Stock Return Predictability in a Monetary Economy

Abraham Lioui∗ and Jesper Rangvid‡

November 2008

∗Department of Economics, Bar Ilan University, 52900 Ramat Gan, Israel. Tel: 972-5318940, liouia@mail.biu.ac.il.
‡Department of Finance, Copenhagen Business School, Solbjerg Plads 3, DK-2000 Frederiksberg, Denmark. Phone: (45) 3815 3615, fax: (45) 3815 3600, e-mail: jr.fi@cbs.dk.

We would like to thank Yakov Amihud, Doron Avramov, Greg Bauer, John Y. Campbell, Long Chen, Ilan Cooper, Hui Guo, Harrison Hong, Hanno Lustig, Luboš Pástor, Loriana Pelizzon, Astrid Schornick, Mike R. Wickens, participants at seminars at Copenhagen Business School, ESSEC, Frankfurt University, Haifa University, Stockholm University, and participants at the Arne Ryde Workshop, the Caesarea Center 4th Annual Conference, the D-CAF Return Predictability Conference, the CEPR European Summer Symposium in Financial Markets, the European Finance Association, and the FMA meeting for comments and suggestions.
Stock Return Predictability in a Monetary Economy

Abstract

In a purely real economy, only real variables determine expected returns. In a monetary economy, nominal variables also matter. In this paper, we allow money to be present in the economy. We set up an affine theoretical model that allows us to study the implications of money for the time-variation in expected returns. Our key theoretical result is that expected excess returns are determined by a financial ratio and the nominal interest rate. The nominal interest accounts for the nominal side of the economy. Empirically, we find strong evidence that the interest rate matters for excess return predictability together with a valuation ratio. We work with two main empirical specifications. One is based on the assumption that consumption is equal to dividends, which is true in a pure exchange economy without leverage. The other one allows consumption to differ from dividends, due to leverage. We find even stronger evidence for excess returns predictability in the latter specification.

Keywords: Monetary economy, short interest rate, return predictability

JEL-classification: E44, E52, G12
1 Introduction

Since Fama and French (1988, 1989) and Campbell and Shiller (1988a, 1988b) showed that excess returns are predictable, models that trace and examine the origins of return predictability in rational economies have been developed. These models specify the preferences of the investor/consumer, the relevant budget constraints, and the laws of motions of the underlying economic state variables in order to derive observable variables that contain information about the time-variation in expected excess returns; see, for instance, Menzly et al. (2004), Santos and Veronesi (2006), Lustig and van Nieuwerburgh (2005), and Piazzesi et al. (2007).

The models that are used to derive predictor variables on the basis of first principles are generally (i) models of real economies, i.e. economies that disregard the role of money, or (ii) models that only indirectly include nominal features, i.e. without explicitly modeling the nominal part of the economy. Reflecting on a similar observation, Cochrane (2006) notes:

...Having said ‘macroeconomics’, ‘risk’, and ‘asset prices’, the reader will quickly spot a missing ingredient: money.

Our paper introduces this missing ingredient to the return-predictability literature. We provide a theoretical and an empirical investigation. Theoretically, we find that in an economy where people hold money, expected excess returns are determined by a joint combination of a stock market valuation ratio and the nominal risk-free interest rate. Empirically, we show that the evidence on excess-return predictability is significantly improved if augmenting a valuation ratio with the short nominal interest rate when running predictive regressions.

Theory. We recognize that standard asset pricing equations can be written as predictive regressions. We also note, however, that some of the terms that enter such predictive regressions – the state variables – are unobservable. We assume in this paper that two unobservable state variables exist in the economy: A real state variable and, as the novel feature of our model, a nominal state variable. Using standard specifications for the law of motions of the state variables and the observable variables, we show that expected excess returns can be written as a linear combination of the two state variables. Moreover, the state variables themselves can be written as linear combinations of two observable variables: the deviation of the log share price to consumption ratio from its long-run mean and the deviation of the nominal interest rate from its long-run mean. Consequently,
we show that expected excess returns can be written as a linear combination of the deviations from the means of two observable variables: a valuation ratio (the share price to consumption ratio) and the nominal interest rate. We stress that the presence of the nominal interest rate is a direct consequence of the assumption that money plays a role in the economy and, hence, that there is a nominal state variable in the economy.

Having derived our basic predictive regression equation, we notice that in a pure exchange economy, consumption will be equal to output, i.e., one can use either the share price to consumption ratio or the share price to GDP ratio. Both ratios should be augmented by the interest rate when trying to capture expected excess returns in a monetary economy. Dividends will also be equal to consumption in a monetary exchange economy. Abel (1999) argues, though, that leverage in the economy can cause consumption to differ from dividends. When dealing with dividends, we thus take two approaches. First, we investigate what happens if augmenting the price-dividend ratio with the interest rate (i.e., we assume that consumption is equal to dividends in a monetary economy) and, second, we estimate a cointegration relation, which we call the \( \hat{pdR} \)-ratio, that relaxes the assumption that consumption is equal to dividends.

**Empirical findings.** We estimate the predictive regressions using post-1947 U.S. quarterly data. Our main empirical findings are as follows. First, we show that commonly used valuation ratios (the price-dividend, the price-consumption, and the price-output ratio) only have weak predictive ability for excess returns. In contrast, when we use both valuation ratios and the nominal interest rate, strong evidence of predictability exists. For instance, we find a \( R^2 \) of only 0.43% at a quarterly horizon when we predict with the price-consumption ratio (1.61% when we use the price-dividend ratio). When we augment the otherwise univariate regression with the interest rate, however, the \( R^2 \) increases to approximately 6.33% (4.83% if using the price-dividend ratio and the interest rate). The \( t \)-statistics also increase considerably. Finally, when we relax the assumption that consumption is equal to dividends, and use the estimated \( \hat{pdR} \)-ratio, we find even stronger evidence of predictability: an \( R^2 \) of 8.64% and a very high \( t \)-statistic (close to 5). Likewise, we do long-horizon tests, where we again find strong support for our theoretical predictions.

We do a number of robustness tests. For instance, we compare our new variable to other variables like the \( \hat{c\hat{q}} \)-ratio of Lettau & Ludvigson (2001), the total payout ratio of Boudoukh et al. (2007), the slope of the term structure, and the default spread. We find that the \( \hat{pdR} \)-ratio generally captures a considerably higher fraction of the variation of future excess returns than these other variables do. We also address the potential look-ahead bias that exists when using the estimated \( \hat{pdR} \)-ratio as a predictor by generating
true out-of-sample predictions. We find that the out-of-sample predictions from the $pdR$-ratio are generally better than the updated historical average that Goyal and Welch (2008) argue is “difficult to beat”. In these and other robustness tests, our overall result that one predicts significantly better if adding the interest rate to a valuation ratio carries through strongly.

In order to facilitate comparison to the literature, most of our in-sample predictability regressions are run under the assumption of a constant steady-state mean of the predictors. Lettau and van Nieuwerburgh (2008) introduced time-varying steady states in predictability regressions. Hence, we also conduct tests where we let the steady-state mean change. First, we confirm that the price-dividend ratio predicts much better if allowing for a change in its steady-state mean in 1991. We also find, however, that our main hypothesis (that adding the interest rate to a valuation ratio helps in capturing the variation in expected returns) goes through, also in this setting, which allows for a structural break in the mean of the price-dividend ratio (and in the mean of the other predictors).

To better understand what the shift in the mean of the steady state implies for our predictive regressions, we present results from recursive estimates of the predictive regressions. We show that the predictive coefficient to the $pdR$-ratio is considerably more stable than the predictive coefficient is to the standard price-dividend ratio. Interestingly, the reason why the predictive coefficient to the $pdR$-ratio is stable is that the interest rate matters more (the coefficient to the interest rate in the $pdR$-ratio increases) exactly during the periods where the price-dividend ratio loses its predictive ability. This result highlights the importance of the nominal interest rate in predictive regressions.

As the final empirical issue, we evaluate how the structural parameters of our model relates to the behavior of returns. We use the framework of Hansen-Jagannathan (1991) to show that the stochastic discount factors we propose are volatile enough to account for the empirical behavior of returns, even for reasonable values of the preference parameters. We show that introducing money helps explain the equity premium puzzle for reasonable values of the preference parameters by bringing up the volatility of the stochastic discount factor. This helps us understand why the introduction of money improves upon the prediction of volatile excess returns.

Related literature. The paper most closely related to ours is Ang and Bekaert (2007); AB hereafter. AB show that the short-horizon predictive ability of the dividend yield is much improved when the interest rate is added to the otherwise univariate regression. There are three main differences between our model and AB’s. First, AB assume that the
nominal interest rate is a state variable in the economy and use non-linear present-value models to explain why the nominal interest rate matters for excess returns predictability, whereas we derive a predictive regression from first principles, i.e. our predictive regression is an endogenous equilibrium outcome of a model where agents have preferences for holding money. Second, AB study the price-dividend ratio (or, to be precise, the dividend yield) and the interest rate. In our model, a more natural valuation ratio is the price-consumption ratio or the price-output ratio, i.e. we study both the price-dividend ratio augmented by the interest rate, as do Ang and Bekaert, but also other valuation ratios augmented by the interest rate. We show that this is an empirically important distinction. Finally, we provide more positive evidence for long-horizon predictability than Ang and Bekaert (2007). The reason why we find more evidence of predictability than they do, even when AB, like us, use a combination of stock prices, dividends, and the nominal interest rate, is that we allow for leverage in the economy in the spirit of Abel (1999). Empirically, this is captured by a non-unitary estimate of the coefficient to dividends in the $\hat{pdR}$-ratio. Our estimate suggests a leverage of 2.225, which is close to the value of 2.74 that Abel (1999) uses.

Our paper is also related to the more recent literature that shows that a larger fraction of return variation can be captured when theoretical models are used to expand an information set otherwise consisting of the pure lagged price-dividend ratio only. For instance, Binsbergen and Koijen (2008) use a structural model to show that the whole history of price-dividend ratios and dividend growth rates should matter for returns. In this sense, Binsbergen and Koijen (2008) argue, like we do, that an expected return regression would suffer from an omitted variable problem if only the lagged price-dividend ratio is used as a predictive variable.

We should also stress that we are – of course – not the first to use the interest rate in a return-predicting regression. Indeed, initial evidence that expected returns carry a predictable component that can also be captured by the short interest rate, by interest rate spreads between government bonds of different maturities and corporate bonds of different quality, or by the relative interest rate has been provided previously by Fama and Schwert (1977), Campbell (1987, 1991), Keim and Stambaugh (1986), and Fama and French (1989). The point we would like to draw attention to in this paper is that the interest rate should naturally matter together with a valuation ratio (a linear combination of share prices and dividends) in a monetary economy.

Finally, it should be mentioned that Harvey (1989) considered the “dividend yield spread” as a predictor variable in his test of a conditional asset-pricing model. As the dividend yield spread combined the information in a valuation ratio with that of the short interest rate, the dividend yield spread is related to our proposed predictor variables.
Harvey (1989) implicitly constrained the coefficient to the interest rate and dividend to be equal to one; our results extend upon his findings by providing an economic motivation for why this coefficient could be different from one and by providing empirical evidence that the variation in expected returns can be better captured if allowing the coefficient to be different from one.

**Structure of the paper.** The outline of this paper as follows. In the next section, we derive the predictive relation that appears in an economy where investors have preferences for holding money and nominal shocks drive the economy. Sections 4 and 5 contain the empirical analysis, with section 4 describing the data we use and, in particular, the construction of the \( pdR \) cointegration residual, and section 5 containing the empirical evidence on predictability. Section 6 concludes.

### 2 Predictors and predictive regressions in a monetary economy

Our objective in this section is to explain why predictive regressions using financial valuation ratios should be augmented by the nominal interest rate in an economy where agents use money to carry out transactions.

We proceed in two steps. In section 2.1, we discuss the general link that exists between the pricing kernel implied by any asset pricing model and the expected excess return on a risky asset. In section 2.2, we use this general framework to show the predictors that one obtains in real and monetary economies.

#### 2.1 An asset pricing perspective on predictive regressions

Our starting point is the Euler equation for consumption of a representative investor who has a time-additive utility function:

\[
u'(c_t)p_t = E_t[\delta u'(c_{t+1}) (d_{t+1} + p_{t+1})]
\]

where \( p \) is a risky asset’s real price, \( d \) the asset’s real dividend, \( c \) real per capita consumption, \( \delta \) the subjective discount factor, and \( u() \) is the utility function.

Following the usual steps of, e.g. Cochrane (2005), and remembering that the real risk-free rate is given as \( 1 + r_{f,t+1}^R = 1/E_t[\delta u'(c_{t+1})/u'(c_t)] \), the expected excess return on an asset can be written as an asset pricing equation relating the asset’s expected excess
return to its systematic risk:

$$E_t [r_{t+1}] - r^f_t = - \left( 1 + r^f_t \right) \text{cov}_t \left[ \frac{\delta u' (c_{t+1})}{u' (c_t)} ; \frac{d_{t+1} + p_{t+1}}{p_t} \right].$$

(2)

The right hand side of this equation corresponds to the risk premium for the asset and its intuition is well understood; see, e.g. Cochrane (2005).

We introduce a slight modification of Eq. (2) and express it as:

$$E_t [r_{t+1}] - r^f_t = - \left( 1 + r^f_t \right) \text{cov}_t \left[ \delta u' (c_{t+1}) ; d_{t+1} + p_{t+1} \right] \frac{1}{p_t u' (c_t)}. $$

(3)

This way of writing Eq. (2) reveals its return-predicting properties, as Eq. (3) relates the expected excess return from time $t$ to time $t+1$ to three time $t$ terms: $\text{cov}_t [\cdot]$, $p_t$, and $u' (c_t)$.

The degrees to which these three variables can be observed in the data is very different, though. Indeed, $p_t$ is directly observable in the data and $u' (c_t)$ will be observable once the functional form of the preferences has been specified. The conditional covariance, on the other hand, is not directly observable, as it is a function of both observable variables, including $p_t$ and $u' (c_t)$, and non-observable state variables:

$$\text{cov}_t [\delta u' (c_{t+1}) ; d_{t+1} + p_{t+1}] = g [p_t, u' (c_t), x_t]$$

(4)

for some function $g[\cdot]$ and a vector of time $t$ state variables $x_t$.

It is difficult to identify the state variables contained in $x_t$. For this reason, we consider situations where the variables contained in $p_t$ and $u' (c_t)$ are sufficient statistics for the unobservable variables contained in $x_t$. In this case, we write Eq. (4) as:

$$\text{cov}_t [\delta u' (c_{t+1}) ; d_{t+1} + p_{t+1}] = h [p_t, u' (c_t)]$$

(5)

for some function $h[\cdot]$, yielding a predictive relation:

$$E_t [r_{t+1}] - r^f_t = k [p_t, u' (c_t)]$$

(6)

where the shift between the $h [\cdot]$ and the $k [\cdot]$ function, from Eq. (5) to Eq. (6), means that all terms on the right-hand side of Eq. (3) can be expressed as functions of $p_t$ and $u' (c_t)$.

Obviously, it is valuable from an empirical point of view if we can use a predictive equation such as Eq. (6), as the unobservable state variables contained in $x_t$ do not enter the predictive regression and, thus, do not have to be identified. However, before proceeding further, one must check that a relationship such as Eq. (5), and thus Eq. (6),
holds under a reasonable set of assumptions. Indeed, at first glance, Eqs. (5) and (6) may seem highly restrictive. It turns out, however, that in many situations, predictive regressions will have the form suggested by Eq. (6).\textsuperscript{1} In the next section, we will explore in detail one specific situation where a relationship such as Eq. (6) holds, i.e. a situation where agents have preferences for holding money.

2.2 Predictive regressions in a monetary economy

A real economy will be driven by real shocks. In a monetary economy, there will be both real and nominal shocks. To make the following discussion tractable, we assume that there are two pure sources of risk/state variables in the economy: one pure real shock and one pure nominal shock.

The setting of the model we work with is affine.\textsuperscript{2} Hence, we proceed in the usual way for these kinds of models. First, we show that the two pure sources of risk (the state variables) can be written as affine functions of two observable variables. Second, we show that the excess return on an equity can be written as a linear function of the two pure sources of risk in the economy. Consequently, excess returns can also be written as a linear function of the two observables. The result of these derivations will be that a representation such as Eq. (6) holds in equilibrium.

We start by deriving the predictive equation for the aggregate market and then discuss the extension of the basic result to an individual risky asset. We also discuss how to proceed if there are more than two pure sources of risk.

2.2.1 Preferences and the real pricing kernel. We consider a representative agent in a monetary economy who, as is standard (see, e.g. Walsh, 2003), is assumed to obtain utility from consumption and his holdings of money (real balances), i.e. the money in the utility function (MIUF) approach. More precisely, we assume that the representative investor’s utility function is:

\[ U(c_t, m_t) = u(c_t)\phi v(m_t)^{1-\phi} \]  

\textsuperscript{1}For instance, the standard share-price to dividends ratio, the share-price to consumption ratio of Santos & Veronesi (2006), or the share-price to GDP ratio of Rangvid (2006) can be viewed as special cases of Eq. (6). We will demonstrate this in the next section.

\textsuperscript{2}For a rigorous definition of an affine setting, including references, see Eraker (2008) and Eraker and Shaliastovich (2008).
where
\[ u(c_t) = \frac{c_t^{1-\gamma_c}}{1 - \gamma_c} \]  
\[ v(m_t) = \frac{m_t^{1-\gamma_m}}{1 - \gamma_m} \]
where \( m_t \) stands for real money balances, \( \phi \) is the weight to consumption in the agent’s utility function, and \( \gamma_c \) and \( \gamma_m \) are the curvatures of the utility function with respect to consumption and real balances, respectively. Obviously, \( \gamma_c \) also corresponds to the parameter of relative risk aversion. We keep the model general here by allowing \( \gamma_c \) to differ from \( \gamma_m \), as seems to be empirically relevant; see e.g. Bouakez et al. (2005) and Christiano et al. (2005).

