A Stock Return Decomposition Using Observables

Benjamin Knox[1] and Annette Vissing-Jorgensen[2]

[1]Federal Reserve Board

[2]Federal Reserve Board and CEPR

Abstract

We propose a method to decompose realized stock returns period by period. First, we argue that one can directly estimate expected stock returns from securities available in modern financial markets (using the real yield curve and the Martin (2017) equity risk premium). Second, we derive a new decomposition of realized returns into components due to changes in real yields, equity risk premia, expected dividends, and the realized dividend yield. The stock price elasticities with respect to these inputs are determined by dividend strip weights which can be estimated from dividend futures prices. We apply our method to decompose quarterly and annual realized S&P500 returns going back to 2004Q4. Change to real yields generate a large positive return component in periods of crisis and a large negative return component during periods of monetary tightening, while we find a substantial negative return component from higher equity risk premia in crisis periods. Adding to the growing literature on the stock market during COVID, we furthermore provide a detailed decomposition of the realized stock return for 2020. Risk premium changes drove much of the crash and rebound in the S&P500 in 2020 while a fall in long-term real yields drove a strong positive return for 2020 as a whole.

This version: December 27, 2022

Emails: ben.knox@frb.gov and Annette.Vissing-Jorgensen@frb.gov. Benjamin Knox acknowledges support from the FRIC Center for Financial Frictions (grant no. DNRF102). We are grateful to discussants Michael Bauer (CEPR), Jules van Binsbergen (NBER Summer Institute), Tolga Cenesizoglu (EFA), Niels Gormsen (FRB), Valentin Haddad (Red Rock Finance), Lars Lochstoer (WFA), and seminar participants at Copenhagen Business School, UT Austin, UCSD, UIUC, New York University, Johns Hopkins University, Indiana University, Cornell University, Carnegie-Mellon University, University of Hong Kong, Bundesbank, Russian Central Bank/HSE/NES, and the SEM conference for very useful feedback. We thank Magnus Dahlquist and Markus Ibert for sharing their data on asset manager return expectations. The views in this paper should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or any other person associated with the Federal Reserve System.

I. Introduction

A central theme in asset pricing is what types of news drive realized asset returns. A large literature combines the log-linearization of Campbell and Shiller (1988) with a VAR approach as in Campbell (1991) to decompose stock return variance into components coming from cash flow news, discount rate news, and the covariance of the two. In this paper we propose a decomposition not of return variance but of the realized return for a given period. This allows for an interpretation of the movement of the stock market period by period. It can also be used to assess potential risks to the market ex-ante, similarly to the way investors use duration to assess risks ex-ante in the bond market.

Our approach relies on two main ideas. First, we make the simple observation that in modern financial markets, a lot of information about the real yield curve and the equity risk premium is observable. The term structure of the real riskless rate can be measured out to 30 years using the nominal rate on interest rate swaps combined with inflation swaps. The term structure of the equity risk premium is not directly observable, but Martin (2017) provides a lower bound on the equity risk premium based on S&P500 index options. We extend Martin's empirical evidence that this lower bound is approximately tight and thus is close to the actual equity risk premium.[1] The bound can be calculated out to around 2 years in recent years, based on available S&P500 options. Because fluctuations in the equity risk premium have a large transitory component, fluctuations in the first two years will account for a substantial part of equity risk premium news. We supplement Martin's equity risk premium approach with data for asset managers' equity risk premium estimates out to year $t + 10$, obtained from Dahlquist and Ibert (2021).

Second, to utilize the availability of rich discount rate data, we develop a simple stock return decomposition for the overall stock market which is straightforward to map to available data. Our approach starts by expressing the stock price as the sum of the present values of each of the dividends. At each point in time, there are dividends one, two, three, ... years out. The realized capital gain on the stock market in a given period is therefore a weighted average of (1) how the value of "generic" dividend strips (i.e., dividend strips that pay off a certain number of periods out) evolve over time, and (2) the initial weights of each of the generic dividend strips in the overall

[1] We also supplement Martin's analysis with theoretical analysis of how the *change* in the Martin lower bound relates to the true change in the equity risk premium. We argue that to the extent the bound is not exact, it is likely that our approach will understate the role of risk premium changes for realized returns.

market (the dividend strip weights). Under the assumption that the expected return on a given dividend strip is proportional to the expected return on the overall market, we can decompose the price changes of generic dividend strips, and in turn the capital gain on the overall market, into components from changes to the (riskfree) real yield curve, changes to the stock market risk premium, and changes to expected dividends. The decomposition shows that the effect of a change to a given input (real yields, risk premia or expected dividends) on the stock market capital gain (what we refer to a stock price elastiticies) is driven by the dividend strips weights.

The elasticity of the stock price with respect to the expected return (from riskless yield, or from risk premium) in a future year k years out is driven by the dividend strip weights of dividends to be received k years out or later. This is intuitive, since the present values of dividends to be received earlier than k years out are unaffected by increased discounting k years out. Changes in expected returns at longer maturities therefore have a smaller stock price impact than changes in expected returns at nearer-term maturities. The elasticity of the stock price with respect to the expected dividend k years out is simply the dividend strip weight for that dividend. Intuitively, the effect of a given percentage change in a (generic) expected dividend on the stock price depends on how large that expected dividend is prior to the change, measured by its weight in the overall stock market. Crucially, dividend strip weights can be calculated from dividend futures prices. We calculate dividend strip weights from dividend futures out to year 10 and show how dividend strip weights can be estimated past year 10 if expected returns and expected growth are approximately constant past year 10. Market data on dividend futures then be used to estimate the market's perception today of the ratio G/R at the terminal date, where G denotes the gross expected growth rate and R the gross expected return, both past year 10.

Armed with observable measures of changes to the real yield curve and changes to equity risk premia, and with observable elasticities, we are able to implement our decomposition of the realized stock return for a given period into components from (a) real yield curve changes to year 30, (b) equity risk premium changes to year 2, (c) changes to expected cash flows or to real yields/risk premia past the horizons of observable data, and (d) the realized dividend yield.

One could also exploit our idea of using observables inputs to decompose returns using the Campbell and Shiller (1988) (CS) log-linearization. In that approach, the inputs to the decomposition would be changes to expected log returns (the log riskless yields and the log risk premia) and expected log dividends. Appendix F lays out how this would work, emphasizing how to use

observable dividend futures prices to guide the choice of the log-linearization parameter ρ in the CS log-linearization. For brevity, we present results only from our decomposition in our empirical applications as we find it more intuitive to think of realized returns drivers being changes to expected returns and expected dividends rather than changes to expected log returns and expected log dividends.

As discussed above, by implementing a stock market decomposition using observables, one can measure the contributions of various return components for the realized stock return in a particular time period. A standard VAR estimation also produces time series of cash flow news and discount rate news. Could one use a VAR-based analysis to understand stock market developments in a given time period? We think the answer is no, in the sense that this would be less informative. Specifically, a VAR-based return decomposition is typically used to estimate whether stock market fluctuations on average, over a long estimation sample, tend to be driven more by cash flow news or discount rate news. This objective aligns well with the assumptions made. One assumed a time-invariant VAR model structure, with constant regression coefficients. The drawback of using a VAR for a period-by-period decomposition of returns is that VARs interpret all movements in the predictors equally. For example, suppose the price-earnings ratio falls in a given period of interest and that declines in the price-earnings ratios predict higher stock returns over the VAR estimation sample. A VAR approach to understanding the realized return for the period will then suggest that there was a negative discount rate component to the realized stock return even if the decline in the price-earnings ratio in this particular period was due mainly to negative news about future cash flows. By contrast, since our approach does not rely on regressions it does not impose any constraints on the mix of discount rate news and cash flow news in a given period.[2] Effectively, because our decomposition is based on observable data and does not require any regressions, it allows equity market researchers to decompose realized stock returns in a way that is conceptually similar to that used in event studies of yield changes in bond markets. See, e.g., Krishnamurthy and Vissing-Jorgensen (2011) for a decomposition of yield changes around quantitative easing announcement dates into components.

We provide two applications of our new approach. Our first application is a decomposition of realized annual returns and realized quarterly returns on the S&P500 over the period 2004Q4-2022Q2. We find a large contribution from changing real yields out to year 30 to realized stock

[2]Recent work in the VAR literature allows for time-varying coefficients, see e.g. Bianchi (2020).

returns. The contribution in annual data (with only a half year of data available for 2022) ranges from +37% in 2008 to -27% in 2022. Across years, large positive contributions from changes to real yields are seen in periods of crisis, while large negative contributions are seen in periods of monetary tightening. The contribution to realized stock returns from changing risk premia for year 1 and 2 is substantial in crisis times, with a -11% contribution in 2008 and contributions of -6% or worse in 2010Q2 (the start of the Greek debt crisis), 2011Q3 (the spread of the Greek debt crisis to other European countries), and 2020Q1 (the onset of the COVID crisis). The component of realized stock returns driven by changes to expected cash flows and to long-term discount rates is also substantial. This component of realized stock returns is strongly negatively correlated with the component capturing the effects of changing real yields implying that adverse changes to cash flows and/or risk premia tend to occur during times of falling real yields.

As our second application we provide a detailed decomposition of the S&P500 return over the year 2020. The question of understanding stock market movements in a given period gained particular interest during the COVID crisis. The compounded S&P500 return was -30.4% from the start of 2020 up to March 23, 2020. This was followed by a sharp recovery and the market ended the year up by 18.4%. The dramatic moves in the stock market led to a series of questions. What drove the sharp decline in the market? Why did the market recover so fast despite the lingering health crisis? Why was the market recovery so strong, with a large gain for the year? The application of our decomposition to 2020 reveals two key facts (we review related literature in section V.A).

First, we estimate that from the start of the year up to March 18, the equity risk premium for the one-year horizon increased from under 3% to 15.3%, with further increases in the year-2 risk premium. The increase in the risk premium for the first two years contributed -12.1% to the stock return up to March 18 (which was -25.4% at that point). Using quarterly data on asset manager equity risk premium expectations out to the 10-year horizon, combined with Nelson-Siegel curve estimation for the equity premium term structure, we estimate that equity premium news out to year 10 can account for all of the market decline in 2020Q1. During the subsequent recovery period, equity risk premia decline quickly. An "A-shaped" pattern for the equity risk premium thus helps explain the V-shaped pattern of the stock price.

Second, with the exception of an upward spike in long rates from March 9-18, real riskless rates drop dramatically across all maturities and do not recover by the end of the year. The 10-year

real riskless rate declines over 100 bps over the year and real forward rates fall even out to the 30-year horizon. For the year 2020, the decline in the term structure of real rates out to year 30 contributes a 23.7% increase in the stock market. There is some negative returns from changes to expected cash flows/remaining long-term discount rate news in the crash but almost none for the year as a whole.

The outline of the paper is as follows. Section II derives our stock return decomposition. Section III describes how to map the decomposition to observable inputs. Section IV provides results from decomposing annual and quarterly returns. Section V shows the application of our method to understand realized stock returns in 2020. Section VI concludes. The Appendix contains all proofs along with a series of supplementary results.

II. A new stock return decomposition

We derive a simple decomposition of the realized stock market return (over a short period) into real yield curve news, equity premium news and cash flow news. Our approach relies on calculating first derivatives and elasticities of the stock price to expected returns and expected dividends at various maturities. The percentage change in an expected return or expected dividend multiplied by the elasticity of the stock price to this variable is then its contribution to the aggregate stock price movement.

A. Background definitions

We begin with some definitions of prices, dividends and returns. Unless otherwise noted, all variables are in real terms. Start from the present value formula of the stock market

$$P_t = \sum_{n=1}^{\infty} P_t^{(n)} = \sum_{n=1}^{\infty} \frac{E_t\left[D_{t+n}\right]}{E_t\left(R_{t+1}^{(n)} R_{t+2}^{(n-1)} ... R_{t+n}^{(1)}\right)} \tag{1}$$

where $P_t^{(n)}$ is the value of the n^{th} dividend strip (i.e., the present value at time t of the expected dividend paid out at time $t+n$), $R_{t,n} = E_t\left(R_{t+1}^{(n)} R_{t+2}^{(n-1)} ... R_{t+n}^{(1)}\right)$ is the n-period cumulative gross discount rate at time t for discounting $E_t\left[D_{t+n}\right]$ back from $t+n$ to t, and $R_{t+1}^{(n)} = \frac{P_{t+1}^{(n-1)}}{P_t^{(n)}}$ is the

one-period gross return of the n^{th} dividend strip where $P_{t+1}^{(0)} = D_{t+1}$.[3]

The one-period gross return on the market can be expressed as the value-weighted average of the one-period gross returns on all dividend strips

$$R_{t+1} = \sum_{n=1}^{\infty} w_t^{(n)} R_{t+1}^{(n)} \tag{2}$$

where

$$w_t^{(n)} = \frac{P_t^{(n)}}{P_t} \tag{3}$$

is the weight of the n^{th} dividend strip (the present value of the expected dividend paid out at time $t+n$ relative to the overall stock market value).

B. Imposing structure on expected returns

Our objective is to provide a return decomposition that can be mapped to observables. The expected cumulative return in the denominator in equation (1) is not an observable variable. However, if we can (a) take the expectation in $E_t\left(R_{t+1}^{(n)} R_{t+2}^{(n-1)} ... R_{t+n}^{(1)}\right)$ inside the bracket, and (b) express expected returns on dividends strips in terms of expected returns on the market, then we can derive price sensitivities of today's stock price with respect to the expected return on the market (at various horizons) and using empirical measures of the market expected return to map the result to the data. We start with (a) and prove the following Lemma.

Lemma 1 (**the expected cumulative dividend strip return**). *For any joint distribution of the one-period dividend strip returns for a given dividend strip,* $R_{t+1}^{(n)}$, $R_{t+2}^{(n-1)}$,...,$R_{t+n}^{(1)}$, *the expected cumulative dividend strip return can be expressed*

$$E_t\left(R_{t+1}^{(n)} R_{t+2}^{(n-1)} ... R_{t+n}^{(1)}\right) = a_t^{(n)} E_t(R_{t+1}^{(n)}) E_t(R_{t+2}^{(n-1)}) ... E_t(R_{t+n}^{(1)}) \tag{4}$$

where the constant $a_t^{(n)}$ *is known at time* t *and depends on the joint distribution of one-period dividend strip returns for a given dividend strip:*

(i) If $R_{t+1}^{(n)}$, $R_{t+2}^{(n-1)}$,...,$R_{t+n}^{(1)}$ *are independent, then* $a_t^{(n)} = 1$.

[3] The value of the n^{th} dividend strip $P_t^{(n)} = \frac{E_t[D_{t+n}]}{E_t\left(R_{t+1}^{(n)} R_{t+2}^{(n-1)} ... R_{t+n}^{(1)}\right)}$ follows from the fact that the only payment from a dividend strip is the payment at maturity. The cumulative discount rate $R_{t,n}$ is therefore the expected hold-to-maturity return $E_t[D_{t+n}]/P_t^{(n)}$, which itself can be expressed as the expected product of one period returns.

(ii) If the joint distribution of $R_{t+1}^{(n)}$, $R_{t+2}^{(n-1)}$,...,$R_{t+n}^{(1)}$ *is conditionally log-normal, then* $a_t^{(n)} = exp\left(\sum_{i=1}^{n-1}\sum_{k=1}^{n-i} cov_t\left(ln\left(R_{t+i}^{(n+1-i)}\right), ln\left(R_{t+i+k}^{(n+1-i-k)}\right)\right)\right)$.

(iii) Without any restrictions on the joint distribution of $R_{t+1}^{(n)}$, $R_{t+2}^{(n-1)}$,...,$R_{t+n}^{(1)}$, *then* $a_t^{(n)}$ *is a function of the variances, covariances and higher-order moments of the one-period dividend strip returns.*

Proof: All proofs are in Appendix A.

Moving to (b), we make the following assumption.

Assumption 1. *The expected gross return on a dividend strip is proportional to the expected gross return on the market*

$$E_t\left(R_{t+k}^{(n)}\right) = b_t^{(n)} E_t\left(R_{t+k}\right) \qquad where \qquad \sum_{n=1}^{\infty} w_t^{(n)} b_t^{(n)} = 1 \tag{5}$$

and $b_t^{(n)}$ *is a constant known at* t.

Assumption 1 rules out that the expected return on a given dividend strip changes over time in a way that is unrelated to changes in the expected return on the market (notice how there is no intercept in (1)).[4]

Under Lemma 1 and Assumption 1, we have the following result for the cumulative return on the nth dividend strip, which we denote "Condition R" (R for return).

Condition R.

$$E_t\left(R_{t+1}^{(n)} R_{t+2}^{(n-1)} ... R_{t+n}^{(1)}\right) = c_t^{(n)} E_t(R_{t+1}) E_t(R_{t+2}) ... E_t(R_{t+n}) \tag{6}$$

where $E_t(R_{t+k})$ *denotes the expected return on the overall stock market in year* $t+k$ *and* $c_t^{(n)}$ *is a constant known at* t *given by* $c_t^{(n)} = a_t^{(n)} b_t^{(n)}$.

[4] Assumption 1 allows for time-variation in expected returns, and expected returns for different maturities can update in a correlated fashion. Assumption 1 also allows for any slope of the equity term structure. The equity term structure is flat if $b_t^{(n)} = 1$ for all n, downward sloping if $b_t^{(n)} > b_t^{(m)}$ for $n < m$, and upward sloping if $b_t^{(n)} < b_t^{(m)}$ for $n < m$. van Binsbergen et al. (2012) provide evidence for a downward sloping equity term structure while Bansal et al. (2021) argue for an upward sloping equity term structure. Assumption 1 also allows for time series variation in the equity term structure slope (van Binsbergen et al. (2013), Gonçalves (2021b), Gormsen et al. (2021)) as $b_t^{(n)}$ can vary with both t and n.

Condition R implies that

$$P_t^{(n)} = \frac{E_t\left[D_{t+n}\right]}{E_t\left(R_{t+1}^{(n)} R_{t+2}^{(n-1)} ... R_{t+n}^{(1)}\right)} = \frac{E_t\left[D_{t+n}\right]}{c_t^{(n)} E_t(R_{t+1}) E_t(R_{t+2}) ... E_t(R_{t+n})} \tag{7}$$

where we can write the expected stock market return in terms of the riskless rate, $E_t(R_{t+k}^F)$ and the risk premium $E_t(R_{t+k}^{RP}) = \frac{E_t(R_{t+k})}{E_t(R_{t+k}^F)}$ (with F denoting riskfree and RP equity premium) as

$$E_t(R_{t+k}) = E_t(R_{t+k}^F) E_t(R_{t+k}^{RP}) \tag{8}$$

In terms of notation, R_{t+k}^F is a one-year real forward rate for year $t+k$. This forward rate will change over time. We use $E_t(R_{t+k}^F)$ to denote the forward rate one can get at date t for investing in year $t+k$ (this rate is known at t so one could drop the E_t and at a t subscript instead).

C. Decomposing capital gains and returns

Consider the capital gain over a period from t to $t+1$

$$\frac{P_{t+1}}{P_t} = \frac{P_{t+1}^{(1)}}{P_t} + \frac{P_{t+1}^{(2)}}{P_t} + ... = \frac{P_t^{(1)}}{P_t}\frac{P_{t+1}^{(1)}}{P_t^{(1)}} + \frac{P_t^{(2)}}{P_t}\frac{P_{t+1}^{(2)}}{P_t^{(2)}} + ... = w_t^{(1)}\frac{P_{t+1}^{(1)}}{P_t^{(1)}} + w_t^{(2)}\frac{P_{t+1}^{(2)}}{P_t^{(2)}} + ... \tag{9}$$

Expressing the capital gain on the market as in (9) leads to straightforward decompositions because one simply needs to keep track of how the value of "generic" dividend strips (i.e., dividend strips that pay off a certain number of periods out, as opposed to on particular calendar dates) evolve over time. For example, the factor $\frac{P_{t+1}^{(1)}}{P_t^{(1)}}$ captures how the value of a dividend strip that pays off one year out differs from date t to $t+1$.[5]

Using (9), we can express the one-period capital gain as follows.

