estimated parameters. We report average OLS slope coefficients and R^2 across simulations. S.E. are standard deviations of coefficients across simulations. All LHS returns are annual in percent. Univariate regressions are run separately:

$$\frac{12}{h}\sum_{i=1}^{h} r_{t+i} = \beta_0 + \beta_X X_t + \epsilon_{t+h}$$

	1m	2m	3m	6m	9m	1y	3y	5y
				$X = \log G/O$				
$\beta_{\log G/O}$	15.83	15.55	15.28	14.50	13.79	13.10	9.05	6.48
s.e.	[5.40]	[5.40]	[5.35]	[5.30]	[5.25]	[5.22]	[5.05]	[4.74]
R^2 (%)	1.82	3.46	4.93	8.54	11.20	13.18	17.54	15.80
				$X = \log G \cdot O$				
$\beta_{\log G \cdot O}$	-4.98	-4.86	-4.74	-4.40	-4.10	-3.82	-2.27	-1.36
s.e.	[4.04]	[4.03]	[4.04]	[4.05]	[4.06]	[4.08]	[4.21]	[4.04]
R^2 (%)	0.66	1.25	1.79	3.16	4.24	5.12	8.85	10.54
				$X = \log G$				
$\beta_{\log G}$	3.92	3.94	3.98	3.98	3.98	3.96	3.61	3.12
s.e.	[11.4]	[11.4]	[11.4]	[11.4]	[11.4]	[11.3]	[10.9]	[10.5]
R^2 (%)	0.35	0.67	0.99	1.87	2.68	3.41	7.88	10.96
				$X = \log O$				
$\beta_{\log O}$	-10.59	-10.38	-10.16	-9.56	-9.00	-8.49	-5.52	-3.69
s.e.	[5.35]	[5.34]	[5.31]	[5.22]	[5.14]	[5.10]	[4.99]	[4.72]
R^2 (%)	1.20	2.26	3.22	5.56	7.30	8.62	12.19	12.07

Table 25. Univariate Dividend Growth Predictability: Model I

Univariate monthly dividend growth predictive regressions under data simulated from the model. The model is simulated 1,000 times at estimated parameters. We report average OLS slope coefficients and R^2 across simulations. S.E. are standard deviations of coefficients across simulations. All LHS dividend growth rates are annual in percent. Univariate regressions are run separately:

$$\frac{12}{h}\sum_{i=1}^{h}\Delta d_{t+1} = \beta_0 + \beta_X X_t + \epsilon_{t+h}$$

	1m	2m	3m	6m	9m	1y	3y	5y
				$X = \log G/O$				
$\beta_{\log G/O}$	-2.78	-2.75	-2.73	-2.66	-2.60	-2.54	-2.09	-1.76
s.e.	[1.03]	[1.03]	[1.03]	[1.03]	[1.03]	[1.02]	[1.01]	[1.02]
R^2 (%)	3.46	6.33	8.75	14.17	17.79	20.25	25.30	24.09
				$X = \log G \cdot O$				
$\beta_{\log G \cdot O}$	2.33	2.31	2.29	2.23	2.17	2.11	1.71	1.39
s.e.	[0.44]	[0.44]	[0.44]	[0.44]	[0.45]	[0.45]	[0.49]	[0.53]
R^2 (%)	5.99	10.97	15.18	24.57	30.70	34.87	41.61	36.34
				$X = \log G$				
$\beta_{\log G}$	3.72	3.69	3.66	3.56	3.45	3.36	2.67	2.10
s.e.	[1.71]	[1.72]	[1.71]	[1.72]	[1.73]	[1.73]	[1.77]	[1.75]
R^2 (%)	3.18	5.78	7.97	12.78	15.86	17.97	21.80	19.78
				$X = \log O$				
$\beta_{\log O}$	3.22	3.20	3.17	3.09	3.01	2.94	2.39	1.96
s.e.	[0.61]	[0.60]	[0.60]	[0.60]	[0.61]	[0.62]	[0.66]	[0.72]
R^2 (%)	6.04	11.06	15.31	24.81	31.06	35.30	42.52	37.66

Table 26. Variance Decomposition: Model II

	Panel A		
	Dis rate (μ_t, θ_t)	Div growth (g_t)	Covariance
$Var(\log G)$	103%	181%	-183%
$Var(\log O)$	28%	27%	45%
$Var(\log G/O)$	180%	28%	-108%
$Var(\log G \cdot O)$	8%	102%	-11%
$Var(pd)$	107%	2%	-9%
	Panel B		
$Var(r)$	Expected ($\mu_t + \theta_t$)	Unexpected ($\varepsilon_{r,t}$)	
100%	1.7%	98.3%	
$Var(\Delta d)$	Expected (g_t)	Unexpected ($\varepsilon_{d,t}$)	
100%	4.9%	95.1%	

