Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Financial Economics 81 (2006) 595–624
JOURNAL OF Financial ECONOMICS
www.elsevier.com/locate/jfec

Output and expected returns ☆

Jesper Rangvid*

Copenhagen Business School, Solbjerg Plads 3, DK-2000 Frederiksberg, Denmark

Received 14 June 2004; received in revised form 24 June 2005; accepted 9 July 2005
Available online 22 March 2006

Abstract

This paper shows for 1929–2003 U.S. data and also for international G-7 data that the ratio of share prices to GDP tracks a large fraction of the variation over time in expected returns on the aggregate stock market, capturing more of that variation than do price–earnings and price–dividend ratios and often also providing additional information about excess returns. The price–output ratio tracks long-term U.S. cumulative stock returns almost as well as the *cay*-ratio of Lettau and Ludvigson [2001a. Journal of Finance 56, 815–849, 2005. Journal of Financial Economics 76, 583–626], although the *cay*-ratio tracks variation in U.S. excess returns better. The price–output ratio, however, involves no parameter estimation and is easily constructed for non-U.S. countries.

JEL classification: F30; G15

Keywords: Share prices; GDP; Return predictability

1. Introduction

Do share prices scaled by fundamentals contain useful information about future movements in the aggregate stock market? Fama and French (1988) provide initial evidence that prices normalized by dividends or earnings can be used to capture time

☆ The author gratefully acknowledges comments from an anonymous referee as well as comments from Andrew Ang, Tom Engsted, Søren Johansen, Christian Lundblad, Richard Sweeney, and participants at the CEPR/NYSE ESSFM, the EEA meeting, the EFA meeting, the European FMA meeting, the Tor Vergata conference, and seminars and workshops at the Aarhus School of Business, Copenhagen Business School, Danmarks Nationalbank, University of Leuven, Lund University, and the Norwegian School of Management.

*Corresponding author. Tel.: +45 3815 3615; fax: +45 3815 3600.
E-mail address: jr.fi@cbs.dk.

0304-405X/$ - see front matter
doi:10.1016/j.jfineco.2005.07.010

variation in expected returns, while Campbell and Shiller (1988a,b) use the definition of returns to show why the price–dividend ratio should be expected to forecast long-run stock-market movements in particular. During the last decade of the 20th century, however, movements in aggregate stock prices and consequently returns were much different from what earnings and especially dividends would seem to have implied (see, e.g., Shiller, 2000, and others discussed later), raising doubts as to whether stock returns are at all predictable (Goyal and Welch, 2003, 2005).

New evidence has emerged, however, that certain macroeconomic variables contain information about future returns in addition to that revealed by financial valuation ratios. In particular, Lettau and Ludvigson (2001a) use the budget constraint of a representative consumer to show why fluctuations in the aggregate consumption–wealth ratio should forecast movements in the aggregate stock market. They also provide empirical evidence that the $\widehat{cay}$-ratio, an estimated consumption–wealth ratio, predicts U.S. excess returns well and captures a considerably larger fraction of the variation in expected returns than the price–dividend ratio and the dividend–earnings ratio. Julliard (2004), who elaborates upon the ideas of Lettau and Ludvigson (2001a), combines the $\widehat{cay}$-ratio with future labor income growth to predict stock returns. Santos and Veronesi (2006) develop a model where fluctuations in the consumption–labor income ratio imply that the risk premium required by investors to hold financial assets will vary over time, and they show that a ratio of labor income to consumption is a good predictor of U.S. returns. Earlier work by Cochrane (1991) relates the consumption-based asset-pricing models to production-based asset-pricing models, showing that an investment–capital ratio predicts U.S. returns; Cochrane (2005) contains a recent survey of the literature linking movements in the real economy with asset markets.

The impetus of this paper is the literature that scales share prices by fundamentals in return-predicting regressions as well as the newer findings on the importance of macroeconomic variables for capturing variation in expected returns. In particular, the paper will scale aggregate share prices with GDP (a macroeconomic variable) in return-predicting regressions, i.e., replace aggregate dividends or earnings with GDP as a price-normalizing variable. The theoretical inspiration for doing so is the large class of models linking movements in aggregate dividends to those of aggregate output in the economy (e.g., Lucas, 1978). When GDP replaces earnings or dividends as a price-normalizing variable, mean reversion in the price-to-GDP ratio implies that future returns are expected to be low if current stock prices are high in relation to the current GDP.

There are some a priori advantages of the price-to-GDP ratio over, for instance, the $\widehat{cay}$-ratio of Lettau and Ludvigson (2001a). GDP data are available for a longer time period than the wealth data necessary to construct the $\widehat{cay}$-ratio. GDP data are also readily available for many countries, whereas wealth data often are not, implying that international price-to-GDP ratios can easily be constructed. Furthermore, the $\widehat{cay}$-ratio and the $\widehat{lr}$-ratio of Julliard (2004) are estimated cointegration relations, whereas the creation of the price-to-GDP ratio does not involve any parameter estimation. Finally, by focusing on GDP, this paper shows that total output also carries information about expected returns, and, in this way, the paper contributes to the literature that use consumption-based variables to predict returns (e.g., Lettau and Ludvigson, 2001a; Julliard, 2004; Santos and Veronesi, 2006).

The most important result of this paper is that the ratio of share prices to GDP captures a considerable fraction of the variation especially in stock returns, but also in excess returns, both in-sample and out-of-sample. The relation between expected returns and the ratio of share prices to GDP is found to be economically and statistically significant when measured over a long period of time, using, for example, U.S. annual data for 1929–2003. The price-to-GDP ratio is also a convincing predictor variable using data from, for instance, the G-7 countries. Because the ratio of share prices to GDP tracks expected return variation over long periods of time, the results reported in this paper are general and not due to the extraordinary behavior of the stock market during the 1990s. It is also noteworthy that the price-to-GDP ratio captures a considerably larger fraction of the variation in future returns than the price–earnings ratio and the price–dividend ratio. Compared to the $\widehat{cay}$-ratio of Lettau and Ludvigson (2001a, 2005), stock returns are generally well predicted by both the price-to-GDP ratio and the $\widehat{cay}$-ratio, whereas the $\widehat{cay}$-ratio predicts excess returns better than the price-to-GDP ratio. It should also be mentioned that the price-to-GDP ratio generally captures a higher fraction of the variation in stock returns than the variation in excess returns. The reason for this is that the price-to-GDP ratio is also correlated with future interest rates. A negative shock to the price-to-GDP ratio (due, for instance, to a high level of GDP in the economy) predicts both higher stock returns and higher interest rates. These effects offset one another when subtracting the interest rate from stock returns to calculate excess returns.

How can the results be explained? Output, earnings, dividends, and share prices are generally expected to follow the same growth trend over long horizons. When empirically forecasting returns over business cycle horizons, however, this paper shows that the variable used to scale share prices matters. In other words, the decision to use either earnings, dividends, or output as a price-normalizing variable is important, even when all three variables eventually follow the same long-term growth trend.[1] Two recent papers specifically discuss when the price–dividend ratio does not capture the full amount of expected return variation. First, Menzly et al. (2004) show that time variation in expected dividend growth rates can offset the relation between dividend yields and expected returns that otherwise arises from time-varying risk aversion. Second, Lettau and Ludvigson (2005) provide empirical evidence that the price–dividend ratio reveals only part of the variation in returns because forecasts of returns and dividend growth rates are positively correlated, and such positive correlations have offsetting effects on the price–dividend ratio. To investigate whether these explanations are relevant for understanding the information about future returns contained in the price-to-GDP ratio, growth rates in earnings are regressed on a GDP-to-earnings ratio and growth rates in dividends are regressed on a GDP-to-dividends ratio. It turns out that the GDP-to-earnings ratio predicts future growth rates in earnings, and the GDP-to-dividends ratio predicts future growth rates in dividends. This is interesting because it lends support to the hypothesis that GDP fluctuations matter for stock market valuations in the sense that financial fundamentals, such as earnings and dividends, mean-revert towards GDP. It is also interesting because earnings and dividends do not mean-revert towards prices, as reported in this paper and in Cochrane (1992, 2001) and Lettau and Ludvigson (2005).

Following these introductory remarks, the rest of the paper is organized as follows. The next section lays out the theoretical motivation for why the price–output ratio should

[1]Similarly, Lamont (1998) shows that there is different information to be gained when scaling share prices by earnings or dividends.

predict returns. Section 3 discusses the data and presents results from cointegration analyses of relations between share prices and GDP, dividends, and earnings, respectively. The analysis of the long-term U.S. data is contained in Sections 4 and 5. Section 4 presents the results from regressions of stock returns and excess returns on the price–output ratio, the price–earnings ratio, the price–dividend ratio, and other control variables, as well as Monte Carlo experiments on long-horizon regressions. Results from out-of-sample exercises are reported in Section 5. After some interpretation of the results for the U.S. in Section 6, Section 7 briefly presents the results from an analysis of international G-7 data. In the final section, the paper is summarized and concluded.

2. The theoretical motivation

This section briefly explains the motivation for the tests performed in this paper. The "dynamic Gordon model," developed by Campbell and Shiller (1988b), is referred to in order to see how the present paper fits into that framework. Campbell and Shiller use the general definition of returns to show that the price–dividend ratio can be written as

$$p_t - d_t = \mathrm{E}_t \sum_{j=0}^{\infty} \rho^j (\Delta d_{t+1+j} - r_{t+1+j}) + \frac{k}{1-\rho}, \tag{1}$$

where p_t is the log of the period t price of the share, d_t is the log of the dividends that the share pays out, r_{t+1} is the log return, Δ is the difference operator, and $k = \ln(1 + \exp^{\overline{p-d}}) - \rho(\overline{p-d})$ with $\overline{p-d}$ as the mean log price–dividend ratio and $\rho = \exp^{\overline{p-d}}/(1+\exp^{\overline{p-d}}) < 1$. When applying (1) to the aggregate stock market, p_t measures the period t value of a share price index and d_t the period t value of the dividends paid out by the firms included in the index.

Eq. (1) has strong implications. Knowing that it is based on the definition of returns, a log-linear approximation, and the ruling out of bubbles (see Campbell and Shiller, 1988b), (1) shows how it is possible to trace the expectations of stock market participants by examining the variation in the price–dividend ratio. If stocks trade at a higher price for given dividends, (1) shows that this is the case because stock market participants expect future discount rates (the required returns on stocks) to be low if the growth in dividends is relatively constant.

As mentioned in the introduction, the implications of (1) have turned out to be less clear in the recent data (Campbell and Shiller, 2001; Lettau and Ludvigson, 2001a, 2005; Goyal and Welch, 2003; Ang and Bekaert, 2004). The question this paper analyzes is whether there is another fundamental factor that, in combination with stock prices, can be used to predict future movements in the aggregate stock market. Motivated especially by the evidence presented in Lettau and Ludvigson (2001a, 2005) that consumption (a macroeconomic variable) combined with financial wealth tracks an important fraction of the variation in returns, this paper investigates whether share prices scaled by GDP can be used to predict returns from the aggregate stock market.

In order to illustrate the main idea of this paper, it is assumed that the nonstationary behavior of dividends comes from output in the economy, $d_t = y_t + v_t$, with y_t as output and v_t as a mean zero stationary disturbance term. If the nonstationary part of dividends arises from output, (1) can be written as

$$p_t - y_t = \mathrm{E}_t \sum_{j=0}^{\infty} \rho^j (\Delta y_{t+1+j} - r_{t+1+j}) + \frac{k}{1-\rho} + v_t. \tag{2}$$

Eq. (2) conveys an essential idea in this paper. Variation over time in the price–output ratio, the left-hand side of (2), captures variation over time in expected returns if output is not expected to be too volatile. The interpretation of Eq. (2) matches up with the interpretation of Eq. (1). When prices are high for a given level of output, investors are willing to pay much for the stocks either because they expect the economy to perform well in terms of how much is produced or because they expect future required rates of return to be low. It should also be noted that, from (2), variations in the price–output ratio reflect changes in expectations about returns over the next many periods. In other words, there is reason to believe that the price–output ratio especially captures long-horizon returns.