Result 1 (**One-period capital gain**). *If Condition R holds, then the (gross) capital gain resulting from changes in expected dividends, real forward rates and forward equity risk premia over one*

[5] An alternative way to write the capital gain in (9) is as

$$\frac{P_{t+1}}{P_t} = \frac{P_{t+1}^{(1)}}{P_t} + \frac{P_{t+1}^{(2)}}{P_t} + ... = \frac{P_t^{(2)}}{P_t}\frac{P_{t+1}^{(1)}}{P_t^{(2)}} + \frac{P_t^{(3)}}{P_t}\frac{P_{t+1}^{(2)}}{P_t^{(3)}} + ... = w_t^{(2)}\frac{P_{t+1}^{(1)}}{P_t^{(2)}} + w_t^{(3)}\frac{P_{t+1}^{(2)}}{P_t^{(3)}} + ... \tag{10}$$

This has the interpretation of tracking the capital gains on a set of investments in dividend strips at date t, with each dividend strip aging by a year by date $t+1$. This approach is much more cumbersome to use for decompositions.

period is:

$$\frac{P_{t+1}}{P_t} = w_t^{(1)} \frac{G^D_{1,t+1}}{G^F_{1,t+1} G^{RP}_{1,t+1}} + w_t^{(2)} \frac{G^D_{2,t+1}}{G^F_{1,t+1} G^F_{2,t+1} G^{RP}_{1,t+1} G^{RP}_{2,t+1}} + ... \tag{11}$$

where for $k = 1, 2, ...$

$$G^D_{k,t+1} = \frac{E_{t+1}\left[D_{t+k+1}\right]}{E_t\left[D_{t+k}\right]}, \qquad G^F_{k,t+1} = \frac{E_{t+1}\left[R^F_{t+k+1}\right]}{E_t\left[R^F_{t+k}\right]}, \qquad G^{RP}_{k,t+1} = \frac{E_{t+1}\left[R^{RP}_{t+k+1}\right]}{E_t\left[R^{RP}_{t+k}\right]} \tag{12}$$

This one-period capital gain can approximately be decomposed as:

$$\begin{aligned} \frac{P_{t+1}}{P_t} \approx 1 \quad &+ \quad \left[w_t^{(1)} \frac{1}{G^F_{1,t+1}} + w_t^{(2)} \frac{1}{G^F_{1,t+1} G^F_{2,t+1}} + ... - 1 \right] \\ &+ \quad \left[w_t^{(1)} \frac{1}{G^{RP}_{1,t+1}} + w_t^{(2)} \frac{1}{G^{RP}_{1,t+1} G^{RP}_{2,t+1}} + ... - 1 \right] \\ &+ \quad \left[w_t^{(1)} G^D_{1,t+1} + w_t^{(2)} G^D_{2,t+1} + ... - 1 \right] \end{aligned} \tag{13}$$

with the G-factors defined in (12).

In (12), $G^D_{k,t}$ captures how the value of the k^{th} generic dividend strip evolves from date t to $t+1$. Similarly, $G^F_{k,t}$ and $G^{RP}_{k,t}$ capture how the k^{th} generic real forward rate (sometimes called the k^{th} constant maturity forward rate) and the kth generic forward equity premium changes from date t and $t+1$. In (13), the term omitted in the approximation captures interactions of the three types of news. To a first approximation it is zero (see the appendix).

Result 1 takes the c-factors from Condition R and (7) as constant over time such that discount rate news comes from movements in the expected return on the market, not from changes in the sensitivities of each generic dividend strips' expected return to the expected return on the market. Drawing a parallel to the CAPM, this is similar to modeling discount rate news for a given stock on a given day as being driven by the beta on that day times the change in the market risk premium, but not by changes in beta. It would be straightforward to add G-factors for changes to the c-factors, $G^C_{k,t+1} = \frac{c^{(k)}_{t+1}}{c^{(k)}_t}$, to the denominators in (11). As the literature evolves and researchers get more precise knowledge about the time-series evolution of the c-factors, this could be explored empirically as an extension.

To gain intuition for Result 1, it is informative to consider special cases of the capital gain decomposition where only one input changes. Suppose that only the generic real forward rate for year k forward changes, so only $G^F_{k,t}$ differs from one. Then the price elasticity with respect to the

k^{th} forward rate is:

$$\psi_{t,k}^{F} = \frac{P_{t+1}/P_t}{G_{k,t+1}^{F}} = w_t^{(1)} \frac{P_{t+1}^{(1)}/P_t^{(1)}}{G_{k,t+1}^{F}} + w_t^{(2)} \frac{P_{t+1}^{(2)}/P_t^{(2)}}{G_{k,t+1}^{F}} + ... \approx w_t^{(k)} + w_t^{(k+1)} + ... \tag{14}$$

This expression emerges because a change in the k^{th} generic forward rate has no effect on the value of generic dividend strips that pay off before the k^{th} year forward, while a change in the k^{th} generic forward rate shrinks the present value of each generic dividend strip that pay off the k^{th} year forward or later by the same factor,. Therefore, $\frac{P_{t+1}^{(n)}/P_t^{(n)}}{G_{k,t+1}^{F}}$ is zero for $n < k$ and one for $n \geq k$ from which (14) emerges. Similarly, the price elasticity with respect to the k^{th} generic forward equity premium is the same as $\psi_{t,k}^{F}$:

$$\psi_{t,k}^{RP} = \frac{P_{t+1}/P_t}{G_{k,t+1}^{RP}} \approx w_t^{(k)} + w_t^{(k+1)} + ... \tag{15}$$

Finally, the price elasticity with respect to the k^{th} generic expected dividend is:

$$\psi_{t,k}^{D} = \frac{P_{t+1}/P_t}{G_{k,t+1}^{D}} \approx w_t^{(k)} \tag{16}$$

$G_{k,t+1}^{D}$ is one plus the percentage change in the generic k^{th} expected dividend. This result thus states that the effect of a change in the generic k^{th} expected dividend leads to a larger capital gain if the k^{th} expected dividend initially was large and thus accounted for a larger fraction $w_t^{(k)}$ of the date t stock price.

Further clarifying Result 1, it is informative to express the G-factors as follows.

$$G_{k,t+1}^{D} = \frac{E_t\left[D_{t+k+1}\right]}{E_t\left[D_{t+k}\right]} \frac{E_{t+1}\left[D_{t+k+1}\right]}{E_t\left[D_{t+k+1}\right]} = G_{k,t+1}^{D,roll} G_{k,t+1}^{D,news} \tag{17}$$

$$G_{k,t+1}^{F} = \frac{E_t\left[R_{t+k+1}^{F}\right]}{E_t\left[R_{t+k}^{F}\right]} \frac{E_{t+1}\left[R_{t+k+1}^{F}\right]}{E_t\left[R_{t+k+1}^{F}\right]} = G_{k,t+1}^{F,roll} G_{k,t+1}^{F,news} \tag{18}$$

$$G_{k,t+1}^{RP} = \frac{E_t\left[R_{t+k+1}^{RP}\right]}{E_t\left[R_{t+k}^{RP}\right]} \frac{E_{t+1}\left[R_{t+k+1}^{RP}\right]}{E_t\left[R_{t+k+1}^{RP}\right]} = G_{k,t+1}^{RP,roll} G_{k,t+1}^{RP,news} \tag{19}$$

where the "roll" and "news" components refer to the first and second fractions in each equation. The "roll" component is the effect of moving one calendar period forward, implying that each object in the PV formula now refers to a later period. This component is known as of t. The "news" component is due to changing expectations from t to $t+1$ about a particular object for a

given time period.

The expected capital gain is approximately

$$\begin{aligned} E_t\left(\frac{P_{t+1}}{P_t}\right) \approx 1 \quad &+ \quad \left[w_t^{(1)}\frac{1}{G_{1,t+1}^{F,roll}} + w_t^{(2)}\frac{1}{G_{1,t+1}^{F,roll}G_{2,t+1}^{F,roll}} + ... - 1\right] \\ &+ \quad \left[w_t^{(1)}\frac{1}{G_{1,t+1}^{RP,roll}} + w_t^{(2)}\frac{1}{G_{1,t+1}^{RP,roll}G_{2,t+1}^{RP,roll}} + ... - 1\right] \\ &+ \quad \left[w_t^{(1)}G_{1,t+1}^{D,roll} + w_t^{(2)}G_{2,t+1}^{D,roll} + ... - 1\right] \end{aligned} \tag{20}$$

and the unexpected capital gain is approximately[6]

$$\begin{aligned} \frac{P_{t+1}}{P_t} - E_t\left(\frac{P_{t+1}}{P_t}\right) \quad &= \quad \left[w_t^{(1)}\frac{1}{G_{1,t+1}^{F,news}} + w_t^{(2)}\frac{1}{G_{1,t+1}^{F,news}G_{2,t+1}^{F,news}} + ... - 1\right] \\ &+ \quad \left[w_t^{(1)}\frac{1}{G_{1,t+1}^{RP,news}} + w_t^{(2)}\frac{1}{G_{1,t+1}^{RP,news}G_{2,t+1}^{RP,news}} + ... - 1\right] \\ &+ \quad \left[w_t^{(1)}G_{1,t+1}^{D,news} + w_t^{(2)}G_{2,t+1}^{D,news} + ... - 1\right] \end{aligned} \tag{21}$$

Turning to a realized (gross) return, we can express it either in terms of the capital gain and dividend yield, or in terms of the expected return plus the unexpected capital gain and unexpected dividend yield.

Result 2 (**One-period realized return**). *If Condition R holds, then the one-period (gross) return on the stock market can be decomposed as either (a) or (b) below.*

(a)

$$R_{t+1} = \frac{P_{t+1}}{P_t} + \frac{D_{t+1}}{P_t} \tag{22}$$

with $\frac{P_{t+1}}{P_t}$ *from equation (13).*

(b)

$$R_{t+1} = E_t\left(R_{t+1}\right) + \left[\frac{P_{t+1}}{P_t} - \frac{E_t\left(P_{t+1}\right)}{P_t}\right] + \left[\frac{D_{t+1}}{P_t} - \frac{E_t\left(D_{t+1}\right)}{P_t}\right] \tag{23}$$

with $\frac{P_{t+1}}{P_t} - E_t\left(\frac{P_{t+1}}{P_t}\right)$ *from equation (21).*

[6]We provide the proofs of the expressions for the expected and unexpected capital gain as part of the proof of Result 2 in the appendix.

C.1. A sub-period of a year

Result 1 and 2 hold for any period length. However, in practice one typically starts from inputs that are in years (e.g., constant maturity yields for maturities that have one-year increments). Therefore, we provide a decomposition result where t denotes years but we consider a sub-period of a year that is expressed in terms of such inputs. In our empirical applications, we will use this result to decompose quarterly returns as well as to illustrate the evolution of the realized stock return and its components within the year 2020.

With this motivation, consider a sub-period of a year going from a fraction s into the year to a fraction s' into the year, with $0 \leq s \leq 1$, $0 \leq s' \leq 1$, and $s' > s$. Denote these two points in time by $t+s$ and $t+s'$. Date t is thus the beginning of a calendar year, $t+s$ is a fraction s into the calendar year, and $t+s'$ is a later date in the same calendar year. We write the capital gain from $t+s$ to $t+s'$ as:

$$\frac{P_{t+s'}}{P_{t+s}} = \frac{P^{(1)}_{t+s'}}{P_{t+s}} + \frac{P^{(2)}_{t+s'}}{P_{t+s}} + ... = \frac{P^{(1)}_{t+s}}{P_{t+s}}\frac{P^{(1)}_{t+s'}}{P^{(1)}_{t+s}} + \frac{P^{(2)}_{t+s'}}{P_{t+s}}\frac{P^{(2)}_{t+s'}}{P^{(2)}_{t+s}} + ... = w^{(1)}_{t+s}\frac{P^{(1)}_{t+s'}}{P^{(1)}_{t+s}} + w^{(2)}_{t+s}\frac{P^{(2)}_{t+s'}}{P^{(2)}_{t+s}} + ... \quad (24)$$

where "(1)" refers to a dividend strip paying off one year from the date considered. This expression assumes that $P_{t+s'} = P^{(1)}_{t+s'} + P^{(2)}_{t+s'} + ...$ At any given point in time (not just at the end of each year, but also on a given day within the year), we thus implicitly assume that the next dividend is one year away. In practice, dividends are paid throughout a given year. The annual dividends in our formulas can be thought of as the future value of dividends to be paid over the next 1-year period, with the future value calculated as of one year out. Our formula with annual cash flows then holds not only at the start of each year but also on any day within the year.[7]

With then have the following result stating the decomposition for the sub-period.

Result 3 (**Capital gain for a sub-period of a year**). *If Condition R holds, then the (gross) capital gain resulting from changes in expected dividends, real yields and risk premia over a sub-period of a year is:*

$$\frac{P_{t+s'}}{P_{t+s}} = w^{(1)}_{t+s}\frac{G^D_{1,t+s'}}{G^F_{1,t+s'}G^{RP}_{1,t+s'}} + w^{(2)}_{t+s}\frac{G^D_{2,t+s'}}{G^F_{1,t+s'}G^F_{2,t+s'}G^{RP}_{1,t+s'}G^{RP}_{2,t+s'}} + ... \quad (25)$$

[7] We do not account for the fact that moving all dividends over a 1-year period out to the end of that period involves an interest rate to calculate the future value. This is an approximation that should have only a very small effect on our results. Specifically, at each point in time, this approach overstates the duration of the stock market by 1/2 year, which is very small relative to the very long duration of the market.

where for $k = 1, 2, ...$

$$G^D_{k,t+s'} = \frac{E_{t+s'}\left[D_{t+s'+k}\right]}{E_{t+s}\left[D_{t+s+k}\right]}, \qquad G^F_{k,t+s'} = \frac{E_{t+s'}R^F_{t+s',t+s'+1}}{E_{t+s}R^F_{t+s,t+s+1}}, \qquad G^{RP}_{k,t+s'} = \frac{E_{t+s'+1}R^{RP}_{t+s',t+s'+1}}{E_{t+s}R^{RP}_{t+s,t+1}}, \tag{26}$$

This sub-period capital gain can approximately be decomposed as:

$$\frac{P_{t+s'}}{P_{t+s}} \approx 1 \quad + \quad \left[w^{(1)}_{t+s}\frac{1}{G^F_{1,t+s'}} + w^{(2)}_{t+s}\frac{1}{G^F_{1,t+s'}G^F_{2,t+s'}} + ... - 1\right] \tag{27}$$

$$+ \quad \left[w^{(1)}_{t+s}\frac{1}{G^{RP}_{1,t+s'}} + w^{(2)}_{t+s}\frac{1}{G^{RP}_{1,t+s'}G^{RP}_{2,t+s'}} + ... - 1\right] \tag{28}$$

$$+ \quad \left[w^{(1)}_{t+s}G^D_{1,t+s'} + w^{(2)}_{t+s}G^D_{2,t+s'} + ... - 1\right] \tag{29}$$

with the G-factors defined in (26).

C.2. Instantaneous price change

An interesting limit of Result 3 emerges as s' goes to s. Then Result 3 simply states the effect of instantaneous changes to expected cash flows, real yields or the equity risk premium, similar to how duration analysis for bonds allows one to assess the capital gain on a bond resulting from an instantaneous shift in the yield curve. As s' goes to s, the expected (gross) capital gain is approximately $E_t\left(\frac{P_{t+s'}}{P_{t+s}}\right) = 1$ and the capital gain in Result 3 is the unexpected capital gain from instantaneous changes to the inputs.

III. Implementing the stock return decomposition with observables

In this section, we layout how to implement our return decomposition utilizing observable data. First, we address the issue that the main inputs (expected returns and expected dividends) are not available to infinite maturities. Second, we show how to use dividend futures prices to attain a dividend weights for all maturities. Finally, we set out the available observable data for implementation on the S&P index, which is the index we focus on for applications.

A. *Finite observable data maturities*

In practice, data for real forward rates, forward equity risk premia and cashflows is not observed to infinite maturities (we detail data in detail in Subsection D). This limitation on data availability means that in empirical implementations one sets the G-factors past available data horizons to 1 and calculate cash flow news and discount rate news past available data horizons as the difference between the observed capital gain and the observed discount rate return components. Equation (13) is then expressed as

$$\frac{P_{t+1}}{P_t} \approx 1 \quad + \underbrace{\left[\sum_{i=1}^{\infty} w_t^{(i)} \frac{1}{\prod_{k=1}^{i} G_{k,t+1}^{F}} - 1\right]}_{\text{Yield curve return component}_{t+1}} + \underbrace{\left[\sum_{i=1}^{\infty} w_t^{(i)} \frac{1}{\prod_{k=1}^{i} G_{k,t+1}^{RP}} - 1\right]}_{\text{Equity risk premium return component}_{t+1}}$$

$$+ \quad \text{Cashflow and long-term discount rate return component}_{t+1} \tag{30}$$

where $G_{k,t+1}^{Z} = 1$ for $k > N_Z$, with $Z = \{F, RP\}$ and N_Z denoting the maximum horizon of observed real risk-free rates and equity risk premium respectively. The cash flow and long-term discount rate return component is then calculated as

$$\left(\frac{P_{t+1}}{P_t} - 1\right) - \text{yield curve return component}_{t+1} - \text{equity risk premium return component}_{t+1}.$$

B. *Estimating dividend strip weights*

In implementing Result 1-3, a central element is the dividend strip weights $w_t^{(n)}$. It is well known that dividend strips (which are not traded) can be valued from dividend futures (e.g. van Binsbergen et al. (2013)). Since dividend futures pay off at maturity $(t+n)$, dividend strips and dividend futures prices are related by

$$P_t^{(n)} = F_{n,t} / \left(1 + y_{t,n}^{\text{nom}}\right)^n \tag{31}$$

where $F_{n,t}$ denotes the date t price of a dividend future paying the nominal dividend for period $t+n$ at $t+n$ and $y_{t,n}^{\text{nom}}$ is the riskless nominal yield at date t for a n-period investment. In this expression, $F_{n,t}$ is nominal (since actual dividend futures contracts pay the nominal dividend) and

therefore they are discounted using the nominal yield $y_{t,n}^{\text{nom}}$. Using the dividend futures prices, we can then express the dividend strip weights as

$$w_t^{(n)} = \frac{P_t^{(n)}}{P_t} = \frac{F_{k,t} / \left(1 + y_{t,k}^{\text{nom}}\right)^k}{P_t}. \tag{32}$$

This highlights how Result 3 is easily mapped to data when dividend future prices are available.

In practice, dividend futures are not traded on the S&P500 index out to an infinite maturity, and are instead observed to a maximum maturity of N_D years. The total value of long-term dividends beyond the maximum maturity is

$$L_t = P_t - \sum_{k=1}^{N_D} P_t^{(n)}, \tag{33}$$

and thus the total weight of long-term dividends as a fraction of the stock market value

$$w_t^{(L)} = \sum_{k=N_{D+1}}^{\infty} w_t^{(k)} = \frac{L_t}{P_t}, \tag{34}$$

is observed. However, the individual dividend strip weights beyond year N_D that sum to $w_t^{(L)}$ are unobservable and assumptions are therefore needed to estimate these dividend strip weights. We assume a Gordon growth model for dividends beyond year $t + N_D$, with G_t^L denoting the expected annual gross dividend growth rate past year $t + N_D$ and R_t^L denoting the expected annual gross return past year $t + N_D$. Both rates are, as of time t, assumed to be constant in expectation across all periods past year $t + N_D$. Under this assumption, we have the following result.

Result 4 (**Long-maturity dividend strip weights estimates**)**.**
Assuming a Gordon growth model for dividends beyond year N_D, and using available data on dividend futures, dividend weights beyond N_D years in maturity can be expressed as:

$$w_t^{(k)} = w_t^{(N_D)} \left(\frac{G_t^L}{R_t^L}\right)^{k-N_D} \qquad \text{for } k > N_D \tag{35}$$

Furthermore, $L_t = P_t^{(N_D)} \left(\frac{G_t^L}{R_t^L - G_t^L}\right)$ so the ratio of the expected long-run dividend growth rate to the expected long-run return

$$\frac{G_t^L}{R_t^L} = \frac{1}{1 + \left(P_t^{(N_D)} / L_t\right)} \tag{36}$$

can be calculated from the observed values of long-term dividend strips and the price of dividend strip at the longest traded maturity.

Under constant expected dividend growth and constant expected returns, $P = D\frac{G}{R-G}$ and thus $\frac{G}{R} = \frac{1}{1+(D/P)}$. Therefore, the market implied value of $\frac{G_t^L}{R_t^L}$ calculated in equation (36) reveals the market's view as of date t of what the dividend-price ratio is likely to be N_D years from date t (and all years subsequent to N_D years).

C. Share repurchases and issuance

Our decomposition is based on price elasticities with respect to expected returns and expected dividends. As we have shown, these elasticities are calculated from dividend strip weights. Is this correct, even in cases with share repurchases or share issuance? The answer is yes, with the following clarification.

As is well known, the stock price per share is the present value of dividends per share even in the presence of share repurchases/issuance. Furthermore, the S&P500 dividend futures from which we calculate the dividend strip weights are adjusted for share issues/repurchases so the contracts always reflect the dividends per share at the time the contract pays off.[8]

The subtlety is that if a company is expected to repurchase/issue shares (or repurchase/issue debt), then changes to expected returns (the cost of capital) may have not only a discount rate effect but also an effect on expected future dividends. In our decomposition, any such effects will enter the cash flow news component, as they should, but it is nonetheless relevant for thinking about the underlying drivers of cash flow news whether such effects are present.