Table 27. Univariate Predictability Using Filtered States: Model II

Using filtered states to predict returns and dividend growth, monthly from Jan 1975 to Dec 2022. S.E. are Newey-West HAC robust standard errors with 3 more lags than forecasting horizon. All LHS variables are annual in percent. Univariate regressions are run separately:

$$\frac{12}{h}\sum_{i=1}^{h}\left(r_{t+i}-r^f_{t+i}\right)=\beta_0+\beta_{E^{Short}[r]}(\mu_t+\theta_t)+\epsilon_{t+h}$$

$$\frac{12}{h}\sum_{i=1}^{h}\left(r_{t+i}-r^f_{t+i}\right)=\beta_0+\beta_{E^{Long}[r]}(B_1\mu_t+B_2\theta_t)+\epsilon_{t+h}$$

$$\frac{12}{h}\sum_{i=1}^{h}\Delta d_{t+i}=\beta_0+\beta_g g_t+\epsilon_{t+h}$$

	1m	2m	3m	6m	9m	1y	3y	5y
Returns by short-term discount rate								
$\beta_{E^{Short}[r]}$	0.88	0.80	0.79	0.74	0.71	0.66	0.30	0.22
s.e.	[0.28]	[0.28]	[0.28]	[0.30]	[0.28]	[0.27]	[0.19]	[0.13]
R^2 (%)	1.47	2.51	3.84	7.01	9.57	11.22	8.79	8.73
Returns by long-term discount rate								
$\beta_{E^{Long}[r]}$	13.60	14.25	13.89	13.92	13.97	14.06	12.26	12.06
s.e.	[8.47]	[8.02]	[7.75]	[7.35]	[7.09]	[6.81]	[3.91]	[1.49]
R^2 (%)	0.26	0.96	1.58	3.67	5.72	8.01	25.74	44.33
Dividend growth								
β_g	1.00	1.01	1.02	1.02	1.00	0.94	0.60	0.61
s.e.	[0.31]	[0.32]	[0.33]	[0.36]	[0.40]	[0.44]	[0.56]	[0.40]
R^2 (%)	5.39	5.74	6.03	6.53	6.59	6.22	4.86	11.09

Table 28. Univariate Return Predictability: Model II

Univariate monthly return predictive regressions under data simulated from the model. The model is simulated 1,000 times at estimated parameters. We report average OLS slope coefficients and R^2 across simulations. S.E. are standard deviations of coefficients across simulations. All LHS returns are annual in percent. Univariate regressions are run separately:

$$\frac{12}{h}\sum_{i=1}^{h} r_{t+i} = \beta_0 + \beta_X X_t + \epsilon_{t+h}$$

	1m	2m	3m	6m	9m	1y	3y	5y
				$X = \log G/O$				
$\beta_{\log G/O}$	19.16	18.78	18.41	17.43	16.49	15.57	9.76	6.19
s.e.	[6.18]	[6.05]	[6.00]	[5.86]	[5.70]	[5.56]	[4.99]	[4.56]
R^2 (%)	1.68	3.22	4.63	8.19	10.90	12.88	16.19	13.36
				$X = \log G \cdot O$				
$\beta_{\log G \cdot O}$	-5.23	-5.02	-4.83	-4.26	-3.71	-3.19	-0.13	1.45
s.e.	[4.97]	[4.95]	[4.92]	[4.81]	[4.73]	[4.63]	[3.96]	[3.55]
R^2 (%)	0.40	0.75	1.07	1.83	2.36	2.74	4.36	7.31
				$X = \log G$				
$\beta_{\log G}$	8.15	8.23	8.29	8.53	8.78	8.97	9.65	9.31
s.e.	[9.16]	[9.14]	[9.20]	[9.15]	[9.01]	[8.92]	[8.23]	[7.61]
R^2 (%)	0.30	0.61	0.92	1.83	2.75	3.67	10.64	16.12
				$X = \log O$				
$\beta_{\log O}$	-13.89	-13.50	-13.15	-12.13	-11.16	-10.23	-4.55	-1.35
s.e.	[6.92]	[6.84]	[6.78]	[6.59]	[6.46]	[6.32]	[5.31]	[4.66]
R^2 (%)	1.02	1.93	2.74	4.68	6.00	6.85	6.64	5.69