3. Data

The analysis uses annual data covering 1929–2003. Current-value (nominal) GDP data were downloaded from the homepage of the U.S. Bureau of Economic Analysis.[2] The U.S. financial data were downloaded from Robert Shiller's homepage and consist of, for a given year, the January value of the Standard and Poor Composite Stock Price Index, the associated dividends and earnings series, and the short interest rate, which is "the total return to investing for six months in January at the January 4–6 month prime commercial paper rate."[3]

In the following, p denotes the log of the S&P stock price index, y denotes the log of GDP, d denotes the log of dividends, and e denotes the log of earnings. Using this notation, the price–output ratio series is calculated as $py_t = p_t - y_{t-1}$, the price–dividend ratio as $pd_t = p_t - d_{t-1}$, and the price–earnings ratio as $pe_t = p_t - e_{t-1}$. The share prices of January in a given year are thus scaled by GDP (or earnings or dividends) from the previous year, as is the convention in the literature.

The three ratios are shown in Fig. 1. The general pattern shows that all ratios increase during the 1950s, fall during the 1970s, increase again during the 1980s, and reach high levels during the 1990s. There are, however, some important differences to point out. For example, the py-ratio fell more during the first part of the sample than did the two financial ratios (the pe-ratio and the pd-ratio are fairly volatile during the 1930s and 1940s), and, perhaps particularly noteworthy, the py-ratio did not increase as much during the boom period of the 1990s as the pd- and pe-ratios did.

This paper examines both stock returns and excess returns. The continuously compounded annual stock return is denoted by $r_t = \ln[(P_t + D_{t-1})/P_{t-1}]$, where P_t is the January value of the share price index in a particular year, P_{t-1} is the value of the share price index in January last year, and D_{t-1} is the dividends paid out throughout the previous year. The log excess return is calculated as $er_t = r_t - i_{t-1}$, where i_{t-1} is the interest rate of January the previous year. Both the nominal stock return, r_t, and the nominal interest rate, i_t, include inflation. The excess return is thus approximately real. The advantage of

[2]Available at http://www.bea.gov/bea/dn/gdplev.xls.

[3]Available at http://www.econ.yale.edu/~shiller/data.htm. Information on the construction of the series is available from Shiller's homepage.

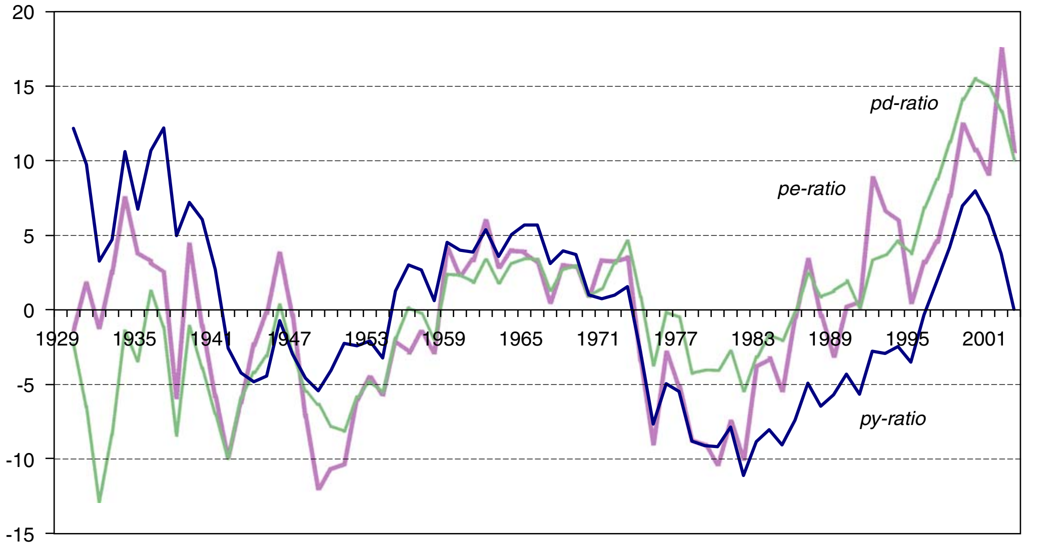

Fig. 1. The figure shows the de-meaned price–output ratio (*py*), the de-meaned price–earnings ratio (*pe*), and the de-meaned price–dividend ratio (*pd*) during the sample period 1929–2003. All series are normalized to one in 1970.

presenting results for both nominal stock returns and approximately real excess returns is that nominal stock returns capture the actual movements in the stock market, whereas excess returns capture the return that the stock market provides in excess of the "risk-free" security.

Table 1, Panel A, provides the means and standard deviations of the series. The average annual return from equity is approximately 11.5% per annum with a standard deviation of 19.57%. The average annual risk premium is 6.78%. Furthermore, the ratios are not as volatile as the returns are. Tests (not shown) reveal that returns, excess returns, the *py*-ratio, and the *pe*-ratio are normally distributed, which is not the case for the price–dividend ratio.

The correlations between the ratios (*py*, *pe*, and *pd*) are shown in Panel B, Table 1. The correlation between the price–earnings ratio and the price–dividend ratio is 0.80 and the correlation between the *pe*-ratio and the *py*-ratio 0.59. The *pd*-ratio and the *py*-ratio are the least correlated. Finally, the correlation between returns and excess returns is very high as a consequence of the low volatility of the short interest rate (not shown). The correlations between the financial ratios and annual returns are fairly low.

3.1. Cointegration

Campbell and Shiller (1988b) and Lettau and Ludvigson (2001a) explain why an important statistical implication of a present-value relation, such as the one in Eq. (2), is that the left-hand side of (2) should be covariance-stationary when returns and changes in output are covariance-stationary. In other words, the output series (y_t) should cointegrate with the price series (p_t) such that the series $p_t - y_t$ is stationary. Similarly, according to Eq. (1), share prices should be cointegrated with dividends. In order to test these and other

Table 1
Summary statistics
U.S. data: 1929–2003

	py	*pe*	*pd*	*r*	*er*
Panel A: Means and standard deviations					
Mean	−2.46	2.67	3.27	11.48	6.78
Std.	0.43	0.41	0.44	19.57	19.80
Panel B: Correlations					
py	1.00				
pe	0.59	1.00			
pd	0.30	0.80	1.00		
r	−0.02	0.12	0.20	1.00	
er	0.11	0.16	0.26	0.97	1.00
Panel C: Univariate unit root and cointegration tests					
ADF	−2.56[a]	−3.16**	−2.42	−7.64**	−8.05**
PP	−2.54[a]	−3.20**	−2.24	−7.74**	−8.16**
ρ_1	0.87	0.77	0.86	0.07	0.06
80% interval for ρ_1	(0.76, 0.99)	(0.65, 0.89)	(0.79, 1.01)	–	–
95% interval for ρ_1	(0.71, 1.03)	(0.59, 0.98)	(0.74, 1.03)	–	–
Panel D: Horvath and Watson cointegration tests					
VAR(1)	49.21**			–	–
VAR(4)	50.30**			–	–

Notes: Row one in Panel A shows the sample means and row two shows the standard deviations for the price–output ratio (*py*), the price–earnings ratio (*pe*), the price–dividend ratio (*pd*), the annual return to equity (*r*), and the excess return (*er*). Panel B shows the correlations between the series. Panel C shows the results from Augmented Dickey–Fuller (ADF) and Phillips–Perron (PP) tests for a unit root in the first two rows. The null hypothesis in the ADF tests and the PP tests is that the series is non-stationary, and the critical values to test these hypotheses are 2.58 at the 10% level and 3.51 at the 1% level. Row three of Panel C shows the largest autoregressive root ρ_1 of the series, and rows four and five contain the 80% and 95% confidence intervals for ρ_1 based on Stock's (1991) procedure. Panel D reports the results from the Horvath and Watson (1995) tests for no cointegration between the time series of *p*, *e*, *d*, and *y* against the alternative of three stationary cointegration ratios *py*, *pe*, and *pd*. The result in the *VAR*(1) row is based on a *VAR* with one lag and the result in the *VAR*(4) row is based on a *VAR* with four lags. The critical values for the Horvath and Watson (1995) tests are 41.08 at the 1% level and 32.33 at the 10% level. Significant statistics at the 1% level are indicated by ** and at the 10% level by *. "a" is used to indicate that the ADF-statistic or PP-statistic is smaller than −2.50 and thus close to the 10% critical value.

stationarity hypotheses, Panels C and D of Table 1 show the results from different unit root tests and cointegration tests.

The first two rows of Panel C provide the results from standard augmented Dickey–Fuller (ADF) tests and Phillips–Perron tests. The null hypotheses of nonstationary series are rejected for the two return series, the *pe*-ratio, and the *py*-ratio (the null hypothesis that the *py*-ratio is a nonstationary time series is rejected at a level of significance just above 10%). On the other hand, the null hypothesis of a nonstationary price–dividend ratio cannot be rejected.

The third row of Panel C shows the largest autoregressive roots of the series and rows 4 and 5 show the 80% and 95% confidence intervals for the largest autoregressive root of each series. Stock's (1991) procedure is used to calculate the confidence intervals for the largest autoregressive root. The largest autoregressive roots for the valuation ratios are all larger than 0.75, i.e., the ratios are rather persistent, whereas the return series are characterized by very low autoregressive roots, i.e., they are clearly stationary.[4] The confidence intervals for the largest autoregressive root indicate that the *pe*-ratio is stationary and that the *py*-ratio is marginally stationary whereas the *pd*-series contains a unit root in the confidence interval.

An alternative to the univariate tests is to evaluate whether all four variables (p, y, e, and d) are driven by one single stochastic trend in such a way that any two variables are cointegrated with unitary coefficients. Horvath and Watson (1995) propose such a test.[5] This test is a joint test of the hypothesis that known one-to-one ratios of prices to the normalizing series — output, dividends, or earnings — are stationary. The Horvath and Watson (1995) test evaluates whether the three ratios (*py*, *pe*, and *pd*) can be excluded from the right-hand side of a four-dimensional vector autoregressive system of Δp, Δe, Δd, and Δy. Given that the full empirical model is used (all four variables and equations are simultaneously investigated, in contrast to the univariate tests that only consider bivariate pairs of variables) and that prices, dividends, earnings, and output should be expected to eventually follow the same stochastic growth trend, this test is a more efficient way of testing for cointegration. The results are shown in Panel D of Table 1 for VARs with one lag and four lags. Interestingly, these tests provide a strong case: the null hypothesis of no cointegration is rejected in favor of the alternative of three known cointegration vectors.

Hence, the overall picture is as follows: The univariate tests indicate that the *py*-ratio and the *pe*-ratio are stationary whereas the *pd*-ratio is not. Furthermore, the more efficient multivariate Horvath and Watson tests show that prices, dividends, earnings, and output all share the same stochastic growth trend, such that ratios of any two of these four variables are stationary.

4. Predicting U.S. annual returns in-sample

This paper aims chiefly at answering two overall questions: (i) does the price–output ratio contain information about expected returns and (ii) does the price–output ratio contain information about expected returns beyond that of especially price–earnings and price–dividend ratios?

The regressions that are run to answer these questions are either univariate regressions,

$$x_{t,t+K} = \kappa + \alpha z_t + \varepsilon_t, \tag{3}$$

where κ is a constant and z_t indicates one of the predictor candidates, or multivariate regressions,

$$x_{t,t+K} = Z_t' \Psi + \varepsilon_t, \tag{4}$$

[4] Stock (1991) tabulates the relevant statistics needed to calculate the confidence bounds for ADF-test statistics larger than −5.70 (for ADF-test statistics smaller than −5.70, it is safe to conclude that the series are stationary).