Whether discount rate effects on expected future dividends emerge depends on whether repurchase/issuance activity affects dividends. In Appendix B, we show that there will be no discount rate effect on expected future dividends if repurchases are funded by debt issuance while there will be such effects if firms are expected to fund share repurchases/issuance by changing dividends. Asness et al. (2018) document that in the last few years leading up to their publication, net repurchases were substantial for firms in the Russell 3000 but the were roughly equal to net debt issuance. This suggests that any effects of discount rates on expected future dividends due to repurchases are likely to be modest.

[8] See https://www.spglobal.com/spdji/en/documents/additional-material/faq-sp-500-dividend-points-index.pdf

D. Data summary

Before turning to applications, this section sets out the data availability for implementation of Results 1-3 using observable inputs on the S&P500 index. Table I provides a summary of the data sources, available horizons and available sample periods, and Figure 1 plots the time series of the key inputs over the sample period 2004-2022, which is where all the main inputs are simultaneously available. The subsections below include explanations of each data component in more detail.

D.1. Realized return and realized dividend yield

Our starting point is the realized return we would like to understand. We focus on the realized S&P500 index return using the total return index provided by Bloomberg (SPXT Index). This index includes returns due to realized dividends. The realized dividend yield is then computed as the return difference between the total return index and capital gain index (also provided by Bloomberg - SPX Index) over the period in question. The bottom right panel of Figure 1 graphs the annual dividend yield based from this calculation, each day computing the total return differential between the two index over the proceeding 365 days.

D.2. Dividend strip weight data

As discussed in Section II.B, a central element for implementing Result 3 is the dividend strip weights $w_t^{(n)} = P_t^{(n)}/P_t$. We use S&P500 dividend futures prices adjusted for the risk-free rate for the numerator, using equation (31), and overall value of the S&P500 index for the denominator. Daily dividend futures prices are obtained from Bloomberg and are available from 2016 (for maturity 1-year through to 5-year) and from 2017 (for maturity 6-year through to 10-year). Proprietary dividend futures data from various banks have previously been used in the literature (starting with van Binsbergen et al. (2013)) and provide dividend futures prices back to the early 2000s. However, this earlier data is not widely available.

Dividend strip values can also be inferred from option data in the absence of arbitrage opportunities (van Binsbergen et al. (2012)). The put-call parity relationship for European call options leads to:

$$P_t^{(n)} = p_{t,n} - c_{t,n} + P_t - Xe^{r_{t,n}^f} \tag{37}$$

where $p_{t,n}$ is the price of a put option with a strike price X and maturity $t+n$, and $c_{t,n}$ is the

price of a call option with strike X and maturity $t+n$. We use daily data from OptionMetrics, available from 1996. Minute-level data is available from CBOE. The maximum maturity of equity options has increased over the sample period, starting at 2 years and increasing to 3 years in 2005 and to 5 years in 2021.

When implementing our return decomposition prior to year 2017, we use the (van Binsbergen et al. (2012)) approach to calculate dividend strip weights. Due to availability of options data, these only go out to two years, corresponding to $N_D = 2$ in Result 4. For both data inputs (dividend futures or option prices), on each date we use linear interpolation across available maturities on the traded assets in order to calculate constant maturity dividend strip weights (this is a standard approach in the literature: see van Binsbergen et al. (2012), van Binsbergen et al. (2013), Martin (2017), van Binsbergen and Koijen (2017), Gormsen et al. (2021). Figure 1 top left graph presents the 1-year dividend weight and sum of dividend weights from year 1 to year 10 over the sample period 2004-2022, marking with a vertical line where we shift from option implied dividend weights to dividend future weights. The option-implied dividend weights require implementation of Result 4 to compute dividend weights from maturity 3-10.

D.3. Real yield curve data

As a baseline approach we propose calculating real interest rates as the nominal interest rate on interest rate swaps (the fixed rate in a fixed-for-floating interest rate swap) minus the rate on inflation swaps. Both series are available out to 30 years, with data obtained from Bloomberg. We convert the term-structure of real zero-coupon yields into annual forward real risk-free rates enabling us to implement Result 1-3 in which the expected return for a given future year is changed. For robustness, we consider real interest rates calculated from nominal Treasury rates minus inflation swaps, with nominal Treasury zero-coupon yields obtained from the Federal Reserve.[9] We also consider real interest rates from TIPS, with data from FRED. Both alternatives are also available to the 30-year maturity.

Observability of real interest rates by either of the three methods is a relatively recent phenomenon. Inflation swap data is available from Bloomberg beginning in July 2004 and TIPS yields from FRED start in 2003. The inflation swap series has a maximum maturity of at least 30-years throughout the sample period, and from April 2008 to January 2021 the maximum maturity

[9] https://www.federalreserve.gov/data/nominal-yield-curve.htm

increased to 40-years.

D.4. Equity risk premium data: The Martin lower bound

For the equity risk premium we use the methodology of Martin (2017) who calculates a lower bound on the risk premium using prices of stock market index options and argues that this lower bound is approximately equal to the true risk premium. Martin's data covers the period 1996-2012. We extend his series to June 2022 using data from OptionMetrics. We are able to almost exactly replicate Martin's series over his sample period. The description below provides a recap of Martin's approach and Appendix C details our data construction.

Martin (2017) starts from the fact that the time t price of a claim to a cash flow X_T at time T can either be expressed using the stochastic discount factor M_T or using risk-neutral notation

$$\text{Price}_t = E_t\left(M_T X_T\right) = \frac{1}{R_{f,t}} E_t^*\left(X_T\right)$$

where the expectation E_t^* is defined by

$$E_t^*\left(X_T\right) = E_t\left(R_{f,t} M_T X_T\right).$$

The return on an investment can similarly be written in terms of the SDF or using risk-neutral notation

$$1 = E_t\left(M_T R_T\right) = \frac{1}{R_{f,t}} E_t\left(R_{f,t} M_T R_T\right) = \frac{1}{R_{f,t}} E_t^*\left(R_T\right).$$

The conditional risk-neutral variance can be expressed as

$$\text{var}_t^* R_T = E_t^* R_T^2 - \left(E_t^* R_T\right)^2 = R_{f,t} E_t\left(M_T R_T^2\right) - R_{f,t}^2$$

The risk premium expressed as a function of the risk-neutral variance is then

$$\begin{aligned} E_t R_T - R_{f,t} &= \left[E_t\left(M_T R_T^2\right) - R_{f,t}\right] - \left[E_t\left(M_T R_T^2\right) - E_t R_T\right] \\ &= \frac{1}{R_{f,t}} \text{var}_t^* R_T - cov_t\left(M_T R_T, R_T\right) \\ &\geq \frac{1}{R_{f,t}} \text{var}_t^* R_T \text{ if } cov_t\left(M_T R_T, R_T\right) \leq 0 \end{aligned}$$

Thus $\frac{1}{R_{f,t}}\text{var}_t^* R_T$ provides a lower bound on $E_t R_T - R_{f,t}$ if $cov_t\left(M_T R_T, R_T\right) \leq 0$, denoted the "negative correlation condition" (NCC).

Martin (2017) shows that the lower bound $\frac{1}{R_{f,t}}\text{var}_t^* R_T$ can be calculated from put and call option prices as follows

$$\frac{1}{R_{f,t}}\text{var}_t^* R_T = \frac{2}{S_t^2}\left[\int_0^{F_{t,T}} \text{put}_{t,T}\left(K\right) dK + \int_{F_{t,T}}^{\infty} \text{call}_{t,T}\left(K\right) dK\right] \tag{38}$$

where S_t is the stock price at t, $F_{t,T} = E_t^*\left(S_T\right)$ is the forward stock price, and K denotes the option strike price. On any date, it is therefore possible to extract a lower bound estimate for each available maturity of expiring options. Consistent with Martin (2017), we use linear interpolation to calculate constant maturity lower bounds, which post-2006 allows estimates out to about two years and 6 months.[10] It is also important to note that equity index options out to 5 years have been traded since 2021 and will provide additional data for calculating the Martin lower bound going forward.

We provide novel support for the use of the Martin lower bound as a measure of the equity risk premium by comparing it to equity premium estimates of asset managers. Dahlquist and Ibert (2021) collect a dataset of equity risk premium estimates of asset managers (and some of the largest investment consultants). The data is based on asset managers' capital market assumptions, posted publicly or provided to clients. For the period 2005-2020, their data has 541 observations of equity risk premia from a total of 45 asset managers such as J.P. Morgan, BlackRock, Franklin Templeton and AQR. The most typical horizons in their dataset are 10 years (45% of the data), 7 years (13%), and 5 years (10%). Dahlquist and Ibert (2021) show that asset managers' expectations appear a lot more rational than expectations of households. Asset managers' perceived equity risk premium is high when the P/E ratio is low – consistent with statistical predictive relations – whereas the literature on behavioral finance tends to find that households' perceived equity risk premium is high when the P/E ratio is high.

Figure 2, Panel A, graphs the 1-year Martin measure and the 10-year asset manager equity premia, with one graph for each asset manager that contributes at least five data points for this investment horizon. The equity risk premium estimates of various asset managers are highly correlated with the Martin (2017) lower bound of the equity risk premium that is computed from

[10] the 3-year maturity is issued once every year in December, meaning that the maximum maturity declines from 3-years to 2-years each calendar year.

option prices. In terms of levels, there are substantial cross-asset manager differences in perceived equity premium levels. We construct an overall asset manager equity premium series (for the 10-year horizon) by taking out asset manager fixed effects. We estimate the following relation

$$EP^{10}_{m,t} = \alpha + \Sigma^{T}_{t=1}\beta_t D(\text{date} = t) + \Sigma^{M}_{m=1}\delta_m D(\text{manager} = m) + u_{m,t} \tag{39}$$

where m denotes manager and calculate the predicted value, excluding manager fixed effects, for each date t, $\alpha + \Sigma^{T}_{t=1}\beta_t D(\text{date} = t)$. Figure 2, Panel B, graphs the resulting asset manager 10-year equity premium series along with the 1-year Martin measure. The correlation is high at 0.72.[11] The high correlation lends credence to the Martin measure and also suggests that asset manager equity premium perceptions are consistent with actual asset prices. In Section V.F, we will exploit the fact that the asset manager equity premia are available for a longer maturity than the Martin series to assess the effect of equity premium news past year 2 for our stock return decomposition for 2020.

The Appendix contains further analysis of the validity of the Martin lower bound as a risk premium estimate. We update Martin's results that predict the realized risk premium with the lower bound and provide theoretical analysis of the how the change in the bound relates to the change in the true equity risk premium.

D.5. Near-term dividends data

The relation between dividend futures and dividend strip prices was discussed in Section II.B with:

$$\frac{F_{n,t}}{\left(1 + y^{\text{nom}}_{n,t}\right)^n} = P^{(n)}_t = \frac{E_t\left[D_{t+n}\right]}{R_{t,n}}. \tag{40}$$

A change in a discounted dividend futures price therefore reflects either a change in expected dividends or a change in the cumulative discount rate. We can rearrange to express the expected real dividend as a function of the cumulative expected real return and the dividend strip price:

$$E_t D_{t+n} = R_{t,n} P^{(n)}_t \tag{41}$$

[11] Regressing the 10-year asset manager series on the 1-year Martin measure results in a regression coefficient of 0.41 with a t-statistic of 9.6.

For years 1 and 2 we can estimate $R_{t,n}$ as the real yield plus the Martin lower bound for the equity risk premium under the assumption (which is stronger than Assumption 1) that the first two dividend strips have $b_t^{(n)} = 1$.[12] If the Martin lower bound is tight, (41) will give the expected real dividends. To the extent the Martin lower bound is not tight, (41) will provide a lower bound on expected real dividends.

Because (41) can only be implemented out to year 2, we do not separate out the return component resulting from changes in $E_t D_{t+1}$ and $E_t D_{t+2}$ in our decomposition. This component is simply included in the component resulting from changes in any dividends and from changes to real yields or risk premia past the horizons of observable data. Instead, we use the time series for $E_t D_{t+1}$ and $E_t D_{t+2}$ to calculate the dividend growth factors capturing changes in the first two generic dividends

$$G^D_{1,t+1} = \frac{E_{t+1}\left[D_{t+2}\right]}{E_t\left[D_{t+1}\right]}, \qquad G^D_{2,t+1} = \frac{E_{t+1}\left[D_{t+3}\right]}{E_t\left[D_{t+2}\right]}.$$

We will graph these to series to help assess whether the cash flow and long-term discount rate return component in a given period is likely to be dominated by cash flow news or not.

IV. Results: Decomposition of annual and quarterly returns

We illustrate our new return decomposition methodology with an application to realized returns on the S&P500 total return index (i.e., including dividends). Our methodology can be applied from 2004Q4 to present. The last quarter of our data set is 2022Q2. As noted, dividend strip weights are based on futures data from 2017, but must be calculated from equity options prices before that. Figure 1 shows the decomposition of annual returns at the top and the decomposition of quarterly returns at the bottom. We provide results from approach (a) in Result 2. The results for 2022 are based on the first half of the year.

Figure 3 illustrates the actual S&P500 return using the black line and the four return components in the colored bars. The blue bars illustrate the component of the realized returns that is due to changes (roll and news) about real yields out to maturity 30 years. Green bars illustrate

[12] Gormsen et al. (2021) take a different approach to calculate the risk premium for each dividend strip. They express the expected return on dividend strip n as $\hat{R}_{t,n} = R^F_{t,n} + \frac{1}{R^F_{t,n}} cov^*_t\left(R_{t,n}, R^{(n)}_{t,n}\right)$. To estimate the risk-neutral covariance of the n-period return on n-period dividend strip with the n-period return on the market they estimate the risk-neutral standard deviation of the dividend strip return (which is available based on short-dated EuroStoxx dividend future options) and consider alternative assumptions for the correlation of the dividend strip return with the market return. They implement this on EuroStoxx data but dividend future options are not available for the US so we cannot use their approach for the US.

realized returns due to changes to equity risk premia out to 2 years. The realized returns from changes to expected cash flows and from changes to real yields or risk premia past the horizons with available data are illustrated in red. The final return component is the realized dividend yield, illustrated in orange. As one would expect, this last component is positive but small, contributing between 0.6% and 3% to annual returns.

The stock return component from observed real yield curve changes in blue is large and positive in crisis times (as real yields fell during those times). This component is 37% for 2008 (the financial crisis), 25% for 2011 (the European sovereign debt crisis), and 24% in 2020 (the COVID crisis). By contrast, the real yield curve component is large and negative during periods of monetary tightening, notably 2022 (-27%) and 2013 (-23%), which included the taper tantrum period in 2013Q2.

The component of realized stock returns due to changes to observed equity risk premia out to 2 years (green) is substantial and negative in crisis times, with the largest effect in annual data seen in 2008 at -11%. The quarterly data reveals a substantial negative equity risk premium component during 2010Q2 (the start of the Greek debt crisis), 2011Q3 (the spread of the Greek debt crisis to other European countries), and 2020Q1 (the onset of the COVID crisis). Because the equity risk premium component is based only on changes to equity risk premia for the next two years, the total risk premium component is likely substantially larger than the green bars shown. In our detailed analysis of 2020 below, we supplement the Martin equity risk premium data with the Dahlquist-Ibert equity risk premium estimates of asset managers which are available out to the 10-year horizon. For 2020Q1, we find that the stock return component based on equity risk premium changes out to year 10 is about 3 times larger than the stock return component based on equity risk premium changes out to year 2.

The red bars in Figure 3 capture the return component resulting from changes to expected cash flows as well as from changes to real yields and the equity risk premia past the horizons observed. This component is strongly negatively correlated with the real yield component, as is visible in the figures and the blue and red bars tend to be of opposite signs. The correlation is -78% in the annual data and -85% in quarterly data, shown in Table II. This makes it clear that changes to real yields past year 30 are not the dominant part of the red bars which are thus dominated by a mix of changes to expected dividends and changes to equity risk premia past year 2.

In periods with little change to equity risk premia, and thus small green bars, the red bars

are therefore likely to be substantially affected by cash flow changes. Year 2021 exemplifies such periods and the contrast between what drove strong stock market performance in years 2020 and 2021 is impressive: 2020 was dominated by capital gains resulting from falling real yields, while 2021 returns were boosted by capital gains from positive changes to expected cash flows. The latter is consistent with the narrative in the financial press that the high-inflation period after the reopening of the economy post-COVID led to large realized and expected corporate profits, despite the tight labor market and rising wages. This narrative is confirmed if we look at the first two dividend growth factors, $G^D_{1,t+1} = \frac{E_{t+1}[D_{t+2}]}{E_t[D_{t+1}]}$ and $G^D_{2,t+1} = \frac{E_{t+1}[D_{t+3}]}{E_t[D_{t+2}]}$ as described in section III.D.5. These are illustrated in the top graph in Figure 4, and are 10% (first dividend) and 14% (second dividend) for 2021. The first two real dividends account for only about 4% of the stock market based on the first two dividend strip weights. By itself, growth in these two real dividends therefore contributed less than 1/2% to the 2021 return, as illustrated in the bottom figure in Figure 4. However, if markets expected later real dividends to grow by the same factor, then expected real dividend changes could account for about 10-14% of the 29% stock return in 2021.

V. Results: Decomposition of returns over 2020

Figure 5 graphs the cumulative total return on the S&P500 index over the year 2020. The market ended the year up by 18.4%, with dramatic price moves throughout the year. At its lowest point on March 23rd 2020, the cumulative return year-to-date was negative 30.4%. However, the market quickly rebounded, and as of June 8th 2020 it had recovered all losses from the start of the year. The stock market rally then continued through the second half of the year.

A. Related literature on the stock return in 2020

An emerging literature is studying stock market movements in 2020.

Landier and Thesmar (2020) analyze analyst earnings forecasts up to May 2020. They estimate a counterfactual path for the stock market which assumes unchanged discount rates (and payout ratios) and uses dividend expectations updated daily based on updates to analyst forecasts for 2020-2022 earnings. The 2022 earnings are central for their terminal value calculation. They find that dividend news was modest, around -5% by March 23, 2020, and became more negative past March 23, 2020. This is in sharp contrast to the large crash and fast recovery of the actual stock market.

Cox, Greenwald, and Ludvigson (2020) study the COVID crisis using the estimated structural asset pricing model of Greenwald, Lettau, and Ludvigson (2019) in which the value of the stock market is expressed as GDP $\times$ [corporate profits/GDP] $\times$ [stock market value/corporate profits]. They also conclude that it is difficult to explain the V-shaped trajectory of the stock market over the COVID crisis with cash flow news. A central argument is that, based on data from the Survey of Professional Forecasters as of May 2020, GDP was expected to fall by about 10% in 2020Q2, but was expected to increase in 2020Q3. The COVID shock to GDP was thus expected to be quite transitory, implying that GDP (and thus earnings) changes can explain only a small fraction of the crash and recovery.

Using the method of Cieslak and Pang (2021), Cieslak estimates (in presentation slides for their paper posted on her webpage) that growth news account for a stock market drop of about 10% out to March 23. Gormsen and Koijen (2020) study dividend futures. They show that changes to the value of dividends out to year 7 can account for little of the stock market crash (given their modest weight in the market and the realized decline in dividend futures values) and none of the recovery up to July 20, 2020. They argue that longer-maturity dividends are likely to be only modestly affected by the COVID crisis, implying that changes to their present value and thus to the overall market may have been driven mostly by discount rate news.

We infer from these papers that, to the extent it is possible to measure cash flow news, such news does not appear able to explain the majority of the stock market decline or recovery in 2020. Is this due to the difficulty of measuring expected cash flows, or can market movements instead be explained by discount rate news? Our decomposition allows us to shed light on this question by taking what is, in essence, the opposite approach of that taken by prior work on 2020. We try to measure discount rate news rather than cash flow news. The focus on discount rates also allows us to provide a more granular decomposition of discount rate effects by distinguishing between the impact of real yield curve news versus the impact of equity risk premium news.

B. Dividend strip weights and elasticities in 2020

To illustrate the relation between dividend strip weights and stock price elasticities, Figure 6 plots the dividend strip weights on the first day of 2020 in the top left panel, and the cumulative sum of weights in the top right panel. Dividends up to 10 (30) years had a combined weight of 17% (44%) in the total stock market value on this date. This highlights the large fraction of stock

market value that is generated from long-maturity dividends, and thus the long duration of the stock market overall (a point also made in van Binsbergen (2020) and Gonçalves (2021a)). As of the start of 2020, the duration of the stock market was 46 years based on the dividend strip weights in Figure 6.

Figure 6 plots the dividend elasticity of the stock price in the bottom left panel and the return elasticity of the stock price in the bottom right panel. These figures show how percentage changes in expected dividends and expected (gross) returns at various maturities affect the stock price and thus the realized market return. The dividend elasticity of the stock price is the same as the dividend strip weight in the panel above. The return elasticity of the stock price is $\psi^R_{t,k} = -\sum_{n=k}^{\infty} w_t^{(n)} = -\left(1 - \sum_{n=1}^{k} w_t^{(n-1)}\right) = \sum_{n=1}^{k} w_t^{(n-1)} - 1$ and therefore equals the value in the figure above, minus one. Increases in expected future returns lead to negative capital gains.