The MIUF approach is standard in the monetary economics literature. For instance, the first chapter in the Walsh’s (2003) benchmark textbook on monetary economics deals with the MIUF approach. In addition, the extensive literature dealing with issues such as consumer price dynamics, money demand, exchange rates, monetary aggregation, and optimal monetary policy is often cast within the MIUF approach. There are also papers that analyze different asset-pricing implications of accounting for the nominal features in the economy using the MIUF assumption; see, for instance, Stulz (1986), Bakshi and Chen (1996), and Buraschi and Jiltsov (2005). To the best of our knowledge, however, ours is the first to study explicitly the implications of a monetary economy for return predictability, and to investigate this empirically.

We choose the MIUF approach because of its generality. Indeed, since Feenstra (1986), transaction-cost models and shopping-time models are known to be functionally equivalent to MIUF models. In addition, cash-in-advances models are extreme versions of transaction-cost models, i.e. also this type of model can be mapped into the general MIUF approach.

It should also be mentioned that the assumption that the agent/investor controls his holdings of real money balances means that the Central Bank can be specified to control either the nominal money supply or the interest rate through a Taylor rule in the MIUF approach.\(^3\) In other words, the choice of the MIUF framework does not imply that we implicitly assume that the Central Bank controls the nominal money supply as its operating instrument. The Central Bank’s choice of operating procedure is not necessarily determined by choosing the MIUF approach.

The main purpose of this paper is to investigate the implications for return predictabil-

\(^3\)Under some circumstances, the equilibrium outcome under money supply or interest rate rules may be the same. For a recent contribution on this issue, see Bruckner and Schabert (2006).
ity from allowing the agents in the economy to hold money. In order to investigate this, consider the pricing kernel in the monetary economy. The real pricing kernel $\Lambda_t$ is the marginal utility of consumption:

$$\Lambda_t = \phi e^{-\rho t} \frac{c_t^{1-\gamma_c} \phi^{-1}}{(1-\gamma_c)^{\phi-1}} v(m_t)^{1-\phi}. \quad (10)$$

We use the “fundamental relation of monetary economies” (the so-called “portfolio balance equation”):$^4$

$$\frac{U_m}{U_c} = R_t, \quad (11)$$

where $U_m$ is the marginal utility of the representative investor from his holdings of real balances and $U_c$ is the marginal utility of consumption, to replace the holdings of real balances $m$ in Eq. (10) with the interest rate.

Relation (11) states that, in equilibrium, the rate of substitution between real balances and real consumption is equal to the opportunity cost of holding real balances. This cost is $R_t$, the nominal short term interest rate. Note that Eq. (11) consistently relates flows from money holdings with consumption flows. In particular, $U_c$ is the marginal utility obtained from the flow of consumption during period $t$ and $U_m/R_t$ is the flow of services from holding real balances during period $t$.\(^5\)

An alternative way for understanding why the nominal interest rate matters for the trade-off between real consumption and real money holdings is the following. Consider an agent who gives up one unit of consumption worth one dollar (consumption is the numeraire) and holds the dollar in cash. One dollar today is also one dollar tomorrow, i.e. in nominal terms it is risk free to carry one dollar until the next period. The present value of one dollar next period is thus $1/(1 + R_t)$, where $R_t$ is the nominal interest rate. Therefore, the net cost of holding 1 extra unit of real money balances is $1 - 1/(1 + R_t) = R_t/(1 + R_t)$. In equilibrium, this net cost of holding one extra unit of real balances must correspond to the utility loss the agent suffers from giving up one unit of consumption today in relation to the utility gain he gets from holding one more unit of real money today. Finally, Carlstrom and Fuerst (2001) show that if one assumes that the money the

\(^4\)Eq. (10) is found by combining the first-order conditions from the maximization of the utility function with respect to real consumption and real money balances. Eq. (10) holds in more general settings than those with time additive utility function and frictionless economies. For example, a similar relationship holds under external habit and nominal rigidities (see Christiano et al., 2005, equation 24, page 18).

\(^5\)As a simple example, consider the log case: $U = (1 - \phi) \ln(m) + \phi \ln(c)$. In this case, Eq. (11) implies that $c = \frac{\phi}{1-\phi} Rm$, meaning that the weighted flow of services from holding money ($Rm$) equals the flow of consumption services at the optimum, where the weight to $Rm$ is the relative weight of real money balances in the utility function.
agent gets utility from is that which is held in the beginning of the period, the opportunity cost of holding the money reduces to $R_t$.

In continuous time, $R_t$ is the opportunity cost of holding money and the discounting simply cancels out, which yields Eq. (11).

Using Eq. (11) and the utility function specification given above, one obtains:

$$m_t = \frac{1 - \gamma_m}{1 - \gamma_c} \frac{1 - \phi}{\phi} c_t$$

(12)

Substituting Eq. (12) into Eq. (10) yields:

$$\Lambda_t = A e^{-\rho t} R_t^{\delta_R} c_t^{\delta_c}$$

(13)

where:

$$A = \phi \left( \frac{1 - \phi}{\phi} \right)^{(1 - \gamma_m)(1 - \phi)} \left( \frac{1 - \gamma_c}{1 - \gamma_m} \right)^{\gamma_m(1 - \phi)}$$

$$\delta_c = (1 - \gamma_c) \phi + (1 - \gamma_m) (1 - \phi) - 1$$

$$\delta_R = -(1 - \gamma_m) (1 - \phi).$$

Eq. (13) is important because it reveals a key implication of a monetary Economy. A nominal variable, the *nominal* short term interest rate, affects the real pricing kernel.

Indeed, it is through its effect on the real pricing kernel that the nominal short interest rate will affect asset prices, and, hence, be useful when predicting excess returns.

Note also that a purely real economy is obtained when $\phi = 1$. In other words, when the agents derive utility from consumption only, and not from holding money, the nominal interest rate will not matter for asset pricing and return predictability.

---

6 We would like to emphasize that $R_t$ is risk free in nominal terms but risky in real terms, because the rate of inflation is uncertain. Even if this is not very important for the first-order conditions we work with in the main body of the paper, it is important in a general equilibrium model because it is exactly this feature that generates money non-neutrality.

7 From Eq. (13), it appears as though the interest rate would disappear from the pricing kernel if the utility function is separable in consumption and money holdings. It turns out, however, that even under such preferences, the pricing kernel will be driven by the nominal rate (and thus affect asset prices) if money is not neutral. In an appendix available on our webpages, we derive the general equilibrium of a monetary economy where we show that even under a log separable utility function, the pricing kernel is still driven by the nominal rate and output in the economy (instead of consumption).

8 In general, the impact of money on real variables will be both direct (as shown by the presence of the nominal interest rate in the pricing kernel) and indirect (through the impact on consumption itself when money is not neutral, which is generally believed to be the case at least in the short term).
Discussion. We prefer to work with Eq. (13) in the empirical implementation, even if – from a theoretical point of view – one could work equally well with Eq. (10). The reason why we have preferences for working with the interest rate (as in Eq. (13)) instead of the money supply (as in Eq. (10)) is that the interest rate is measured empirically without systematic errors, in contrast to the money supply. In particular, there is no consensus on the empirical measurement of real money balances. To calculate real money balances, one needs a nominal money supply and a price deflator. Regarding the nominal money supply, some authors use M1 (or even M0), M2, or even M3. In addition, theory does not point clearly to the price deflator that should be used to convert nominal money to real money. Should it be the consumer price index, the producer price index, the GDP deflator, or some other measure of price movements? The implications for asset pricing may be sizable; see Balvers and Huang (2008) for a discussion of this issue. If, on the other hand, the nominal interest rate is used, the choice is clear. It should be the interest rate that is risk-free in nominal terms over the frequency at which the model is specified. This makes it preferable to work with the interest rate and not the money supply.

It should also be mentioned that most of the qualitative analysis above and hereafter hold under more sophisticated utility functions.\(^9\) For tractability, we have decided to work within the time-additive framework.

2.2.2 Asset prices. There are two pure sources of risk in the monetary economy. In order to relate the two pure sources of risk to observable variables, we need two financial assets. The two assets are equity and one unit of the currency.

Consider first the equity. The equity guaranties the optimal consumption path of the representative investor from now to infinity. Therefore, the price of the equity is:

\[ p_t = E_t \left[ \int_t^{\infty} \frac{\Lambda_s}{\Lambda_t} c_s ds \right]. \tag{14} \]

Substituting Eq. (13) into Eq. (14) yields:

\[ \frac{p_t}{c_t} = E_t \left[ \int_t^{\infty} e^{-\rho(s-t)} \left( \frac{R_s}{R_t} \right)^{\delta_R} \left( \frac{c_s}{c_t} \right)^{\delta_{c+1}} ds \right]. \tag{15} \]

\(^9\)For example, our results can be extended to the Epstein-Zin framework, although at a cost of a more burdensome notation. See Bufman and Leiderman (1993) and the recent results in Eraker (2008) for an affine equilibrium in an Epstein-Zin setting.
Consider now the second asset: one unit of the currency. Holding one unit of the currency in the bank (in the risk-free asset) returns the nominal interest rate $R_t$ in each period. Therefore, in real terms, we have:

$$ \frac{1}{\Pi_t} = E_t \left[ \int_t^\infty \Lambda_s \frac{R_s}{\Lambda_t \Pi_s} dS \right] \tag{16} $$

where $\Pi$ is the general price level. That is, the real value today of one unit of the currency is equal to the discounted value of $R_s/\Pi_s$ real units of the currency at each point $s$ in time until infinity. In other words, if the investor deposits one unit of money in the bank today, the investor gets the entire future stream of $R_s$. The future $R_s$ are unknown today, and hence risky, implying that the price of one unit of the currency today is the future risk-free rates discounted by the stochastic discount factor $\Lambda_s/\Lambda_t$. Dividing by $\Pi_s$ expresses it in real terms. Manipulating this equation and substituting for the real pricing kernel (13) yields:

$$ \frac{1}{R_t} = E_t \left[ \int_t^\infty e^{-\rho(s-t)} \left( \frac{R_s}{R_t} \right)^{\delta_R} \left( \frac{c_s}{c_t} \right)^{\delta_c} \left( \frac{\Pi_s}{\Pi_t} \right)^{-1} ds \right]. \tag{17} $$

The left-hand side, $1/R$, corresponds to the “price/dividend” ratio of the (locally) riskless asset, since its current nominal price is 1 and its dividend is the nominal rate $R_t$.

We thus have two pricing equations, which under well-behaved growth rates for consumption and the nominal interest rate as well as for the price level (inflation), can be written as:\textsuperscript{10}

$$ \frac{p_t}{c_t} = \int_t^\infty e^{-\rho(s-t)} E_t \left[ \left( \frac{R_s}{R_t} \right)^{\delta_R} \left( \frac{c_s}{c_t} \right)^{\delta_c} \right] ds \tag{18} $$

$$ \frac{1}{R_t} = \int_t^\infty e^{-\rho(s-t)} E_t \left[ \left( \frac{R_s}{R_t} \right)^{\delta_R+1} \left( \frac{c_s}{c_t} \right)^{\delta_c} \left( \frac{\Pi_s}{\Pi_t} \right)^{-1} \right] ds. \tag{19} $$

### 2.2.3 Deriving the predictive regressions.

We now use the pricing equations (18) and (19) to isolate the state variables in the economy. We develop the process heuristically. Appendix A provides a rigorous proof.

We make three assumptions regarding the law of motion of consumption, the nominal interest rate, and inflation:

\textsuperscript{10}We use Fubini’s theorem to switch between the integral and the expectation.
1. We assume that the conditional expectations of the growth rates of consumption, the nominal interest rate, and inflation are time varying and depend linearly on the two pure shocks in the economy.

2. We assume that the uncertainty that affects the growth rates of these three variables is also time varying and depends upon the pure shocks.

3. The data generating processes of the state variables are AR(1) processes.

These assumptions allow us to remain within the affine setting. In Appendix A, we explain that these assumptions are often used in the return predictability, conditional asset pricing, and dynamic asset allocation literature.

Using these assumptions, Appendix A shows that a log-linearization of equations (18) and (19) allows the two state variables to be expressed as an affine combination of two observable variables: \( \ln \frac{p_t}{c_t} \) and \( \ln R_t \). \( \ln \frac{p_t}{c_t} \) is the deviation of the log share price/consumption ratio from its long-run mean and \( \ln R_t \) is the deviation of the nominal interest rate from its long-run mean. In addition, the affine setting allows us to show that the expected return on the equity is a linear function of the two state variables in the economy.

In summary, given that the two state variables can both be written as affine combinations of \( \ln \frac{p_t}{c_t} \) and \( \ln R_t \), and given that the expected return on the equity is a linear function of the two state variables, the expected return on the equity is a linear function of \( \ln \frac{p_t}{c_t} \) and \( \ln R_t \). Our final result is:

\[
E_t [r_{t+1}] - r_{t+1}^f = \alpha_0 + \alpha_1 \ln \frac{p_t}{c_t} + \alpha_2 \ln R_t
\]

where \( \alpha_1 \) and \( \alpha_2 \) are parameters of the underlying structural parameters, i.e. of the preference parameters and the parameters of the dynamics of the growth rates of consumption, inflation, and the interest rate. The precise expressions for the parameters \( \alpha_1 \) and \( \alpha_2 \) are provided in Appendix A.

A key insight from our model is that the interest rate would disappear from the right-hand side of Eq. (20) in a real economy, as shown explicitly in Appendix A. In other words, the appearance of the interest rate on the right-hand side of Eq. (20) is a result of our assumption that both real and nominal shocks drive the economy. In essence, this is probably the main result of our model.

**Comparing Eq. (20) with Eq. (6)**. Since \( c \) and \( R \) are the two variables that enter the pricing kernel in Eq. (13) in the monetary economy, the result in Eq. (20) matches
the form conjectured in Eq. (6). To illustrate, note first that Eq. (13) showed that
\[ u'(c_t) = AR_t^\gamma e^\delta_t \] and therefore
\[ h[p_t, u'(c_t)] = h[p_t, AR_t^\delta e^\delta_t], \] i.e. the three endogenous variables that enter \( h() \) are \( p_t, R_t, \) and \( c_t, \) which is exactly the three variables that also enter Eq. (20).

However, comparing Eq. (20) with Eq. (6), it becomes clear that the functional form for \( h() \) implied by Eq. (20) is not trivial. For instance, if one had – completely ad hoc – specified the functional form for \( h \) as separable and linear, this would not deliver a predictive regression similar to Eq. (20) since (in logarithms)
\[ h(p_t, u'(c_t)) = \ln p_t + \delta \ln c_t, \] which is obviously different from Eq. (20). In other words, one has to take great care when specifying the functional form of \( h() \) and, in particular, whether or not a relationship such as Eq. (6) holds under a reasonable set of assumptions has to be checked.

It should be mentioned that the price we pay for using theory to consistently derive the predictive regression in Eq. (20) is that the structural parameters cannot be backed out, i.e. it is not possible to back out, for instance, \( \gamma_m \) and \( \gamma_c \) from \( \alpha_1 \) and \( \alpha_2. \) On the other hand, the great benefit we gain from using theory to derive Eq. (20) is that we do not have to identify the state variables in the economy. For the empirical investigation, this is of course a valuable result.

2.2.4 Extensions. In the above derivation, we have relied on the assumption that there are only two sources of pure risks. What if there is more than two state variables? Given the procedure pursued above, it follows that more than two sources of risk can be handled by using additional marketable assets (like the equity was used to price the real source of risk and the return on the bank account was used to price the nominal source of risk). Indeed, the pricing equations of such additional variables allows the identification of additional pure sources of risk.

The previous analysis could also be extended to individual stocks. One option could be to assume that the cash flows of an individual stock are a fraction of the cash flows of the aggregate market (aggregate consumption). If the uncertainty about this fraction is driven only by idiosyncratic risks, and the latter are not priced (for example because markets are complete), then Eq. (20) holds also for an individual risky asset. If there is a systematic source of risk that determines the fraction of each risky asset in the aggregate cash flow (à la Santos and Veronesi, 2006), this case can be made equivalent to the presence of an additional source of risk. In this case, one would follow the procedure just outlined: find an additional pricing equation for each pure source of risk.
2.3 Empirical implementations

In this section, we discuss how we implement Eq. (20), as well as different variants thereof, in our empirical analysis. Before doing so, though, we point out two basic issues regarding the construction of the variables entering Eq. (20). First, if the long-run means of \( \ln p_t/c_t \) and \( \ln R_t \) are constant through time, it does not matter whether one uses the deviations from the (constant) mean or the variables themselves. This is important to note as the variables that have been used in the empirical return-predictability literature are generally the variables themselves. The second general issue to note is that it does not matter whether one uses real or nominal share prices and consumption when the price-consumption ratio (or other valuation ratios) is employed as a predictor, as the price deflator cancels out. In other words, if the nominal share price is given as \( P_t = p_t \Pi_t \), where \( \Pi_t \) is the price level in the economy, and nominal consumption is given as \( C_t = c_t \Pi_t \), it follows that \( P_t/C_t = p_t \Pi_t/c_t \Pi_t = p_t/c_t \).

Predictive regressions in real and monetary economies. The empirical part of the paper includes the following kinds of regressions:

- First, we run a bivariate predictive regression explaining excess returns with the price-consumption ratio and the nominal interest rate, i.e. the regression in Eq. (20) under the assumption of a constant steady state. The key point here will be to compare the results from such a regression to the case where we run a univariate regression using only the price-consumption ratio (the pc-ratio) as a predictor. This our first comparison because the most important aspect of Eq. (20) to note is that the presence of the nominal interest rate is a direct consequence of the nominal nature of the economy. In other words, the predictor that arises in a real economy is \( \ln(p_t/c_t) \) (assuming a constant mean), i.e. excluding the nominal interest rate. The predictive ability of the price-consumption ratio has already been investigated by Santos and Veronesi (2006). Therefore, we start out by comparing the differences in results achieved when the nominal interest rate is added to the pc-ratio.