[5] Horvath and Watson (1995) argue that their test for known cointegration vectors is "significantly more powerful than the test that does not impose the cointegration vector," such as Johansen (1991) tests, and find that the power gains in the bivariate example they analyze "correspond to sample size increases ranging from 40 to 70% for a test with power equal to 50%." Lamont (1998) argues for the use of the more efficient Horvath and Watson (1995) tests.

where Z_t is the vector of relevant predictor variables and Ψ contains the estimated parameters. Both stock returns and excess returns are analyzed, i.e., in (3) and (4), $x_{t,t+K}$ is the sum of either continuously compounded stock returns or excess returns over the next K years, where $K = 1, 2, 3, 4, 5$, or 6.

4.1. Results

Table 2 reports the first results. The top panel shows results from univariate regressions of stock returns on the *py*-ratio, the *pe*-ratio, and the *pd*-ratio respectively; the center panel shows results from univariate regressions of excess returns on the ratios; and the bottom panel shows results from multivariate regressions of excess returns on the ratios. The table

Table 2
Full sample regressions of long-horizon cumulative returns

Horizon *K*:		1	2	3	4	5	6
Panel A: Stock return univariate regressions							
py	coef.	−0.17**	−0.32**	−0.37**	−0.44**	−0.51**	−0.56**
	t-stat.	(3.23)	(3.26)	(3.87)	(5.64)	(6.81)	(7.42)
	$\overline{R}^2$	**0.15**	**0.25**	**0.35**	**0.35**	**0.42**	**0.49**
pe		−0.11*	−0.17*	−0.17	−0.21	−0.27	−0.30
		(2.38)	(1.99)	(1.16)	(1.21)	(1.45)	(1.27)
		0.04	**0.04**	**0.02**	**0.03**	**0.06**	**0.07**
pd		−0.05	−0.11	−0.09	−0.12	−0.19	−0.17
		(0.77)	(0.95)	(0.54)	(0.64)	(0.94)	(0.67)
		0.00	**0.01**	**−0.00**	**0.00**	**0.02**	**0.01**
Panel B: Excess return univariate regressions							
py		−0.14*	−0.25*	−0.27*	−0.29*	−0.32*	−0.33*
		(2.26)	(2.18)	(2.14)	(2.38)	(2.43)	(2.21)
		0.08	**0.13**	**0.12**	**0.12**	**0.12**	**0.12**
pe		−0.09*	−0.15	−0.14	−0.16	−0.21	−0.23
		(1.89)	(1.44)	(0.79)	(0.79)	(0.91)	(0.82)
		0.02	**0.02**	**0.01**	**0.01**	**0.02**	**0.02**
pd		−0.06	−0.15	−0.16	−0.22	−0.34	−0.38
		(0.98)	(1.23)	(0.89)	(1.05)	(1.44)	(1.25)
		0.01	**0.03**	**0.02**	**0.03**	**0.08**	**0.08**
Panel C: Excess return multivariate regressions							
py		−0.13*	−0.31*	−0.33*	−0.37*	−0.48**	−0.49**
		(2.02)	(2.14)	(2.13)	(2.54)	(3.00)	(3.12)
pe		0.03	0.25	0.29	0.37	0.56*	0.57*
		(0.30)	(1.30)	(1.18)	(1.44)	(2.22)	(1.94)
pd		−0.06	−0.38*	−0.38*	−0.54**	−0.77**	−0.88**
		(0.58)	(1.69)	(2.05)	(2.71)	(4.03)	(3.72)
		0.06	**0.16**	**0.18**	**0.24**	**0.32**	**0.34**

Notes: The table shows parameter estimates with *t*-statistics based on Newey and West's (1987) standard errors, in parentheses below (the weights are truncated at lag $K+1$), and $\overline{R}^2$s, indicated in bold. Panels A and B show results from univariate regressions of annual and cumulative stock returns (in Panel A) and excess returns (in Panel B) on the price–output ratio (*py*), the price–earnings ratio (*pe*), and the price–dividend ratio (*pd*). Panel C shows results from multivariate regressions of excess returns on all three ratios. Significant statistics at the 1% level are indicated by ** and at the 10% level by *.

presents the parameter estimates, with t-statistics in parentheses below the parameter estimates, and the $\overline{R}^2$s indicated in bold. The t-statistics are calculated using Newey–West's (1987) autocorrelation and heteroskedasticity-consistent standard errors with the weights truncated at $K+1$.[6]

Consider the regressions using annual returns first ($K=1$). The price–output ratio captures around 15% of the variation in annual U.S. stock returns and 8% of the variation in excess returns. Neither the price–earnings ratio nor the price–dividend ratio capture nearly as much of the variation in stock returns and excess returns. In the multivariate regressions of annual ($K=1$) excess returns on the ratios, only the *py*-ratio is significant.

The variation in the *py*-ratio is both statistically and economically significant. In the annual stock return regression, the coefficient on the *py*-ratio is −0.17. Table 1 reveals that the standard deviation of the *py*-ratio is 0.43, which implies that a one standard deviation increase in the *py*-ratio corresponds to a 731 basis point (around 7.3%) change in expected annualized returns. This can be compared to the roughly 9% increase in the S&P 500 index that a one standard deviation increase in the $\widehat{cay}$-ratio gives rise to (Lettau and Ludvigson, 2001a).

The signs of the coefficients to the ratios are right. From the theoretical model in (2), positive deviations from $p_t - y_t = py_t$ (and from pe_t and pd_t) should lead to decreasing returns, and indeed the coefficients on the ratios are all negative.

4.1.1. Long-horizon returns

When considering the longer-horizon regressions ($K = 2, 3, \ldots 6$), there are three major conclusions. First, the *py*-ratio tracks long-horizon stock returns to an interestingly high extent and is very significant. The *py*-ratio captures as much as 40–50% of the variation in five- and six-year cumulative stock returns and the t-statistic in the regression of six-year cumulative stock returns on the *py*-ratio is more than seven.

Second, neither the *pe*-ratio nor the *pd*-ratio captures nearly as much of the variation in long-horizon stock returns as does the price–output ratio. The *pe*-ratio and the *pd*-ratio capture 7% and 1%, respectively, of the variation in six-year cumulative stock returns. Furthermore, both the *pd*-ratio and the *pe*-ratio are insignificantly related to long-horizon stock returns.

Third, the $\overline{R}^2$s from regressions of long-horizon excess returns on the *py*-ratio are not as high as when regressing stock returns on the *py*-ratio. In other words, the *py*-ratio captures movements in stock returns better than it captures movements in excess returns, as will also be made more clear in the following sections.

4.2. Statistical issues with long-horizon regressions

The results presented in Table 2 provide evidence that the *py*-ratio captures more of the variation especially in stock returns than do the *pe*-ratio and the *pd*-ratio. Subtle statistical issues arise when using long-horizon regressions, however. Long-horizon returns ($K = 2, 3, 4, 5$, and 6) are calculated by summing the (continuously compounded) annual returns. This summation implies that the observations on long-horizon returns overlap, which possibly biases the different test statistics towards rejecting the null hypothesis of no

[6]In order not to lose sight of the big picture, the t-statistics in this section are adjusted only by the Newey–West factor. In the following section, several kinds of simulations are performed to evaluate the robustness of the long-horizon regressions.

Table 3
$t/\sqrt{T}$-statistics and Implied R^2 s from full sample regressions of long-horizon cumulative returns on the *py*-, the *pe*-, and the *pd*-ratio

	Horizon:	1	2	3	4	5	6
Panel A: Stock returns							
py	$t/\sqrt{T}$	−0.42**	−0.57*	−0.63*	−0.71**	−0.81**	−0.94**
	Implied R^2	0.14**	0.24**	0.33**	0.40**	0.46**	0.51**
	$(R^2_{0.90}, R^2_{0.99})$	(0.07, 0.14)	(0.09, 0.19)	(0.11, 0.23)	(0.13, 0.25)	(0.14, 0.26)	(0.14, 0.27)
pe	$t/\sqrt{T}$	−0.23	−0.24*	−0.19	−0.22	−0.26*	−0.28*
	Implied R^2	0.07*	0.11*	0.13*	0.14**	0.15*	0.16*
	$(R^2_{0.90}, R^2_{0.99})$	(0.07, 0.13)	(0.07, 0.16)	(0.08, 0.18)	(0.08, 0.19)	(0.08, 0.19)	(0.08, 0.19)
pd	$t/\sqrt{T}$	−0.12	−0.16	−0.11	−0.12	−0.18	−0.16
	Implied R^2	0.03	0.04	0.06	0.07	0.08	0.08
	$(R^2_{0.90}, R^2_{0.99})$	(0.07, 0.13)	(0.08, 0.17)	(0.10, 0.21)	(0.11, 0.23)	(0.12, 0.24)	(0.12, 0.25)
Panel B: Excess returns							
py	$t/\sqrt{T}$	−0.30*	−0.39*	−0.38*	−0.37*	−0.37*	−0.37*
	Implied R^2	0.09*	0.15*	0.20*	0.25**	0.29**	0.32**
	$(R^2_{0.90}, R^2_{0.99})$	(0.07, 0.13)	(0.09, 0.18)	(0.11, 0.22)	(0.13, 0.24)	(0.14, 0.26)	(0.14, 0.27)
pe	$t/\sqrt{T}$	−0.19	−0.19	−0.14	−0.15	−0.18	−0.19
	Implied R^2	0.12*	0.18**	0.23**	0.27**	0.29**	0.31**
	$(R^2_{0.90}, R^2_{0.99})$	(0.07, 0.13)	(0.08, 0.16)	(0.09, 0.19)	(0.10, 0.21)	(0.10, 0.22)	(0.10, 0.22)
pd	$t/\sqrt{T}$	−0.14	−0.21	−0.18	−0.21	−0.30*	−0.31*
	Implied R^2	0.12*	0.18*	0.24*	0.29*	0.34**	0.39**
	$(R^2_{0.90}, R^2_{0.99})$	(0.08, 0.15)	(0.10, 0.22)	(0.13, 0.27)	(0.15, 0.30)	(0.17, 0.32)	(0.18, 0.34)

Notes: The table reports results from regressions of cumulative long-horizon stock returns on the price–output ratio (*py*), the price–earnings ratio (*pe*), and the price–dividend ratio (*pd*) in Panel A, and from excess return regressions in Panel B. The "$t/\sqrt{T}$" rows contain Valkanov's (2003) rescaled *t*-statistics. The "Implied R^2" rows show implied R^2s from $VAR(1)$ models with no overlapping observations, while the "$(R^2_{0.90}, R^2_{0.99})$" rows contain the critical values. The critical values for the implied R^2s are found from the empirical distributions of the statistics based upon simulations of $VAR(1)$ models under the null of no predictability. Significant statistics at the 1% level are indicated by ** and at the 10% level by *.

predictability more often than is correct. In order to guard against these potential biases, two different tests are implemented.

The first test is based on Valkanov (2003), who discusses the distribution of a rescaled *t*-statistic, $t/\sqrt{T}$, with t as the traditional *t*-statistic and T as the number of observations. Unlike the standard *t*-statistic, Valkanov (2003) shows that the rescaled *t*-statistic has a well-defined distribution that can be simulated given the value of a nuisance parameter c and the correlation δ between the innovations to returns and the forecasting variable.[7] The critical values of Valkanov (2003, Table 3), case 1 with $c = -10$ and $\delta = -0.90$, are used to evaluate the significance of the $t/\sqrt{T}$ statistics (the choices of c and δ are based on results reported in Rangvid, 2005). The second test is based on work done by Hodrick (1992), who

[7] c can be found using the procedure of Stock (1991). The Stock procedure, and thus the value of c, was also used to calculate the confidence intervals for the largest autoregressive root reported in Table 1.

shows how to calculate the implied R^2s for long-horizon regressions based on the estimation of the parameters and the variance-covariance matrix of a $VAR(1)$ model, i.e., a model that includes only non-overlapping annual returns. In order to evaluate the significance of the implied R^2s, empirical critical values based on 5,000 simulations of the $VAR(1)$ model under the null hypothesis of no predictability are used. In each simulation, the implied R^2s are calculated. From the resulting 5,000 observations (for each forecast horizon) of the implied R^2s, the relevant critical values for testing the null hypotheses of no predictability are found. Details on the simulations used to find the empirical critical values for the implied R^2s can be found in Rangvid (2005) which also presents empirical critical values based on simulations of Valkanov's $t/\sqrt{T}$ statistics.