Our dividend strip weights and elasticities are based on dividend futures prices. These are a relatively new product and liquidity may be an issue. In Appendix E, we analyze the robustness of our main results to the liquidity of the contracts. We also consider the sensitivity of results to the maximum maturity of observed dividend futures used in the implementation of Result 4.

C. Yield curve return component in 2020

Figure 7, top graph, shows the evolution of the 10-year and 30-year real rates estimated from interest rate swaps and inflation swaps. Real yields fall dramatically over 2020, with a 114 bps decline in the 10-year real rate and an 80 bps decline in the 30-year real rate. Figure 8 top panel illustrates the role of interest rate swaps (the nominal interest rate) and inflation swaps (capturing expected inflation) separately. Both fell dramatically in at the onset of the COVID-crisis, with a larger drop in the nominal rate up to March 9 driving the decline in real rates up to this point. After a short-lived spike in nominal yields from March 9 to 18, nominal yields fall and do not recover fully over the year. By contrast, inflation swap rates start increasing in late March, resulting in falling real rates up to August. As a side note, it is interesting to observe that financial markets initially thought the COVID crisis would reduce inflation while in fact it turned out to be inflationary (perhaps due to stronger than expected fiscal and monetary stimulus).

The bottom panel of Figure 7 shows how changes in real yields affected the stock market over the year via real yield curve changes based on our decomposition. The figure is based on our baseline approach of using changes in real yields out to 30 year maturity based on interest rate

swaps and inflation swaps. The figure shows that falling real yields had a very important role in the rise in the stock market in 2020. In fact, the yield curve return component of the S&P500 return, taken in isolation, explains a 23.7% increase in the stock market (that increased 18.4% overall). Given that the stock market is a long-duration asset, the large fall in long-term real yields had a big impact on its price level.

Figure 8 considers robustness to using the two alternative approaches to calculate real rates. The decline in the 30 year real rate is around 80bps whether we compute real yields as the difference between interest rate swaps and inflation swaps, the difference between Treasury nominal yields and inflation swaps, or from inflation indexed Treasuries (TIPS). The real yields are about 50bps lower in terms of levels when measured based on interest rates swaps and inflation swaps. Klingler and Sundaresan (2019) argue that low rates on interest rate swaps is due to certain pension funds preferring to get interest rate exposure off-balance sheet. However, the level difference is not important for our results given that it is the changes (and not levels) of these yields that generate the real yield curve change component of stock returns. Indeed, the bottom right panel of Figure 8 shows that the stock return component from changes to the real yield curve is broadly similar for the year whether we use interest rate swaps or Treasury yields for nominal yield component of our real yields (and using real yields out to the 30-year maturity).

In Figure 8, it is noticeable that the decline in real yields is interrupted by a sharp spike in real long yields from March 9 to March 18. The spike is particularly dramatic for real yields based on nominal Treasuries or TIPS. Vissing-Jorgensen (2021) and He, Nagel, and Song (2022) study this spike which led to Federal Reserve purchases of over $1T of Treasuries in 2020Q1 in order to stabilize Treasury markets. The spike is associated with heavy selling by bond mutual funds, foreign central banks and hedge funds and reverses as the Federal Reserve increases its daily purchases sharply starting on March 19. Although there was price-pressure pass-through into interest rate swaps, the spike in these were less dramatic (perhaps due to derivatives absorbing less balance sheet space than holding Treasuries outright). It is possible that the March spike in real yields is disconnected from the stock market in the sense that stock market investors viewed it as driven by urgent liquidity needs that may have been perceived as less relevant for the fundamental value of stocks. If so, our real yield curve news component based on interest rate swaps will be the most accurate but we will nonetheless overestimate the negative return effect of the spike on realized stock returns in March 2020. This issue will not affect our decomposition for

the full year, nor our assessment of the role of the risk premium for the crash, nor our estimate of the role of real rates outside of the March period.

D. Equity risk premium return component in 2020

The top panel of Figure 9 shows the equity risk premia for the first two years while the bottom panel shows how changes in these equity risk premia impacted the stock return over 2020. The upward spike in equity risk premia in March generates a negative realized return effect which accounts for minus 12.1 percentage point of the realized return of minus 25.4 percent up to March 18. The equity risk premium component of the stock return recovers somewhat from the height on crisis, but still ends the year negative, contributing -3 percentage points to the overall 2020 stock return.

Providing more detail on the evolution of the equity risk premium, Figure 10 top panel shows our estimated equity risk premia (the lower bounds) over 2020 using three constant-maturity series (10-day, 1-month and 1-year). All risk premia shown in the figure are annualized. While the 10-day equity risk premium peaked closed to 100 percent in March, the 1-year equity risk premium increased from around 3 percent at the start of the year to just above 15 percent on March 18. The much larger spike to shorter-maturity equity risk premia suggests that investors perceived some of the risk from COVID to resolve fairly quickly.

As a supplementary way to describe the term structure of equity risk premia, the bottom panel in Figure 10 graphs the cumulative equity risk premium by maturity for the beginning and end of the year as well as for March 18, they day risk premia peaks. Higher annualized risk premia at shorter horizons translate in to a concave cumulative equity risk premium. At the peak of the crisis, investors required a risk premium of 3.9 percent to invest for a 30-day period and a risk premium of 15.3 percent to invest over the next year.

In Section V.F we explore whether the equity risk premium moved past year two and the implication of such changes for our decomposition results.

E. Total observable return component and total cash-flow/long-term discount rate return component in 2020

To summarize our results for 2020 so far, the total return on index over 2020 was 18.4%, of which the realized dividend yield contributed +2.1%. The decline in real riskless rates is central for

understanding the market's impressive performance for the year, with the yield curve component of the S&P500 return generating a plus 23.7 percent return for 2020 as a whole. A spike in equity risk premia out to two years accounts for minus 12.1 percentage points of the realized return of minus 25.4 percent up to March 18, and for -3 percentage points of realized returns over the full year.

Figure 11 pulls all the results described above together to show the total return generated by observables in the top panel and the total return from changes to expected dividends and long-term discount rates the bottom panel. The latter series captures the effect of changes to dividend expectations and any changes to risk premia past year 2 and real riskless rates past year 30. It is small for the year as a whole, but large in the crisis, contributing about 40 percentage points to the crash and a roughly equal amount to the recovery. We turn to the drivers of this component in the next subsection.

F. What drove the cash-flow/long-term discount rate component in 2020?

F.1. Movement in expected cash flows

To understand the cash-flow/long-term discount rate component in 2020, we first consider the possible effect of cash flow changes for the realized stock return. The left graph in Figure 12 graphs a sharp fall in near-term dividend futures prices in the spring of 2020. The largest declines were in the 1st and 2nd year dividend futures prices, with less pronounced declines in 5-year and 10-year dividend futures prices. These changes in dividends futures prices are lower bounds of underlying changes in expected real dividends as equity risk premia increased and expected inflation decreased in the early part of the COVID crisis (see Figure 8 and Figure 10).

For the 1-year and 2-year maturities, we can compute a more precise estimate of expected real dividends as we have inflation and risk premium data available. The middle and right graphs in Figure 12 therefore explore the first 2-years of expected real dividends in more detail. The figure graphs the 1-year and 2-year expected real dividends which fell by roughly 30& and 35% respectively. Therefore, for cash flow changes to fully explain the long-term return component of 40%, investors would have had to expect these declines in cash flows to be permanent. This is contrary to the evidence, with the longer-term dividends futures prices falling by less than the near term prices (consistent with the argument from Cox, Greenwald, and Ludvigson (2020) that professional forecasters in expected the stock to be transitory, as of spring 2020). This suggests a

role for discount rates, which we consider next.

F.2. Movements in the equity risk premium past year 2

Figure 9, top, illustrates the time series for the equity premium for year 1 and the forward equity premium for year 2 (both annualized). The forward equity premium for year 2 moves up in March, though much less than the equity premium for year 1. Both remain substantially higher at the end of 2020 compared to their values at the start of 2020.

Given that there is some increase in the risk premium even for year 2, it is likely that risk premia increased to some extent even past this horizon. The asset manager expectations dataset of Dahlquist and Ibert (2021) provides us with longer-dated equity risk premium estimates and we use the 10-year asset manager equity premium series to assess the importance of equity premium movements past year 2. We estimate a Nelson and Siegel (1987) curve for the equity risk premium term-structure

$$EP_t(\tau) = \beta_{1t} + \beta_{2t}\left(\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau}\right) + \beta_{3t}\left(\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau} - e^{-\lambda_t\tau}\right) \tag{42}$$

where the parameter λ_t governs the rate of exponential decay in the term structure at date t, and, as shown by Diebold and Li (2006), the parameters β_{1t}, β_{2t} and β_{3t} are dynamic latent factors that can be interpreted as the level, slope and curvature of the curve.

To implement the Nelson-Siegel curve, we use the full cross section of Martin (2017) equity risk premium maturities greater than 1 month,[13] supplemented with the 10-year asset manager risk premium time series. We focus this exercise on quarter-end values because 2/3 of the asset manager data are as of the end of the quarter, making data for other dates less reliable.[14]

To fit the yield curve, we estimate the parameters $\theta_t = \{\beta_{1t}, \beta_{2t}, \beta_{3t}, \lambda_t\}$ by nonlinear least squares for each quarter end t. The top panel of Figure 13 shows the fitted yield curves at the end of 2019 and for each quarter end of 2020. As can be seen, we find that the equity risk premium term structure can be well fitted with the relatively parsimonious Nelson-Siegel curve. Furthermore, the curve is able to replicate a variety of yield curve shapes across dates. The bottom panel of Figure 13 shows the fit of the curve over time for the 1-year risk premium, the 2-year risk premium

[13] We remove the shortest maturity options so that the estimation better fits the long-run dynamics of the curve.

[14] This is partly due to Dahlquist and Ibert (2021)'s (entirely reasonable) assumption that data are end of the prior month when the asset manager only states a year and month and not an exact date.

and the 10-year risk premium.

Figure 14 shows our return decomposition in quarterly data accounting for equity premium changes out to year 10. For each year 1 to 10, we use the fitted value for the equity premium from the Nelson-Siegel estimation and translate these to annual forward equity risk premia. Accounting for equity premium changes out to year 10 results in a substantially larger (more negative) return component from equity risk premium changes in 2020Q1 (bottom left) and therefore a substantially smaller component from cash flow changes/remaining long-term discount rate changes. At the end of March, the cumulative effect of the risk premium news is -20.6% (compared to -7.2% in the baseline estimation). Risk premium news for year 1 to 10 can thus account for all of the decline in the stock market in 2020Q1 which amounts to -19.5%. The return component from cash flow changes/remaining long-term discount rate changes remains negative at the end of 2020Q1 because the real yield curve return component is substantially positive for 2020Q1. A negative but modest component from cash flow changes/remaining long-term discount rate changes up to the end of 2020Q1 is consistent with the prior cash-flow focused literature documenting some but modest changes to expected cash flows up to this point.

For the full calendar year of 2020, extending the risk premium has a smaller impact relative to 2020Q1 only, as risk premia mean-revert significantly after the crisis. The return component from equity risk premium changes is -7.5% (compared to -3.5% in the baseline estimation) and the component from cash flow/remaining long-term discount rate changes is 6.1%. Overall, introducing longer-term risk premium changes strengthen the role of equity risk premium changes for explaining the crash and immediate post-crash recovery of the stock market, but falling real yields remain the dominant reason the market ended the year up almost 20%.[15]

F.3. Movements in real rates beyond 30 years

Figure A.2 helps assess whether changes to real rates past year 30 are likely to be a substantial driver of the stock return component from cash flow changes/long-term discount rate changes. We graph the real (annualized) 10-year forward rates for each of the next 3 decades (and 4th decade when using real rates based on interest rate swap data in the first panel). The real forward rate

[15]It is important to recognize that our extension does not account for risk premium changes past year 10. The asset manager data-set has little data past year 10 and, inline with the spirit of the paper, we have chosen not to extend the risk premium curve beyond observed maturities for the results reported in Figure 14. The fitted Nelson-Siegel curve estimates parameters that in theory provide risk premia at all maturities. However, as the curve is extended past the observed data maturities, the validity the estimates becomes questionable.

for the 3rd and 4th decade from now fall over the year, though less than the real forward rates for the first decades. Given the decline in the longest observable real rates, it is possible that real rates changed somewhat even past year 30. Indeed, in the UK, inflation-indexed bonds are traded with 50-year maturity and we find that the real forward rate for years 31-50 fell about 40 bps for the year.

To understand how much real yield curve moves past year 30 could have affected the stock market over 2020, we consider to alternatives to the baseline results. First, we recalculate the yield curve return component using interest rate and inflation swap data out to 40 years.[16] Second, we use the average 2020 real yield growth rate, $G^F_{k,t+1}$, across horizons $k = 31$ to $k = 40$, and then recalculate the yield curve return component assuming this average long-horizon real yield growth rate, which is .9906, occurred across all horizons from 40-years out to 60-years (doubling the maximum maturity relative to the 30-year baseline). With the two extensions, yield curve news is 29.9 and 40.3 respectively, compared to the baseline 23.7 percent. To the extent that these ultra long-dated yields impact the stock return decomposition for 2020, it leaves a more important negative role for long-term risk premium and expected cashflows.

VI. Conclusion

The paper contributes to answering a core question in asset pricing: what is the role of discount rate news versus cash flow news. We focus on decomposing stock returns period by period and contribute two main ideas.

First, we provide a simple decomposition of the capital gain for a given period. This allows a decomposition of realized stock returns into components from (i) real yield curve changes, (ii) equity risk premium changes, (iii) expected dividend changes, and (iv) the realized dividend yield. In our approach, the weight of dividend strips of various maturities in today's stock price drive the elasticities of the stock price to changes in real yields, risk premia and expected dividends. Second, we argue that to implement the decomposition, a lot can be observed about the real yield curve and the equity risk premium in modern financial markets.

We illustrate our approach by providing decompositions of annual and quarterly realized returns on the S&P500 as well a detailed analysis of the S&P500 return during the COVID crisis

[16] We use out to 30-year in the baseline results to be consistent with other periods in the full sample where the 30-year to 40-year horizons are unavailable.

in 2020. Movements in the equity risk premium had a substantial role in the market crash and rebound in March and April, while a fall in real risk-free yields far out the yield curve was a key driver of the stock market ending the year with strong positive returns.

References

Asness, C., T. Hazelkorn, and S. Richardson (2018). Buyback derangement syndrome. *Journal of Portfolio Management 44*.

Bansal, R., S. Miller, D. Song, and A. Yaron (2021). The term structure of equity risk premia. *Journal of Financial Economics 142*.

Bansal, R. and A. Yaron (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance 59*.

Barro, R. J. (2006). Rare disasters and asset markets in the twentieth century. *Quarterly Journal of Economics 121*.

Bianchi, F. (2020). The great depression and the great recession: A view from financial markets. *Journal of Monetary Economics 114*, 240–261.

Bordalo, P., N. Gennaioli, R. La Porta, and A. Shleifer (2020). Expectations of fundamentals and stock market puzzles. *Working Paper* (May).

Campbell, J. Y. (1991). A Variance Decomposition for Stock Returns. *The Economic Journal 101*(405), 157–159.

Campbell, J. Y. and J. Ammer (1993). What Moves the Stock and Bond Markets. *The Journal of Finance 48*(1), 3–37.

Campbell, J. Y. and J. H. Cochrane (1999). By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy 107*.

Campbell, J. Y. and R. J. Shiller (1988). The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors. *Review of Financial Studies 1*(3), 195–228.

Campbell, J. Y. and T. Vuolteenaho (2004). Bad Beta, Good Beta. *American Economic Reivew 94*(5), 1249– 1275.

Chen, L. and X. Zhao (2009). Return decomposition. *Review of Financial Studies 22*(12), 5213–5249.

Cieslak, A., A. Morse, and A. Vissing-Jorgensen (2019). Stock returns over the fomc cycle. *The Journal of Finance 74*(5), 2201–2248.

Cieslak, A. and H. Pang (2021). Common shocks in stocks and bonds. *Journal of Financial Economics 142*.

Cochrane, J. H. (2011). Presidential Address: Discount Rates. *Journal of Finance 66*(4), 1047–1108.

Cox, J., D. L. Greenwald, and S. C. Ludvigson (2020). What Explains the COVID-19 Stock Market? *MIT Sloan Working Paper*.

Dahlquist, M. and M. Ibert (2021). How Cyclical Are Stock Market Return Expectations? Evidence from Capital Market Assumptions. *Swedish House of Finance Research Paper No. 21-1* (April).

De la O, R. and S. Myers (2021). Subjective Cash Flow and Discount Rate Expectations. *Journal of Finance, Forthcoming* (0).

Diebold, F. X. and C. Li (2006). Forecasting the term structure of government bond yields. *Journal of Econometrics 130*.

Fama, E. F. and K. R. French (2002). The equity premium. *The Journal of Finance 57*(2), 637–659.

Gao, C. and I. W. Martin (2021). Volatility, valuation ratios, and bubbles: An empirical measure of market sentiment. *Journal of Finance 76*(6).

Gonçalves, A. (2021a). What Moves Equity Markets? A Term Structure Decomposition for Stock Returns. *University of North Carolina Working Paper*.

Gonçalves, A. S. (2021b). Reinvestment risk and the equity term structure. *Journal of Finance 76*.

Gonçalves, A. S. (2021c). The short duration premium. *Journal of Financial Economics 141*.

Gormsen, N. J. (2021). Time Variation of the Equity Term Structure. *Journal of Finance 76*(4), 1959–1999.

Gormsen, N. J. and R. S. J. Koijen (2020). Coronavirus: Impact on Stock Prices and Growth Expectations. *The Review of Asset Pricing Studies*, 1–34.

Gormsen, N. J., R. S. J. Koijen, and I. Martin (2021). Implied Dividend Volatility and Expected Growth. *AEA Papers and Proceedings, forthcoming*.

Greenwald, D., M. Lettau, and S. Ludvigson (2019). How the Wealth Was Won: Factor Shares as Market Fundamentals. *NBER Working Paper Series* (510), 1–9.

He, Z., S. Nagel, and Z. Song (2022). Treasury inconvenience yields during the covid-19 crisis. *Journal of Financial Economics 143*(1).

Klingler, S. and S. Sundaresan (2019). An explanation of negative swap spreads: Demand for duration from underfunded pension plans. *Journal of Finance 72*(2), 675–710.

Krishnamurthy, A. and A. Vissing-Jorgensen (2011). The effects of quantitative easing on interest rates: Channels and implications for policy. *Brookings Papers on Economic Activity* (2).

Landier, A. and D. Thesmar (2020). Earnings expectations during the covid-19 crisis. *Review of Asset Pricing Studies 10*(4).

Martin, I. (2017). What is the expected return on the market? *Quarterly Journal of Economics 132*(1), 367–433.

Nelson, C. R. and A. F. Siegel (1987). Parsimonious modeling of yield curves. *The Journal of Business 60*.

Rietz, T. A. (1988). The equity risk premium a solution. *Journal of Monetary Economics 22*.

van Binsbergen, J., M. Brandt, and R. Koijen (2012). On the timing and pricing of dividends. *American Economic Review 102*(4).

van Binsbergen, J. H. (2020). Duration-Based Stock Valuation: Reassessing Stock Market Performance and Volatility. *Wharton Working Paper*.

van Binsbergen, J. H., W. Hueskes, R. S. J. Koijen, and E. B. Vrugt (2013). Equity Yields. *Journal of Financial Economics 110*(3), 503–519.

van Binsbergen, J. H. and R. S. Koijen (2017). The term structure of returns: Facts and theory. *Journal of Financial Economics 124*(1), 1–21.

Vissing-Jorgensen, A. (2021). The Treasury Market in Spring 2020 and the Response of the Federal Reserve. *Journal of Monetary Economics 124*, 19–47.