- A second predictor is obtained if one considers a pure exchange economy where consumption is equal to output. When consumption is equal to output, one obtains the price-output ratio introduced by Rangvid (2006). In a pure monetary exchange economy, the relevant bivariate predictor will be one that uses the py-ratio (the share price to GDP ratio) and the nominal short interest rate as explanatory variables. Hence, we will compare the bivariate regression (using the py-ratio and the interest rate) to the univariate case using only the py-ratio.
• Assuming that consumption is equal to dividends, one obtains the price-dividend ratio and the predictors in the monetary economy will be the \(pd\)-ratio and the interest rate. We compare the results we get here with the ones we get in the real economy, where the standard \(pd\)-ratio is the only variable used. The analysis we conduct here is closely related to the analysis in Ang and Bekaert (2007).

• Recently, it has been advocated that the assumption that consumption and dividend are equal is a highly restrictive equilibrium condition.\(^{11}\) One possible way to relax this assumption is to use Abel’s (1999) suggestion to set:

\[ c_t = d_t^\lambda \]  

(21)

where \(d\) is the asset’s real dividend and \(\lambda\) is interpreted by Abel as a leverage parameter.\(^{12}\) In this case, the predictor we use is a combination of the equity price, the dividend, and the nominal short interest rate, where the coefficient to dividends can be different from one. It is important to note that because \(p_t\) and \(d_t\) are included on their own here (and not in the form of a ratio such as the \(pd\)-ratio), this has two consequences for the predictions. The first is that the price deflator does not cancel out when using nominal values. The second is that \(p_t\) is non-stationary in levels and \(d_t\) is non-stationary in levels, i.e. when using these variables together, we have to estimate a cointegration relation between the variables to make sure that the predictor we use is stationary. In other words, we will use a cointegration relation called \(\hat{pdR}\), where \(\hat{pdR}\) will be given as:

\[ \hat{pdR}_t = \ln P_t - \lambda \ln D_t + (\lambda - 1) \ln \Pi_t + \frac{\alpha_2}{\alpha_1} \ln R_t \]  

(22)

and where the most important issue to notice is that \(\lambda\) can be different from one.

3 Data and construction of the predictive variable

We estimate the model using U.S. data. The main variables we use are the aggregate nominal share price index \((P_t)\), the dividends paid out by the firms included in the index \((D_t)\), the nominal output in the economy \((Y_t)\), the nominal consumption in the economy \((C_t)\), the nominal interest rate \((R_t)\), and the level of prices in the economy \((\Pi_t)\) as the data we use for robustness checks at later stages in the paper. Consequently, we only describe

\(^{11}\)See, for example, Santos and Veronesi (2006) and the discussion therein.

\(^{12}\)Abel’s (1999) model has been used recently to address several empirical stylized facts regarding the behavior of asset markets. See, for example, Lettau, Ludvigson and Wachter (2008).
the most important features of the data here.

Consumption and GDP data are available quarterly from 1947 and onwards. As a result, the sample we use spans the period from the first quarter of 1947 to the fourth quarter of 2007. The measure of the aggregate stock market we use is the S&P 500 (later, as a robustness check, we also predict the returns from an even broader measure of the stock market). Given that we use quarterly data, the nominal risk-free interest rate we use is the three-month Treasury Bill rate. As the measure of the price level in the economy, we use the CPI.

3.1 The predictors

The variables we use to predict excess returns are valuation ratios (the pd-ratio, the pc-ratio, and the py-ratio), the interest rate, and the pdR-ratio. From a “construction” point of view, the pdR-ratio is different from the other variables, as there is no estimation involved in the construction of the valuation ratios or the interest rate whereas the parameters entering the pdR-ratio must be estimated. This implies that the description of the construction of the valuation ratios can be kept brief, as this is standard, whereas the construction of the pdR-ratio requires some explanation. The specific estimation procedure used when estimating the parameters of the pdR-ratio is described in Appendix C.

The valuation ratios. The valuation ratios express the valuation of the stock market in relation to a fundamental. Because we want to use the valuation ratios to predict returns, we scale the level of share prices in the beginning of a quarter with the level of the fundamental known in the beginning of the quarter, such that current information only is used to predict future returns. For instance, when predicting excess returns from period \( t \) to period \( t + 1 \), the pd-ratio we use is \( pd_t = \ln(P_t/D_t) \), where \( P_t \) is the level of nominal stock prices in the beginning of period \( t \) and \( D_t \) the level of dividends known when entering period \( t \). The same is true when scaling with consumption. We scale the nominal stock price in the beginning of a quarter with the level of nominal consumption, \( C_t \), known in the beginning of the quarter, i.e. \( pc_t = \ln(P_t/C_t) \). Likewise, \( py_t = \ln(P_t/Y_t) \), with \( Y_t \) being the level of nominal output known in the beginning of the quarter.

The pdR-ratio. We also introduce a pdR-ratio in order to capture movements in expected excess returns.

\footnote{To be precise, \( D_t \) is the level of dividends paid out during the last 12 months up until the beginning of quarter \( t \).}
Returns are stationary. In order for the return-predicting regression to be balanced, the predictive variables must be stationary, too. Because of growth in the economy, dividends, share prices, and the CPI (that all enter into the construction of the $pdR$-ratio, as seen in Eq. (22)) increase stochastically over time, however. To deal properly with the non-stationarity of the regressors, we estimate the relation between share prices and relevant variables using a cointegration estimation.

We use the Stock-Watson (1993) Dynamic OLS estimation procedure to estimate the cointegration parameters. Appendix C describes the procedure we use to estimate the coefficients in the $pdR$-ratio in more detail. The result from the estimation of the cointegration parameters based on the 1947:1–2007:4 data is:

$$
\text{pd}R_t = \ln(P_t) - 9.550 - 2.225 \ln(D_t) + 7.757 \ln (1 + R_t) + 1.089 \ln(\Pi_t)
$$

where the numbers in parentheses below the coefficient estimates are standard errors adjusted for long-run variance, as described in Stock-Watson (1993).

Our model tells us that the coefficient to dividends is the level of leverage in the economy (see Eq. (22)), i.e. we find an estimate of the leverage in the economy of $-\lambda = -2.225 \Rightarrow \lambda = 2.225$. Abel (1999) uses a value of 2.74 to calibrate the moments of asset returns. Given the standard error of 0.542, an estimated leverage of 2.225 is close to Abel’s preferred value of $\lambda$.

Our theoretical model tells us that the coefficient to the CPI is equal to $\lambda - 1$. We find that the estimate to the CPI is 1.089, i.e., close to $2.225 - 1 = 1.225$, which, with a standard error of the coefficient to the CPI of 0.609, is within the confidence band. Hence, we run the predictive regressions using the unconstrained cointegration relation as our predictor variable.

Finally, the estimate to the interest rate implies that if there is a one percent increase in the interest rate, stock prices will fall “in the long run” (cointegration relations are often interpreted as displaying the “long-run relationship” between the variables) by around eight percent.

3.2 Summary statistics

In Table 1, panel A, we collect the means and the standard deviations of the variables we focus on in the following empirical analyses. In addition to the $pd$-ratio, the $pc$-ratio, the $py$-ratio, and the $pdR$-ratio, we show the means and the standard deviations of the nominal interest rate.
The continuously compounded excess returns is the dependent variable in the following analyses. The average annualized quarterly excess returns over the 1947:1–2007:4 period is 6.52% with an annualized standard deviation of 13.45% per quarter. These numbers are well in line with previous studies that report an annualized equity premium between 6% and 8%.

Panel B shows how the predictors are correlated. Most important to note is the fact that the \( pc \)- and \( py \)-ratios are almost perfectly correlated, i.e. it does not matter much whether one scales share prices with GDP or consumption, as also discussed in Rangvid (2006). It should also be noted that the contemporaneous correlations between the valuation ratios and the interest rate are negative, i.e. when the interest rate is increased, the valuation of the stock market in terms of a fundamental (dividends, consumption, or GDP) is reduced.

4 Predictions of stock returns and excess returns

We now turn to the main empirical part of our paper where we run predictive regressions of excess returns, continuously compounded over \( k \) periods, on the predictors contained in \( z_t \):

\[
er_{t,t+k} = \alpha_k + \kappa'_{z,k} z_t + u_{t+k}
\]

where \( \kappa'_{z,k} \) is a vector of coefficients and \( z_t \) is a vector of variables containing either one variable, for instance, the \( pd \)-, the \( pc \)-, the \( py \)-, or our newly constructed \( pdR \)-ratio, or two variables: one of the valuation ratios and the short-term risk-free rate.\(^{14}\)

4.1 Results

There are two main implications of our model that we would like to investigate. First, our model indicates that movements in excess returns should be captured better when the interest rate is included combined with a valuation ratio in a predictive regression. Second, if there is leverage in the economy, relaxing the restriction that the coefficient to dividends is equal to one should help when predicting excess returns.

In this section, we present results from models that are based on the assumption of a constant steady state. We first show results using quarterly, and hence non-overlapping,

\(^{14}\)Even if the \( pdR \)-ratio is a generated predictor, one does not have to adjust the standard errors of \( \kappa \), the reason being that the estimates of the cointegration parameters in the \( pdR \)-ratio are “super consistent”, i.e. they converge to their limiting distribution faster than is normally the case. For this reason, the \( pdR \)-ratio can be treated as known in the second-stage regression. For more on this, see, for instance, Lettau & Ludvigson (2001) or, for the original treatment, Stock (1987).
returns. After this, we show the results we get when longer-horizon returns are used. In section 4.4, we investigate how our predictive regressions performs if we allow the steady state to change over time.

4.1.1 Quarterly non-overlapping excess returns. Table 2 shows the first results. Looking at the estimated coefficient Compared to the traditional pd-ratio (row 1.), one would conclude that quarterly excess returns can only slightly be forecast, as the t-statistic is barely above its 95% critical level and the $R^2$ is only 1.61%. Looking at the pc-ratio or the py-ratio, excess returns look even less forecastable, as the two valuation ratios are insignificant. The interest rate is not a convincing predictor either, even if it is barely significant ($t$-statistic = $-2.35$ and an $R^2$ of 2.51%).

This overall picture of only slightly predictable quarterly excess returns changes dramatically when panels B and C are considered. Panel B shows results from bivariate regressions, while panel C shows results using estimated cointegration relations. In Panel B, the pc-ratio and the py-ratio are now strongly significant and the interest rate is also clearly significant ($t$-statistics $-3.86$ and $-3.81$ compared with $-2.35$ when the interest rate was included on its own), as seen in rows 6 and 7. In addition, and perhaps even more importantly, the explanatory power (the $R^2$) has increased substantially to 6.33% (or 6.28% using the py-ratio). The fact that the $R^2$ increases so much tells us that the increase in the $t$-statistic is not due to the correlation that exists between the pc-ratio (or the py-ratio) and the interest rate.

Regarding the estimated signs of the coefficients, we find that when the interest rate increases, excess returns are expected to fall. From a monetary policy point of view, we believe this is an important point as it implies that a tightening of monetary policy, i.e. an increase in the interest rate, will reduce expected excess returns for a given valuation of the stock market. In this sense our findings are in accordance with those in Bernanke and Kuttner (2005). Likewise, we find that for a given interest rate level, an increase in the valuation of the stock market leads to a reduction in next period’s expected excess returns.

In row 5, we show the results from predicting with the pd-ratio and the interest rate. The increase in explanatory power in the bivariate regression, compared to the univariate regressions in lines 1 and 4, is visible, but it is not as overwhelming as it is for the pc- and the py-ratio. This makes sense. Our theory indicates that it is only when the leverage in the economy is assumed to be equal to one that the pd-ratio can be used together with the

---

15Nevertheless, we want to emphasize that even if the results are not as strong as when the pc-ratio and the py-ratio are used, we do find that the pd-ratio predicts short-horizon returns better when augmented by the interest rate, as first reported by Ang & Bekaert (2007).
interest rate. In other words, when changing the regression from using the \( pc \)-ratio (or the \( py \)-ratio) together with the interest rate to using the \( pd \)-ratio together with the interest rate, we implicitly impose the restriction that consumption/output is equal to dividends. What happens if we relax this assumption? This can be seen in row 8, where we show the result from the regression that uses the \( \hat{pdR} \)-ratio to predict. The result is astonishing. The \( t \)-statistic to the \( \hat{pdR} \)-ratio is very high, close to \(-5\), and the \( R^2 \) is \( 8.64\% \).

To put the predictive power of the \( \hat{pdR} \) into perspective, the results achieved here can be compared to the ones achieved when the \( \hat{cay} \)-ratio of Lettau and Ludvigson (2001) is used.\(^{16}\) The \( \hat{cay} \) ratio is also a very strong predictor of excess returns; the \( t \)-statistic is above 4, as shown in row 9. In terms of predictive power, however, the \( \hat{cay} \)-ratio does not capture as large a fraction of the variation in excess returns as does the \( \hat{pdR} \)-ratio.

One can also assess the economic importance of fluctuations in the \( \hat{pdR} \)-ratio. The coefficient to the \( \hat{pdR} \)-ratio is estimated to be a negative \(-0.091\). The standard deviation of the \( \hat{pdR} \)-ratio is \( 0.221 \). A one standard deviation change in the \( \hat{pdR} \)-ratio thus leads to a change in expected excess returns of close to \( 200 \) basis points, corresponding to a change in annualized expected excess returns of around \( 8\% \). Given that the average annualized excess return is \( 6.52\% \), the fluctuations in the \( \hat{pdR} \)-ratio are thus also economically important. This number perhaps seems big at first sight, but it is in line with what Lettau and Ludvigson (2001) found.\(^{17}\)

To make sure that the empirical performance of the \( \hat{pdR} \)-ratio is not driven by a correlation between the residual of the predictive regression and the innovation in the \( \hat{pdR} \)-ratio, i.e. the Stambaugh (1999) bias, we correct for this. The Stambaugh correction does not readily allow hypotheses about the bias-adjusted coefficient to be tested. Building on the work of Stambaugh, Amihud and Hurvich (2004) provides a simple augmented regression method to bias adjust the predictive coefficient and test hypotheses.\(^{18}\) Using the procedure in Amihud and Hurvich (2004), we find that predictive power of the \( \hat{pdR} \)-ratio remains after correcting for the bias. The bias-adjusted regression coefficient using the \( \hat{pdR} \)-ratio is \(-0.082 \) and the bias-adjusted \( t \)-statistic is \(-4.104 \). In other words, the predictive power of the \( \hat{pdR} \)-ratio is strong also after adjusting for the Stambaugh bias.\(^{19}\)

---

\(^{16}\) We mention the performance of additional predictors in section 4.7. We find it relevant to compare with \( \hat{cay} \) already here, though, as \( \hat{cay} \) is generally found to perform better than most other predictors. In addition, the \( \hat{cay} \) ratio is, like the \( \hat{pdR} \)-ratio, also based on a cointegration framework. The updated data on \( \hat{cay} \) are quarterly data spanning the 1951:4–2006:4 period.

\(^{17}\) Lettau & Ludvigson (2001) report that a one standard deviation increase in \( \hat{cay} \) result in “roughly a nine percent increase at an annual rate” in excess returns.

\(^{18}\) Amihud, Hurvich & Wang (2008) show that the augmented regression method performs well compared to other procedures used to obtain bias-adjusted standard errors in predictive regressions.

\(^{19}\) Compared to the bias in, for instance, the \( pd \)-ratio, the bias in the \( \hat{pdR} \)-ratio is small. Indeed,
Let us briefly summarize the results of this section. We have shown that standard valuation ratios do not predict the movements in the excess returns on the stock market particularly well.\textsuperscript{20} Neither does the interest rate. Combining a valuation ratio with the interest rate, however, one can capture a much larger share of the movements in expected excess returns. We emphasize that this is exactly what our theory predicts. In a monetary economy, one should use the valuation ratio together with the interest rate. In addition, we find that if one relaxes the assumption that dividends are equal to consumption, one predicts even better; the $pdR$ ratio captures a larger fraction of variation than, for instance, the $cy$-ratio does. Finally, the variation in expected returns that the $pdR$-ratio captures is economically large and the statistical significance of the $pdR$ ratio is robust towards a correction for the Stambaugh bias.

\textbf{4.1.2 Predictability of long horizon excess returns.} We now turn to the results we get when regressing long-horizon excess returns on our predictors.

One advantage of using long-horizon returns is that the noise inherent in short-horizon returns is reduced relative to the low-frequency movements in returns. Moreover, slow-moving predictor variables such as those used here and elsewhere in the literature consequently often appear to better capture the low-frequency movements of returns. A drawback of long-horizon regressions, however, is that the observations and residuals are overlapping, which affects the $t$-statistics and $R^2$s in erroneous ways, in particular when the return-forecasting horizon is large relative to the sample size.

Ang and Bekaert (2007) argue strongly that one should use Hodrick (1992) standard errors when testing for the significance of predictive variables in long-horizon regressions. Indeed, Ang and Bekaert (2007) use comprehensive simulations to show that the asymptotic distributions underlying the normally used Newey-West (1987) and Hansen and Hodrick (1992) standard errors are not well-suited for the small samples that finance researchers have at hand. On the other hand, Ang and Bekaert (2007) also show that the standard errors developed by Hodrick (1992) retain the correct size in small samples. For this reason, we follow Ang and Bekaert (2007) and use Hodrick (1992) standard errors throughout our study. Appendix D briefly repeats how to calculate the Hodrick standard errors and explains the difference between the Hodrick (1992) and the Newey-West bias-adjusted regression coefficient to the $pd$-ratio is $-0.007$ and its bias-adjusted $t$-statistic is only $-0.607$. Therefore, the bias in the coefficient to the $pd$-ratio is $68\%$ of its estimated value (the bias is $0.0158$ which is compared to its estimated value of $0.022$), whereas the bias in the estimated coefficient to the $pdR$-ratio is $0.0074$, corresponding to approximately $8.3\%$ of the estimated coefficient. In addition, the $pdR$-ratio remains significant after bias adjusting, whereas the $pd$-ratio does not.