4.2.1. Results

In Table 3, the "$t/\sqrt{T}$" rows show the rescaled t-statistics, the "Implied R^2" rows show the implied R^2s from $VAR(1)$ models, and the corresponding simulated critical values at the 90% and 99% levels are shown in rows ($R^2_{0.90}, R^2_{0.99}$).

First, note that the *py*-ratio is a strongly significant predictor for long-horizon stock returns and a marginal predictor of excess returns, which was also true in the tests, reported in Table 2, that did not adjust for biases in the long-horizon statistics. For instance, the implied R^2s are all outside the confidence bounds calculated from the empirical distributions of the implied R^2s found from $VAR(1)$ models simulated under the null of no predictability. The fractions of return variation captured by the *py*-ratio are thus statistically significant and should not be due to biases in the test statistics caused by overlapping observations. This last result is confirmed by the fact that the rescaled t-statistics are all highly significant when analyzing the predictability of stock returns, and marginally significant when analyzing excess returns. Furthermore, the finding that the implied R^2s rise "faster" with the horizon than do the corresponding values simulated under the null of no predictability implies that the long-horizon regressions provide additional evidence on return predictability.

When evaluating the $t/\sqrt{T}$ statistics, neither the *pe*-ratio nor the *pd*-ratio are systematically related to stock returns, which is also true for their relation to future excess returns. This is in accordance with the results shown in Table 2.

4.3. Subsample results and comparisons with the $\widehat{cay}$-ratio

Sections 4.1 and 4.2 have verified that the *py*-ratio contains information about expected returns for the full sample period 1929–2003. In this section, three subsamples are investigated: 1948–2003, 1929–1966, and 1967–2003. The 1948–2003 sample is the sample Lettau and Ludvigson (2001a) use to verify the predictive power of $\widehat{cay}$, i.e., use of this sample period permits a comparison between results from regressions using the *py*-ratio and regressions using $\widehat{cay}$ when data on $\widehat{cay}$ are available. $\widehat{cay}$ is an estimated cointegration relation linking consumption and labor income (macroeconomic variables) with asset wealth (a financial variable). Lettau and Ludvigson (2001a) show that $\widehat{cay}$ captures future movements in excess returns significantly better than standard predictor variables do.[8] The 1929–1966 and 1967–2003 subsamples are based on splitting the full sample into two equally sized subsamples.

[8]The annual $\widehat{cay}$ data (used in Lettau and Ludvigson, 2005) are downloadable from Martin Lettau's homepage: http://pages.stern.nyu.edu/~mlettau/data/cay_a_01.txt. The $\widehat{cay}$ data only date back to 1948 as the wealth data used to construct $\widehat{cay}$ are not available before then.

Table 4
Regressions of long-horizon cumulative returns and excess returns on *py*, *pe*, *pd*, and $\widehat{cay}$. 1948–2003

Horizon K:		1	2	4	6
Stock returns					
py	coef.	-0.12^{**}	-0.24^{**}	-0.37^{**}	-0.57^{**}
	t-stat.	(3.03)	(3.52)	(4.38)	(4.28)
	$\bar{R}^2$	**0.08**	**0.17**	**0.26**	**0.39**
pe		-0.13^{**}	-0.24^{**}	-0.34^{**}	-0.42^{*}
		(2.88)	(3.30)	(2.90)	(1.97)
		0.13	**0.19**	**0.22**	**0.21**
pd		-0.13^{*}	-0.24^{*}	-0.36^{*}	-0.54^{*}
		(2.12)	(2.35)	(2.54)	(2.26)
		0.12	**0.18**	**0.20**	**0.26**
$\widehat{cay}$		5.43^{**}	10.15^{**}	12.92^{**}	17.55^{**}
		(4.72)	(7.06)	(4.73)	(5.72)
		0.26	**0.46**	**0.41**	**0.46**
Excess returns					
py		-0.09^{*}	-0.17^{*}	-0.26^{*}	-0.43^{*}
		(1.81)	(1.89)	(1.66)	(1.68)
		0.02	**0.06**	**0.07**	**0.09**
pe		-0.11^{*}	-0.21^{*}	-0.29	-0.39
		(2.26)	(2.15)	(1.53)	(1.30)
		0.07	**0.11**	**0.10**	**0.10**
pd		-0.13^{*}	-0.25^{*}	-0.39^{*}	-0.69^{*}
		(2.06)	(2.19)	(2.06)	(2.23)
		0.09	**0.15**	**0.17**	**0.25**
$\widehat{cay}$		6.32^{**}	11.93^{**}	15.84^{**}	22.20^{**}
		(4.79)	(7.39)	(6.56)	(6.64)
		0.29	**0.50**	**0.43**	**0.48**

Notes: The table shows parameter estimates with *t*-statistics based on Newey and West's (1987) standard errors, in parentheses below (the weights are truncated a lag $K+1$), and $\bar{R}^2$s, indicated in bold, from regressions of stock returns (in the upper part of the table) and excess returns (in the lower part of the table) on the price–output ratio (*py*), the price–earnings ratio (*pe*), the price–dividend ratio (*pd*), and Lettau and Ludvigson's (2001a) $\widehat{cay}$-ratio. $\widehat{cay}$ is an estimated cointegration relation linking consumption and labor income with asset wealth. Significant statistics at the 1% level are indicated by ** and at the 10% level by *.

Before presenting the results, it is instructive to visually compare the movements in $\widehat{cay}$ with those of the *py*-ratio. This is done in Fig. 2. The correlation between $\widehat{cay}$ and the *py*-ratio is not that high. The reasons are the following: GDP is highly correlated with labor income and consumption (the correlation coefficient in each case is 0.99), but the correlation between asset wealth and stock prices is "only" 0.78. Furthermore, the relation between c, a, and labor income in $\widehat{cay}$ depends upon two estimated cointegration parameters, whereas there is no estimation involved in the construction of the *py*-ratio. All of this implies that the correlation between *py* and $\widehat{cay}$ is "only" -0.44. Differences between the results from the return-predicting regressions with *py* and $\widehat{cay}$ as explanatory variables are thus not unexpected.

Consider first the period 1948–2003. Table 4 shows that the *t*-statistics are high in the regressions of stock returns on the *py*-ratio but only marginally significant in the excess

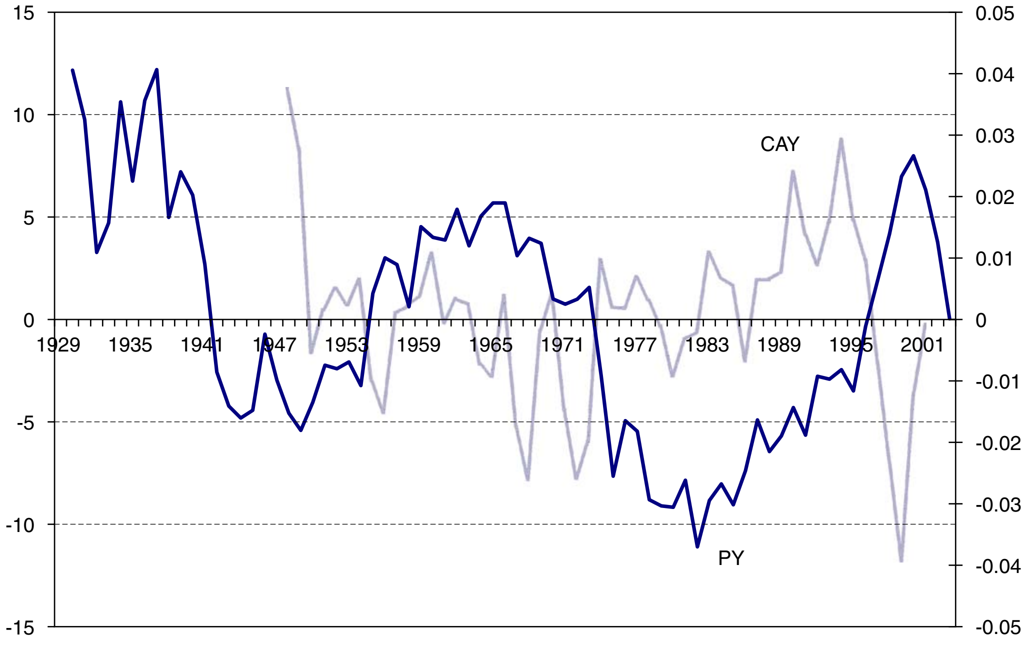

Fig. 2. The figure shows the price–output ratio (*py*-ratio) and Lettau and Ludvigson's (2001a) $\widehat{cay}$-ratio. $\widehat{cay}$ is an estimated cointegration relation linking consumption and labor income with asset wealth.

return regressions (to save space, Table 4 and the following tables show standard Newey–West based *t*-statistics and results for $K = 1, 2, 4$, and 6 only). Both the *pd*-ratio and the *pe*-ratio predict stock returns. During the 1948–2003 period, the *pd*-ratio captures a higher fraction of the variation in excess returns than does the *py*-ratio. Overall, $\widehat{cay}$ does an impressive job in predicting both stock returns and excess returns with *t*-statistics all exceeding four and $\overline{R}^2$s all exceeding 0.25.

Next, consider results based on the 1929–1966 and 1967–2003 periods. The results appear in Table 5 using excess returns as the dependent variable. The ability of the *py*-ratio to predict excess returns in the early subsample is comparable to, or even better than, the ability of the *py*-ratio to predict excess returns for the full sample. The $\overline{R}^2$s rise to as much as 70%. During the early subsample, the price–dividend ratio and the price–earnings ratio are not significant. During the 1967–2003 sample, the *py*-ratio is significant, whereas the *pd*-ratio and the *pe*-ratio are not. In other words, during both the 1929–1966 and the 1967–2003 subperiods, the *py*-ratio is significantly related to future excess returns, whereas the *pd*-ratio and *pe*-ratio are insignificant. The *py*-ratio is closely related to stock returns during both the 1929–1966 sample and the 1967–2003 sample (not shown to save space). Actually, for the 1967–2003 sample, the *py*-ratio predicts long-horizon stock returns just as well as the $\widehat{cay}$-ratio does. The *pe*-ratio and the *pd*-ratio, on the other hand, are only marginally related to stock returns for the 1967–2003 sample.