VII. Figures

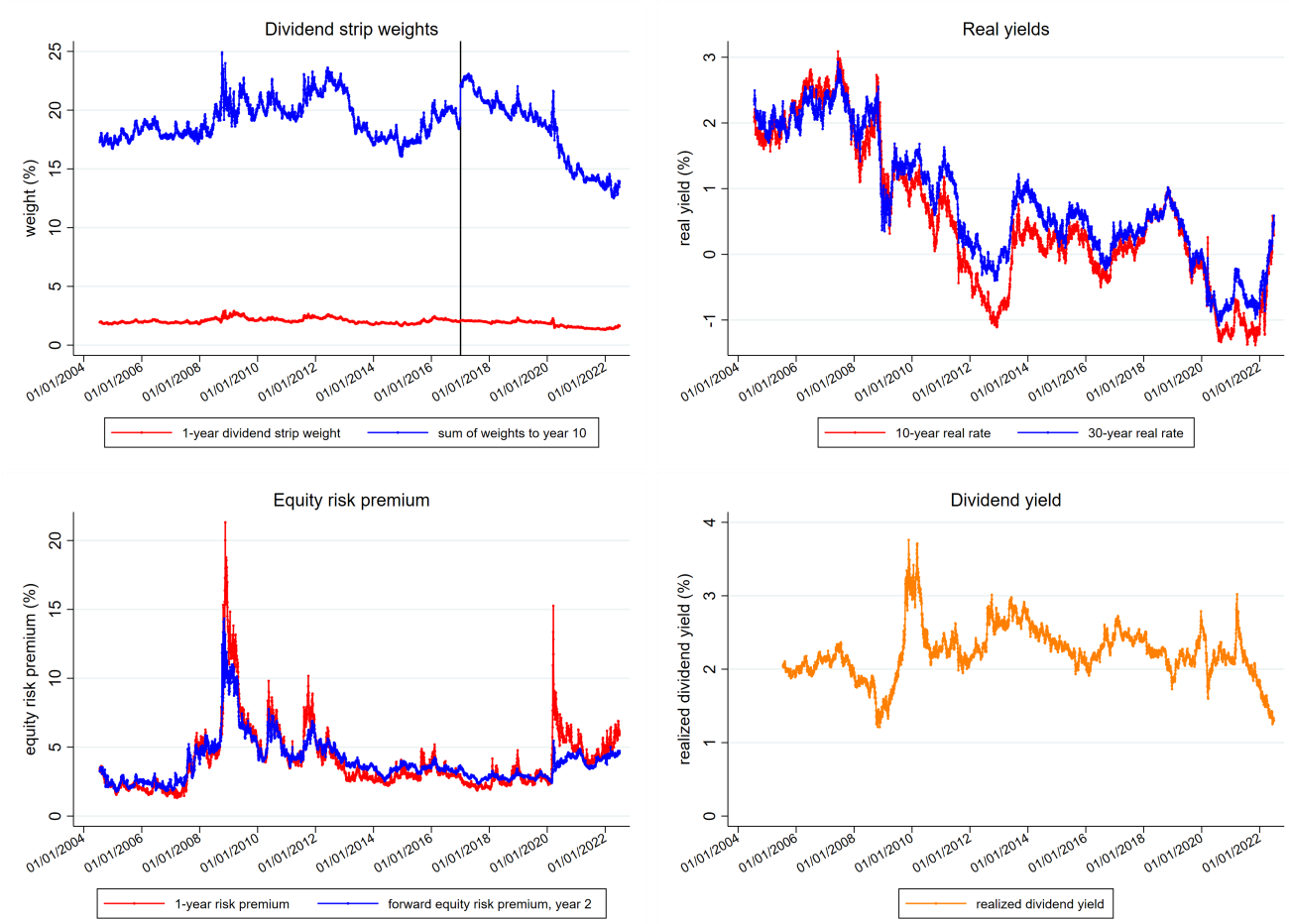
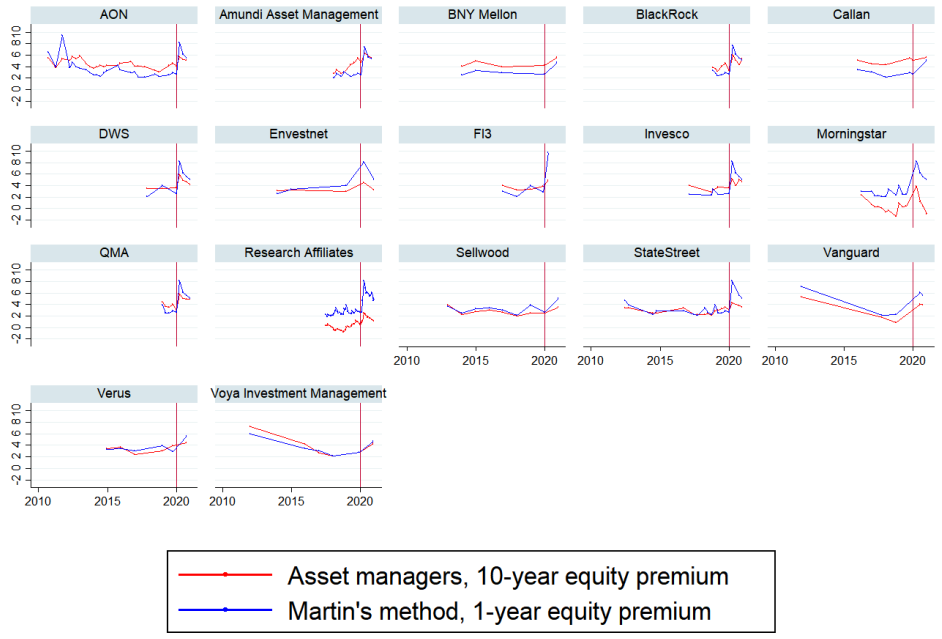


Figure 1. S&P 500 return decomposition inputs, 2004-2022.

This figure shows observable inputs for a return decomposition over the full sample of available data (July 2004 to July 2022). The top left panel presents dividend strip weights from maturity of 1-year and for the sum of maturities from 1-10 years. The black vertical lines indicate the start of 2017, from which point dividend futures for maturities 1-year to 10-years are available from Bloomberg. The top right panel presents the real yield measured via interest rate swaps and inflation swaps at the 10-year and 30-year maturity. The bottom right panel show the Martin (2017) lower bound of the 1-year equity risk premium and the forward 1-year equity risk premia in 1-year. The bottom right panel shows the dividend yield on the S&P 500, computed as the difference between the return on the total return index and the price index over the proceeding 12 months.

Panel A.



Panel B.

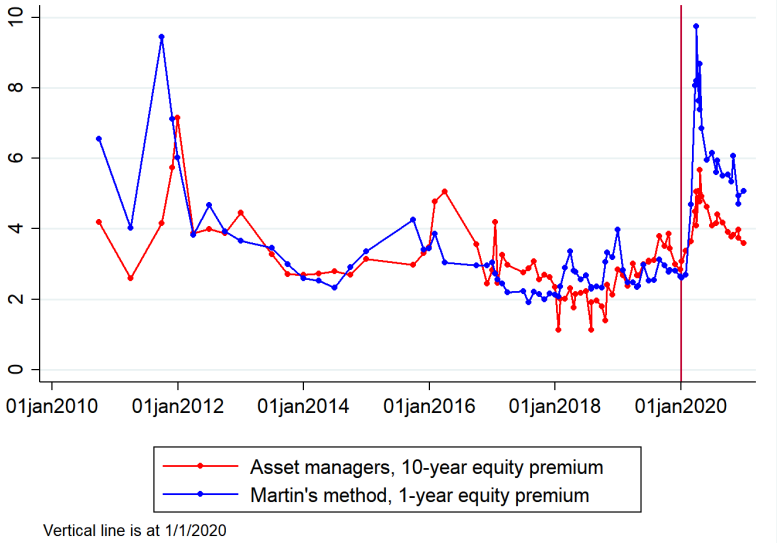


Figure 2.
Asset manager expectations and the option-based equity risk premium estimates

Panel A. This panel shows the 10-year equity risk premium estimates of various asset managers and compares this to the 1-year Martin (2017) lower bound of the equity risk premium that is computed from option prices. Each figure is a time-series plot for one asset manager. Panel B shows a combined series for all asset managers along with the 1-year Martin (2017) lower bound.

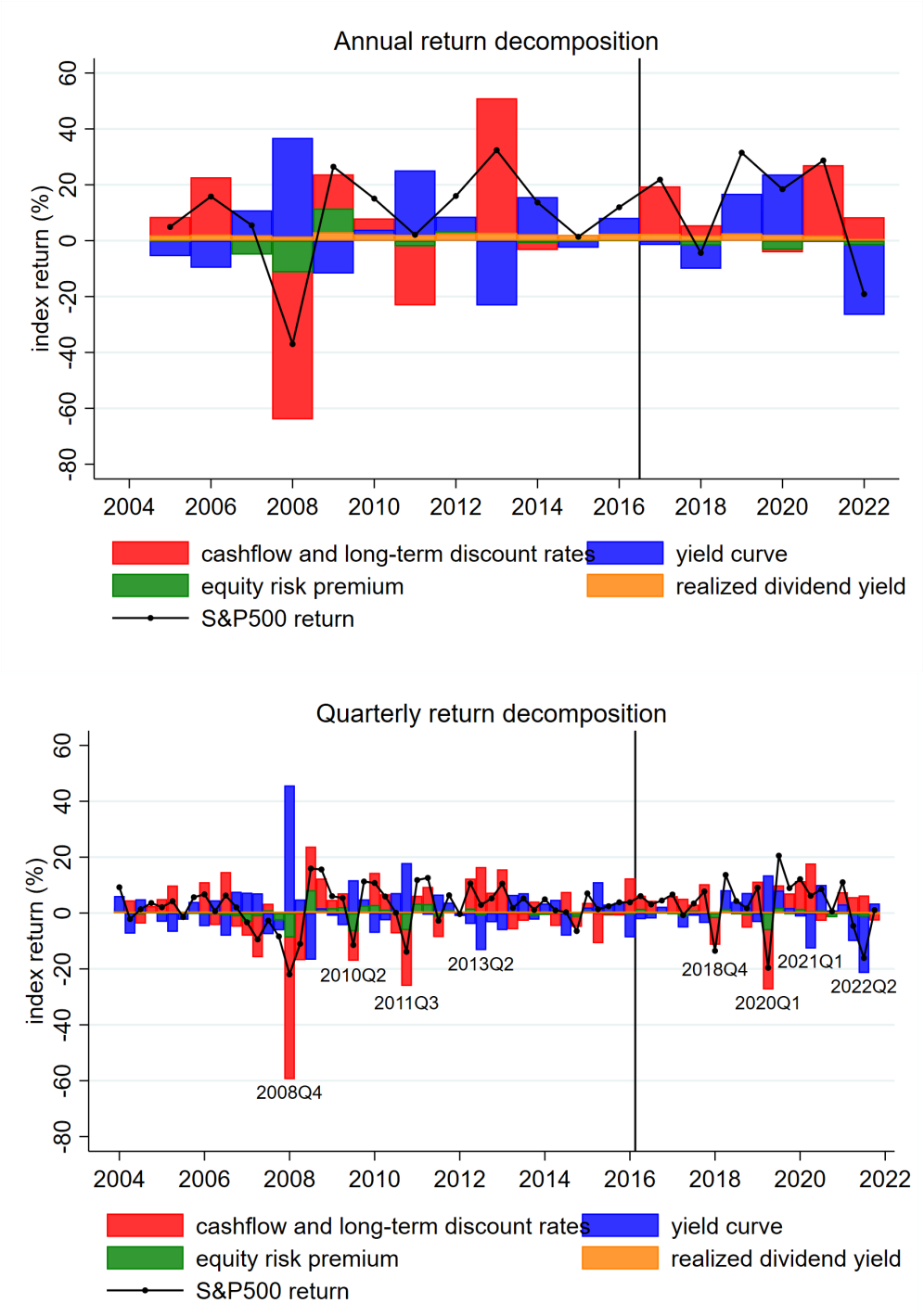


Figure 3. S&P500 return decomposition, 2004-2022.

This figure shows our observable-based return decomposition over the full sample of observable data (July 2004 to July 2022). The top panel presents the return decomposition on an annual frequency and bottom panel shows the return decomposition on a quarterly frequency. The black vertical lines indicate the start of 2017, from which point dividend futures are available on Bloomberg. We use them to compute dividend strip weights in the decomposition calculation.

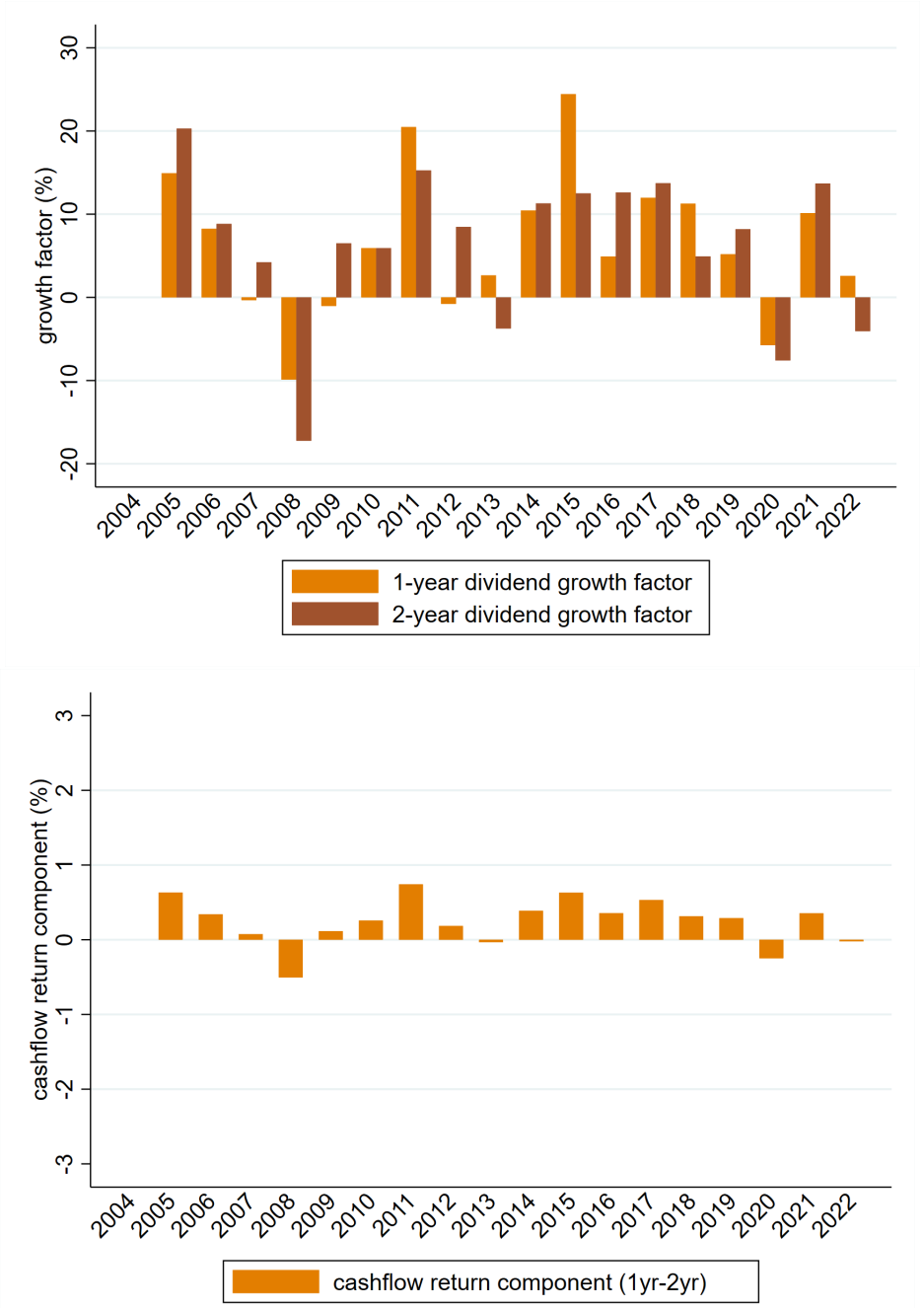


Figure 4.
Near-term dividend, 2004-2022: growth factor and S&P500 return component.

This figure shows near-term expected real dividends over the full sample (2004-2022) and the impact these have on the overall S&P500 return. The top panel presents the 1-year and 2-year real dividend growth factors at an annual frequency. The bottom panel shows how these 1-year and 2-year factors translate into the cashflows return component using

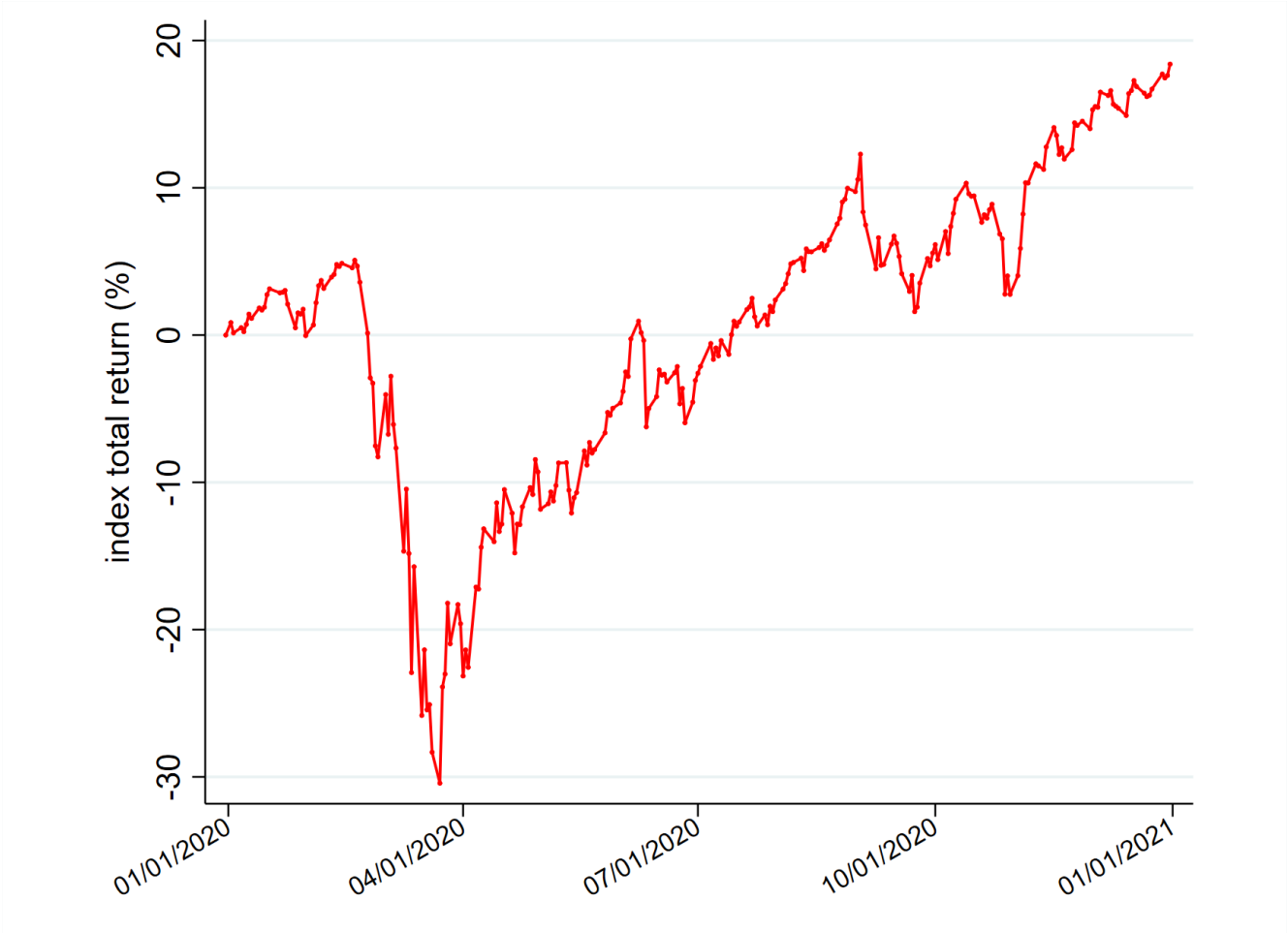


Figure 5. Total return on the S&P500 index, 2020.

This figure shows the total compounded return of the S&P500 index in 2020. It includes price level changes and dividend income.At its lowest point on March 23rd 2020, the cumulative return year-to-date was negative 30.4%. As of June 8th 2020, the market had recovered all losses and the cumulative return was once again positive. The stock market rally continued through the second half of the year and the cumulative return over the full calendar year of 2020 was positive 18.4%.