\textsuperscript{20}Rangvid (2006) finds that the $py$-ratio captures stock returns better than the $pd$-ratio does. Rangvid also notes, though, that the $py$-ratio does not capture excess returns, as also reported here.
(1987) standard errors. We also present the Implied $R^2$ of Hodrick (1992) and Campbell (1991). These $R^2$s are based on VAR(1) models, and, hence, are not subject to overlapping observations. Appendix D also briefly repeats how to find the Implied $R^2$s.

Table 3, which shows results from regressions of excess returns on the different Predictors, lists the coefficient estimates $\hat{c}_z$, three different kinds of $t$-statistics, and the $R^2$s and Implied $R^2$s from regressions of excess returns over one quarter and accumulated over one to six years ($k = 1, 4, 8, 12, 16, 20, \text{ and } 24$). The three kinds of $t$-statistics are the basic OLS $t$-statistics, the Newey-West (1987) statistics, and the Hodrick (1992) statistics.\footnote{We use $k + 1$ lags of $C_T(j)$ when we calculate the Newey-West statistics (see Appendix D).} We show three $t$-statistics to make the consequences of using Hodrick’s $t$-statistics clear by comparing them with the non-adjusted basic OLS $t$-statistics and the Newey-West (1987) $t$-statistics that are often used in long-run return regressions.

The structure of Table 3 follows the structure of Table 2. Part 1 shows results from the long-horizon univariate regressions, i.e. from regressions of excess returns on either one of the valuation ratios or on the interest rate.\footnote{To save space, we only show results using the $py$-ratio and, hence, not the $pc$-ratio. The reason is that the $py$-ratio and the $pc$-ratio are highly correlated, as showed in Table 1, and, the results achieved are very similar, as shown in Table 2, for the quarterly regressions.} Part 2 shows results from bivariate regressions and part 3 shows results from univariate regressions using cointegration residuals as regressors.

Before commenting on the economic content of our results, a brief word on the consequences of the use of Hodrick standard errors is appropriate. Consider the results, shown in panel B, part 1, from predictions of excess returns using the standard $pd$-ratio. The basic non-adjusted OLS $t$-statistics increase with the horizon due to the overlapping residuals. The Newey-West (1987) adjustment brings down the $t$-statistics somewhat. The Hodrick (1992) adjustment brings down the $t$-statistics even more. For instance, when using the traditional $pd$-ratio to predict, the basic non-adjusted OLS $t$-statistic at a six-year horizon is $-10.365$ compared to its Newey-West value of $-3.707$, i.e. if correcting for the overlapping observations with the normally used Newey-West standard errors, one would conclude that the $pd$-ratio is a strong predictor of long-horizon excess returns. However, when the Hodrick (1992) $t$-statistics that Ang and Bekaert (2007) report to have better small-sample properties are used, we can confirm Ang and Bekaert’s (2007) finding that the traditional $pd$-ratio is not significant, or only marginally significant, for long-term excess-returns forecasts.

In the following, we focus on the results obtained when we use the Hodrick standard errors. The overall picture is similar to the one shown in Table 2, where we presented the results from the quarterly regressions. Note first that the univariate regressions in part
of Table 3 are generally not significant, or at least only marginally so, when evaluated using the Hodrick standard errors. In other words, one does not significantly capture the variation in long-horizon excess returns when one of the standard valuation ratios or the interest rate is used. This overall picture changes when using the py-ratio combined with the interest rate; see panel E of Table 3. Indeed, the interest rate and the py-ratio are now both highly significant when predicting long-horizon returns.

We also find, as first reported in Ang and Bekaert (2007), that the pd-ratio and the interest rate are not significant predictors for long-horizon returns; see panel D. In other words, when we use the pc-ratio or the py-ratio to predict returns together with the interest rate, we find that the variables are significant. When we impose the extra assumption that dividends are equal to consumption/output, we find the Ang and Bekaert (2007) result that excess returns are not predictable for longer horizons. Section 4.2.1 discusses in even more detail why we find results that differ from Ang and Bekaert’s.

Finally, part 3 of Table 3 shows results from using estimated cointegration relations to predict returns. Again, we find that long-horizon excess returns are predictable. Indeed, even when using the Hodrick (1992) t-statistics Ang and Bekaert (2007) advocate, the pdR ratio is still significant for longer horizons. It is not more significant than for the shorter horizons, but it is significant.23

4.2 Understanding the results

We have shown that the short-term interest rate matters together with a valuation ratio when predicting excess returns. We have also shown that if we allow for leverage in the economy, an even higher fraction of the variation in returns is captured. Finally, we have confirmed Ang and Bekaert’s (2007) results that augmenting the pd-ratio with the interest rate does not help for long-horizon predictability.

We would like to understand these results more closely. In particular, we would like to compare our long-horizon return regressions exactly to those of Ang and Bekaert (2007) in order to understand the ways in which we depart from Ang and Bekaert (2007) and why we find more evidence in favor of long-horizon predictability than they do in some

23 We follow Ang and Bekaert (2007) and adjust the OLS standard errors in the long-run regressions with the Hodrick (1992) corrections. Other procedures are available, though. For instance, Valkanov (2003) proposes using a rescaled t-statistic when running long-horizon regressions. We find a value of Valkanov’s rescaled t-statistic (t/√T) equal to −0.595 from the regression of, for instance, six-years cumulative excess returns on the pdR-ratio. This can be compared to the critical values provided in Table 4 in Valkanov (2003). The relevant critical values to compare with in Valkanov are case 1 with c = −10 and δ = 0.9. The critical value at the one percent level is −0.405, i.e. the pdR-ratio is a significant predictor, also according to Valkanov’s rescaled t-statistic. The same is true for the one-year excess returns, two-year excess Returns, etc.
4.2.1 Comparison with Ang and Bekaert (2007): Results for the 1952:01-2001:4 period. Ang and Bekaert (2007) find that a bivariate regression including the dividend yield and the short interest rate predicts excess returns at short horizons but not at long horizons. We find that the $\hat{pdR}$-ratio predicts excess returns at long horizons. Why do we report results that are different than Ang and Bekaert’s (2007)?

There are two main differences between our empirical approach and that of Ang and Bekaert. (i) Ang and Bekaert use the dividend yield and the short interest rate in bivariate regressions and (ii) our data extend longer than Ang and Bekaert’s. In this section, we want to compare our results directly with those of Ang and Bekaert (2007).

To do so, we first construct a $\hat{PDR}$-ratio, which is our $\hat{pdR}$ ratio excluding the level of the CPI, the reason being that Ang and Bekaert use a combination of the nominal share price and dividend (the dividend yield) and the short interest rate, i.e. excluding the CPI, as their predictors. To compare directly with Ang and Bekaert, we construct a $\hat{PDR}$-ratio that uses the same variables (share prices, dividends, and the interest rate) as the ones that Ang and Bekaert use. We call this the $\hat{PDR}$-ratio to emphasize that it is based on nominal variables. The estimated $\hat{PDR}$-ratio, based on the Stock-Watson (1993) procedure, i.e. with standard errors adjusted for long-run variances, takes the form:

$$\hat{pdR}_t = \ln(P_t) - 5.247 - 1.385 \ln(D_t) + 7.508 \ln(1 + R_t).$$

Compared to the $\hat{pdR}$-ratio, the only main difference is that the point estimate of the coefficient to dividends drops from 2.225 to 1.385. However, the long-run response of share prices to a change in dividends is still estimated to be significantly larger than one.

Table 4 shows results from regressions using the $pd$-ratio (in panel A), from bivariate regressions using the $pd$-ratio and the short interest rate (as in Ang and Bekaert, 2007), and from regressions using the $\hat{PDR}$-ratio (in panel C). We show these results from regressions using data for the 1952:1–2001:4 period that Ang and Bekaert also use. To save space, we only show $t$-statistics based Hodrick (1992) standard errors.

The results reported in Table 4 confirm the main finding of Ang and Bekaert (2007) that the price-dividend ratio is generally not significant on its own, and particularly not at long horizons, but that adding the short interest rate provides some evidence for predictability, though primarily on the shorter horizons.\footnote{Note that the $pd$-ratio is not significant on its own over this shorter 1952:1–2001:4 period, whereas Table 2 reports that the $pd$-ratio is marginally significant over the full 1947:1–2007:4 period.} On the other hand, the $\hat{PDR}$-
ratio is strongly significant at all horizons. Given that the same variables (the share price index, the dividend series, and the short interest rate) enter the regressions in both panels B and C of Table 4, the only difference between the regressions in panels B and C is the way these variables are combined. In panel B, they are combined by regressing excess returns on the \( pd \)-ratio and the short interest rate in a bivariate framework, i.e. without allowing for a non-unitary coefficient to dividends. In panel C, on the other hand, the three variables (\( \ln(P) \), \( \ln(D) \), and \( \ln(1 + R) \)) are combined in an estimated cointegration framework with free coefficients. The results in Table 4 thus make clear that the evidence of predictability can be significantly enhanced if the coefficient to dividends is allowed to be different from one.

**Is it enough to allow the coefficient to dividends to be different from one?** We have now made clear that the reason why the \( PDR \)-ratio predicts better than the \( pd \)-ratio on its own augmented by the nominal interest rate is that we allow the coefficient to dividends to differ from one. This result could make one wonder whether the main reason why the \( pdR \)- or the \( PDR \)-ratio significantly predicts returns (whereas the \( pd \)-ratio does not) comes from allowing the coefficient to dividends to be different from one. If this was the case, there would be no need to focus on the monetary side of the economy (and include the interest rate), as we propose doing in this paper.

Is it then enough to allow \( \lambda \neq 1 \) and disregard the short interest rate? In order to evaluate this, we construct a price-dividend ratio where \( \lambda \) is allowed to be different from one. We call it the \( PD\lambda \) ratio. Continuing with the 1952:1–2001:4 period that Ang and Bekaert also use, the estimate (using the Stock-Watson procedure) of the \( PD\lambda \) ratio is:

\[
PD\lambda_t = \ln(P_t) - \left[ 4.941 + 1.260\ln(D_t) \right]
\]

with standard errors in parentheses below coefficient estimates.

We can now compare the predictive performance of this \( PD\lambda \)-ratio to that of the standard \( pd \)-ratio and the \( PDR \)-ratio. We find:

\[
\begin{align*}
err_{t,t+1} & = 0.121 - 0.022pd_t & \bar{R}^2 & = 0.885\%, \\
err_{t,t+1} & = 0.010 - 0.028PD\lambda_t & \bar{R}^2 & = 1.103\%, \\
err_{t,t+1} & = 0.007 - 0.081PDR_t & \bar{R}^2 & = 8.508\%,
\end{align*}
\]

where we have reported the numerical values of the \( t \)-statistics in parentheses below the coefficient estimates.
It is clear that the \( \hat{PD} \lambda \)-ratio does not predict much better than the \( \hat{pd} \)-ratio, and that the \( \hat{PDR}_t \) is significantly better at capturing the movements in quarterly excess returns. The results for long-horizon return regressions support this picture.\(^{25}\) We think this is strong evidence for our theory. It is not enough to allow the coefficient to be different from one. The interest rate is necessary.

4.3 Plots

By plotting the different regressors against time, we can provide more intuition. In Figure 1, we show the \( \hat{pd}R \)-ratio and the \( \hat{pd} \)-ratio and in Figure 2 we show the \( \hat{pd}R \)-ratio and the \( py \)-ratio. Shaded areas are NBER recessions. Several issues are noteworthy. The \( \hat{pd}R \)-ratio has a clear tendency to peak right before a recession and fall during the recession. Second, the \( \hat{pd}R \)-ratio is less persistent compared to the \( \hat{pd} \)-ratio or the \( py \)-ratio. Third, the high valuation of the stock market in 2001 is common to all three regressors. Finally, in relation to GDP or dividends, the valuation of the stock market started increasing after the 1982 recession, whereas the valuation of the stock market started increasing only after the 1991 recession when looking at the \( \hat{pd}R \)-ratio.

4.4 Time-variation in the steady state

Our model tells us that it is the deviations from the steady state that should predict returns. Above, we have assumed that the steady state is constant in order to compare most easily with the empirical return prediction literature. In this section, we investigate what happens if we follow up on the suggestion of Lettau and van Nieuwerburgh (2008) and allow for a time-varying steady state.

Lettau and van Nieuwerburgh (2008) show empirically that the price-dividend ratio does not predict on its own, but does so if one allows for a structural break in its mean. Lettau and van Nieuwerburgh focus on the year 1991 as the specific period in time where the mean of the price-dividend ratio changed.\(^{26}\) In this section, our starting point is

\(^{25}\)To save space, the long-horizon results are not shown here, but they are of course available upon request.

\(^{26}\)Lettau & van Nieuwerburgh (2008) discuss briefly several reasons why the mean of the price-dividend ratio could have changed. The volatility of aggregate consumption had declined (Lettau, Ludvigson & Wachter, 2008), risk sharing among households or regions has increased (Lustig & van Nieuwerburgh, 2008), the tax code could have changed (McGrattan & Prescott, 2005), or new entrants to the equity market could have depressed the expected equity premium (Vissing-Jørgensen, 2002). We show in an equilibrium framework why a change in the mean of a predictive regressors has direct implications for the predictive regression, whereas Lettau & van Nieuwerburgh rely on the definition of returns, and the Campbell-Shiller approximation thereof, to illustrate why a change in the mean of the predictor should have implications for future expected returns.
simply the thorough analysis of Lettau and van Nieuwerburgh. We also investigate 1991 as the point in time where a structural break in the steady state mean potentially has occurred.

The empirical analysis starts with Figure 3 showing the time series of the demeaned predictors, where we use mean calculated from the total sample period to demean the series. We show the demeaned series, as the valuation ratios and the \( \frac{pd}{R} \)-ratio can then be plotted on graphs with the same scales on the axes. In addition, we calculate the means of the predictor during two sub periods: the 1947:1–1991:4 period and during the 1992:1–2007:4 period. The graph with the \( pd \)-ratio clearly shows the remarkable change in the \( pd \)-ratio. The mean of the \( pd \)-ratio changes from \(-0.20\) during the 1947:1–1991:4 period to \(0.57\) during the 1992:1–2007:4 period. One way to scale this change, is to divide it with the range over which the \( pd \)-ratio has moved. We find that \( \frac{0.57 - (-0.20)}{\text{max}(pd) - \text{min}(pd)} = 42.73\% \), where \( \text{max}(pd) \) is the maximum value of the \( pd \)-ratio during the complete sample period (1947:1–2007:4) and \( \text{min}(pd) \) is the minimum value. In the same way, we find the change in the mean of the \( py \)-ratio (compared to the range over which the \( py \)-ratio has moved) to be \( 19.95\% \), while the mean of the \( \frac{pd}{R} \)-ratio has changed with \( 2.90\% \), and the mean of the interest rate has fallen with \( 8.33\% \). In other words, the change in the \( pd \)-ratio is big, whereas the change in the other predictors we use is not that big.

4.4.1 Regression results. Table 5 presents the results obtained if regressing excess returns on the predictors that have been demeaned by the means over the two different subsamples; 1947–1991 and 1992:1–2007:4, as in Lettau and van Nieuwerburgh (2008). The results in Table 5 can be directly compared to the results in Table 2 based on a constant mean.

The first issue to notice is the replication of the finding of Lettau and van Nieuwerburgh (2008):\(^{27}\) The \( pd \)-ratio captures a significantly larger fraction of the variation in excess returns if allowing for a break in the mean of the \( pd \)-ratio. Indeed, the \( R^2 \) of the regression is now \( 4.84\% \), which can be compared to the \( 1.61\% \) shown in Table 2. Likewise, the \( t \)-statistic has increased from being barely below \(-2\) (in Table 2) to being a now clearly significant \(-3.81\).

Compared to Lettau and van Nieuwerburgh (2008), the additional information provided in Table 5 is two-fold. First, we show that an even larger fraction of excess returns variation is captured if both the (demeaned) price-dividend ratio and the (demeaned)...

---

\(^{27}\)Lettau & van Nieuwerburgh (2008) use annual 1927-2004 CRSP data. We use quarterly 1947-2008 S&P 500 data. We thus confirm that also for these data, the Lettau & van Nieuwerburgh (2008) finding goes through.
interest rate are used compared to the univariate regression using only the (demeaned) \(pd\)-ratio (the \(R^2\) goes from 4.84% to 7.20%). In other words, the general finding that we want to promote in this paper – that adding the interest rate to a valuation ratio helps in capturing the variation in expected returns – does not arise because we fail to take into account the structural break in the \(pd\)-ratio that Lettau and van Nieuwerburgh (2008) identify. Our findings simply become stronger if we allow for this break.

Second, it is only the \(pd\)-ratio that is affected by the break in 1991. The performances of the other predictors basically do not change when allowing for the structural break, as can be seen by comparing the results for the other predictors in Tables 2 and 5.

4.4.2 Stability over time in the predictive regressions. In order to better understand the importance of the interest rate, figure 4 shows how the estimates of the coefficient to the \(pd\)-ratio (the standard \(pd\)-ratio, i.e. the one not allowing for a shift in the steady state) and its \(t\)-statistic behave during the 1990s. As the graph clearly demonstrates, there is thus extensive time variation in the coefficient to the \(pd\)-ratio. Table 6 shows that allowing for a structural break in the mean eliminates the time-variation in the coefficient.