Table 29. Bivariate Return Predictability: Model II

Bivariate monthly return predictive regressions under data simulated from the model. The model is simulated 1,000 times at estimated parameters. We report average OLS slope coefficients and R^2 across simulations. S.E. are standard deviations of coefficients across simulations. All LHS returns are annual in percent. Regression:

$$\frac{12}{h}\sum_{i=1}^{h} r_{t+i} = \beta_0 + \beta_{\log G/O} \log G_t/O_t + \beta_{pd} pd_t + \epsilon_{t+h}$$

	1m	2m	3m	6m	9m	1y	3y	5y
$\beta_{\log G/O}$	18.04	17.63	17.26	16.13	15.06	14.05	7.88	4.27
s.e.	[6.94]	[6.87]	[6.76]	[6.58]	[6.37]	[6.18]	[4.81]	[3.86]
β_{pd}	-14.67	-14.81	-14.89	-15.07	-15.23	-15.35	-15.37	-14.43
s.e.	[9.39]	[9.34]	[9.28]	[9.04]	[8.74]	[8.47]	[6.27]	[4.73]
R^2 (%)	2.33	4.53	6.60	12.19	16.95	21.06	41.04	51.90

Table 30. Univariate Dividend Growth Predictability: Model II

Univariate monthly dividend growth predictive regressions under data simulated from the model. The model is simulated 1,000 times at estimated parameters. We report average OLS slope coefficients and R^2 across simulations. S.E. are standard deviations of coefficients across simulations. All LHS dividend growth rates are annual in percent. Univariate regressions are run separately:

$$\frac{12}{h}\sum_{i=1}^{h}\Delta d_{t+1} = \beta_0 + \beta_X X_t + \epsilon_{t+h}$$

	1m	2m	3m	6m	9m	1y	3y	5y
				$X = \log G/O$				
$\beta_{\log G/O}$	-1.82	-1.79	-1.78	-1.73	-1.68	-1.63	-1.28	-1.03
s.e.	[1.03]	[1.03]	[1.03]	[1.02]	[1.02]	[1.00]	[0.95]	[0.89]
R^2 (%)	1.48	2.77	3.93	6.73	8.79	10.29	14.34	14.26
				$X = \log G \cdot O$				
$\beta_{\log G \cdot O}$	2.27	2.23	2.20	2.12	2.04	1.96	1.44	1.05
s.e.	[0.53]	[0.52]	[0.52]	[0.52]	[0.51]	[0.51]	[0.52]	[0.54]
R^2 (%)	4.15	7.73	10.85	18.08	23.03	26.39	30.73	24.94
				$X = \log G$				
$\beta_{\log G}$	3.04	3.00	2.95	2.81	2.69	2.58	1.79	1.21
s.e.	[1.35]	[1.35]	[1.36]	[1.37]	[1.37]	[1.37]	[1.38]	[1.41]
R^2 (%)	2.04	3.80	5.33	8.82	11.21	12.83	15.59	14.08
				$X = \log O$				
$\beta_{\log O}$	3.13	3.09	3.05	2.94	2.84	2.75	2.07	1.56
s.e.	[0.74]	[0.75]	[0.75]	[0.74]	[0.74]	[0.74]	[0.76]	[0.76]
R^2 (%)	3.97	7.40	10.40	17.41	22.29	25.67	30.82	25.82

Table 31. Bivariate Predictability, Gold and Oil: Model II

Bivariate monthly return predictive regressions under data simulated from the model. The model is simulated 1,000 times at estimated parameters. We report average OLS slope coefficients and R^2 across simulations. S.E. are standard deviations of coefficients across simulations. All LHS variables are annual in percent. Bivariate regressions:

$$\frac{12}{h}\sum_{i=1}^{h} r_{t+i} = \beta_0 + \beta_{\log G}\log G_t + \beta_{\log O}\log O_t + \epsilon_{t+h}$$

$$\frac{12}{h}\sum_{i=1}^{h} \Delta d_{t+i} = \beta_0 + \beta_{\log G}\log G_t + \beta_{\log O}\log O_t + \epsilon_{t+h}$$