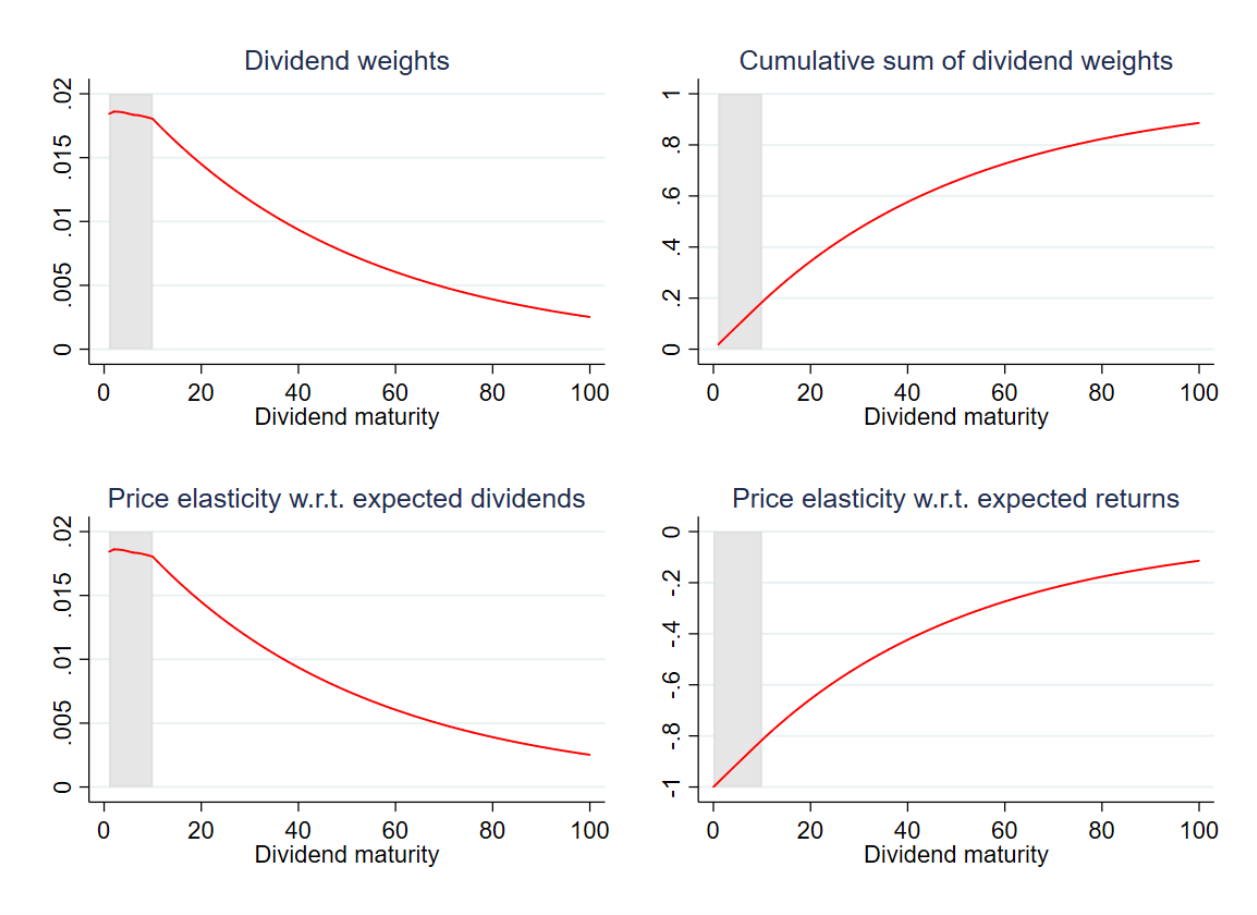


Figure 6. Dividend weights and stock price elasticities, 2020.

This figure shows dividend strips weights and how these weights translate into the stock price elasticity with respect to expected dividends and expected returns of various maturities. The top left panel shows the dividend strip weight, $w_t^{(n)}$, for the dividend payment in n years. The weight is the dividend strip's present value as a fraction of the overall stock market valuation. The top right panel shows the cumulative sum of dividend strip weights to maturity n. The bottom left panel shows the stock price elasticity with respect to expected dividend in period $t+k$:

$$\psi_{t,k}^{D} = w_t^{(k)}$$

and the bottom right panel shows the stock price elasticity with respect to the expected return in period $t+k$ (from F or RP):

$$\psi_{t,k}^{F} = \psi_{t,k}^{RP} = -\sum_{n=k}^{\infty} w_t^{(n)} = -\left(1 - \sum_{n=1}^{k} w_t^{(n-1)}\right).$$

Weights are computed from traded dividend futures prices, which are available to a maturity of 10 years (the shaded region of the figures). From this maturity, dividend weights are estimated using dividend futures prices to year 10 and assuming a Gordon growth model beyond year 10 (see Result 4). All panels use dividend futures prices as of January 2nd 2020.

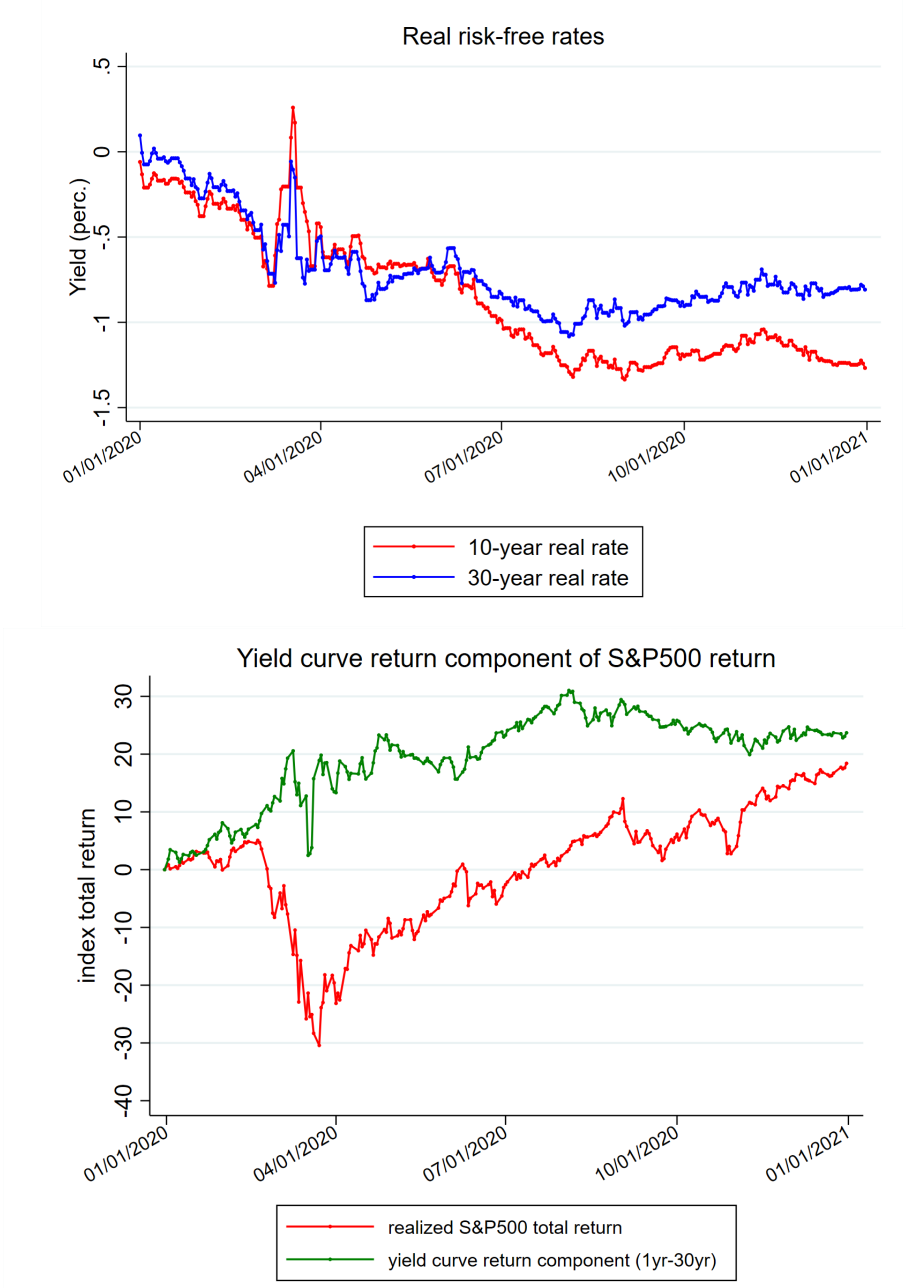


Figure 7. Real risk-free yields and yield curve return component, 2020.

This figure shows 10-year and 30-year real yields over the calendar year 2020 (top panel) and the yield curve returns estimated with forward real yield growth rates from 1-year through to 30-year maturity (bottom panel). Real yields are calculated as the difference between nominal interest rate swap rates and inflation swap rates.

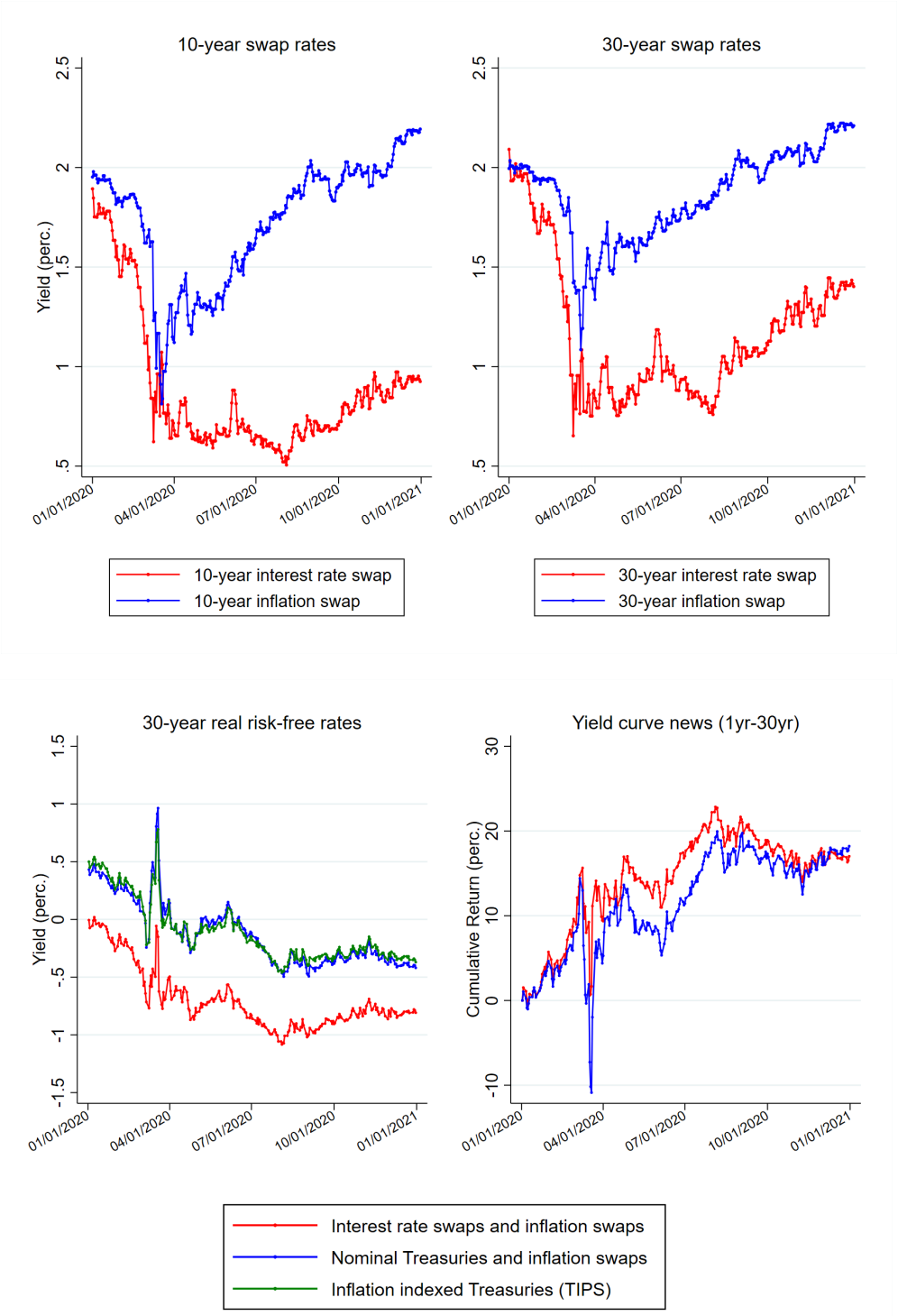


Figure 8. Real risk-free yields in 2020: Inputs and robustness.

The top panel shows the nominal interest rate swap and inflation swap yields over the calendar year 2020 for the 10 and 30-year maturities. The bottom panel shows that the yield curve return component result is robust to different measures of the real yields. The bottom left figure shows the 30 year real yield over 2020 with the real yield calculated using (a) interest rate swaps and inflation swaps, (b) nominal Treasury yields and inflation swaps, and (c) inflation indexed Treasuries (TIPS) yields. The bottom right figure shows the yield curve return component using the full yield curve of (a) and (b) out to a 30-year maturity.

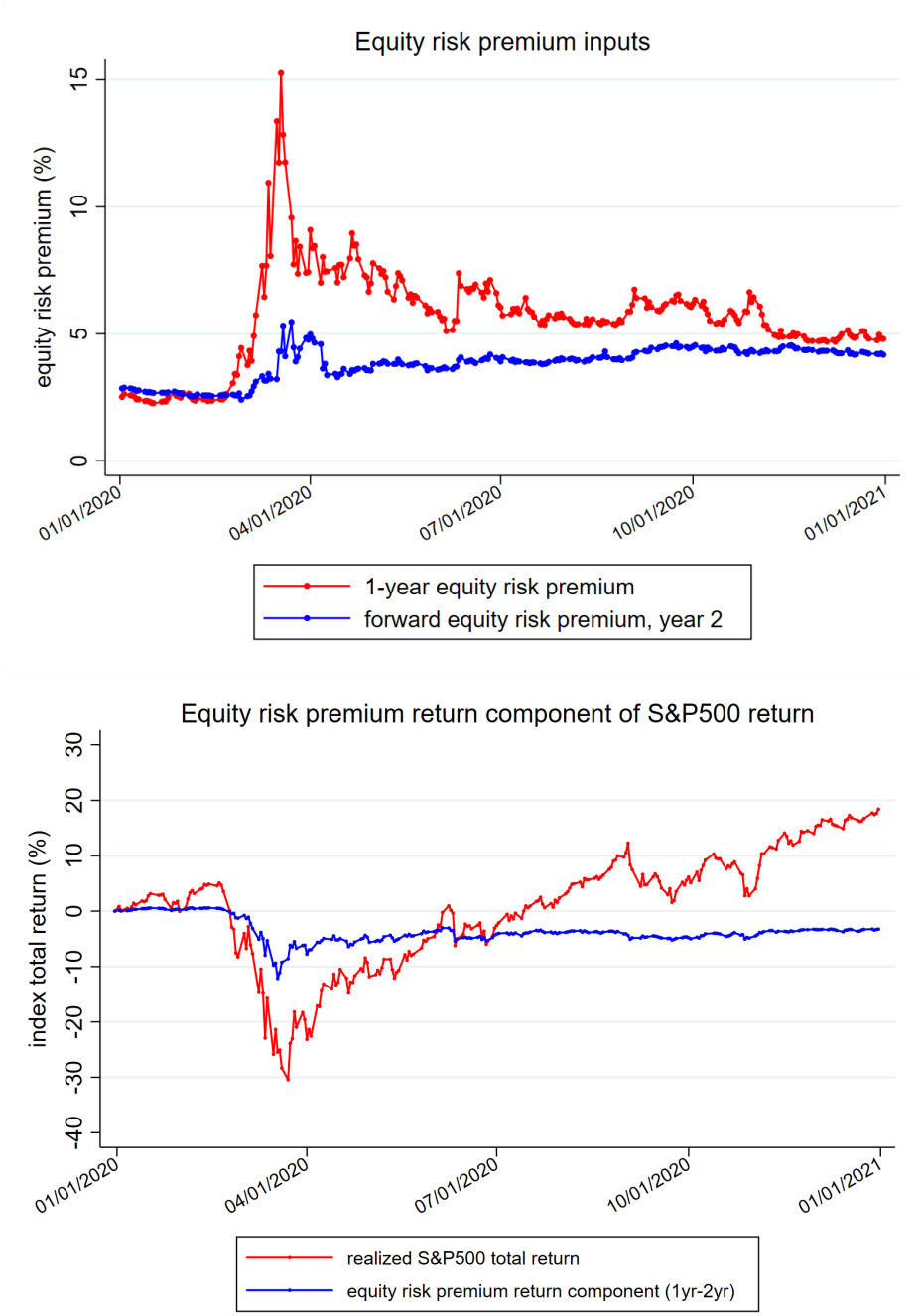


Figure 9. Equity risk premium and risk premium return component, 2020.

The top figure shows the 1-year equity risk premium and the forward 1-year equity risk premium in 1 year over the calendar year 2020 (top panel) and the cumulative sum of daily equity risk premium return component. The daily equity risk premium return component is calculated from changes in the 1-year and forward 1-year risk premium (i.e. using maturities 1-2 years). Equity risk premium is measured using Martin (2017) methodology and is computed from SP500 option prices that are taken from OptionMetrics.

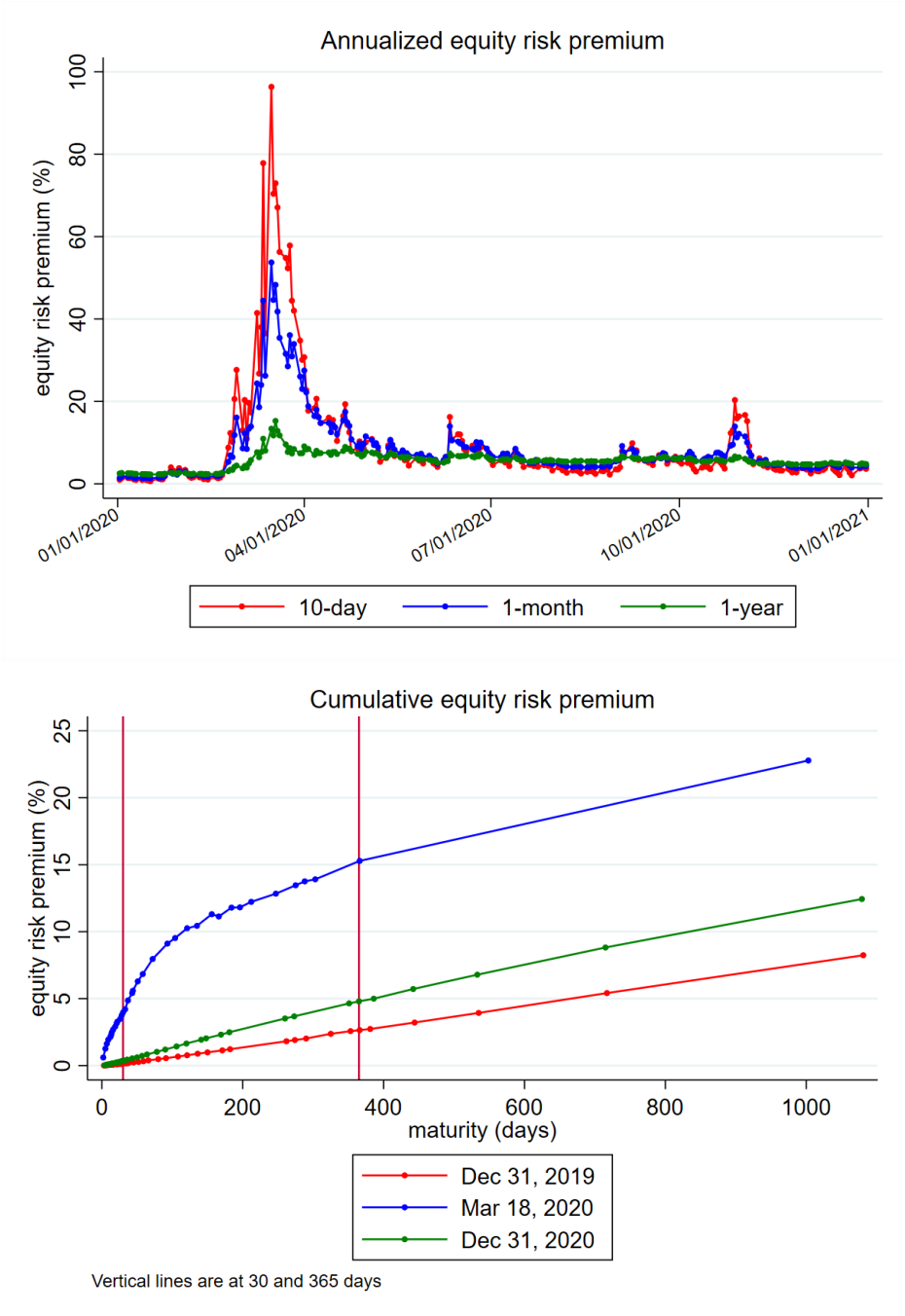


Figure 10. Equity risk premium in 2020, horizon analysis.

This figure shows the Martin (2017) lower bound of the equity risk premia during 2020. The top panel shows the time series of constant maturity risk premium estimates with maturities 10 days, 1 month and 1 year. All risk premia in the top panel are annualized. The bottom panel shows the unanualised Martin (2017) lower bound of the equity risk premia estimates plotted against holding period (expressed in days). The three lines show the cumulative risk premium curve on individual days across 2020 (the first day, the peak of the equity risk premia spike (18th March), and the last day).