Table 6 also demonstrates that allowing for a structural break in the mean of the \(pd\)-ratio does not change the predictive information contained in the \(pd\)-ratio. Again this can be made even clearer if one considers recursive estimates of the coefficient to the \(pd\)-ratio. These are shown in figure 5 (in this figure, we also update the cointegration coefficients in the \(pd\)-ratio each quarter; we comment on the cointegration estimates below).

Why is the time-variation in the coefficient to the \(pd\)-ratio much smaller than the time-variation in the coefficient to the \(pd\)-ratio? To provide some insights into this, consider figure 6, which shows the recursive estimates of the coefficients in the \(pd\)-ratio (except the constant - the estimates of the constant are very stable over time). As shown, there is only little variation over time in the estimate of the cointegration coefficient to the CPI. The coefficient to the level of dividends does not change much either. On the other hand, the coefficient to the interest rate changes somewhat; in particular, it increases (in absolute size) when updating the coefficient for 1995-2002. This is interesting because the coefficient increases exactly during the period where the predictive coefficient to the \(pd\)-ratio (figure 4) gets smaller and smaller. We can thus conclude that the reason why the \(pd\)-ratio remains significant when including data for the 1990s, where the commonly used \(pd\)-ratio loses its predictive power, is that the interest rate becomes more important during the late 1990s. Hence, the \(pd\)-ratio becomes insignificant during a period where
the interest rate plays an increasingly important role. The \( \text{pd}R \)-ratio, on the other hand, allows the interest rate to play a role and consequently remains significant during the 1990s. We believe that this is an important finding in our paper.

### 4.5 Out-of-sample tests

Most of the empirical analysis above is based on full-sample estimations. As argued in Lettau and Ludvigson (2008), this is reasonable if one wants to determine whether a particular theory is backed by the data or not. In other words, when doing an out-of-sample analysis, one automatically loses information, as only the last estimation of a recursive estimation uses the whole dataset, and all other estimations are based on limited samples. On the other hand, the parameters that are used to construct the \( \text{pd}R \)-ratio are based on the full sample, and, hence, there is a potential look-ahead bias in the predictive regressions using the \( \text{pd}R \)-ratio.

To address this issue, Table 6 shows out-of-sample \( R^2 \)s that Goyal and Welsh (2008) and Campbell and Thompson (2008) have introduced as a convenient metric to evaluate out-of-sample forecastability. The out-of-sample \( R^2 \) is calculated as

\[
R^2_{\text{OOS}} = 1 - \frac{\sum_{t=1}^{T}(er_t - \hat{er}_t)^2}{\sum_{t=1}^{T}(er_t - \overline{er})^2},
\]

where \( er_t \) is the period \( t \) excess return, \( \hat{er}_t \) is the predicted period \( t \) excess return, and \( \overline{er} \) is the historical average excess return estimated through period \( t - 1 \). A negative \( R^2_{\text{OOS}} \) thus implies that the model used to generate the predictions \( \hat{x}_t \) predicts worse (has a higher mean squared forecast error) than a simple model that uses the updated historical mean of excess return as its predictor.

When doing out-of-sample forecasts, one has to pick an initial sample upon which the first estimation can be done. To make sure that our results are not driven by a particular choice of an initial sample period, we present results based upon four different initial starting points. We also present results from two kinds of \( \text{pd}R \) ratios. One where we fix the cointegration parameters used to calculate the \( \text{pd}R \) ratios at their full sample values and one where we recursively estimate the cointegration parameters, i.e. we use only information available at the point in time where the forecasts are generated. The difference between the two sets of models tells us something about how much a practitioner loses in out-of-sample accuracy compared to the hypothetical situation where the practitioner knows the full-sample values of the parameters used to construct the \( \text{pd}R \)-ratio.

The results from the out-of-sample tests are relatively clear. Regardless of the specific initial period chosen for the out-of-sample tests, the \( R^2_{\text{OOS}} \) from the models where the \( \text{pd}R \)-ratio is used as a predictor are positive. In other words, the \( \text{pd}R \)-ratio generates more accurate predictions than the updated historical mean does that Goyal and Welsh (2008) argue is difficult to “beat”. Table 8 also shows, as expected, that the predictive accuracy
is lower when using the model where the parameters in the \( \hat{p}dR \)-ratio are recursively updated compared to the model where they are fixed at their full-sample values. We stress, though, that for the model where the parameters are recursively updated, the \( R^2_{OOS} \) are also positive. Finally, we see that the \( \text{cay} \) ratio often does not predict better than an updated historical mean.

A concluding remark concerning the out-of-sample tests is that even if the \( \hat{p}dR \)-ratio fares relatively well in terms of out-of-sample predictions, it should be remembered that Cochrane (2008) argues that a high (or a low) \( R^2_{OOS} \) does not give stronger statistical evidence on return predictability than in-sample tests. What is reassuring here, though, is that both the in-sample tests and the out-of-sample tests point towards the \( \hat{p}dR \)-ratio being a reasonable predictor.

### 4.6 Explorations into some asset pricing implications.

Taking into account the monetary dimension of the economy improves the ability of an otherwise standard asset pricing model to predict returns. Yet, one cannot assess from the previous analysis whether the improvement due to money is obtained for reasonable values of the structural parameters. To address this issue indirectly, we calibrated the model to check whether money helps in getting the stochastic discount factor to fall within the Hansen-Jagannathan (1991) variance bounds.

Using the stochastic discount factor \( \Lambda_t \) as given in Eq. (13), we computed the ratio of the standard deviation \( \sigma(\Lambda) \) to the mean \( E(\Lambda) \) of the stochastic discount factor. Our purpose is to compare the behavior of \( \sigma(\Lambda)/E(\Lambda) \) with the Sharpe ratios of the assets for different values of the structural parameters entering Eq. (13); for instance, the real economy is the particular case where the weight of money in the utility function is set to 0 (i.e. \( \phi = 1 \)).

The Sharpe ratio of the S&P 500 is 0.24. The Hansen-Jagannathan bound specifies that \( \sigma(\Lambda)/E(\Lambda) \) should be greater than the Sharpe ratio attained by any portfolio. Table 7 shows the results of our calibration.

The first – standard – point to note is the complete failure of the real economy. Indeed, the first column of Table 9 shows values of \( \sigma(\Lambda)/E(\Lambda) \) for the real economy (the economy with \( \phi = 1 \)) and these values are all far from 0.24, even for a risk aversion coefficient of 10 (\( \gamma_c = 10 \)). This feature of the real economy is reminiscent of the existence of an equity premium puzzle.

Does the picture change if one allows for a monetary economy, i.e. if \( \phi \) is allowed to be smaller than one? We find that including money improves the picture considerably as
\( \sigma(\Lambda)/E(\Lambda) \) is above the lower bound (above 0.24) for values of the parameter measuring the curvature of the utility function relative to real balances \((\gamma_m)\) above five combined with values of the parameter measuring the share of real balances in the utility function \((1 - \phi)\) above 20%; for instance, for \( \phi = 0.7 \) and \( \gamma_m = 5 \), \( \sigma(\Lambda)/E(\Lambda) = 0.234 \) for \( \gamma_c = 1 \). Even for higher values of \( \phi \) (lower weight to money in the utility function), the picture improves compared to the real economy case. For instance, for \( \phi = 0.9 \) and \( \gamma_m = 10 \), one has come a long way in explaining the equity premium puzzle.

Are such values economically reasonable? The results reported in Christiano et al. (2005) are illuminating when trying to answer this question. They estimated the curvature parameter for real balances for eight specifications of the utility function and/or the economic environment (including habit, investment adjustment costs, etc.) and obtained a robust estimate for the real balances curvature parameter of around 11. With respect to \( 1 - \phi \) (the parameter measuring the share of real money balances in the utility function), a recent calibration by Balduzzi (2007) yields a value as low as 0.31.

It seems fair to conclude that the improvement obtained by introducing money does not come at the cost of a new monetary equity premium puzzle. The stochastic discount factor in a monetary economy is sufficiently volatile to explain asset prices for reasonable values of the preference parameters in contrast to the equity-premium puzzle that characterizes a real economy.

### 4.7 Robustness of our results

We have in different ways evaluated how robust our results are. We mention only the main findings here. An appendix on our webpages provides detailed investigations and results.

In order to evaluate whether our main findings hold for portfolios other than the market portfolio, we regressed the 25 Fama-French portfolios’ excess returns on, respectively, the \( pd \)-ratio, the \( pd \)-ratio and the interest rate, the \( py \)-ratio and the interest rate, and the \( pdR \)-ratio. The overall predictability patterns we found were the same as mentioned earlier in this paper, i.e. our results hold also for other portfolios than the market portfolio. In addition, Ken French’s website provides the return on a market portfolio consisting of all NYSE, AMEX, and NASDAQ firms, in contrast to the market return we have used until now (i.e. the returns on the S&P 500). We find the same kinds of results for this definition of the market as for the S&P 500 used in this paper.

We also controlled with the slope of the term structure and the default spread. Adding one of these variables to an otherwise univariate regression did not improve the fit of excess
returns. This is in line with what we would expect. Our theoretical framework says that it is the interest rate alone that should matter together with valuation ratios – not other measures of the situation on the credit market together with valuation ratios. We have shown strong empirical evidence that the interest rate indeed does matter together with valuation ratios. In the robustness tests just mentioned, we show that other measures of the situation on the credit market are not significant.

Finally, upon examining the total payout ratio of Boudoukh et al. (2007), we find that it is marginally significant in a univariate regression. The total payout ratio augmented by the interest rate basically captures the same fraction of the variation in returns as does the $pd$-ratio augmented with the interest rate.

5 Conclusion.

This paper shows that the nominal interest rate, through its function as the relative price of holding money, matters theoretically and empirically for asset-return predictability. We show empirically that financial ratios augmented by the nominal rate perform better in a predictive regression than a standard univariate predictive regression does using only a financial ratio. We also note that the crucial assumption for the dividend yield to be a predictor is that consumption is equal to dividend. We show that the $pdR$ cointegration residual (which allows for leverage in the economy) predicts even better, and, in particular, better than the $pd$-ratio and the $pd$-ratio augmented by the interest rate.

The results we report are robust and appear across a wide range of empirical specification and sample periods. In addition, we note that the monetary economy not only generates new interesting predictor variables but also helps explain the equity-premium puzzle for reasonable values of the structural parameters of the model.

The next most natural step to take is to use the insights provided in this paper to investigate their implications for conditional asset pricing models and dynamic asset allocation.
Appendix A

In this Appendix, we provide the steps that lead to Eq. (20).

Consider the two pricing equations for the equity and the riskless asset that we repeat here for convenience:

\[ \frac{p_t}{c_t} = \int_t^\infty e^{-\rho(s-t)} E_t \left[ \left( \frac{R_s}{R_t} \right)^{\delta_R} \left( \frac{c_s}{c_t} \right)^{\delta_c+1} \right] ds \]  

\[ \frac{1}{R_t} = \int_t^\infty e^{-\rho(s-t)} E_t \left[ \left( \frac{R_s}{R_t} \right)^{\delta_R+1} \left( \frac{c_s}{c_t} \right)^{\delta_c} \left( \frac{\Pi_s}{\Pi_t} \right)^{-1} \right] ds \]

Log-linearizing these two equations, we get:

\[ cte + cte \cdot (\ln p_t - \ln c_t) \approx \int_t^\infty e^{-\rho(s-t)} \kappa_s E_t \left[ 1 + (\delta_c + 1) \Delta \ln c_s + \delta_R \Delta \ln R_s \right] ds \]  

and

\[ cte + cte \cdot \ln \widehat{R}_t \approx \int_t^\infty e^{-\rho(s-t)} \lambda_s E_t \left[ 1 + \delta_c \Delta \ln c_s + (\delta_R + 1) \Delta \ln R_s - \Delta \ln \Pi_s \right] ds \]

where we use the following notation:

\[ \Delta \widehat{x}_s = x_s - x_t \quad \text{for } s > t \]

c\text{te}, \kappa_s and \lambda_s are generic notation for log linearization constants that one does not have to specify. Hence, we have:

\[ cte + cte \cdot (\ln p_t - \ln c_t) \approx \int_t^\infty e^{-\rho(s-t)} \kappa_s E_t \left[ 1 + (\delta_c + 1) E_t \left[ \Delta \ln c_s \right] + \delta_R E_t \left[ \Delta \ln R_s \right] \right] ds \]

and

\[ cte + cte \cdot \ln \widehat{R}_t \approx \int_t^\infty e^{-\rho(s-t)} \lambda_s \left[ 1 + \delta_c E_t \left[ \Delta \ln c_s \right] + (\delta_R + 1) E_t \left[ \Delta \ln R_s \right] - E_t \left[ \Delta \ln \Pi_s \right] \right] ds \]

\[ cte + cte \cdot (\ln p_t - \ln c_t) \approx \int_t^\infty e^{-\rho(s-t)} \kappa_s \left[ 1 + (\delta_c + 1) E_t \left[ \Delta \ln c_s \right] + \delta_R E_t \left[ \Delta \ln R_s \right] \right] ds \]

\[ cte + cte \cdot \ln \widehat{R}_t \approx \int_t^\infty e^{-\rho(s-t)} \lambda_s \left[ 1 + \delta_c E_t \left[ \Delta \ln c_s \right] + (\delta_R + 1) E_t \left[ \Delta \ln R_s \right] - E_t \left[ \Delta \ln \Pi_s \right] \right] ds. \]

These constants are related to the steady state values of the variables as usual.
We specify the dynamics of consumption growth, the change in the interest rate, and inflation as being driven by a real state variable $x_t^r$ and a nominal state variable $x_t^N$ in the following ways:

\[
\Delta \ln c_s = \rho_{0,s-t}^c + \rho_{1,s-t}^c x_t^r + \rho_{2,s-t}^c x_t^N + \sigma_{r,s-t}^c \sqrt{\tilde{x}_t^c} \varepsilon_s^r + \sigma_{N,s-t}^c \sqrt{\tilde{x}_t^N} \varepsilon_s^N
\]

\[
\Delta \ln R_s = \rho_{0,s-t}^R + \rho_{1,s-t}^R x_t^r + \rho_{2,s-t}^R x_t^N + \sigma_{r,s-t}^R \sqrt{\tilde{x}_t^r} \varepsilon_s^r + \sigma_{N,s-t}^R \sqrt{\tilde{x}_t^N} \varepsilon_s^N
\]

\[
\Delta \ln \Pi_s = \rho_{0,s-t}^\Pi + \rho_{1,s-t}^\Pi x_t^r + \rho_{2,s-t}^\Pi x_t^N + \sigma_{r,s-t}^\Pi \sqrt{\tilde{x}_t^r} \varepsilon_s^r + \sigma_{N,s-t}^\Pi \sqrt{\tilde{x}_t^N} \varepsilon_s^N
\]

where $\varepsilon_s^r$ and $\varepsilon_s^N$ are two pure shocks in the economy, a real and a nominal, and the $\rho$-parameters are deterministic functions of time. By specifying the dynamics of $\Delta \ln c_s$, $\Delta \ln R_s$, and $\Delta \ln \Pi_s$ these ways, we assume (i) that the uncertainties characterizing the growth rates of the variables are time-varying and (ii) that the expected growth rates of the variables are linear functions of the state variables in the economy.

We discuss our assumption of time-varying uncertainty first. Time-varying uncertainty is by now a well accepted-ingredient in many asset pricing models if these models should be able to capture stylized empirical facts. For the variables we specify here, we are motivated by Bansal and Yaron (2004) and Lettau et al. (2008) for the consumption growth rate process, by Bekaert and Grenadier (2001) for the inflation rate process, and by Dai and Singleton (2002) for the nominal interest rate process. In more detail: Empirical support for a consumption growth process where the conditional volatility is time varying can be found in Bansal and Yaron (2004) and Lettau et al. (2008) who show that changing economic uncertainty (captured by time varying consumption growth volatility) is a key ingredient toward an explanation of traditional asset pricing puzzles (e.g., the equity premium puzzle, the risk-free rate puzzle etc.). Similarly, time varying consumption volatility guaranties time varying market prices of risk which leads to predictability of excess returns. Regarding the interest rate uncertainty, Dai and Singleton (2002), for example, were able to match a number of term structure puzzles by allowing for time varying uncertainty of the state variables in the economy. This time varying uncertainty leads to time varying uncertainty of the short term interest rate itself; a feature we incorporate in Eq. (33). Finally, Bekaert and Grenadier (2001) advocate the introduction of stochastic inflation uncertainty in affine asset pricing when pricing simultaneously bonds and stocks.

Regarding our second assumption (that the expected growth rates of the variables are linear functions of the state variables), it is obvious why the real factor influences the growth rate of real consumption and why the nominal factor influences the rate of inflation and the nominal interest rate. What we need to discuss is why the nominal factor influences real consumption growth and why the real factor influences inflation and the nominal interest rate. Regarding consumption, we first note that if money is non neutral,
pure nominal factor will influence real variables such as GDP and consumption growth rates. One key assumption in the New Keynesian literature (the – today – standard way of modelling monetary policy and its effects) is exactly that money is non neutral. In other words, this non neutrality is at the heart of monetary policy activism. Empirical evidence that nominal factors influence the growth rate of real variables such as GDP and consumption can be found in New Keynesian models, such as those in Bouakez et al. (2005) or Christiano et al. (2005).

Concerning the interest rate dynamics, the literature on Taylor rules shows that the behavior of interest rates is better captured if including both a real and a nominal factor.\(^{29}\) Indeed, Eq. (33) can be seen as an extreme case of a Taylor rule where the coefficient on the lagged interest rate has been set equal to 1. Extensive evidence on interest rate smoothing by central banks estimate the coefficient on the lagged interest rate in a Taylor rule type regression as being very close to 1 (see the literature review in Sack and Wieland, 2000). Finally, inflation dynamics are in New-Keynesian models often specified as dependent upon both nominal variables (such as inflation itself) and real variables, such as the output gap; for one recent example, see Bekaert et al. (2006). All in all, we follow the literature in assuming that the three conditional expectations will in general be driven by both sources of risk.