Hence, the overall picture is as follows: The *py*-ratio not only captures movements in stock returns well no matter which subsample is considered, it also does so better than the *pe*- and *pd*-ratios normally used to predict returns. In the 1967–2003 subsample, the

Table 5
Subsample regressions of long-horizon cumulative excess returns

Horizon K:		1	2	4	6
1929–1966					
py	coef.	−0.27**	−0.50**	−0.63**	−0.69**
	t-stat.	(2.80)	(2.64)	(5.06)	(7.80)
	$\overline{R}^2$	**0.21**	**0.34**	**0.48**	**0.70**
pe		−0.09	−0.14	−0.16	−0.27
		(1.34)	(0.92)	(0.51)	(0.79)
		−0.00	**−0.01**	**−0.01**	**0.02**
pd		−0.01	−0.11	−0.16	−0.18
		(0.14)	(0.68)	(0.50)	(0.51)
		−0.03	**−0.02**	**−0.01**	**−0.02**
1967–2003					
py		−0.12*	−0.22*	−0.31**	−0.47*
		(1.93)	(2.38)	(2.73)	(2.73)
		0.05	**0.12**	**0.14**	**0.21**
pe		−0.08	−0.13	−0.09	−0.04
		(1.32)	(1.20)	(0.56)	(0.15)
		0.02	**0.02**	**−0.02**	**−0.03**
pd		−0.08	−0.15	−0.12	−0.22
		(1.09)	(1.16)	(0.66)	(0.67)
		0.02	**0.04**	**−0.01**	**−0.01**
$\widehat{cay}$		6.55	12.70	17.89	23.86
		(3.98)	(7.22)	(5.43)	(9.48)
		0.30	**0.58**	**0.56**	**0.60**

Notes: The table shows results from regressions of future cumulative excess returns on the price–output ratio (*py*), the price–earnings ratio (*pe*), the price–dividend ratio (*py*), and Lettau and Ludvigson's (2001a) $\widehat{cay}$-ratio when data on $\widehat{cay}$ are available. $\widehat{cay}$ is an estimated cointegration relation linking consumption and labor income with asset wealth. The sample periods are 1929–1966 and 1967–2003. The table shows parameter estimates with t-statistics based on Newey and West's (1987) standard errors, in parentheses below (the weights are truncated at lag $K+1$), and $\overline{R}^2$s, indicated in bold. Significant statistics at the 1% level are indicated by ** and at the 10% level by *.

py-ratio captures long-horizon stock return variation as well as does $\widehat{cay}$ (results not shown to save space). Generally, the *py*-ratio captures less of the variation in excess returns relative to its ability in capturing stock return variation, but, at the same time, it also captures a higher fraction of the variation in excess returns than do the *pe*-ratio and the *pd*-ratio (with the exception of the 1948–2003 subsample, where the *pd*-ratio captures more of the variation in excess returns than the *py*-ratio).

Rangvid (2005) conducts unit root tests on different subsamples of the full sample. The general impression is that there is more evidence of cointegration between share prices and output for subsamples starting in 1929 compared to subsamples starting in 1948 or 1967; for the subsamples starting in 1948 and 1967, only the $\widehat{cay}$-ratio is stationary. Furthermore, the results for the *pd*-ratio are rather sensitive depending on whether 1929 or 1931 is used as the starting year; the *py*-ratio, on the other hand, is not sensitive to using either 1929 or 1931 as the starting year. All in all, if starting in 1929 or 1931, most of the test statistics indicate that the *py*-ratio is stationary.

4.4. Additional tests

4.4.1. Does the py-ratio predict interest rates?

The *py*-ratio in general captures more of the variation in stock returns than in excess returns, especially for more recent samples. The difference between stock returns and excess returns is that the risk-free rate is subtracted from stock returns to get excess returns. One way to understand the differences in predictability results using stock returns and excess returns would be to study whether the *py*-ratio is related to variation in future interest rates. Table 6 shows that the *py*-ratio is significantly related to future cumulative short interest rates for the full sample and marginally significantly related for the 1948–2003 period (for $K = 1, 2$, and 4). The *py*-ratio is not significant during the pre-1948 period. The *pe*-ratio and the *pd*-ratio are related to future interest rate movements for the short pre-1948 period only, whereas $\widehat{cay}$ is significantly related to long-term (six-year) interest rates.

To understand the implication of these results, note that Eq. (2) relates the *py*-ratio to future stock returns (the sum of excess returns and the risk-free rate). It does not, however, necessarily relate the *py*-ratio to excess returns only, just as Eq. (1) relates the *pd*-ratio to future stock returns, but not necessarily exclusively to future excess returns. It is interesting to note how the *py*-ratio is closely related to future stock returns, as the "dynamic Gordon model" implies, and more weakly related to future excess returns, especially for the more recent samples. Table 6 shows that a negative shock to the *py*-ratio (from a low p_t or a high y_t) is associated with higher interest rates and (from Table 2) with higher stock returns in the future. In other words, the *py*-ratio predicts both stock returns and interest rates (for the full 1929–2003 sample), but these predictions will have offsetting effects on the predictions of excess returns. As the *pe*-ratio and the *pd*-ratio do not predict interest rates, such offsetting effects are not present when using these variables. In addition, note that the *py*-ratio is not related to future interest rates for the early subsample, which makes sense because, as mentioned in Section 4.3, the *py*-ratio predicts both stock returns and excess returns for the early subsample—when the *py*-ratio does not predict interest rates, there is no offsetting effect.

Compared to other variables often reported to predict interest rates, the *py*-ratio captures variation in future interest rates better than, for instance, a relative interest rate (Campbell, 1991), where the average of the last three years of interest rates is subtracted from the current interest rate. On the other hand, there are at the same time many other variables that are better at predicting future interest rates than the *py*-ratio is. For instance, the persistence in interest rates is high, implying that a lagged interest rate captures a higher fraction of the variation in future interest rates than does the *py*-ratio. In regressions of cumulative future interest rates on the lagged *py*-ratio and a lagged interest rate, the *py*-ratio is only marginally significant, whereas the lagged interest rate is highly significant. The aim of the results in Table 6 are thus not necessarily to find a new variable that predicts interest rates to a higher extent than standard interest rate predictors do, but rather to understand why the *py*-ratio predicts stock returns better than it predicts excess returns.

4.4.2. Scaling the total value of the stock market with GDP

Up to this point, the *py*-ratio is derived from relating the Standard and Poor's Composite stock price index with nominal GDP. The stock price index summarizes the

Table 6
Regressions of cumulative interest rates on *py*, *pe*, *pd*, and $\widehat{cay}$

Horizon K:		1	2	4	6
1929–2003					
py	coef.	−0.04**	−0.08**	−0.16**	−0.24**
	t-stat.	(3.18)	(3.03)	(2.96)	(2.82)
	$\overline{R}^2$	**0.23**	**0.26**	**0.29**	**0.29**
pe		−0.01	−0.01	−0.01	−0.01
		(0.32)	(0.29)	(0.16)	(0.06)
		−0.01	**−0.01**	**−0.01**	**−0.01**
pd		0.01	0.04	0.10*	0.21**
		(1.31)	(1.39)	(1.95)	(2.91)
		0.02	**0.04**	**0.08**	**0.13**
1929–1947					
py		0.01	0.01	0.02	0.03
		(1.35)	(1.19)	(1.06)	(1.16)
		0.08	**0.09**	**0.05**	**0.02**
pe		−0.01*	−0.02*	−0.03*	0.04*
		(1.98)	(1.80)	(2.02)	(3.07)
		0.16	**0.23**	**0.22**	**0.38**
pd		−0.01	−0.02	−0.03*	−0.04**
		(1.63)	(1.60)	(1.83)	(2.65)
		0.09	**0.02**	**0.05**	**0.12**
1948–2003					
py		−0.04*	−0.07*	−0.12*	−0.15
		(2.40)	(2.24)	(1.64)	(1.17)
		0.17	**0.19**	**0.13**	**0.07**
pe		−0.02	−0.03	−0.06	−0.02
		(0.80)	(0.82)	(0.61)	(0.14)
		0.02	**0.03**	**0.01**	**−0.02**
pd		−0.01	−0.01	0.02	0.13
		(0.33)	(0.25)	(0.19)	(0.91)
		−0.00	**−0.02**	**−0.02**	**0.03**
$\widehat{cay}$		−0.55*	−0.75	−1.49	−3.29**
		(1.94)	(1.25)	(1.60)	(2.95)
		0.04	**0.01**	**0.01**	**0.06**

Notes: The table shows results from the regressions of future cumulative short interest rates on the price–output ratio (*py*), the price–earnings ratio (*pe*), the price–dividend ratio (*pd*), and Lettau and Ludvigson's (2001a) $\widehat{cay}$-ratio when data on $\widehat{cay}$ are available. $\widehat{cay}$ is an estimated cointegration relation linking consumption and labor income with asset wealth. The sample periods are 1929–2003, 1929–1947, and 1948–2003. The table shows parameter estimates with t-statistics based on Newey and West's (1987) standard errors, in parentheses below (the weights are truncated at lag $K+1$), and $\overline{R}^2$s, indicated in bold. Significant statistics at the 1% level are indicated by ** and at the 10% level by *.

return for those firms that are included the index, i.e., it tracks the value of a hypothetical investment in a fixed number of firms.[9] GDP, on the other hand, is a measure of the total

[9]While there are of course changes in the composition of the index over time, with some firms leaving the index and others being added, the number of firms in the index remains constant. To be precise, the Standard and Poor's Index was introduced in 1923 and covered 233 companies. The index, as it is known today, was introduced in 1957 when it was expanded to include 500 companies. The S&P 500 index used here tracks returns for "500 leading companies in leading industries of the U.S. economy." It also states on the S&P homepage that "although the

Table 7
Regressions of long-horizon cumulative returns on py^{total}

Horizon K:		1	2	4	6
Returns					
py^{total}	coef.	-0.10^{*}	-0.19^{*}	-0.30^{*}	-0.40^{*}
	t-stat.	(2.27)	(2.41)	(2.33)	(2.24)
	$\bar{R}^2$	**0.06**	**0.10**	**0.12**	**0.17**
Excess returns					
py^{total}		-0.13^{**}	-0.25^{**}	-0.41^{**}	-0.60^{*}
		(2.72)	(2.87)	(2.88)	(2.99)
		0.09	**0.15**	**0.21**	**0.32**

Notes: The table shows results from regressions of stock returns and excess returns on the py^{total}-ratio. The py^{total}-ratio is calculated by scaling the total value of the stock market with GDP. The table shows parameter estimates with t-statistics based on Newey and West's (1987) standard errors, in parentheses below (the weights are truncated at lag $K+1$), and $\bar{R}^2$s, indicated in bold. Significant statistics at the 1% level are indicated by ** and at the 10% level by *.

value of production in an economy. For this reason, it is pertinent to attempt to relate GDP to the total value of all firms in the economy, i.e., to the total value of the stock market. The total value of the U.S. stock market can be calculated by summing the total value of firms listed on the New York Stock Exchange (NYSE), the American Stock Exchange (AMEX), and Nasdaq. Relating the log of the total value of firms to GDP, a py^{total}-ratio can be calculated.

The py^{total}-ratio is not stationary. The ADF-test for stationarity for the full sample is -1.50 and the PP test is -1.34. The py^{total}-ratio is also not stationary when unit root tests are run on different subsamples.

If annual stock returns are regressed on the py^{total}-ratio, the t-statistic is 2.27 and the $\bar{R}^2$ is 0.06, whereas in a regression of long-horizon six-year stock returns on the py^{total}-ratio, the t-statistic is 2.24 and the $\bar{R}^2$ is 0.17. The figures appear in Table 7. For excess returns, the $\bar{R}^2$ increases from 0.09 to 0.32 at the six-year horizon while the t-statistic increases from 2.72 to 2.99. Because returns are stationary whereas the py^{total}-ratio is not, these regressions should be interpreted in the same way as results from standard regressions in the literature of returns on dividend yields and pd-ratios, which are often reported to be nonstationary, too; see, e.g., Table 1 and Campbell and Yogo (2006).

Why is the py-ratio stationary while the py^{total}-ratio is not? As mentioned previously, the stock price index tracks capital gains to a hypothetical investment in the large-cap firms included in the index, whereas the total market value of all firms includes firms added to (and deleted from) the stock market as well as equity expansions (and contractions, e.g., share repurchases) of existing firms. The cointegration results of this paper reveal that the stock price index of existing companies follows underlying output in the economy during the 1929–2003 period, whereas the total stock market value (which includes the added values of IPOs and other equity expansions) increases at a rate higher than GDP. It is difficult to fully explain this difference. There is, however, much debate on whether IPO

(*footnote continued*)
S&P 500 focuses on the large-cap segment of the market, with over 80% coverage of U.S. equities, it is also an ideal proxy for the total market.''

volumes can be explained by rational market behavior (Pastor and Veronesi, 2005) or whether IPO volumes are related to market efficiencies (e.g., Pagano et al., 1998).