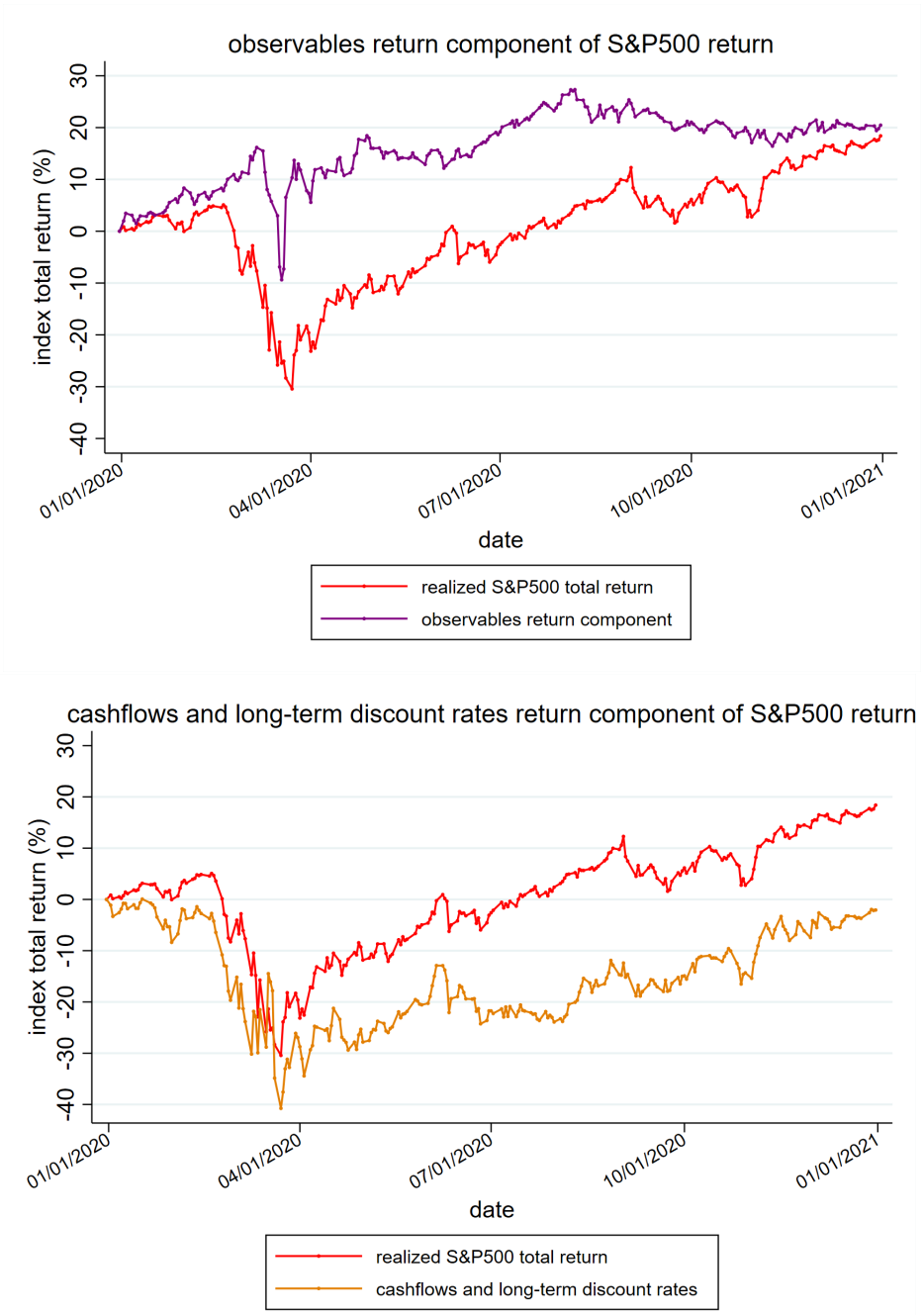


Figure 11. Observable and long-term return components, 2020.

This figure shows the return on the SP500 in 2020 that is explained by observable return component (top panel) and the return that is not unexplained by the observable return component (bottom panel). We call the return not explained by observables *long-term return component* as it captures captures the effect of dividend expectations past year 10 and any changes to risk premium past year 2 and real riskless rates past year 30.

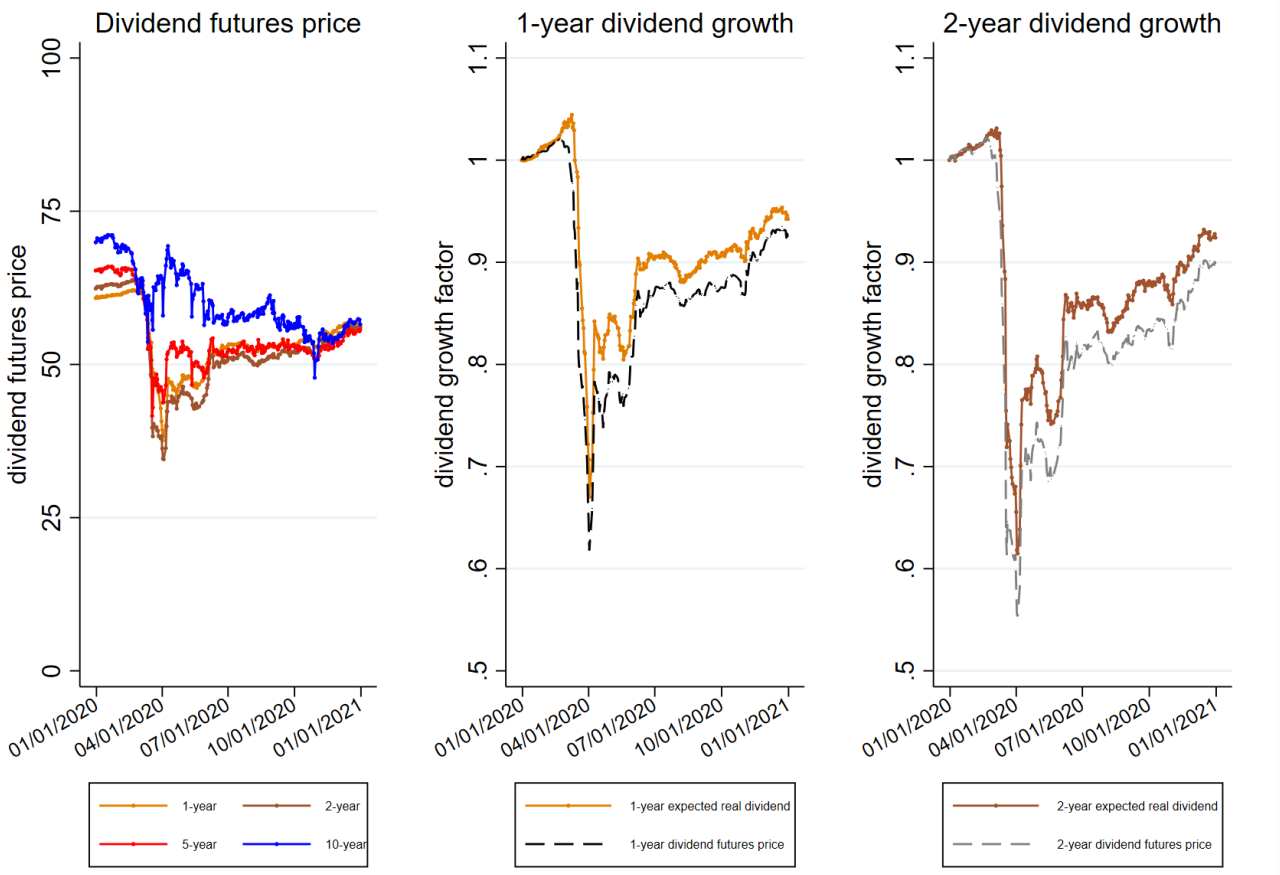


Figure 12. Near-term dividends, 2020

This figure shows near-term dividends through 2020. The left panel shows dividend futures prices for various constant maturities (1-year, 2-year, 5-year and 10-year). The middle and right panels show dividend growth factors for the 1-year and 2-year maturity respectively. In each of these two panels, the full line shows the growth factor if one uses dividend futures directly, and the dashed line uses growth factors using more price estimates for real expected dividends, adjusting the futures price for equity risk premium and expected inflation (see Section III.D.5 for details on the methodology).

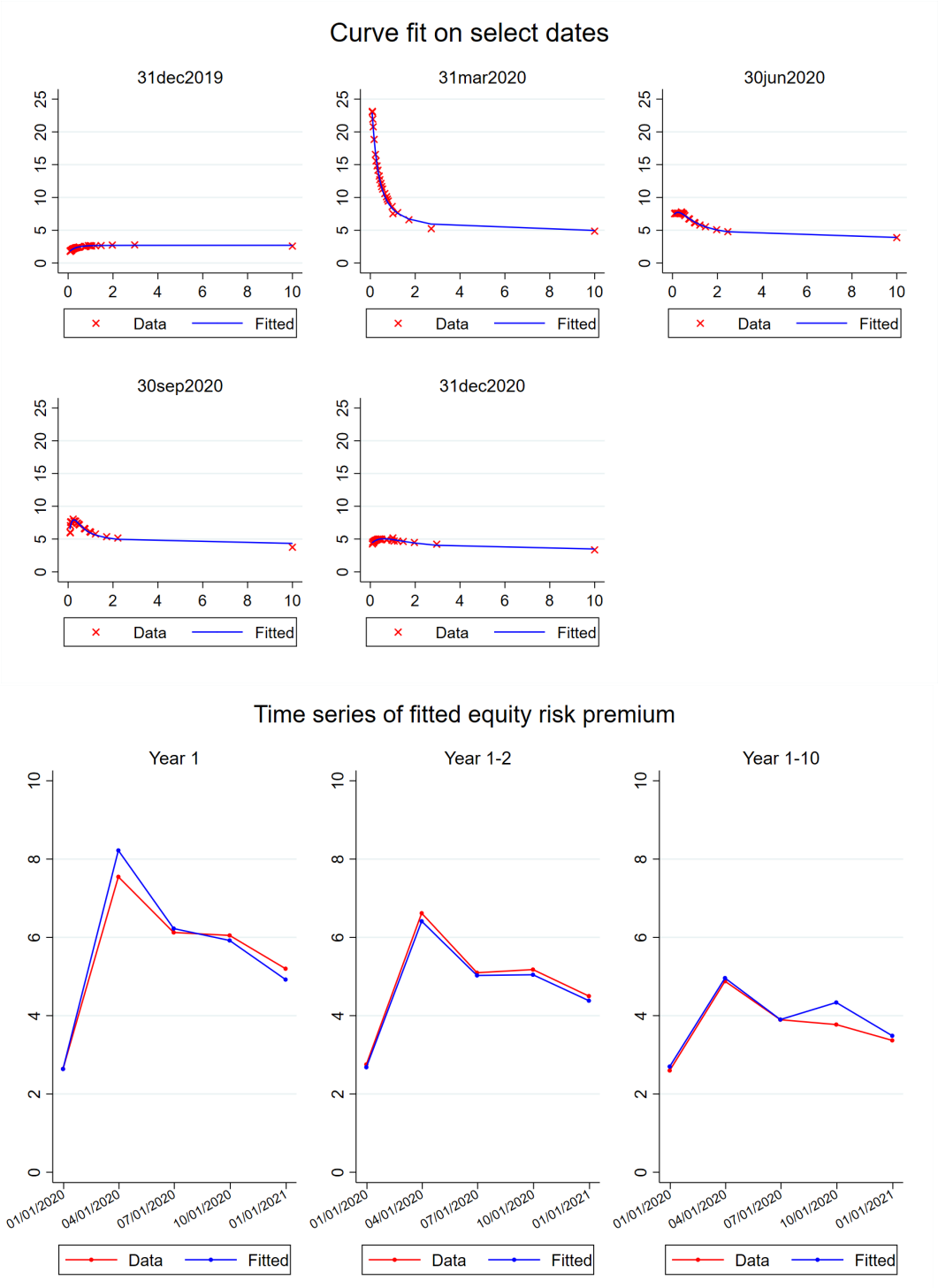


Figure 13. Nelson-Siegel estimation of equity risk premium term-structure.

This figure shows longer-term equity risk premium estimates. The top panel shows one year and one year forward equity risk premiums based from the Martin measure and at a daily frequency. The bottom panel, on a quarterly frequency, shows these variables as well as the 10 year risk premium obtained from manager expectations and the forward risk premium implied through years 3 to 10. The forward risk premium are calculated such that the short-dated Martin measure estimates and the 10-year manager expectations are consistent.

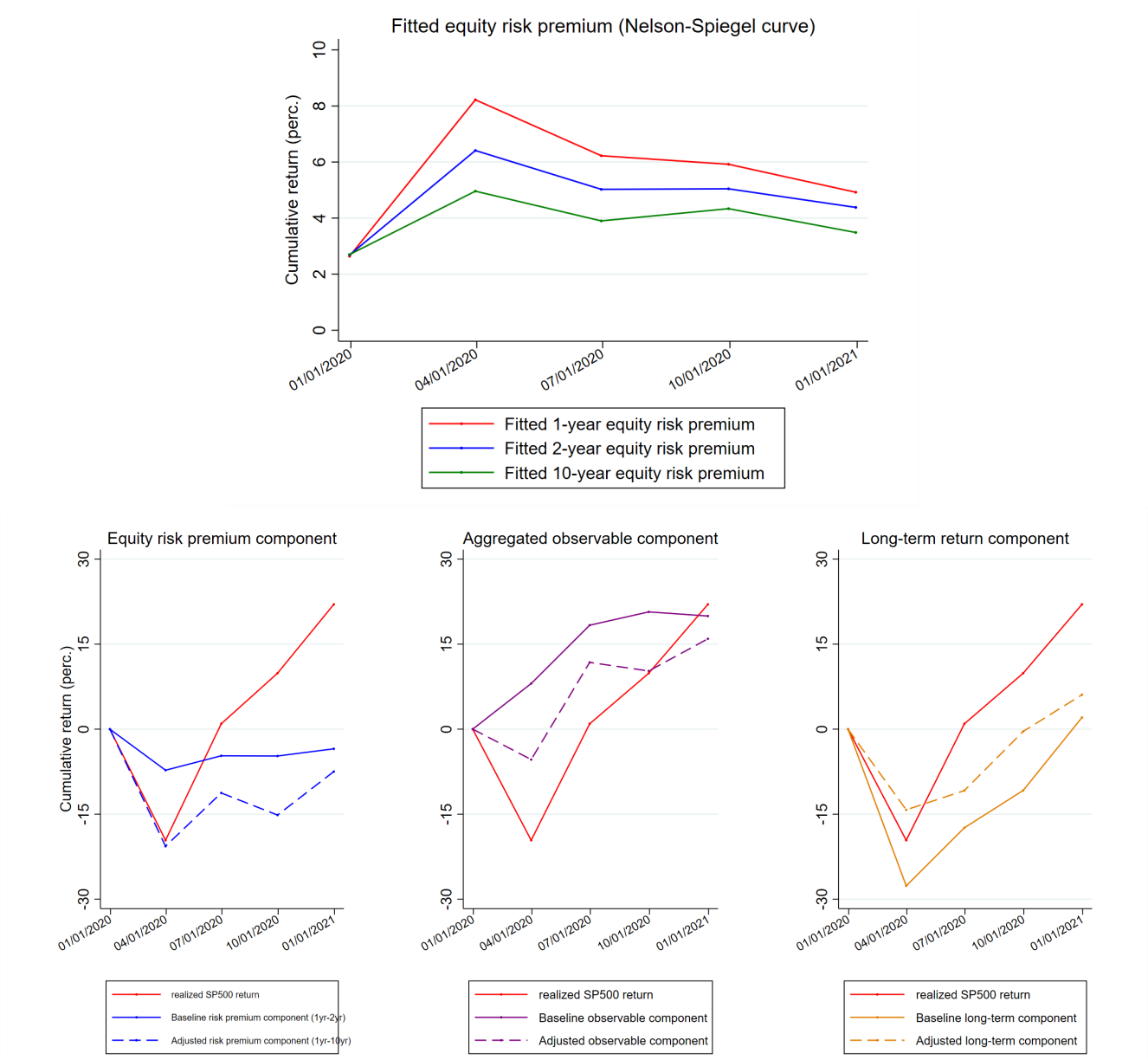


Figure 14. Adjusted stock return decomposition with 10-year risk premium.

This figure shows how incorporating longer-term equity risk premium estimates impacts on the decomposition of the cumulative return of the S&P500 index in 2020. The top panel shows the fitted 1-year, 2-year and 10-year equity risk premium Nelson-Siegel curve. The bottom panel shows how incorporating risk premium from 3-years to 10-years impacts on the estimated equity risk premium return component, observable return component and long-term return component respectively.

VIII. Tables

Observable Input	Financial Instrument	Source	Starts	Maximum maturity
Realized total return, R	Bloomberg SPXT Index	Bloomberg	1988	-
Realized dividend yield	Bloomberg SPX and SPXT index	Bloomberg	1998	-
Stock price elasticities, ψ	Dividend futures (adjusted by nominal risk-free rate)	Bloomberg (and Federal Reserve)	2016	5-year (2016-2017), 10-year (2017-)
	S&P500 equity options	OptionMetrics (or CBOE)	1996	2-year (1996-2005), 3-year (2005-2021) 5-year (2021-)
Real risk-free rates, R^F	Interest rate swaps and inflation swaps	Bloomberg	2004	30-year (2004-2008,2021-), 40-year (2008-2020)
	Nominal treasury and inflation swaps	Federal Reserve and Bloomberg	2004	30-year (2004-)
	TIPS	FRED	2003	30-year (2003-)
Equity risk premium, R^{RP}	S&P500 equity options	OptionMetrics (or CBOE)	1996	2-year (1996-2005), 3-year (2005-2021) 5-year (2021-)
Near-term cash flows, D	Dividend futures (adjusted for discount rates)	Bloomberg (Federal Reserve/OptionMetrics)	2016	(5-year), 2017 (10-year)
	S&P500 equity options	OptionMetrics (or CBOE)	1996	2-year (1996-2005), 3-year (2005-2021) 5-year (2021-)

Table I. Data summary.

This table reports available observable data to be implement Results 1-3 on the S&P 500 index. It includes information on the underlying financial instrument, the data source, the start of sample availability and the maximum maturity of the traded instrument.

Panel A: Return correlations (annual frequency)

	S&P500	yield curve	risk premia	cashflow+long-term discount rates	D/P
S&P500	1.00				
yield curve	-0.22	1.00			
risk premia	0.70	-0.53	1.00		
cashflow+long-term discount rates	0.77	-0.78	0.71	1.00	
D/P	0.76	0.06	0.64	0.39	1.00

Panel B: Return correlations (quarterly frequency)

	S&P500	yield curve	risk premia	cashflow+long-term discount rates	D/P
S&P500	1.00				
yield curve	-0.30	1.00			
risk premia	0.79	-0.54	1.00		
cashflow+long-term discount rates	0.76	-0.85	0.75	1.00	
D/P	0.29	0.06	0.29	0.09	1.00

Table II. Correlation matrix of S&P500 return components.

This table shows the correlation matrix of the return components on the S&P500 index. Yield curve measures returns due to real risk-free rate yield changes (1 year to 30 year maturity), risk premia measures returns due to equity risk premia changes (1 year to 2 year maturity), D/P measures the realized dividend yield, and cashflow+long-term discount rates measures returns due to changes in expected dividends as well as changes in discount rates beyond the observed horizons in yield curve and risk premia. Both panels show correlation across the sample period 2004Q4 - 2022Q2, with the top panel using annual returns and the bottom panel using quarterly returns.

	Realized Return				
	1 month	2 month	3 month	6 month	1 year
Martin lower bound	1.43	1.41	1.36	2.04**	1.70**
	(1.00)	(1.03)	(1.26)	(0.86)	(0.82)
Constant	0.00	0.00	0.00	-0.01	-0.00
	(0.00)	(0.01)	(0.01)	(0.02)	(0.04)
R^2 (perc.)	1.14	1.76	2.04	5.45	4.09
Observations	6,618	6,618	6,618	6,595	6,470

Table III. The Martin lower bound as a predictor variable.
This table reports the parameter estimate from the following time series regression: $\frac{1}{T-t}(R_T - R_{f,t}) = a + b \times \frac{1}{T-t}\frac{1}{R_{f,t}}\text{var}_t^* R_T$ together with Newey-West standard errors with lag selection based on the number of overlapping observations. Columns refer to separate estimations with $T - t = 1$, 2, 3, 6 and 12 months respectively. The sample period is January 1996 - June 2022.

A. Proofs

A. *Proofs for Lemma 1*

Lemma 1(i) follows immediately from the definition of independence. To show that independence is not needed, we then prove Lemma 1(ii) for the special case of log-normal returns, and finally turn to the proof of the general case without restrictions on return distributions, Lemma 1(iii).