Substituting these processes Eqs. into (30) and (31) at the relevant positions, we have:

\[
\begin{align*}
cte + cte \ast \left( \ln \frac{p_t}{c_t} \right) & = \xi_0 + \xi_1 x^r_t + \xi_2 x^N_t \tag{35} \\
cte + cte \ast \ln R_t & = \varepsilon_0 + \varepsilon_1 x^r_t + \varepsilon_2 x^N_t \tag{36}
\end{align*}
\]

where

\[
\begin{align*}
\xi_0 & = \int_t^\infty e^{-\rho(s-t)\kappa_s} \left[ 1 + (\delta_c + 1) \rho^c_{0,(s-t)} + \delta_R \rho^R_{0,(s-t)} \right] ds \tag{37} \\
\xi_1 & = \int_t^\infty e^{-\rho(s-t)\kappa_s} \left[ (\delta_c + 1) \rho^c_{1,(s-t)} + \delta_R \rho^R_{1,(s-t)} \right] ds \tag{38} \\
\xi_2 & = \int_t^\infty e^{-\rho(s-t)\kappa_s} \left[ (\delta_c + 1) \rho^c_{2,(s-t)} + \delta_R \rho^R_{2,(s-t)} \right] ds \tag{39}
\end{align*}
\]

\(^{29}\)A Taylor rule states that the nominal interest rate is driven by inflation dynamics and real output dynamics (the deviation of GDP from its potential full employment level).
Using Eq. (13), we also have:

\[
\begin{align*}
\varepsilon_0 &= \int_{t}^{\infty} e^{-\rho(s-t)} \lambda_s \left[ 1 + \delta_c \rho^*_0(s-t) + (\delta_R + 1) \rho^*_0(s-t) - \rho^*_0(s-t) \right] ds \\
\varepsilon_1 &= \int_{t}^{\infty} e^{-\rho(s-t)} \lambda_s \left[ \delta_c \rho^*_1(s-t) + (\delta_R + 1) \rho^*_1(s-t) - \rho^*_1(s-t) \right] ds \\
\varepsilon_2 &= \int_{t}^{\infty} e^{-\rho(s-t)} \lambda_s \left[ \delta_c \rho^*_2(s-t) + (\delta_R + 1) \rho^*_2(s-t) - \rho^*_2(s-t) \right] ds
\end{align*}
\]  

Solving this system yields:

\[
\begin{bmatrix}
x_t^r \\
x_t^N
\end{bmatrix} = \begin{bmatrix}
\xi_1 & \xi_2 \\
\varepsilon_1 & \varepsilon_2
\end{bmatrix}^{-1} \begin{bmatrix}
cte - \xi_0 \\
cte - \varepsilon_0
\end{bmatrix} + \begin{bmatrix}
\xi_1 & \xi_2 \\
\varepsilon_1 & \varepsilon_2
\end{bmatrix}^{-1} \begin{bmatrix}
cte * (\ln p_t / c_t) \\
cte * \ln R_t
\end{bmatrix}
\]  

(41)

(42)

(43)

We implicitly assume that the matrix \( \begin{bmatrix}
\xi_1 & \xi_2 \\
\varepsilon_1 & \varepsilon_2
\end{bmatrix} \) is invertible. We finally have:

\[
x_t^r = \frac{1}{\xi_1 \varepsilon_2 - \xi_2 \varepsilon_1} (\varepsilon_2 (cte - \xi_0) - \xi_2 (cte - \varepsilon_0)) + \frac{1}{\xi_1 \varepsilon_2 - \xi_2 \varepsilon_1} cte * (\ln p_t / c_t) - \frac{1}{\xi_1 \varepsilon_2 - \xi_2 \varepsilon_1} \varepsilon_2 * cte * \ln R_t
\]

(45)

\[
x_t^N = \frac{1}{\xi_1 \varepsilon_2 - \xi_2 \varepsilon_1} (-\varepsilon_1 (cte - \xi_0) + \xi_1 (cte - \varepsilon_0)) - \frac{1}{\xi_1 \varepsilon_2 - \xi_2 \varepsilon_1} \varepsilon_1 * cte * (\ln p_t / c_t) + \frac{1}{\xi_1 \varepsilon_2 - \xi_2 \varepsilon_1} \varepsilon_1 * cte * \ln R_t
\]

(46)

(47)

(48)

At this stage, we already have the first result we need: The state variables in the economy can be written as a linear combination of the two variables \( \ln p_t / c_t \) and \( \ln R_t \).

We now compute the expected excess return on the equity. Log-linearizing Eq. (2), we have:

\[
E_t [r_{t+1} - \bar{r}_t^f] = cte * cov_t [\Delta \ln \Lambda_{t+1}; \Delta \ln p_{t+1}].
\]

(49)

Using Eq. (13), we also have:

\[
\ln \Lambda_t = \ln A - \rho t + \delta_c \ln c_t + \delta_R \ln R_t
\]

\[
\Delta \ln \Lambda_{t+1} = \rho + \delta_c \Delta \ln c_{t+1} + \delta_R \Delta \ln R_{t+1}.
\]

(50)

(51)

Using the processes for consumption growth and the change in the interest rate given in
Eqs. (32) and (33), the dynamics of the pricing kernel are given as:

\[
\Delta \ln \Lambda_{t+1} = \rho + \delta_c \left[ \rho^c_{0,1} + \rho^c_{1,1} x^r_t + \rho^c_{2,1} x^N_t + \sigma^c_{r,1} \sqrt{x^r_t} \epsilon^r_t + \sigma^c_{N,1} \sqrt{x^N_t} \epsilon^N_{t+1} \right] + \delta R \left[ \rho^R_{0,1} + \rho^R_{1,1} x^r_t + \rho^R_{2,1} x^N_t + \sigma^R_{r,1} \sqrt{x^r_t} \epsilon^r_{t+1} + \sigma^R_{N,1} \sqrt{x^N_t} \epsilon^N_{t+1} \right]
\]

(52)

and therefore:

\[
\Delta \ln \Lambda_{t+1} = \rho + \delta_c \rho^c_{0,1} + \delta R \rho^R_{0,1}
\]

(53)

Using Eq. (35) we have:

\[
\widehat{\ln \ p_t} = \ln c_t + \frac{(\xi_0 - \text{cte})}{\text{cte}} x^r_t + \frac{\xi_1}{\text{cte}} x^r_t + \frac{\xi_2}{\text{cte}} x^N_t
\]

(54)

and therefore:

\[
\Delta \widehat{\ln \ p_{t+1}} = \Delta \widehat{\ln \ c_{t+1}} + \frac{\xi_1}{\text{cte}} \Delta x^r_{t+1} + \frac{\xi_2}{\text{cte}} \Delta x^N_{t+1}.
\]

(55)

We assume that the state variables follow AR(1) processes with stochastic volatility:

\[
x^r_{t+1} = \phi^r_0 + \phi^r_1 x^r_t + \phi^r_2 \sqrt{x^r_t} \epsilon^r_{t+1}
\]

(56)

\[
x^N_{t+1} = \phi^N_0 + \phi^N_1 x^N_t + \phi^N_2 \sqrt{x^N_t} \epsilon^N_{t+1}
\]

(57)

The AR(1) assumption for the state variables is standard in the return predictability, conditional asset pricing, and the dynamic asset allocation literature (see, for example, Stambaugh, 1999; Santos and Veronesi, 2006; and Sangvinatsos and Wachter, 2005). Therefore:

\[
\Delta x^r_{t+1} = \phi^r_0 + (\phi^r_1 - 1) x^r_t + \phi^r_2 \sqrt{x^r_t} \epsilon^r_{t+1}
\]

(58)

\[
\Delta x^N_{t+1} = \phi^N_0 + (\phi^N_1 - 1) x^N_t + \phi^N_2 \sqrt{x^N_t} \epsilon^N_{t+1}.
\]

(59)

We thus have:

\[
\widehat{\Delta \ln \ p_{t+1}} = \rho^c_{0,1} + \frac{\xi_1}{\text{cte}} \phi^r_0 + \frac{\xi_2}{\text{cte}} \phi^N_0
\]

(60)

\[
+ \left[ \rho^c_{1,1} + \frac{\xi_1}{\text{cte}} (\phi^r_1 - 1) \right] x^r_t + \left[ \rho^c_{2,1} + \frac{\xi_2}{\text{cte}} (\phi^N_1 - 1) \right] x^N_t
\]

(61)

\[
+ \left[ \sigma^c_{r,1} + \frac{\xi_1}{\text{cte}} \phi^r_2 \right] \sqrt{x^r_t} \epsilon^r_{t+1} + \left[ \sigma^c_{N,1} + \frac{\xi_2}{\text{cte}} \phi^N_2 \right] \sqrt{x^N_t} \epsilon^N_{t+1}
\]

(62)
It turns out thus that:

\[
E_t [r_{t+1}^r - r_{t+1}^f] = cte \left[ \sigma_{c,1}^r + \frac{\xi_1}{cte} \phi_2^r \right] \left[ \delta_R \sigma_{r,1}^R + \delta_c \sigma_{r,1}^c \right] x_t^c \sigma_{r,1}^r + cte \left[ \sigma_{c,1}^N + \frac{\xi_2}{cte} \phi_2^N \right] \left[ \delta_c \sigma_{c,1}^c + \delta_R \sigma_{N,1}^R \right] x_t^N \sigma_{c,1}^N.
\]

As a consequence, using (44), we have:

\[
E_t [r_{t+1}^r - r_{t+1}^f] = \alpha_0 + \alpha_1 \left( \ln p_t - \ln c_t \right) + \alpha_2 \ln R_t.
\]

where:

\[
\alpha_0 = cte \left[ \sigma_{c,1}^r + \frac{\xi_1}{cte} \phi_2^r \right] \left[ \delta_R \sigma_{r,1}^R + \delta_c \sigma_{r,1}^c \right] \frac{1}{\xi_1 \varepsilon_2 - \xi_2 \varepsilon_1} \left( \varepsilon_2 (cte - \varepsilon_0) - \varepsilon_2 (cte - \varepsilon_0) \right) \sigma_{r,1}^r + cte \left[ \sigma_{c,1}^N + \frac{\xi_2}{cte} \phi_2^N \right] \left[ \delta_c \sigma_{c,1}^c + \delta_R \sigma_{N,1}^R \right] \frac{1}{\xi_1 \varepsilon_2 - \xi_2 \varepsilon_1} \left( \varepsilon_1 (cte - \varepsilon_0) + \xi_1 (cte - \varepsilon_0) \right)
\]

\[
\alpha_1 = cte \left[ \sigma_{c,1}^r + \frac{\xi_1}{cte} \phi_2^r \right] \left[ \delta_R \sigma_{r,1}^R + \delta_c \sigma_{r,1}^c \right] \frac{1}{\xi_1 \varepsilon_2 - \xi_2 \varepsilon_1} \sigma_{r,1}^r \varepsilon_2 * cte
\]

\[
\alpha_2 = -cte \left[ \sigma_{N,1}^c + \frac{\xi_2}{cte} \phi_2^N \right] \left[ \delta_c \sigma_{c,1}^c + \delta_R \sigma_{N,1}^R \right] \frac{1}{\xi_1 \varepsilon_2 - \xi_2 \varepsilon_1} \sigma_{c,1}^c \varepsilon_2 * cte
\]

In a pure real economy, the second term in the definition of \( \alpha_2 \) is 0 by definition (since for example \( \sigma_{c,1}^c = 0 \)). On the other hand, from (35), \( \xi_2 \) should also be equal to 0 and thus \( \alpha_2 = 0 \) in a real economy. \( \xi_2 = 0 \) in a real economy since \( \kappa_s \) in its definition is the product of the steady state growth rate of consumption and the nominal interest rate. The later is 0 by construction in a real economy.

This completes the proof.
Appendix B: Data description

We describe in this appendix the data series we use and their sources.

LEVEL OF NOMINAL SHARE PRICE ($P_t$): Our source is Robert Shiller’s webpage with Online Data. R. Shiller provides monthly figures for the S&P 500. The S&P 500 is a value-weighted index of 500 leading companies in leading industries of the U.S. economy. The monthly figures are averages of daily closing prices. We use the first month of a quarter as our quarterly series.

LEVEL OF NOMINAL DIVIDENDS ($D_t$): Our source is Robert Shiller’s webpage with Online Data. The nominal dividend series are dividends per share, 12 months moving total adjusted to index. R. Shiller states that these are from a table entitled “Earnings, Dividends and Price–Earnings Ratio — Quarterly.” of Standard and Poor’s Statistical Service Security Price Index Record.

LEVEL OF NOMINAL AGGREGATE CONSUMPTION ($C_t$): Our source is the Federal Reserve Economic Data (FRED®) Database, maintained by the Federal Reserve Bank of St. Louis. The nominal consumption data are seasonally adjusted figures for Personal Consumption Expenditures in billions of current dollars.

LEVEL OF NOMINAL AGGREGATE GROSS DOMESTIC PRODUCT ($Y_t$): Our source is the Bureau of Economic Analysis. The GDP data are seasonally adjusted figures for GDP in billions of current dollars.

PRICE LEVEL ($Π_t$): Our source is Robert Shiller’s webpage with Online Data. The price level series is the Seasonally Adjusted Consumer Price Index – All Urban Consumers. We use the first month of a quarter as our quarterly series.

LEVEL OF NOMINAL SHORT TERM INTEREST RATE ($R_t$): Our source is the Federal Reserve Economic Data (FRED®) Database. The nominal interest rate is the 3-Month Treasury Bill: Secondary Market Rate. Fed maintains a monthly series which consists of monthly averages of business days. We use the first month of a quarter as our quarterly series.

EXCESS RETURNS: Given the data definitions, we calculate the continuously compounded excess returns as $er_t = \ln[1 + (P_t + D_{t-1}/4)/P_{t-1} - R_t/400]$. We divide the dividend series by 4 to convert to figures that can be used in a calculation of quarterly returns and the same goes for the interest rate.

VALUATION RATIOS: We calculate the valuation ratios as $\ln(P_t/X_t)$ where $X_t = D_t, C_t, or Y_t$, with $X_t$ being the level of the fundamental when entering period $t$, such that only past information is used to predict future returns.
TERM SPREAD: Our source is the Federal Reserve Economic Data (FRED®) Database. The term spread is the difference between 10-Year Treasury Constant Maturity Rate and the 3-Month Treasury Bill: Secondary Market Rate. Fed maintains a monthly series which consists of monthly averages of business days. We use the first month of a quarter as our quarterly series.

DEFAULT PREMIUM: Our source is the Federal Reserve Economic Data (FRED®) Database. The default spread is calculated as the difference between Moody’s Seasoned Baa Corporate Bond Yield and Moody’s Seasoned Aaa Corporate Bond Yield. The bonds are long-term bonds with “remaining maturities as close as possible to 30 years”. Fed maintains a monthly series which consists of monthly averages of business days. We use the first month of a quarter as our quarterly series.

25 FAMA-FRENCH PORTFOLIO: Our source is the Data Library on Ken French webpage. We use the 25 portfolios formed on size and book-to-market. Ken French maintains a monthly series, and we use the first month of a quarter as our quarterly series.

Appendix C: Estimation of the \( \widehat{pdR} \) cointegration relation

Following Lettau & Ludvigson (2001, 2005), we use the Stock-Watson (1993) Dynamic OLS estimation procedure to estimate the co-integration parameters. The Stock-Watson procedure adds leads and lags of first-order differences of the explanatory variables to the list of explanatory variables in levels, i.e. the estimated model takes the form:

\[
\ln(P_t) = \beta_0 + \beta_D \ln(D_t) + \beta_{CP} \ln(\Pi_t) + \beta_R \ln(1 + R_t) \\
+ \sum_{i=-l}^{l} \gamma_{1,i} \Delta \ln(D_{t-i}) + \sum_{i=-l}^{l} \gamma_{2,i} \Delta \ln(\Pi_{t-i}) \\
+ \sum_{i=-l}^{l} \gamma_{3,i} \Delta \ln(1 + R_{t-i}) + \varepsilon_t.
\] (65)

We set \( l = 4 \). Our results are not dependent on this choice of \( l \).

Preliminary univariate tests (Augmented Dickey-Fuller tests) for a unit root in the time series of \( P_t, D_t, R_t, \) and \( \Pi_t \) are consistent with the time series being non-stationary. It should be said, though, that it is a borderline case whether the interest rate \( R_t \) is non-stationary or not.

**Number of cointegration vectors.** Before actually estimating Eq. (65), we used the Johansen (1991) procedure to test for the number of co-integration vectors. The Johansen procedure returns a *Trace* statistic from a *VAR*-estimation.\(^{30}\) The *Trace* statistic can be

\(^{30}\)The Johansen procedure also produces a \( \lambda - \max \) test statistic. Lütkepohl et al. (2001) compare the size and power properties of the two kinds of co-integration tests and recommend using the *Trace* test.
used to test for the maximum number of cointegration vectors \((r)\) in a VAR-system of \(p\) variables. We have \(p = 4\) variables here \((\ln(P_t), \ln(D_t), \ln(1 + R_t), \text{and } \ln(\Pi_t))\). We use four lags (as in the Stock & Watson tests) in the VAR. The results from the estimation of the Trace test statistic was:

<table>
<thead>
<tr>
<th>Rank</th>
<th>Trace</th>
<th>95%</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r = 0 )</td>
<td>61.92</td>
<td>47.21</td>
<td>54.46</td>
</tr>
<tr>
<td>( r \leq 1 )</td>
<td>32.12</td>
<td>29.68</td>
<td>35.65</td>
</tr>
<tr>
<td>( r \leq 2 )</td>
<td>12.19</td>
<td>15.41</td>
<td>20.04</td>
</tr>
<tr>
<td>( r \leq 3 )</td>
<td>0.02</td>
<td>3.76</td>
<td>6.65</td>
</tr>
</tbody>
</table>

When comparing the Trace statistic for zero co-integration vectors with the 95% or 99% critical value, the null hypothesis of no co-integration between the four variables is rejected; in other words, there is at least one co-integration vector between the four variables. On the other hand, the Trace test statistic for, at most, one co-integration vector cannot be rejected at the 95-99% level. Hence, we continue working with one cointegration relation between the four variables.