4.4.3. Does the py-ratio predict changes in GDP?

The *py*-ratio does not forecast growth in GDP. Results (not shown for the sake of brevity) show this to be the case for both annual growth and multiperiod growth in GDP. This is comparable to the findings of Lettau and Ludvigson (2001a), whose $\widehat{cay}$-ratio forecasts returns but not changes in consumption. Lettau and Ludvigson also show that the dividend yield forecasts only returns and not dividend growth; see also Campbell (1991) and Cochrane (1992, 2001).[10] Eq. (2) revealed that fluctuations in the *py*-ratio arise from changes in expectations regarding future returns and/or changes in GDP. Given that the *py*-ratio varies over time, while these variations have been disconnected from future changes in GDP, the finding that the *py*-ratio tracks expected returns variation is further strengthened.

4.4.4. Timing of the py-ratio

The *py*-ratio, as explained in Section 3, is calculated by scaling the value of the share price index in January with the GDP of the previous year, i.e., share prices from, for instance, January 2003 are divided by the total GDP of 2002. The *py*-ratio thus gives the January "price" for each unit of GDP in the previous year in the same way that the price–dividend ratio gives the January price for a unit of dividends paid out during the previous year.

The final figure for total GDP, however, is published with a lag. In other words, the total GDP from 2002, for instance, is in principle unknown at the beginning of 2003 (even though market participants in January 2003 probably have a good estimate of the total 2002 GDP figure because they know GDP for the first 2–3 quarters of 2002). To gain perspective on this, an alternative *py*-ratio $py^a = p_t - y_{t-2}$ was tested. The results from regressions with the alternative py^a-ratio were mostly similar to those presented in Table 2. The alternative py^a-ratio captured approximately 11% of the variation in the following year's stock returns and the coefficient on the py^a-ratio was -0.16 (t-statistic $= 3.19$). For six-year cumulative returns, the $\overline{R}^2$ was 0.51 and the t-statistic on the py^a-ratio was 6.92. Results using excess returns and other long-horizon returns were also in accordance with those presented earlier. All in all, the results are robust for the calculation of the *py*-ratio as $py = p_t - y_{t-1}$ or $py^a = p_t - y_{t-2}$.

4.4.5. Other ratios/controls

Returns were also regressed on a price–consumption ratio. In Lucas (1978), aggregate dividends equal output that again equals aggregate consumption, i.e., if a price–output ratio predicts returns, then perhaps a price–consumption ratio does so as well. To calculate the price–consumption ratio, the 1929–2003 series for nominal Total Personal Consumption Expenditures was downloaded from the U.S. Bureau of Economic Analysis' database. As expected, the price–output ratio and the price–consumption ratio are highly correlated. Furthermore, the results using the price–consumption ratio are in many regards similar to those using the price–output ratio, i.e., the price–consumption ratio also captures a high

[10] Lettau and Ludvigson (2005) show that $\widehat{cay}$ also does not forecast labor income growth, GDP growth, or investment growth. Dividend–growth predictability will be further analyzed in Section 6.

fraction of the variation in future returns. For the sake of brevity, the results using the price–consumption ratio are not shown.[11]

A consumption–GDP ratio. Even when the correlation between total consumption and total output is high, there are of course fluctuations in the ratio of consumption to GDP. Cochrane (1994) was the first to study bivariate VARs of real GDP and real consumption, imposing cointegration between GDP and consumption (the measure of consumption used in Cochrane was the consumption of services plus non-durable goods), i.e., VAR models including lagged changes in real GDP as well as lagged changes in real consumption of non-durables and services in addition to the cy-ratio (the consumption-GDP ratio).[12] The results from Cochrane-like regressions using the full 1929–2003 sample are in Panel A, Table 8. The cy-ratio predicts future growth rates of real GDP, just as Cochrane (1994) reported. Lagged GDP growth rates are also significant, as in Cochrane (1994), whereas lagged consumption growth rates are not significant in the 1929–2003 sample. In Panel B, cumulative stock returns replace GDP growth rates as the dependent variables, and in Panel C, results with growth rates of dividends as the dependent variables are shown. From Panels B and C the overall conclusion can be drawn that the cy-ratio captures future long-horizon cumulative stock returns (though not excess returns—results not shown) as well as future long-horizon growth rates in dividends. On the other hand, the cy-ratio is not related to one-period stock returns or one-period growth in dividends.

A dividend–earnings ratio. The lagged payout ratio $d - e$ advocated by Lamont (1998) was not significant in a return-predicting regression. Lamont investigates the period of 1947–1994 using quarterly data and estimates the coefficient on the lagged payout ratio as 0.083 and the t-statistic as 2.96 in his examination of excess returns (Lamont's Table IV, regression 4). Lamont's 1947–1994 regression was performed using annual data and the coefficient was 0.32 (approximately four times Lamont's coefficient with quarterly data) and the t-statistic was 2.37. In the regression on the full 1929–2003 sample period, however, the lagged payout ratio turned out to be insignificant. The coefficient estimate for the whole sample period was -0.11, using excess returns, and the t-statistic was -0.81 ($\overline{R}^2 = 0.04$). For stock returns, the coefficient estimate was -0.17 while the t-statistic was -1.39. The lagged payout ratio was also insignificant when regressing cumulative long-horizon returns on the lagged payout ratio and in regressions on the 1947–2003 period. Lettau and Ludvigson (2001a) also find that the coefficient to the lagged payout ratio is insignificant in the sample they investigate.

Other controls. Finally, returns were regressed on lagged returns and on a lagged relative short interest rate (*rrel*). There was no autocorrelation in annual returns (stock returns or excess returns) as the lagged returns were insignificant. The relative interest rate was not significant either.

[11]The results can be obtained upon request.

[12]Specifically, Cochrane (1994) looks at quarterly real GNP. A recomputation of Cochrane's tests shows that there are basically no differences between the use of real GDP and real GNP data in the regressions.

Table 8
Full sample regressions of long-horizon cumulative GDP growth rates (Panel A), returns (Panel B), and dividend growth rates (Panel C) on the lagged *cy*-ratio and lagged one-period growth rates of consumption and GDP

Horizon K:		1	2	4	6
Panel A: GDP growth rate regressions					
cy	coef.	0.13*	0.33*	0.72**	0.95**
	t-stat.	(1.89)	(2.23)	(2.94)	(5.43)
Δy		0.62**	0.87**	0.52*	0.26**
		(4.20)	(2.90)	(2.15)	(2.60)
Δc		−0.06	0.08	0.36	1.01
		(0.26)	(0.14)	(0.60)	(2.61)
	$\overline{R}^2$	**0.34**	**0.30**	**0.37**	**0.56**
Panel B: Stock return regressions					
cy		−-0.35	−0.66	−0.79*	−1.05**
		(1.45)	(1.53)	(2.08)	(3.14)
Δy		0.46	1.06*	1.25*	−0.13
		(1.05)	(1.76)	(2.40)	(0.26)
Δc		0.24	−1.22	−3.42**	−2.41*
		(0.29)	(1.17)	(2.83)	(1.71)
		0.04	**0.07**	**0.08**	**0.07**
Panel C: Dividend growth rate regressions					
cy		−0.14	−0.32	−0.67*	−1.02**
		(0.78)	(1.01)	(2.26)	(2.90)
Δy		−0.17	−0.20	−1.10	−0.61
		(0.66)	(0.44)	(1.78)	(1.20)
Δc		1.46**	1.93*	1.80	0.92
		(3.60)	(2.02)	(1.52)	(0.99)
		0.09	**0.09**	**0.12**	**0.26**

Notes: Panel A shows results from regressions of future changes in GDP on the lagged consumption-GDP ratio as well as lagged changes in GDP and consumption. The regressions are inspired by Cochrane (1994). In Panel B, results from the regressions of stock returns (and in Panel C, dividend growth rates) on the lagged consumption-GDP ratio are shown. As in Cochrane (1994), the coefficients to the lagged changes in GDP and consumption are shown as well. The table shows parameter estimates with t-statistics based on Newey and West's (1987) standard errors, in parentheses below (the weights are truncated at lag $K+1$), and $\overline{R}^2$s, indicated in bold. Significant statistics at the 1% level are indicated by ** and at the 10% level by *.

5. Out-of-sample forecasts

The results presented in the previous sections have all been based on in-sample regressions. This makes sense when taking into account the arguments given in Inoue and Kilian (2004) for the use of in-sample tests when searching for predictability in population. Their argument is that out-of-sample analyses suffer from having less power because they are based on a splitting of the full sample into smaller subsamples. Campbell and Thompson (2005) make a related point. Nevertheless, one would often like to know whether a high $\overline{R}^2$ in-sample also indicates predictability out-of-sample.

Out-of-sample predictions are generally based on the estimation of the model over a base period with T observations from which a first set of forecasts are generated. The model is then reestimated over the base period plus one observation, with a second set of forecasts generated. This process continues until all observations have been used. In the tests presented in this section, out-of-sample forecasts for the period 1980–2003 are considered. Hence, the 1929–1980 period is used to generate the first set of forecasts, the 1929–1981 period generates the next set of forecasts, and so on. The out-of-sample analysis is based on the 1980–2003 period as Ang and Bekaert (2004), Goyal and Welch (2003), and Lettau and Ludvigson (2001a) all report that the variation in returns during this period is hard to capture using standard financial valuation ratios. It is thus interesting to evaluate whether the *py*-ratio does a better job in out-of-sample predictions during this period than financial ratios.

Comparing the forecasts from the different models throughout the out-of-sample period requires a metric that captures the quality of the forecasts in comparison to a benchmark model. Denote the root mean square error of the predictions from an unrestricted model RMS_{UR}. It is common to compare the RMS_{UR} with the root mean square error of the forecasts generated from a restricted benchmark model, RMS_{R}. This metric is generally called "Theil's U" so that $U = RMS_{\mathrm{UR}}/RMS_{\mathrm{R}}$. If $U<1$, the forecasts from the unrestricted model are more accurate than the forecasts generated from the restricted benchmark model. In order to judge whether the root mean square errors of the restricted and unrestricted models differ significantly from each other, the Harvey et al. (1997) modification to the Diebold and Mariano (1995) test statistic is used. Harvey et al. show that one of the several advantages of their modified test statistic is that it can be applied to forecast horizons exceeding one period (long-horizon forecasts).

Table 9 presents the results from evaluating the out-of-sample predictions from different forecasting models. Panel A reports the results for stock returns and Panel B contains the

Table 9
Forecasting U.S. returns out-of-sample using the price–output ratio (*py*), the price–earnings ratio (*pe*), or the price–dividend ratio (*pd*). Out-of-sample forecasting period: 1980–2003

K	*py*			*pe*			*pd*		
	U	*t*-stat	*p*-value	*U*	*t*-stat	*p*-value	*U*	*t*-stat	*p*-value
Panel A: Forecasting stock returns									
1.	0.919	−0.72	0.48	0.966	−0.40	0.69	1.024	0.66	0.52
2.	0.872	−0.79	0.44	1.008	0.07	0.94	1.098	1.11	0.28
4.	0.698**	−4.54	0.00	1.002	0.01	0.99	1.069	0.69	0.50
6.	0.595**	−9.67	0.00	1.118	0.42	0.68	1.149	1.139	0.18
Panel B: Forecasting excess returns									
1.	1.021	0.19	0.85	0.990	−0.16	0.87	1.021	0.44	0.66
2.	1.071	0.32	0.75	1.051	0.53	0.60	1.123	0.79	0.44
4.	1.033	0.15	0.88	1.083	0.74	0.47	1.204	0.97	0.34
6.	1.150	0.33	0.74	1.328*	2.66	0.02	1.501*	2.23	0.13

Notes: The table shows the root mean square forecast error of the *py*-ratio based predictions of returns and cumulative returns in relation to the root mean square forecast error of random-walk-with-a-drift predictions in the first U column (a value less than one indicates that the forecasts generated from the *py*-ratio are more precise than those generated from a random walk), the *t*-statistics for evaluating whether U is equal to one in the *t*-stat column (the test statistics are based on the procedure in Harvey et al., 1997), and the associated *p*-values in the third column. The *pe* columns show the same three statistics (U, *t*-statistic, and *p*-value) when predicting with the price–earnings ratio, while the *pd* columns show the statistics when predicting with the price–dividend ratios. Panel A reports results from the predictions of stock returns, while Panel B reports results for excess returns. Significant statistics at the 1% level are indicated by ** and at the 10% level by *.

results for excess returns. Three unrestricted forecasting models were used to predict annual returns and long-horizon cumulative returns: the *py*-ratio, the *pe*-ratio, and the *pd*-ratio. The forecasts from these models are compared to the forecasts from a random walk with a drift. Table 9 presents Theil's *U*s, the associated *t*-statistics (the degrees of freedom equal the number of observations minus one) as well as the probability values from tests of the hypotheses that the *U*-statistics are significantly equal to one, in which case the forecasts from the models are not significantly better than those of a random walk with a drift.