	1m	2m	3m	6m	9m	1y	3y	5y
				Returns				
$\beta_{\log G}$	19.28	19.09	18.89	18.46	18.05	17.61	14.45	11.89
s.e.	[9.89]	[9.79]	[9.77]	[9.57]	[9.31]	[9.11]	[8.47]	[7.88]
$\beta_{\log O}$	-19.57	-19.10	-18.67	-17.48	-16.34	-15.23	-8.43	-4.46
s.e.	[6.70]	[6.57]	[6.50]	[6.31]	[6.15]	[6.00]	[5.24]	[4.61]
R^2 (%)	1.83	3.51	5.06	9.03	12.12	14.52	22.08	24.38
				Dividend growth				
$\beta_{\log G}$	1.47	1.44	1.41	1.32	1.24	1.17	0.69	0.33
s.e.	[1.22]	[1.21]	[1.21]	[1.22]	[1.22]	[1.22]	[1.28]	[1.35]
$\beta_{\log O}$	2.72	2.69	2.66	2.57	2.48	2.40	1.82	1.40
s.e.	[0.82]	[0.81]	[0.81]	[0.81]	[0.80]	[0.79]	[0.78]	[0.77]
R^2 (%)	4.45	8.30	11.68	19.57	25.09	28.91	35.85	32.02

Table 32. Return Predictability Controlling For Expected Dividend Growth: Model II

Bivariate monthly return predictive regressions under data simulated from the model. The model is simulated 1,000 times at estimated parameters. We report average OLS slope coefficients and R^2 across simulations. S.E. are standard deviations of coefficients across simulations. All LHS returns are annual in percent. Bivariate regression:

$$\frac{12}{h}\sum_{i=1}^{h} r_{t+i} = \beta_0 + \beta_X X_t + \beta_{\log G \cdot O}(\log G_t \cdot O_t) + \epsilon_{t+h}$$

	1m	2m	3m	6m	9m	1y	3y	5y
				$X = \log G$				
$\beta_{\log G}$	38.86	38.20	37.57	35.94	34.39	32.84	22.88	16.35
s.e.	[13.94]	[13.70]	[13.62]	[13.28]	[12.86]	[12.52]	[11.43]	[10.36]
$\beta_{\log G \cdot O}$	-19.57	-19.10	-18.67	-17.47	-16.34	-15.23	-8.43	-4.46
s.e.	[6.70]	[6.57]	[6.50]	[6.31]	[6.15]	[6.00]	[5.23]	[4.61]
R^2 (%)	1.83	0.80	5.06	9.03	12.12	14.52	22.08	24.38
				$X = \log O$				
$\beta_{\log O}$	-38.86	-38.20	-37.57	-35.93	-34.39	-32.84	-22.88	-16.36
s.e.	[13.94]	[13.71]	[13.63]	[13.29]	[12.86]	[12.52]	[11.43]	[10.36]
$\beta_{\log G \cdot O}$	19.28	19.09	18.89	18.46	18.05	17.61	14.45	11.89
s.e.	[9.89]	[9.79]	[9.77]	[9.57]	[9.31]	[9.11]	[8.47]	[7.88]
R^2 (%)	1.83	0.80	5.06	9.03	12.12	14.52	22.08	24.38

Table 33. PC Predictive Regressions

Using PCA to predict returns or dividend growth, monthly Jan 1975 to Dec 2022. S.E. are Newey-West HAC robust standard errors with 3 more lags than forecasting horizon. All LHS variables are annual in percent.

	1m	2m	3m	6m	9m	1y	3y	5y
$\frac{12}{h}\sum_{i=1}^{h}\Delta d_{t+i} = \beta_0 + \beta_1 PC(1)_t + \epsilon_{t+h}$								
β_1	2.23	2.26	2.29	2.34	2.32	2.25	1.38	1.19
s.e.	[0.56]	[0.59]	[0.61]	[0.66]	[0.72]	[0.79]	[1.08]	[0.82]
R^2 (%)	7.28	7.76	8.18	9.20	9.66	9.73	7.20	11.48
$\frac{12}{h}\sum_{i=1}^{h} r_{t+i} = \beta_0 + \beta_2 PC(2)_t + \epsilon_{t+h}$								
β_2	22.24	20.32	20.38	19.22	18.62	17.46	7.90	4.34
s.e.	[8.23]	[8.15]	[8.32]	[8.86]	[8.29]	[7.50]	[4.19]	[3.27]
R^2 (%)	1.41	2.28	3.55	6.44	9.04	10.67	7.43	4.22
$\frac{12}{h}\sum_{i=1}^{h} r_{t+i} = \beta_0 + \beta_5 PC(5)_t + \epsilon_{t+h}$								
β_5	19.80	20.69	20.56	23.79	22.74	24.06	27.48	22.56
s.e.	[15.71]	[13.78]	[12.75]	[11.75]	[12.40]	[12.12]	[8.99]	[5.62]
R^2 (%)	0.27	0.56	0.86	2.33	3.15	4.71	23.93	30.62