**Stationarity of the \( \widehat{pdR} \)-ratio.** The estimated cointegration vector is shown in Eq. (23) in the text. We tested whether the constructed \( pdR \)-ratio is stationary using univariate unit root tests. The ADF test (with 4 lags) was \(-3.688\) and the PP tests \(-3.799\). Due to the fact that the \( \widehat{pdR} \)-ratio is a generated regressor, we cannot use standard Dickey-Fuller critical values. However, we can calculate the critical values from the table of response surfaces provided in MacKinnon (1991). These critical values depend upon the number of observations and the number of time series used in the estimation. The 10% critical value is \(-3.477\) and the 5% critical value is \(-3.777\). The null of a non-stationary \( pdR \)-ratio is thus rejected at a 5%-10% significance level.

**Appendix D: Hodrick’s standard errors and Implied \( R^2 \)s**

The Hodrick (1992) standard errors are calculated from the variance-covariance matrix \( \Omega = Z_0^{-1} S Z_0^{-1} \), where \( Z_0 = E(x_t x'_t) \) with \( x_t \) being a vector of explanatory variables, i.e. \( x_t' = (1, z_t) \). The point in Hodrick (1992) is that \( S \) is created by summing \( x_t x'_{t-j} \) into the past, i.e. estimated as:

\[
\hat{S} = \frac{1}{T} \sum_{t=k}^{T} w_{k,t} w'_{k,t}
\]

where

\[
w_{k,t} = e_{t+1} \left( \sum_{i=0}^{t-1} x_{t-i} \right)
\]

with \( e_{t+1} \) being the one-step ahead forecast error which will be serially uncorrelated under the null hypothesis of no predictability (\( e_{t+1} \) is estimated as the residuals from a regression
of one-period excess returns on a constant). Asymptotically, $\theta' = (\alpha, \kappa)$ will be distributed as $\sqrt{T}(\hat{\theta} - \theta) \sim N(0, \Omega)$.

One can compare the Hodrick standard errors with those of Hansen & Hodrick (1980) and Newey-West (1987) standard errors. The Hansen & Hodrick (1980) and Newey-West (1987) standard errors use a $S$ that is estimated as $S_T = C_T(0) + \sum_{j=1}^{k-1} \omega_j [C_T(j) + C_T(j)']$ where $C_T(j) = (1/T) \sum_{t=j+1}^{T} (w_{t+k}w_{t+k}'$ and $w_{t+k} = u_{t+k}x_t$. Under the null hypothesis, $u_{t+k} = (e_{t+1} + \ldots + e_{t+k})$, i.e. the Hansen-Hodrick and Newey-West standard errors rely on the summation of $e_{t+1}$ into the future, in contrast to the Hodrick (1992) standard errors, which sums into the past. Remember also that the Newey-West estimator is a special case of the Hansen-Hodrick estimator, where weights $\omega_j = 1 - \left(\frac{j}{T-k}\right)$ are attached to the summation of the $C_T$s, i.e. in Hansen & Hodrick $\omega_j = 1$. We use $k + 1$ lags of $C_T(j)$ when we calculate the Newey-West statistics.

$R^2$s will also be erroneously affected by the overlapping observations when the predictor variable is persistent. Hodrick (1992) proposes using first-order vector autoregressive (VAR) models of the demeaned variables to calculate “Implied $R^2$s” for long-horizon regressions. The first-order VAR-model is $Z_{t+1} = AZ_t + u_{t+1}$. Here, the vector $Z_t$ of demeaned variables is $Z_t = [e_{t,t+1} - E(e_{t,t+1}), z_t - E(z_t)]'$. From this VAR(1) model, the implied $R^2$ from the return equation is calculated as

$$\text{Implied } R^2_{\text{Hodrick}}(k) = 1 - \frac{e'W_k e}{e'V_k e}$$

for a horizon $k$, where $e'W_k e$ is the innovation variance of the sum of $k$ consecutive returns and the variance of the sum of $k$ consecutive returns can be selected as $e'V_k e$. The formulas for $W_k$ and $V_k$ can be found in Hodrick (1992).
References


45


Table 1: Summary statistics

<table>
<thead>
<tr>
<th></th>
<th>pd</th>
<th>pc</th>
<th>py</th>
<th>ln(1 + R)</th>
<th>pdR</th>
<th>er</th>
</tr>
</thead>
<tbody>
<tr>
<td>Means</td>
<td>4.83</td>
<td>-2.13</td>
<td>-2.56</td>
<td>4.59%</td>
<td>-0.05</td>
<td>6.52%</td>
</tr>
<tr>
<td>Std.</td>
<td>0.43</td>
<td>0.37</td>
<td>0.37</td>
<td>2.72%</td>
<td>0.22</td>
<td>13.45%</td>
</tr>
</tbody>
</table>

Panel B: Correlations

<table>
<thead>
<tr>
<th></th>
<th>pd</th>
<th>pc</th>
<th>py</th>
<th>ln(1 + R)</th>
<th>pdR</th>
<th>er</th>
</tr>
</thead>
<tbody>
<tr>
<td>pd</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pc</td>
<td>0.59</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>py</td>
<td>0.67</td>
<td>0.99</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln(1 + R)</td>
<td>-0.13</td>
<td>-0.51</td>
<td>-0.53</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pdR</td>
<td>0.44</td>
<td>0.41</td>
<td>0.41</td>
<td>0.14</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>er</td>
<td>0.01</td>
<td>0.09</td>
<td>0.09</td>
<td>-0.20</td>
<td>-0.05</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Table 2: Regressions of quarterly excess returns on predictor variables. Quarterly 1947:1 – 2007:4 data. For $\hat{cay}$, the sample period is 1952:1 – 2006:4.

<table>
<thead>
<tr>
<th></th>
<th>pd</th>
<th>py</th>
<th>pc</th>
<th>ln(1 + R)</th>
<th>pdR</th>
<th>$\hat{cay}$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Univariate regressions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.</td>
<td>Coefficient</td>
<td>−0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.61%</td>
</tr>
<tr>
<td></td>
<td>t-stat</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>2.02</strong></td>
</tr>
<tr>
<td>2.</td>
<td>Coefficient</td>
<td>−0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.31%</td>
</tr>
<tr>
<td></td>
<td>t-stat</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>−1.20</td>
</tr>
<tr>
<td>3.</td>
<td>Coefficient</td>
<td>−0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.43%</td>
</tr>
<tr>
<td></td>
<td>t-stat</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>−1.33</td>
</tr>
<tr>
<td>4.</td>
<td>Coefficient</td>
<td>−0.42</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.51%</td>
</tr>
<tr>
<td></td>
<td>t-stat</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>2.35</strong></td>
</tr>
<tr>
<td><strong>Panel B: Multivariate regressions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>Coefficient</td>
<td>−0.03</td>
<td>−0.47</td>
<td></td>
<td></td>
<td></td>
<td>4.83%</td>
</tr>
<tr>
<td></td>
<td>t-stat</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>2.42</strong></td>
</tr>
<tr>
<td>6.</td>
<td>Coefficient</td>
<td>−0.04</td>
<td>−0.73</td>
<td></td>
<td></td>
<td></td>
<td>6.28%</td>
</tr>
<tr>
<td></td>
<td>t-stat</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>3.26</strong></td>
</tr>
<tr>
<td>7.</td>
<td>Coefficient</td>
<td>−0.04</td>
<td>−0.71</td>
<td></td>
<td></td>
<td></td>
<td>6.33%</td>
</tr>
<tr>
<td></td>
<td>t-stat</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>3.34</strong></td>
</tr>
<tr>
<td><strong>Panel C: Univariate regressions using cointegration residuals</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td>Coefficient</td>
<td>−0.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8.64%</td>
</tr>
<tr>
<td></td>
<td>t-stat</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>4.91</strong></td>
</tr>
<tr>
<td>9.</td>
<td>Coefficient</td>
<td>1.44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7.12%</td>
</tr>
<tr>
<td></td>
<td>t-stat</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>4.29</strong></td>
</tr>
</tbody>
</table>
Table 3: Forecasting regressions of long-horizon excess returns. Three sets of regressions:

Part 1: Univariate regression on the interest rate (Panel A), the pd-ratio (Panel B), and the py-ratio (Panel C).

Part 2: Bivariate regression on the pd-ratio augmented by the interest rate (Panel D) and the py-ratio augmented by the interest rate (Panel E).


<table>
<thead>
<tr>
<th>Horizon:</th>
<th>Part 1: Univariate regressions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 quart.</td>
</tr>
<tr>
<td>Panel A: Interest rate</td>
<td></td>
</tr>
<tr>
<td>$\hat{r}_{RF}$</td>
<td>$t_{OLS}(\hat{r}_{RF})$</td>
</tr>
<tr>
<td>$\hat{r}_{pd}$</td>
<td>$t_{OLS}(\hat{r}_{pd})$</td>
</tr>
<tr>
<td>$\hat{r}_{py}$</td>
<td>$t_{OLS}(\hat{r}_{py})$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3: continued

<table>
<thead>
<tr>
<th>Horizon:</th>
<th>1 quart.</th>
<th>1 year</th>
<th>2 years</th>
<th>3 years</th>
<th>4 years</th>
<th>5 years</th>
<th>6 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel D: pd-ratio + the interest rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{R}_{pd}$</td>
<td>$-0.03$</td>
<td>$-0.11$</td>
<td>$-0.21$</td>
<td>$-0.30$</td>
<td>$-0.38$</td>
<td>$-0.48$</td>
<td>$-0.57$</td>
</tr>
<tr>
<td>$t^{OLS}(\hat{R}_{pd})$</td>
<td>$-2.62$</td>
<td>$-5.07$</td>
<td>$-6.91$</td>
<td>$-8.57$</td>
<td>$-9.59$</td>
<td>$-10.66$</td>
<td>$-11.07$</td>
</tr>
<tr>
<td>$t^{NW}(\hat{R}_{pd})$</td>
<td>$-2.42$</td>
<td>$-2.53$</td>
<td>$-2.65$</td>
<td>$-3.46$</td>
<td>$4.56$</td>
<td>$-5.61$</td>
<td>$-5.04$</td>
</tr>
<tr>
<td>$t^{Hodrick}(\hat{R}_{pd})$</td>
<td>$-2.39$</td>
<td>$-2.50$</td>
<td>$-2.31$</td>
<td>$-2.19$</td>
<td>$-2.07$</td>
<td>$-2.09$</td>
<td>$-2.03$</td>
</tr>
<tr>
<td>$\hat{R}_{f}$</td>
<td>$-0.47$</td>
<td>$-1.44$</td>
<td>$-2.12$</td>
<td>$-3.06$</td>
<td>$-3.84$</td>
<td>$-4.22$</td>
<td>$-4.09$</td>
</tr>
<tr>
<td>$t^{OLS}(\hat{R}_{f})$</td>
<td>$-3.02$</td>
<td>$-4.24$</td>
<td>$-4.61$</td>
<td>$-5.79$</td>
<td>$-6.53$</td>
<td>$-6.30$</td>
<td>$-5.48$</td>
</tr>
<tr>
<td>$t^{NW}(\hat{R}_{f})$</td>
<td>$-2.75$</td>
<td>$-2.61$</td>
<td>$-3.18$</td>
<td>$-3.89$</td>
<td>$-3.78$</td>
<td>$-3.12$</td>
<td>$2.75$</td>
</tr>
<tr>
<td>$t^{Hodrick}(\hat{R}_{f})$</td>
<td>$-2.70$</td>
<td>$-2.10$</td>
<td>$-1.63$</td>
<td>$-1.69$</td>
<td>$-1.69$</td>
<td>$-1.51$</td>
<td>$1.23$</td>
</tr>
<tr>
<td>$R^2$</td>
<td>$0.048$</td>
<td>$0.134$</td>
<td>$0.203$</td>
<td>$0.292$</td>
<td>$0.353$</td>
<td>$0.396$</td>
<td>$0.407$</td>
</tr>
<tr>
<td>Implied $R^2_{Hodrick}$</td>
<td>$0.068$</td>
<td>$0.163$</td>
<td>$0.268$</td>
<td>$0.345$</td>
<td>$0.401$</td>
<td>$0.442$</td>
<td>$0.473$</td>
</tr>
<tr>
<td>Panel E: py-ratio + the interest rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{R}_{py}$</td>
<td>$-0.04$</td>
<td>$-0.18$</td>
<td>$-0.32$</td>
<td>$-0.46$</td>
<td>$-0.59$</td>
<td>$-0.73$</td>
<td>$-0.84$</td>
</tr>
<tr>
<td>$t^{OLS}(\hat{R}_{py})$</td>
<td>$-3.26$</td>
<td>$-6.16$</td>
<td>$-8.24$</td>
<td>$-10.73$</td>
<td>$-12.92$</td>
<td>$-14.79$</td>
<td>$-15.56$</td>
</tr>
<tr>
<td>$t^{NW}(\hat{R}_{py})$</td>
<td>$-3.26$</td>
<td>$-3.41$</td>
<td>$-3.56$</td>
<td>$-4.73$</td>
<td>$-5.61$</td>
<td>$-6.08$</td>
<td>$-5.99$</td>
</tr>
<tr>
<td>$t^{Hodrick}(\hat{R}_{py})$</td>
<td>$-3.32$</td>
<td>$-3.27$</td>
<td>$-2.92$</td>
<td>$-2.87$</td>
<td>$-2.91$</td>
<td>$-3.00$</td>
<td>$-2.96$</td>
</tr>
<tr>
<td>$\hat{R}_{f}$</td>
<td>$-0.73$</td>
<td>$-2.46$</td>
<td>$-3.96$</td>
<td>$-5.76$</td>
<td>$-7.49$</td>
<td>$-9.10$</td>
<td>$-9.99$</td>
</tr>
<tr>
<td>$t^{NW}(\hat{R}_{f})$</td>
<td>$-3.86$</td>
<td>$-3.65$</td>
<td>$-4.19$</td>
<td>$-6.36$</td>
<td>$-6.95$</td>
<td>$-5.49$</td>
<td>$-4.71$</td>
</tr>
<tr>
<td>$t^{Hodrick}(\hat{R}_{f})$</td>
<td>$-3.79$</td>
<td>$-3.21$</td>
<td>$-2.73$</td>
<td>$-2.93$</td>
<td>$-3.18$</td>
<td>$-3.17$</td>
<td>$3.00$</td>
</tr>
<tr>
<td>$R^2$</td>
<td>$0.063$</td>
<td>$0.173$</td>
<td>$0.257$</td>
<td>$0.38$</td>
<td>$0.477$</td>
<td>$0.541$</td>
<td>$0.562$</td>
</tr>
<tr>
<td>Implied $R^2_{Hodrick}$</td>
<td>$0.084$</td>
<td>$0.199$</td>
<td>$0.304$</td>
<td>$0.361$</td>
<td>$0.390$</td>
<td>$0.405$</td>
<td>$0.410$</td>
</tr>
</tbody>
</table>
Table 3: continued

<table>
<thead>
<tr>
<th>Horizon:</th>
<th>1 quart.</th>
<th>1 year</th>
<th>2 years</th>
<th>3 years</th>
<th>4 years</th>
<th>5 years</th>
<th>6 years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Part 3: Using cointegration residuals as predictors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel F: pdR-ratio

<table>
<thead>
<tr>
<th>$\hat{\kappa}_{pdR}$</th>
<th>$t^{OLS}(\hat{\kappa}_{pdR})$</th>
<th>$t^{NW}(\hat{\kappa}_{pdR})$</th>
<th>$t^{Hodrick}(\hat{\kappa}_{pdR})$</th>
<th>$R^2$</th>
<th>Implied $R^2_{Hodrick}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.09</td>
<td>-4.88</td>
<td>-4.91</td>
<td>-4.63</td>
<td>0.086</td>
<td>0.112</td>
</tr>
<tr>
<td>-0.33</td>
<td>-8.49</td>
<td>-4.22</td>
<td>-4.36</td>
<td>0.230</td>
<td>0.268</td>
</tr>
<tr>
<td>-0.47</td>
<td>-8.51</td>
<td>-3.52</td>
<td>-3.26</td>
<td>0.234</td>
<td>0.378</td>
</tr>
<tr>
<td>-0.57</td>
<td>-8.50</td>
<td>-3.48</td>
<td>-2.83</td>
<td>0.237</td>
<td>0.416</td>
</tr>
<tr>
<td>-0.66</td>
<td>-8.57</td>
<td>-3.30</td>
<td>-2.68</td>
<td>0.243</td>
<td>0.419</td>
</tr>
<tr>
<td>-0.84</td>
<td>-9.62</td>
<td>-4.05</td>
<td>-2.98</td>
<td>0.292</td>
<td>0.409</td>
</tr>
<tr>
<td>-0.88</td>
<td>-8.80</td>
<td>-4.01</td>
<td>-2.88</td>
<td>0.260</td>
<td>0.393</td>
</tr>
</tbody>
</table>

Panel G: cay-ratio

<table>
<thead>
<tr>
<th>$\hat{\kappa}_{cay}$</th>
<th>$t^{OLS}(\hat{\kappa}_{cay})$</th>
<th>$t^{NW}(\hat{\kappa}_{cay})$</th>
<th>$t^{Hodrick}(\hat{\kappa}_{cay})$</th>
<th>$R^2$</th>
<th>Implied $R^2_{Hodrick}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.44</td>
<td>4.28</td>
<td>4.29</td>
<td>4.13</td>
<td>0.071</td>
<td>0.111</td>
</tr>
<tr>
<td>5.10</td>
<td>7.01</td>
<td>4.08</td>
<td>3.78</td>
<td>0.184</td>
<td>0.226</td>
</tr>
<tr>
<td>8.98</td>
<td>8.69</td>
<td>4.74</td>
<td>3.50</td>
<td>0.297</td>
<td>0.266</td>
</tr>
<tr>
<td>12.26</td>
<td>11.44</td>
<td>4.76</td>
<td>3.25</td>
<td>0.387</td>
<td>0.256</td>
</tr>
<tr>
<td>13.92</td>
<td>11.92</td>
<td>5.19</td>
<td>3.04</td>
<td>0.411</td>
<td>0.237</td>
</tr>
<tr>
<td>15.45</td>
<td>11.65</td>
<td>6.29</td>
<td>2.92</td>
<td>0.405</td>
<td>0.219</td>
</tr>
<tr>
<td>16.80</td>
<td>11.68</td>
<td>6.75</td>
<td>2.83</td>
<td>0.411</td>
<td>0.204</td>
</tr>
</tbody>
</table>
Table 4: Regressions of long-horizon excess returns on the \(pd\)-ratio (Panel A), on the \(pd\)-ratio and the risk-free rate (Panel B), and on the \(PDR\)-ratio (Panel C). Quarterly 1952:1–2001:4 data.