Panel A in Table 9 reveals that the *U*-statistics are smaller than one when predicting stock returns out-of-sample using the *py*-ratio (for $k > 2$, the *U*-statistics are significantly different from one) whereas the *U*-statistics are all basically larger than one when using either the price–dividend or the price–earnings ratio to predict stock returns. This implies that the stock return forecasts generated from the *py*-ratio are (for $k > 2$, significantly) superior to those generated from a random walk with a drift out-of-sample, whereas the out-of-sample stock return forecasts generated from the *pe*-ratio or the *pd*-ratio are generally worse than those generated from a random walk with a drift.

None of the regressors predict excess returns better than a random walk does. Notably, however, even if the *U*-statistics for predictions of excess returns are not statistically different from one, the root mean square errors from the *py*-ratio-based predictions are numerically smaller than the predictions based upon the *pd*-ratio.

6. Interpretations

How should the results presented thus far be interpreted? Why does the *py*-ratio contain information about expected returns that is not already captured by ratios of prices to earnings and dividends?

A recent theoretical explanation as to why dividend yields and future returns are sometimes only weakly correlated is provided by Menzly et al. (2004), who show how time-varying risk preferences induce a positive correlation between dividend yields and expected stock returns. They also show, however, that if expected dividend growth is also time varying, the otherwise positive relation between dividend yields and future stock returns is offset. They demonstrate that expected dividend growth is indeed time varying. Lettau and Ludvigson (2005) extend the arguments of Menzly et al. (2004) by verifying empirically that forecasts of returns and dividend growth rates are positively correlated. This is an important result because positive correlations between forecasts of returns and dividend growth rates will have offsetting effects on the price–dividend ratio, as in Eq. (1), implying that there is more variation in returns than that implied by the price–dividend ratio.

The procedure followed by Lettau and Ludvigson (2005) is to replace asset wealth in $\widehat{cay}$ with dividends in order to construct a $\widehat{cdy}$-ratio. They then regress future growth rates in dividends on the $\widehat{cdy}$-ratio and confirm that the $\widehat{cdy}$-ratio contains information about expected dividends. In this paper, share prices can be replaced with earnings in the *py*-ratio after which future growth rates in earnings can be regressed on the resulting earnings-to-GDP ratio. Likewise, future growth rates in dividends can be regressed on a dividends-to-GDP ratio. If the hypothesis is true that variation in dividends offsets the relation between returns and the price–dividend ratio, GDP scaled by dividends should be positively correlated with future growth rates in dividends. The same

Table 10
Growth in earnings and dividends regressions

Horizon K:		1	2	4	6
Growth in earnings					
ye	Coef.	0.20*	0.38**	0.50*	0.44*
	t-stat.	(2.40)	(2.83)	(2.74)	(3.54)
	$\overline{R}^2$	**0.10**	**0.18**	**0.22**	**0.22**
pe		0.17*	0.26*	0.40*	0.45*
		(2.02)	(1.88)	(1.65)	(2.04)
		0.07	**0.07**	**0.09**	**0.12**
Growth in dividends					
yd		0.07	0.15*	0.23*	0.26**
		(1.61)	(1.88)	(2.21)	(3.10)
		0.08	**0.15**	**0.20**	**0.27**
pd		0.12	0.13	0.19	0.24
		(1.45)	(1.08)	(0.90)	(1.01)
		0.09	**0.06**	**0.04**	**0.05**

Notes: The table shows results from regressions of future changes in earnings on the GDP-to-earnings ratio and on the price–earnings ratio, and from the regressions of future changes in dividends on the GDP-to-dividends and on the price–dividend ratio. The sample period is 1929–2003. The table shows parameter estimates with t-statistics based on Newey and West's (1987) standard errors, in parentheses below (the weights are truncated at lag $K+1$), and $\overline{R}^2$s, indicated in bold. Significant statistics at the 1% level are indicated by ** and at the 10% level by *.

is true for earnings. Table 10, which shows results from these regressions, also shows results from regressions of growth rates in earnings on the *pe*-ratio and from regressions of growth rates in dividends on the *pd*-ratio. The results verify that GDP scaled by earnings is generally positively correlated with future changes in earnings, while the GDP-to-dividends ratio is generally positively correlated with future changes in dividends. On the other hand, share prices scaled by dividends are not related to future changes in dividends, while share prices scaled by earnings are only marginally related to future changes in earnings.

These findings suggest that growth in earnings and dividends carry a predictable component that is not captured by financial valuation ratios. The predictable component is, however, captured by ratios of GDP to dividends and earnings. In other words, there is mean-reversion in financial fundamentals towards GDP. If GDP is high relative to earnings or dividends (high *ye*- or *yd*-ratios), future earnings and dividends will have a tendency to increase (as revealed by the positive coefficients to the lagged *ye*- and *yd*-ratios in Table 10). Such mean-reversion is not found using the *pd*-ratio and only to a lesser extent using the *pe*-ratio. In this sense, GDP fluctuations are relevant for stock price valuations.

All in all, these findings correspond well with those of Lettau and Ludvigson (2005), who show that dividends mean-revert towards the consumption-dividends-labor income ratio.

6.1. Alternative explanations discussed in the literature

The literature provides additional arguments why share prices can be temporarily disconnected from underlying dividends and earnings. The traditional argument is put forward by Miller and Modigliani (1961), who show that the amount of dividends that firms pay out can be completely disconnected from the true performance of firms. In theory, a firm can pay out an arbitrary level of dividends without influencing the value of the firm, i.e., the dividend policy of the firm is "irrelevant."

More recently, the relation between dividends and share prices has been discussed from a different angle. During the 1990s, stock prices diverged considerably from underlying dividends (Shiller, 2000). As reported elsewhere (Campbell and Shiller, 2001), this divergence was probably due to both a skyrocketing of the stock prices themselves as well as the low dividends that were paid out. Fama and French (2001), for instance, argue that dividends have been "disappearing," although DeAngelo et al. (2004) report that while the number of firms paying out dividends has decreased, the total amount of dividends paid out has actually increased during the 1978–2000 period. Campbell and Shiller (2001) investigate whether the lower payouts can be explained by the fact that firms increasingly pay out profits in terms of share repurchases instead of dividends. They conclude that, even after taking share repurchases into account, the high values of the price–dividend ratio during the 1990s remain puzzling.

What about earnings? The level of earnings is not "irrelevant," and earnings have not been reported to be disappearing. Earnings are based on firms' accounts, however, which are partly dependent upon discretionary decisions that do not always reflect the true fundamental value of the firm. For instance, the way executive bonuses are treated in financial reports can cause earnings to become noisy measures of a firm's true value (Hall and Murphy, 2003). Hall (2001) discusses the interpretation of earnings from a somewhat different angle, focusing on whether investments in intangibles can cause earnings to be biased. He concludes that the behavior during the 1990s of the aggregate stock market in relation to GDP was understandable in terms of rational asset pricing based on discounted cash flows. Campbell and Shiller (2001), on the other hand, evaluate the impact of investments in intangibles on earnings and find that after correcting for these investments, the price–earnings ratios of the 1990s still remain high.

The aim of this short discussion on dividends and earnings is not to say that dividends and earnings are uninformative or that GDP data are perfect.[13] Indeed, it is difficult to imagine economies where output, dividends, earnings, and ultimately prices are not related in the long run (i.e., they follow the same long-term growth trend). Empirically, however, there are temporary noise terms impinging upon dividends, earnings, and GDP, as just discussed, and these variables thus differ in every period, which means they will also contain different information about future returns in every period. This paper has shown that there is reasonable evidence that share prices scaled by GDP contain information beyond that contained by dividends and earnings.

[13]One reason why GDP data are not perfect is that it takes time before GDP has been calculated and the figures released, as also discussed and examined in Section 4.4.4. Furthermore, GDP measures total production in the economy and not only the production of private firms. For instance, if the public sector in the economy has grown, GDP will ceteris paribus increase, too. Whether a larger public sector is good or bad for private firms, and thus share prices, is not a priori clear. In order to gain perspective on the possible consequences of different compositions of GDP for the relation between returns and the *py*-ratio, Section 4.3 showed that there was a strong relation between the *py*-ratio and stock returns, also for different subsamples of the full sample. Finally, there are also measurement problems with GDP data. Notwithstanding these issues with GDP data, the results presented so far indicate that GDP data contain interesting information about expected returns.

7. Results from other countries

It is reasonable to ask whether the findings presented thus far occur in data from the United States only or whether they are more general and found in data from other economies as well. This section presents results for the G-7 countries.

The IMF's International Financial Statistics (IFS) contain long series of Industrial Share Prices. For Canada, Germany, Japan, the U.K., and the U.S., the series are available from 1960. For France, the series is available from 1965 and for Italy it is available from 1970. IFS provides only share prices, i.e., the returns examined in this section are proxied by the log changes in the share price indices (the capital gains). The use of IFS data is fairly common when studying the long-run features of international stock markets and is also used by Goetzmann and Jorion (1999), among others. As in the previous sections, the January share prices are used when calculating annual returns and the January share prices are divided by the GDP of the previous year when calculating the *py*-ratios. The nominal GDP data are also from the IMF's International Financial Statistics as are the interest rates on Treasury bills that are used to calculate excess returns.

7.1. Predicting returns in the G-7 countries

The results from the regressions of annual and long-horizon cumulative stock returns for the different countries on their *py*-ratios are shown in panel (*a*) of Table 11 and the results for excess returns in panel (*b*).

Consider the results for stock returns first (panel (*a*), Table 11). The most important aspects are that the *py*-ratios are significant for all countries in the one-year stock return regressions (where there are no overlapping observations), they rise with the horizon to reach very high *t*-values and $\overline{R}^2$s at the six-year horizon (most notably in Italy with a *t*-statistic of 14.17 and an $\overline{R}^2$ of 0.75), and all signs are negative as expected. In other words, the *py*-ratio predicts annual stock returns in countries other than the U.S.

Table 11, Panel (b) contains the results for excess returns. The results for excess returns are as strong as for stock returns, or almost as strong, for France, Germany, Italy, Japan, and the U.K., i.e. the *t*-statistics and $\overline{R}^2$s from the excess return regressions rise with the horizons and reach high values (for Italy an $\overline{R}^2$ of 0.74 when $K = 6$). On the other hand, as also found in the previous sections focusing on U.S. data, the *py*-ratio is insignificant or only marginally significant in the annual excess return regressions using data from Canada and the U.S. The finding that the *py*-ratio is good at predicting stock returns but predicts excess returns less well in the U.S. during the post-war period is compatible with the findings presented in Tables 4 and 5.

8. Summary and conclusion

The first ratios that were used to predict stock returns were ratios of purely financial indicators, such as the price–dividend ratio, the price–earnings ratio, the dividend yield (Campbell and Shiller, 1988a,b; Fama and French, 1988, 1989),[14] the dividend–earnings ratio (Lamont, 1998), and the ratio of a short interest rate to its historic moving average (Campbell, 1991; Hodrick, 1992).

[14] Fama and French (1988) cite Dow (1920) as the first to use the dividend yield to predict stock returns.