<table>
<thead>
<tr>
<th>Horizon:</th>
<th>1 quart.</th>
<th>1 year</th>
<th>2 years</th>
<th>3 years</th>
<th>4 years</th>
<th>5 years</th>
<th>6 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\hat{\kappa}_{pd})</td>
<td>-0.02</td>
<td>-0.09</td>
<td>-0.13</td>
<td>-0.15</td>
<td>-0.17</td>
<td>-0.298</td>
<td>-0.49</td>
</tr>
<tr>
<td>(t^{Hodrick}(\hat{\kappa}_{pd}))</td>
<td>-1.43</td>
<td>-1.38</td>
<td>-0.87</td>
<td>-0.56</td>
<td>-0.40</td>
<td>-0.52</td>
<td>-0.68</td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.009</td>
<td>0.038</td>
<td>0.038</td>
<td>0.028</td>
<td>0.021</td>
<td>0.05</td>
<td>0.10</td>
</tr>
<tr>
<td>Implied (R^2)</td>
<td>0.052</td>
<td>0.088</td>
<td>0.158</td>
<td>0.224</td>
<td>0.283</td>
<td>0.337</td>
<td>0.385</td>
</tr>
<tr>
<td>Panel B: (pd)-ratio + the interest rate: 1952:1-2001:4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\hat{\kappa}_{Rf})</td>
<td>-0.03</td>
<td>-0.12</td>
<td>-0.17</td>
<td>-0.21</td>
<td>-0.24</td>
<td>-0.40</td>
<td>-0.59</td>
</tr>
<tr>
<td>(t^{Hodrick}(\hat{\kappa}_{Rf}))</td>
<td>-1.95</td>
<td>-1.73</td>
<td>-1.06</td>
<td>-0.74</td>
<td>-0.57</td>
<td>-0.66</td>
<td>-0.80</td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.044</td>
<td>0.103</td>
<td>0.081</td>
<td>0.087</td>
<td>0.086</td>
<td>0.108</td>
<td>0.148</td>
</tr>
<tr>
<td>Implied (R^2)</td>
<td>0.071</td>
<td>0.158</td>
<td>0.252</td>
<td>0.314</td>
<td>0.357</td>
<td>0.387</td>
<td>0.409</td>
</tr>
<tr>
<td>Panel C: (PDR)-ratio: 1952:1-2001:4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\hat{\kappa}_{PDR})</td>
<td>-0.08</td>
<td>-0.28</td>
<td>-0.41</td>
<td>-0.56</td>
<td>-0.71</td>
<td>-0.94</td>
<td>-1.11</td>
</tr>
<tr>
<td>(t^{Hodrick}(\hat{\kappa}_{PDR}))</td>
<td>-4.50</td>
<td>-3.66</td>
<td>-2.46</td>
<td>-2.25</td>
<td>-2.25</td>
<td>-2.49</td>
<td>-2.56</td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.085</td>
<td>0.212</td>
<td>0.227</td>
<td>0.292</td>
<td>0.359</td>
<td>0.474</td>
<td>0.538</td>
</tr>
<tr>
<td>Implied (R^2)</td>
<td>0.106</td>
<td>0.261</td>
<td>0.394</td>
<td>0.457</td>
<td>0.482</td>
<td>0.485</td>
<td>0.480</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$p_d$</th>
<th>$\bar{p}_y$</th>
<th>$\hat{p}_c$</th>
<th>$\ln(1 + R)$</th>
<th>$pdR$</th>
<th>$\bar{cay}$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Univariate regressions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Coefficient</td>
<td>$-0.06$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$4.84%$</td>
</tr>
<tr>
<td>$t$-stat</td>
<td>$-3.81$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Coefficient</td>
<td>$-0.02$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$0.40%$</td>
</tr>
<tr>
<td>$t$-stat</td>
<td>$-1.34$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Coefficient</td>
<td>$-0.02$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$0.48%$</td>
</tr>
<tr>
<td>$t$-stat</td>
<td>$-1.42$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Coefficient</td>
<td>$-0.44$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$2.63%$</td>
</tr>
<tr>
<td>$t$-stat</td>
<td>$-2.46$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel B: Multivariate regressions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Coefficient</td>
<td>$-0.06$</td>
<td>$-0.42$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$7.20%$</td>
</tr>
<tr>
<td>$t$-stat</td>
<td>$-3.69$</td>
<td>$-2.34$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Coefficient</td>
<td>$-0.05$</td>
<td>$-0.73$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$6.37%$</td>
</tr>
<tr>
<td>$t$-stat</td>
<td>$-3.41$</td>
<td>$-3.86$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Coefficient</td>
<td>$-0.04$</td>
<td>$-0.71$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$6.33%$</td>
</tr>
<tr>
<td>$t$-stat</td>
<td>$-3.39$</td>
<td>$-3.80$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel C: Univariate regressions using cointegration residuals</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Coefficient</td>
<td>$-0.09$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$8.70%$</td>
</tr>
<tr>
<td>$t$-stat</td>
<td>$-4.94$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Coefficient</td>
<td>$1.43$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$7.10%$</td>
</tr>
<tr>
<td>$t$-stat</td>
<td>$4.26$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 6: Out-of-sample $R^2$s using different starting points for the out-of-sample investigations.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Re-estimated $pdR$</td>
<td>2.84%</td>
<td>1.73%</td>
<td>0.17%</td>
<td>0.23%</td>
</tr>
<tr>
<td>Full-sample $pdR$</td>
<td>8.71%</td>
<td>5.31%</td>
<td>4.34%</td>
<td>1.88%</td>
</tr>
<tr>
<td>$cay$</td>
<td>5.30%</td>
<td>−3.28%</td>
<td>−2.82%</td>
<td>−1.92%</td>
</tr>
</tbody>
</table>
Table 7: Hansen-Jagannathan bounds.

<table>
<thead>
<tr>
<th>φ</th>
<th>γc = 1</th>
<th>γm = 1</th>
<th>γm = 3</th>
<th>γm = 5</th>
<th>γm = 7</th>
<th>γm = 10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.00</td>
<td>0.011</td>
<td>0.011</td>
<td>0.011</td>
<td>0.013</td>
<td>0.015</td>
<td>0.019</td>
</tr>
<tr>
<td>0.99</td>
<td>0.032</td>
<td>0.032</td>
<td>0.032</td>
<td>0.033</td>
<td>0.034</td>
<td>0.036</td>
</tr>
<tr>
<td></td>
<td>0.054</td>
<td>0.054</td>
<td>0.054</td>
<td>0.055</td>
<td>0.055</td>
<td>0.057</td>
</tr>
<tr>
<td></td>
<td>0.077</td>
<td>0.076</td>
<td>0.076</td>
<td>0.077</td>
<td>0.077</td>
<td>0.078</td>
</tr>
<tr>
<td></td>
<td>0.112</td>
<td>0.111</td>
<td>0.111</td>
<td>0.112</td>
<td>0.112</td>
<td>0.113</td>
</tr>
<tr>
<td>0.95</td>
<td>0.011</td>
<td>0.011</td>
<td>0.020</td>
<td>0.036</td>
<td>0.052</td>
<td>0.078</td>
</tr>
<tr>
<td></td>
<td>0.032</td>
<td>0.031</td>
<td>0.036</td>
<td>0.047</td>
<td>0.061</td>
<td>0.084</td>
</tr>
<tr>
<td></td>
<td>0.054</td>
<td>0.052</td>
<td>0.056</td>
<td>0.063</td>
<td>0.074</td>
<td>0.095</td>
</tr>
<tr>
<td></td>
<td>0.077</td>
<td>0.073</td>
<td>0.076</td>
<td>0.082</td>
<td>0.091</td>
<td>0.109</td>
</tr>
<tr>
<td></td>
<td>0.112</td>
<td>0.106</td>
<td>0.109</td>
<td>0.114</td>
<td>0.121</td>
<td>0.135</td>
</tr>
<tr>
<td>0.90</td>
<td>0.011</td>
<td>0.011</td>
<td>0.036</td>
<td>0.069</td>
<td>0.105</td>
<td>0.164</td>
</tr>
<tr>
<td></td>
<td>0.032</td>
<td>0.030</td>
<td>0.046</td>
<td>0.076</td>
<td>0.109</td>
<td>0.167</td>
</tr>
<tr>
<td></td>
<td>0.054</td>
<td>0.050</td>
<td>0.062</td>
<td>0.086</td>
<td>0.117</td>
<td>0.171</td>
</tr>
<tr>
<td></td>
<td>0.077</td>
<td>0.070</td>
<td>0.079</td>
<td>0.100</td>
<td>0.128</td>
<td>0.179</td>
</tr>
<tr>
<td></td>
<td>0.112</td>
<td>0.101</td>
<td>0.108</td>
<td>0.125</td>
<td>0.148</td>
<td>0.194</td>
</tr>
<tr>
<td>0.80</td>
<td>0.011</td>
<td>0.011</td>
<td>0.069</td>
<td>0.144</td>
<td>0.234</td>
<td>0.423</td>
</tr>
<tr>
<td></td>
<td>0.032</td>
<td>0.028</td>
<td>0.075</td>
<td>0.146</td>
<td>0.234</td>
<td>0.420</td>
</tr>
<tr>
<td></td>
<td>0.054</td>
<td>0.045</td>
<td>0.083</td>
<td>0.151</td>
<td>0.236</td>
<td>0.419</td>
</tr>
<tr>
<td></td>
<td>0.077</td>
<td>0.063</td>
<td>0.095</td>
<td>0.157</td>
<td>0.240</td>
<td>0.418</td>
</tr>
<tr>
<td></td>
<td>0.112</td>
<td>0.091</td>
<td>0.116</td>
<td>0.171</td>
<td>0.248</td>
<td>0.419</td>
</tr>
<tr>
<td>0.70</td>
<td>0.011</td>
<td>0.011</td>
<td>0.105</td>
<td>0.234</td>
<td>0.423</td>
<td>0.962</td>
</tr>
<tr>
<td></td>
<td>0.032</td>
<td>0.026</td>
<td>0.108</td>
<td>0.234</td>
<td>0.421</td>
<td>0.953</td>
</tr>
<tr>
<td></td>
<td>0.054</td>
<td>0.041</td>
<td>0.113</td>
<td>0.235</td>
<td>0.419</td>
<td>0.944</td>
</tr>
<tr>
<td></td>
<td>0.077</td>
<td>0.057</td>
<td>0.120</td>
<td>0.238</td>
<td>0.418</td>
<td>0.935</td>
</tr>
<tr>
<td></td>
<td>0.112</td>
<td>0.080</td>
<td>0.134</td>
<td>0.244</td>
<td>0.418</td>
<td>0.923</td>
</tr>
</tbody>
</table>
Notes for Tables

Notes to Table 1:
In panel A, the table shows summary statistics of the times series used in this paper. The $pd$-ratio is the share price to dividends ratio ($pd = \ln(P_t/D_t)$), the $pc$-ratio is the price-consumption ratio ($pd = \ln(P_t/C_t)$), the $py$-ratio is the price-output ratio ($pd = \ln(P_t/Y_t)$), $R$ is the nominal interest rate, and the $pdR$-ratio is an estimated ratio between share prices, dividends, the nominal interest rate, and the CPI level. The means and the standard deviations of the times series are shown. In panel B, the correlations between the series are shown.

Notes to Table 2:
The table shows estimates from univariate regressions $er_{t,t+1} = \text{constant} + \kappa_z z_{t-1} + \varepsilon_{t+1}$, where $z$ is either the $pd$-ratio, the $pc$-ratio, the $py$-ratio, or the interest rate in Panel A and where $er_{t,t+1}$ is the excess return from period $t$ to $+1$. In Panel B, results from bivariate regressions $er_{t,t+1} = \text{constant} + \kappa_z z_{t-1} + \kappa_r \ln(1 + R_{t-1}) + \varepsilon_{t+1}$ are shown. In Panel C, results from univariate regressions are shown, the predictors being cointegration residuals. $T$-statistics significant at a five percent level are in bold face type.

Notes to Table 3:
The table shows in Part 1 estimates from univariate regressions $er_{t,t+k} = \text{constant} + \kappa_z z_{t-1} + \varepsilon_{t+k}$, where $z$ is either the $pd$-ratio, the $py$-ratio, or the nominal interest rate and where $er_{t,t+k}$ is the sum of continuously compounded excess returns over the next $k$ quarters, with $k = 1, 4, 8, 12, 16, 20, \text{and } 24$. In Part 2, the results from bivariate regressions are shown and in Part 3 from univariate regressions using cointegration residuals as predictors. In the “$\hat{\kappa}_z$” rows, the table reports the estimates of $\hat{\kappa}_z$. Associated $t$-statistics based on three different kinds of standard errors are reported: Standard non-adjusted OLS $t$-statistics are reported in the $t^{OLS}(\hat{\kappa}_z)$-rows, $t$-statistics based on standard errors developed by Newey-West (1987), where the weights are truncated at lag $k + 1$, are shown in the $t^{NW}(\hat{\kappa}_z)$-rows, and $t$-statistics based on standard errors developed by Hodrick (1992) are shown in the $t^{Hodrick}(\hat{\kappa}_z)$-rows. Three measures of fit are reported: The standard $\tilde{R}^2$ and the Implied $R^2$s of Hodrick (1992) are also reported. $T$-statistics significant at a five percent level are in bold face type.
Notes to Table 4:
The table shows estimates from univariate regressions $e_{t,t+k} = \text{constant} + \kappa_{pd}d_{t-1} + \varepsilon_{t+k}$ in panel A, from bivariate regressions $e_{t,t+k} = \text{constant} + \kappa_{pd}d_{t-1} + \kappa_{R} \ln(1 + R_{t}) + \varepsilon_{t+k}$ in panel B, and from univariate regressions $e_{t,t+k} = \text{constant} + \kappa_{pdR}dR + \varepsilon_{t+k}$ in panel C. In the “$\hat{\kappa}_{x}$” rows, the table reports coefficient estimates. Associated $t$-statistics based on two different kinds of standard errors are reported: $t$-statistics based on standard errors developed by Newey-West (1987), where the weights are truncated at lag $k + 1$, are shown in the $t^{NW}(\hat{\kappa}_{x})$-rows, and $t$-statistics based on standard errors developed by Hodrick (1992) are shown in the $t^{Hodrick}(\hat{\kappa}_{x})$-rows. Two measures of fit are reported: The $R^2$ and the implied $R^2$s derived by Hodrick (1992). $T$-statistics significant at a five percent level are in bold face type.

Notes to Table 5:
See explanations to Table 2. The difference between Tables 2 and 5 is that Table 5 uses the deviations from the means of the predictor variables. Two means are used: the mean of the predictor variable over the 1947:1–1991:4 period and the mean over the 1992:1–2007:4 period. $T$-statistics significant at a five percent level are in bold face type.

Notes to Table 6:
The table shows Out-of-Sample $R^2$ using the estimated $\hat{cay}$ ratio of Lettau & Ludvigson (2001) or the estimated $\hat{pdR}$ ratio to generate the out-of-sample forecasts. Two versions of the $\hat{pdR}$ ratio are used: one where the parameters entering the $\hat{pdR}$ ratio are fixed at their full-sample values (“Full-sample $\hat{pdR}$) or one where the parameters are estimated in a recursive fashion, such that only information available at the time of the forecast are used to generate the forecasts. The table shows Out-of-Sample $R^2$'s based on different sample periods ending in 2007:4 but with different starting periods.

Notes to Table 7:
The table shows ratios of standard deviations of stochastic discount factors to means of stochastic discount factors for different values of the structural parameters entering the stochastic discount factor given in Eq. (13). The values should be compared with the Sharpe ratio of the return on the market portfolio (0.24) or the highest Sharpe ratio among the 25 Fama & French portfolios.
Figure 1. The $pdR$-ratio (black, left scale) and the $pd$-ratio (blue)

Figure 2. The $pdR$-ratio (black, left scale) and the $py$-ratio (blue)
Figure 3. Time series of demeaned predictive variables and their 1947-2001 and 2001-2008 means.
Figure 4. Recursive estimates of the coefficient to the $pd$-ratio and its $t$-statistic.

Figure 5. Recursive estimates of the coefficient to the $pdR$-ratio and its $t$-statistic.
Figure 6. Recursive estimates of the cointegration coefficients involved in constructing the $pdR$-ratio.