Table 11

Horizon K:	1	2	4	6
(a) Regressions of stock returns on the py-ratio. G-7 countries				
Canada	−0.07*	−0.17*	−0.24*	−0.35**
	(1.70)	(1.99)	(1.99)	(2.94)
	0.02	**0.09**	**0.18**	**0.31**
France	−0.22**	−0.37**	−0.55**	−0.80**
	(2.69)	(4.02)	(4.58)	(8.67)
	0.14	**0.24**	**0.34**	**0.54**
Germany	−0.21*	−0.30**	−0.55**	−0.70**
	(3.67)	(3.99)	(4.24)	(5.76)
	0.14	**0.30**	**0.42**	**0.55**
Italy	−0.30**	−0.61**	−1.24**	−1.26**
	(3.53)	(4.36)	(8.11)	(14.17)
	0.22	**0.42**	**0.74**	**0.75**
Japan	−0.17*	−0.41**	−0.71**	−1.04**
	(1.78)	(2.96)	(4.28)	(4.86)
	0.05	**0.18**	**0.34**	**0.45**
UK	−0.27**	−0.51**	−0.71**	−0.97**
	(3.45)	(4.46)	(6.23)	(8.07)
	0.16	**0.35**	**0.55**	**0.71**
US	−0.11*	−0.22**	−0.34**	−0.59**
	(2.01)	(2.55)	(3.92)	(4.69)
	0.06	**0.13**	**0.19**	**0.34**
(b) Regressions of excess returns on the py-ratio. G-7 countries				
Canada	−0.03	−0.09	−0.10	−0.16
	(0.67)	(0.96)	(0.70)	(1.00)
	−0.02	**0.00**	**−0.00**	**0.03**
France	0.21*	−0.35**	−0.54**	−0.81**
	(1.92)	(2.78)	(3.06)	(3.92)
	0.09	**0.15**	**0.19**	**0.33**
Germany	−0.19**	−0.37**	−0.50**	−0.65**
	(3.07)	(3.37)	(3.37)	(4.10)
	0.10	**0.22**	**0.31**	**0.45**
Italy	−0.28**	−0.58**	−1.12**	−1.28**
	(2.83)	(3.72)	(7.34)	(9.10)
	0.14	**0.31**	**0.64**	**0.74**
Japan	−0.19*	−0.46**	−0.79**	−1.09**
	(1.89)	(3.21)	(5.52)	(5.82)
	0.07	**0.22**	**0.43**	**0.56**
UK	−0.25*	−0.48**	−0.62**	−0.92**
	(2.51)	(3.16)	(3.40)	(5.74)
	0.09	**0.23**	**0.34**	**0.56**
US	−0.08	−0.17	−0.25*	−0.49**
	(1.25)	(1.61)	(1.97)	(2.64)
	0.01	**0.06**	**0.06**	**0.16**

Notes: The table shows results from the regressions of stock returns, in panel (*a*), and excess returns, panel (*b*), on the *py*-ratios of the G-7 countries. Returns are calculated as log-changes in the Industrial Share Price Indices. The share price series are available from the IMFs IFS. The periods examined are 1960–2003 (Canada, Germany, Japan, the U.K., and the U.S.), 1965–2003 (France), and 1970–2003 (Italy). The *py*-ratios are calculated by scaling share prices by lagged GDP. The table shows parameter estimates with t-statistics based on Newey and West's (1987) standard errors, in parentheses below (the weights are truncated at lag $K+1$), and $\overline{R}^2$s, indicated in bold. Significant statistics at the 1% level are indicated by ** and at the 10% level by *.

If returns are predictable using financial ratios, the immediate question is *why* returns are predictable. One answer could be that markets react to information that should not lead to movements in prices in an efficient market (Cutler et al., 1989). Other reasons, however, could be that required returns and possibly risk aversion (Campbell and Cochrane, 1999) change over the business cycle as a result of time-variation in investment opportunities or that aggregate cash flows contain a small but persistent component (Bansal and Yaron, 2004). If these latter explanations contain some truth, it seems reasonable to conjecture that the macroeconomic variables that ultimately determine investment opportunities should contain information that can be used to predict returns.

The intuitive motivation for investigating whether the price–output ratio predicts returns is thus a simple one. If the predictability of returns is related to the macroeconomic situation and if financial ratios, including stock prices, predict returns because stock prices are mean-reverting towards some fundamental, then perhaps a ratio of stock prices to a macroeconomic variable also predicts returns. This paper relates share prices to aggregate output in the economy. For this reason, the paper relates to the work of Cochrane (1991), who also focuses on production in the economy. GDP also measures total income in the economy. Hence, this paper is also related to Lettau and Ludvigson (2001a, 2005), Santos and Veronesi (2006), and Julliard (2004), who focus on the importance of fluctuations in consumption and labor income for measuring the predictability of stock returns.

The overall result of this paper is that the price–output ratio captures a substantial fraction of the variation in future returns. The finding that the *py*-ratio predicts returns seems to be robust and is particularly true for stock returns but also, though to a lesser extent during the recent decades, for excess returns.

The paper examines whether the price–output ratio captures a high fraction of the variation in returns because earnings and dividends mean-revert towards GDP. It is found that fluctuations in GDP matter with regard to fluctuations in financial fundamentals in the sense that a high GDP relative to earnings or dividends implies that dividends and earnings will have a tendency to increase in the future. Such mean-reversion of earnings and dividends cannot be established when examining the *pd*-ratio and only to a lesser extent when examining the *pe*-ratio. In this regard, this paper is related to Lettau and Ludvigson (2005), who report that dividends mean-revert towards a consumption-dividend-labor income ratio.

Given the results of the paper, one extension in particular relates to the forever-interesting issue of market efficiency. Indeed, when the *py*-ratio predicts returns, one would ultimately like to know whether this predictability is due to some form of irrationality or whether it can be explained by rational asset pricing. In order to investigate such issues, a full-fleshed asset-pricing model is needed. A natural extension of this paper would thus be to use the *py*-ratio in more formal asset-pricing tests and see whether the *py*-ratio is useful in trying to explain the cross-sectional distribution of stock returns as in, for instance, Lettau and Ludvigson (2001b), who show how their $\widehat{cay}$-ratio is able to capture a substantial part of the cross-sectional variation in U.S. returns.

References

Ang, A., Bekaert, G., 2004. Stock return predictability: is it there? Manuscript. Columbia Business School.

Bansal, R., Yaron, A., 2004. Risks for the long run: a potential resolution of asset pricing puzzles. Journal of Finance 59, 1481–1510.

Campbell, J.Y., 1991. A variance decomposition of stock returns. Economic Journal 101, 157–179.
Campbell, J.Y., Cochrane, J.H., 1999. By force of habit: a consumption-based explanation of aggregate stock market behaviour. Journal of Political Economy 107, 205–251.
Campbell, J.Y., Shiller, R.J., 1988a. Stock prices, earnings, and expected dividends. Journal of Finance 43, 661–676.
Campbell, J.Y., Shiller, R.J., 1988b. The dividend-price ratio and expectations of future dividends and discount factors. Review of Financial Studies 1, 195–228.
Campbell, J.Y., Shiller, R.J., 2001. Valuation ratios and the long-run stock market outlook: an update. NBER WP 8221.
Campbell, J.Y., Thompson, S., 2005. Predicting the equity premium out of sample: can anything beat the historical average? Manuscript. Harvard University.
Campbell, J.Y., Yogo, M., 2006. Efficient tests of stock return predictability. Journal of Financial Economics, forthcoming.
Cochrane, J.H., 1991. Production-based asset pricing and the link between stock returns and economic fluctuations. Journal of Finance 46, 207–234.
Cochrane, J.H., 1992. Explaining the variance of price–dividend ratios. Review of Financial Studies 5, 243–280.
Cochrane, J.H., 1994. Permanent and transitory components of GNP and stock prices. Quarterly Journal of Economics 109, 241–265.
Cochrane, J.H., 2001. Asset Pricing Theory. Princeton University Press, Princeton.
Cochrane, J.H., 2005. Financial markets and the real economy. Manuscript. GSB, University of Chicago.
Cutler, D.M., Poterba, J.M., Summers, L.H., 1989. What moves stock prices? Journal of Portfolio Management 15, 4–12.
DeAngelo, H., DeAngelo, L., Skinner, D.J., 2004. Are dividends disappearing? Dividend concentration and the consolidation of earnings. Journal of Financial Economics 72, 425–456.
Diebold, F.X., Mariano, R.S., 1995. Comparing predictive accuracy. Journal of Business and Economic Statistics 13, 253–263.
Dow, C.H., 1920. Scientific stock speculation. The Magazine of Wall Street.
Fama, E.F., French, K.F., 1988. Dividend yields and expected stock returns. Journal of Financial Economics 22, 3–25.
Fama, E.F., French, K.F., 1989. Business conditions and expected returns on stocks and bonds. Journal of Financial Economics 25, 23–50.
Fama, E.F., French, K.F., 2001. Disappearing dividends: changing firm characteristics or lower propensity to pay? Journal of Financial Economics 60, 3–43.
Goetzmann, W.N., Jorion, P., 1999. Global stock markets in the twentieth century. Journal of Finance 54, 935–980.
Goyal, A., Welch, I., 2003. Predicting the equity premium with dividend ratios. Management Science 49, 639–654.
Goyal, A., Welch, I., 2005. A comprehensive look at the empirical performance of equity premium prediction. Yale ICF Working Paper No. 04-11.
Hall, R.E., 2001. Struggling to understand the stock market. American Economic Review 91, 1–11.
Hall, B.J., Murphy, K.J., 2003. The trouble with stock options. Journal of Economic Perspectives 17, 49–70.
Harvey, D., Leybourne, S., Newbold, P., 1997. Testing the equality of prediction mean squared errors. International Journal of Forecasting 13, 281–291.
Hodrick, R., 1992. Dividend yields and expected stock returns: alternative procedures for inference and measurement. Review of Financial Studies 5, 357–386.
Horvath, M.T.K., Watson, M.W., 1995. Testing for cointegration when some of the cointegration vectors are prespecified. Econometric Theory 11, 984–1014.
Inoue, A., Kilian, L., 2004. In-sample or out-of-sample tests: which one should we use? Econometric Reviews 23, 371–402.
Johansen, S., 1991. Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. Econometrica 59, 1551–1580.
Julliard, C., 2004. Labor income risk and asset returns. Manuscript, Princeton University.
Lamont, O., 1998. Earnings and expected stock returns. Journal of Finance 53, 1551–1587.
Lettau, M., Ludvigson, S., 2001a. Consumption, aggregate wealth and expected stock returns. Journal of Finance 56, 815–849.
Lettau, M., Ludvigson, S., 2001b. Resurrecting the CCAPM: a cross-sectional test when risk premia are time-varying. Journal of Political Economy 109, 1238–1287.

Lettau, M., Ludvigson, S., 2005. Expected returns and expected dividend growth. Journal of Financial Economics 76, 583–626.

Lucas, R.E., 1978. Asset prices in an exchange economy. Econometrica 49, 1429–1446.

Menzly, L., Santos, T., Veronesi, P., 2004. Understanding predictability. Journal of Political Economy 112, 1–47.

Miller, M.H., Modigliani, F., 1961. Dividend policy, growth, and the valuation of shares. Journal of Business 34, 411–433.

Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.

Pagano, M., Panetta, F., Zingales, L., 1998. Why do companies go public? An empirical analysis. Journal of Finance 58, 1749–1789.

Pastor, L., Veronesi, L., 2005. Rational IPO waves. Journal of Finance 60, 1713–1757.

Rangvid, J., 2005. Output and expected returns. Manuscript, Copenhagen Business School.

Santos, T., Veronesi, P., 2006. Labor income and predictable stock returns. Review of Financial Studies 19, 1–44.

Shiller, R.J., 2000. Irrational Exuberance. Princeton University Press, Princeton.

Stock, J.H., 1991. Confidence intervals for the largest unit root in U.S. macroeconomic time series. Journal of Monetary Economics 28, 435–459.

Valkanov, R., 2003. Long-horizon regressions: theoretical results and applications. Journal of Financial Economics 68, 201–232.