A.1. Log-normality

Consider three random variables X_1, X_2, X_3 which are jointly log-normal and let $x = \ln X$. Start with the following result

$$\begin{aligned} & \ln\left(E(X_1X_2X_3)\right) \\ = \; & E\left(x_1+x_2+x_3\right)+\frac{1}{2}V\left(x_1+x_2+x_3\right) \\ = \; & E(x_1)+E(x_2)+E(x_3) \\ & +\frac{1}{2}\left[V(x_1)+V(x_2)+V(x_3)+2cov(x_1,x_2)+2cov(x_1,x_3)+2cov(x_2,x_3)\right] \\ = \; & \ln(E(X_1))+\ln(E(X_2)\ln(E(X_3))+cov(x_1,x_2)+cov(x_1,x_3)+cov(x_2,x_3). \end{aligned}$$

Therefore,

$$\begin{aligned} E(X_1X_2X_3) &= e^{\ln(E(X_1X_2X_3))} \\ &= E(X_1)E(X_2)E(X_3)e^{cov(x_1,x_2)+cov(x_1,x_3)+cov(x_2,x_3)} \end{aligned}$$

Applying this to Lemma 1, if $R_{t+1}^{(n)}$, $R_{t+2}^{(n-1)}$, ..., $R_{t+n}^{(1)}$ are jointly log-normal, then

$$E_t\left(R_{t+1}^{(n)}R_{t+2}^{(n-1)}...R_{t+n}^{(1)}\right) = a_t^{(n)}E_t(R_{t+1}^{(n)})E_t(R_{t+2}^{(n-1)})...E_t(R_{t+n}^{(1)})$$

where $a_t^{(n)}$ is a constant known at t given by

$$\ln\left(a_t^{(n)}\right) = \sum_{i=1}^{n-1}\sum_{k=1}^{n-i} cov_t\left(\ln\left(R_{t+i}^{(n+1-i)}\right), \ln\left(R_{t+i+k}^{(n+1-i-k)}\right)\right)$$

with the log constant the sum of covariances between all future log one-period dividend strip returns.

A.2. Beyond log-normality

Denote $\tilde{X} = X - E[X]$ so that the nth central moment of the random variable X is $E\left[\tilde{X}^n\right] = E\left[(X - E[X])^n\right]$. Similarly, with $x = \ln X$, define $\tilde{x} = x - E[x]$. Also use the properties of cumulant-generating functions:

$$\ln E[X] = E[x] + \sum_{n=2}^{\infty} \frac{\kappa_n(x)}{n!}$$

where $\kappa_2(x) = E[\tilde{x}^2]$, $\kappa_3(x) = E[\tilde{x}^3]$, and $\kappa_4(x) = E[\tilde{x}^4] - 3E[\tilde{x}^2]^2$,...etc. Conditionally log-normal random variables are the special case when the higher cumulants $\kappa_n(x)$ are zero for $n \geq 3$.

Consider two jointly distributed random variables X_1 and X_2. In the spirit of the log-normal proof, we will show that the following holds

$$\ln E[X_1 X_2] = E[x_1 + x_2] + \frac{1}{2}E\left[\widetilde{(x_1 + x_2)}^2\right] + \frac{1}{6}E\left[\widetilde{(x_1 + x_2)}^3\right] + ... \tag{A.1}$$

$$= \underbrace{E[x_1] + \frac{1}{2}E\left[\widetilde{x_1}^2\right] + \frac{1}{6}E\left[\widetilde{x_1}^3\right] + ...}_{\ln E[X_1]} + \underbrace{E[x_2] + \frac{1}{2}E\left[\widetilde{x_2}^2\right] + \frac{1}{6}E\left[\widetilde{x_2}^3\right] + ...}_{\ln E[X_2]} + C$$

$$= \ln E[X_1] + \ln E[X_2] + C \tag{A.2}$$

where C is a function of the covariance of X_1 and X_2 and all other higher-order co-movements between the random variables. With this result, we then have the required observation $E[X_1 X_2] = e^{\ln E[X_1 X_2]} = cE[X_1]E[X_2]$, where $\ln c = C$.

To derive equation (A.2), consider each of the right-hand terms in equation (A.1) separately. First,

$$E[x_1 + x_2] = E[x_1] + E[x_2].$$

Next,

$$\begin{aligned} E\left[\left(\widetilde{x_1+x_2}\right)^2\right] &= E\left[\left(x_1+x_2-E\left[x_1+x_2\right]\right)^2\right] \\ &= E\left[\left(x_1-E\left[x_1\right]+x_2-E\left[x_2\right]\right)^2\right] \\ &= E\left[\left(\widetilde{x_1}+\widetilde{x_2}\right)^2\right] \\ &= E\left[\widetilde{x_1}^2\right]+E\left[\widetilde{x_2}^2\right]+\underbrace{2E\left[\widetilde{x_1}\widetilde{x_2}\right]}_{\text{contributes to } C} \end{aligned}$$

where the last term is the covariance between the variables, which will contribute to C in equation (A.2). Finally,

$$\begin{aligned} E\left[\left(\widetilde{x_1+x_2}\right)^3\right] &= E\left[\left(x_1+x_2-E\left[x_1+x_2\right]\right)^3\right] \\ &= E\left[\left(\widetilde{x_1}+\widetilde{x_2}\right)^3\right] \\ &= E\left[\widetilde{x_1}^3\right]+E\left[\widetilde{x_2}^3\right]+\underbrace{3E\left[\widetilde{x_1}\left(\widetilde{x_2}^2\right)\right]+3E\left[\left(\widetilde{x_1}^2\right)\widetilde{x_2}\right]}_{\text{contributes to } C} \end{aligned}$$

where the last two terms will contribute to C in equation (A.2). Similar arguments apply to higher order cumulants and moments in equation (A.1).

Furthermore, the general expression for equation (A.1), extended to the log expected product of many random variables, follows from similar derivations.

B. Proof of Result 1

(11) and (12) follow from (9) and (7). (13) follows from (11), with an omitted term of

$$w_t^{(1)}\left(\frac{G_{1,t}^D}{G_{1,t}^F G_{1,t}^{RP}}-1-\left(\frac{1}{G_{1,t}^F}-1\right)-\left(\frac{1}{G_{1,t}^{RP}}-1\right)-\left(G_{1,t}^D-1\right)\right)$$

$$+\,w_t^{(2)}\left(\frac{G_{2,t}^D}{G_{1,t}^F G_{2,t}^F G_{1,t}^{RP} G_{2,t}^{RP}}-1-\left(\frac{1}{G_{1,t}^F G_{2,t}^F}-1\right)-\left(\frac{1}{G_{1,t}^{RP} G_{2,t}^{RP}}-1\right)-\left(G_{2,t}^D\right)-1\right)+\ldots \quad \text{(A.3)}$$

The first-order Taylor-expansion of $\frac{G_{1,t}^D}{G_{1,t}^F G_{1,t}^{RP}}-1$ around $G_{1,t}^D=1$, $G_{1,t}^F=1$, and $G_{1,t}^{RP}=1$ is

$$1-1+\frac{1}{1^2}\left(\frac{1}{G_{1,t}^F}-1\right)+\frac{1}{1^2}\left(\frac{1}{G_{1,t}^{RP}}-1\right)+1\left(G_{1,t}^D-1\right)=\left(\frac{1}{G_{1,t}^F}-1\right)+\left(\frac{1}{G_{1,t}^{RP}}-1\right)+\left(G_{1,t}^D-1\right)$$

Similarly for the second row in (A.3) and so on. The omitted term in (13) is thus zero to a first order.

C. *Proof of Result 2*

Result 2 exploits the expressions from the expected and unexpected capital gain in equations (20) and (21). These are both derived from (13).

Starting with the expected capital gain, focus on the term $\frac{1}{G^F_{1,t+1}}$ in (13).

$$E_t\left(\frac{1}{G^F_{1,t+1}}\right) = \frac{1}{G^{F,roll}_{1,t+1}} E_t\left(\frac{1}{G^{F,news}_{1,t+1}}\right) \approx \frac{1}{G^{F,roll}_{1,t+1}}$$

since

$$\begin{aligned} E_t\left(\frac{1}{G^{F,news}_{1,t+1}}\right) &= E_t\left(\frac{E_t\left[R^F_{t+2}\right]}{E_{t+1}\left[R^F_{t+2}\right]}\right) = E_t\left(\frac{E_t\left[R^F_{t+2}\right]}{E_{t+1}\left[R^F_{t+2}\right]}\right) \\ &= E_t\left[R^F_{t+2}\right] E_t\left(\frac{1}{E_{t+1}\left[R^F_{t+2}\right]}\right) \approx 1 \end{aligned} \tag{A.4}$$

The approximation in (A.4) uses rational expectations combined with a first-order Taylor expansion of $\frac{1}{E_{t+1}\left[R^F_{t+2}\right]}$ around a value of $E_{t+1}\left[R^F_{t+2}\right]$ of $E_t\left[R^F_{t+2}\right]$:

$$\frac{1}{E_{t+1}\left[R^F_{t+2}\right]} \approx \frac{1}{E_t\left[R^F_{t+2}\right]} - \frac{1}{\left(E_t\left[R^F_{t+2}\right]\right)^2}\left(E_{t+1}\left[R^F_{t+2}\right] - E_t\left[R^F_{t+2}\right]\right)$$

This implies $E_t\left(\frac{1}{E_{t+1}\left[R^F_{t+2}\right]}\right) \approx \frac{1}{E_t\left[R^F_{t+2}\right]}$ since $E_t\left(E_{t+1}\left[R^F_{t+2}\right] - E_t\left[R^F_{t+2}\right]\right) = 0$ by rational expectations. Exploiting similar approximations for the other terms in (13) results in (20).

To derive the unexpected capital gain in (21), consider how $\frac{1}{G^F_{1,t+1}}$ in (13) differs from its expectation:

$$\begin{aligned} \frac{1}{G^F_{1,t+1}} - E_t\left(\frac{1}{G^F_{1,t+1}}\right) &= \frac{1}{G^{F,roll}_{1,t+1}}\frac{1}{G^{F,news}_{1,t+1}} - \frac{1}{G^{F,roll}_{1,t+1}} E_t\left(\frac{1}{G^{F,news}_{1,t+1}}\right) \\ &\approx \frac{1}{G^{F,roll}_{1,t+1}}\left(\frac{1}{G^{Fnews}_{1,t+1}} - 1\right) \approx \left(\frac{1}{G^{F,news}_{1,t+1}} - 1\right) \end{aligned}$$

where we exploit (a) the fact that the roll term is known as of t, (b) the approximation from above

that $E_t\left(\frac{1}{G^{F,news}_{1,t+1}}\right) \approx 1$, and (c) the fact that for x and y close to 1, $\frac{1}{x}\left(\frac{1}{y}-1\right) \approx \left(\frac{1}{y}-1\right)$ (this last approximation is only needed to get a simpler expression for the unexpected capital gain, it is straight forward to keep the roll factors in the expression for the unexpected capital gain). Exploiting similar approximations for the other terms in (13) results in (21).

D. *Proof of Result 4*

Under the gordon growth assumption beyond 10-years, the value of long-term dividends is

$$\begin{aligned} L_t &= \sum_{k=11}^{\infty} \frac{E_t D_{t+k}}{R_{t,k}} \\ &= P_t^{(10)} \sum_{k=11}^{\infty} \left(\frac{G_t^L}{R_t^L}\right)^{k-10} \\ &= P_t^{(10)} \left(\frac{G_t^L}{R_t^L - G_t^L}\right) \end{aligned} \tag{A.5}$$

Rearranging in terms of the ratio of the long-run gross dividend growth rate to the long-run gross expected return,

$$\frac{G_t^L}{R_t^L} = \frac{L_t}{L_t + P_t^{(10)}} \tag{A.6}$$

with L_t observed from the difference between aggregate stock market value and the sum of dividend strip values up to 10 years maturity. The price of dividend strips beyond year 10

$$P_t^{(k)} = P_t^{(10)} \left(\frac{G_t^L}{R_t^L}\right)^{k-10} \qquad \text{for } k > 10 \tag{A.7}$$

are then determined by the observable $\frac{G_t^L}{R_t^L}$, and the weights beyond year 10 follow from dividing through by the value of the overall stock market.

E. *Proof of Result 5*

$E_t\left(M_T R_T\right) = 1$ implies that

$$\begin{aligned} \ln\left(E_t\left(M_T R_T\right)\right) &= E_t\left(\ln M_T + \ln R_T\right) + \frac{1}{2} V_t\left(\ln M_T + \ln R_T\right) \\ &= \left(\mu_{R,t} - r_{f,t} - \frac{1}{2}\sigma^2_{M,t} - \frac{1}{2}\sigma^2_{R,t}\right) + \frac{1}{2}\left(\sigma^2_{M,t} + \sigma^2_{R,t} + 2cov_t\left(\ln R_T, \ln M_T\right)\right) \\ &= \mu_{R,t} - r_{f,t} + cov_t\left(\ln R_T, \ln M_T\right) = 0 \end{aligned}$$

and thus

$$\mu_{R,t} - r_{f,t} = -cov_t \left(\ln R_T, \ln M_T \right). \tag{A.8}$$

$E_t \left(M_T R_T \right) = 1$ furthermore implies that

$$cov_t \left(M_T R_T, R_T \right) = E_t \left(M_T R_T^2 \right) - E \left(R_T \right)$$

Consider each term on the right hand side separately.

$$\begin{aligned} \ln E_t \left(M_T R_T^2 \right) &= E_t \left(\ln M_T + 2 \ln R_T \right) + \frac{1}{2} V_t \left(\ln M_T + 2 \ln R_T \right) \\ &= -r_{f,t} - \frac{1}{2}\sigma_{M,t}^2 + 2\left(\mu_{R,t} - \frac{1}{2}\sigma_{R,t}^2 \right) + \frac{1}{2}\left(\sigma_{M,t}^2 + 4\sigma_{R,t}^2 - 4\left(\mu_{R,t} - r_{f,t} \right) \right) \\ &= r_{f,t} + \sigma_{R,t}^2 \end{aligned}$$

$$\begin{aligned} \ln E_t \left(R_T \right) &= E_t \left(\ln R_T \right) + \frac{1}{2} V_t \left(\ln R_T \right) \\ &= \mu_{R,t} \end{aligned}$$

Combining these two expressions

$$cov_t \left(M_T R_T, R_T \right) = e^{r_{f,t} + \sigma_{R,t}^2} - e^{\mu_{R,t}} \tag{A.9}$$

The derivative with respect to a state variable s_t is

$$\frac{\partial cov_t \left(M_T R_T, R_T \right)}{\partial s_t} = e^{r_{f,t} + \sigma_{R,t}^2} \left[\frac{\partial r_{f,t}}{\partial s_t} + 2\sigma_{R,T} \frac{\partial \sigma_{R,t}}{\partial s_t} \right] - e^{\mu_{R,t}} \left[\frac{\partial \mu_{R,t}}{\partial s_t} \right]$$

If the NCC holds, $cov_t \left(M_T R_T, R_T \right) \leq 0$ and thus $e^{r_{f,t} + \sigma_{R,t}^2} \leq e^{\mu_{R,t}}$. Therefore, it is sufficient for $\frac{\partial cov_t (M_T R_T, R_T)}{\partial s_t} \leq 0$ that $\frac{\partial r_{f,t}}{\partial s_t} + 2\sigma_{R,t} \frac{\partial \sigma_{R,t}}{\partial s_t} \leq \frac{\partial \mu_{R,t}}{\partial s_t}$. Rewrite this sufficient condition as follows

$$\left(\frac{\partial \mu_{R,t}}{\partial s_t} - \frac{\partial r_{f,t}}{\partial s_t} \right) \frac{1}{\sigma_{R,t}} - \frac{\partial \sigma_{R,t}}{\partial s_t} \geq \frac{\partial \sigma_{R,t}}{\partial s_t}$$

Consider now the change in the conditional Sharpe ratio (for log returns):

$$\begin{aligned}\frac{\partial}{\partial s_t}\left[\frac{\mu_{R,t}-r_{f,t}}{\sigma_{R,t}}\right] &= \frac{1}{\sigma_{R,t}^2}\left[\left(\frac{\partial \mu_{R,t}}{\partial s_t}-\frac{\partial r_{f,t}}{\partial s_t}\right)\sigma_{R,t}-\left(\mu_{R,t}-r_{f,t}\right)\frac{\partial \sigma_{R,t}}{\partial s_t}\right] \\ &= \left(\frac{\partial \mu_{R,t}}{\partial s_t}-\frac{\partial r_{f,t}}{\partial s_t}\right)\frac{1}{\sigma_{R,t}}-\frac{\left(\mu_{R,t}-r_{f,t}\right)}{\sigma_{R,t}^2}\frac{\partial \sigma_{R,t}}{\partial s_t} \\ &\geq \left(\frac{\partial \mu_{R,t}}{\partial s_t}-\frac{\partial r_{f,t}}{\partial s_t}\right)\frac{1}{\sigma_{R,t}}-\frac{\partial \sigma_{R,t}}{\partial s_t}\end{aligned}$$

where the last line follows from (1) the fact that $\frac{\left(\mu_{R,t}-r_{f,t}\right)}{\sigma_{R,t}^2} \geq 1$ under the NCC and (2) the assumption that $\frac{\partial \sigma_{R,t}}{\partial s_t} \geq 0$. Thus, it is sufficient for $\frac{\partial cov_t(M_T R_T, R_T)}{\partial s_t} \leq 0$ that the change in the conditional Sharpe ratio is at least as large as the change in the conditional standard deviation

$$\frac{\partial}{\partial s_t}\left[\frac{\mu_{R,t}-r_{f,t}}{\sigma_{R,t}}\right] \geq \frac{\partial \sigma_{R,t}}{\partial s_t}.$$

F. Proof of Result 6

We can exploit (A.8) to get

$$\begin{aligned}\mu_{R,t}-r_{f,t} &= -cov_t\left(\ln R_T, \ln M_T\right) \\ &= \gamma cov_t\left(\ln R_T, \ln\left(C_T/C_t\right)\right).\end{aligned}$$

This implies (A.22),

$$\begin{aligned}\frac{\mu_{R,t}-r_{f,t}}{\sigma_{R,t}} &= \gamma\frac{cov_t\left(\ln R_T, \ln\left(C_T/C_t\right)\right)}{\sigma_{R,t}^2}\sigma_{R,t} \\ &= \gamma\beta_t^C\sigma_{R,t}\end{aligned}$$

where β_t^C is the (potentially time-varying) beta of $\ln\left(C_T/C_t\right)$ with respect to $\ln R_T$. The rest of Result 6 follows directly from (A.22).

B. When do expected repurchases/issuance lead to effects of discount rates on expected future dividends?

Consider a firm that generates a perpetual stream of free cash flows of C at times t = 0, 1, 2,.... The firm has a cost of capital of r and for simplicity assume there is no uncertainty. The firm has N shares outstanding at $t = 0$ and no debt outstanding. Compare the following cases.

Case 1 (base case, no repurchases): The firm pays all free cash flows as dividends, period by period. Dividends per share are $D_t = C/N$ for all t. By the formula for the present value of a perpetuity, the stock price per share at $t = 0$ and at all later dates is

$$P_0 = \frac{D}{r} = \frac{C/N}{r} \tag{A.10}$$

The duration of dividends is

$$\mathbb{D} = \sum_i i \frac{P_0^{(i)}}{P_0} = \frac{1+r}{r}. \tag{A.11}$$

The price elasticity with respect to the cost of capital (in all periods, as opposed to just one period in our earlier derivations) therefore is

$$\Psi^R = \frac{\partial P_0/P_0}{\partial r/(1+r)} = -\frac{1+r}{r}, \tag{A.12}$$

With fixed dividend payments, the base case therefore produces the classic fixed income result that $\Psi^R = -\mathbb{D}$.

Case 2 (repurchase at $t = 1$, funded with debt issuance): At $t = 1$, the firm issues debt with market value of C and uses the proceeds to buy back X shares at the market price P_1. X and P_1 solve:

$$P_1 X = C \tag{A.13}$$

$$P_1(N - X) = \frac{C}{r} - C \tag{A.14}$$

In the second expression, C/r is the present value of free cash flows as of $t = 1$, of which a value of C in present value terms will go to repaying debt. Solving for X and P_1,

$$X = Nr,\ P_1 = \frac{C/N}{r}. \tag{A.15}$$

The price at $t = 1$ is thus unaffected by the repurchase as is the price at $t = 0$ which remains $P_0 = \frac{C/N}{r}$ (because the dividend at $t = 1$ and the price at $t = 1$ are both unaffected by the repurchase). Assuming that the debt issued is perpetual or rolled over perpetually, interest payments are rC at $t = 2, 3, ...$ and dividends per share are C/N at $t = 1$ and $(1-r)C/(N-X) = C/N$ at $t = 2, 3, ...$ Dividends per share are thus the same period by period as in Case 1 and therefore unaffected by the share repurchases. In this case, there is no effect of r on dividends even with repurchases.

Case 3 (repurchases at $t = 1$, funded by reduced dividends): At $t = 1$, the firm pays no dividend and instead spends all free cash flows C on buying back X shares at the market price P_1. X and P_1 solve:

$$P_1 X = C \tag{A.16}$$

$$P_1(N - X) = \frac{C}{r} \tag{A.17}$$

which imply

$$X = Nr\frac{1}{1+r} = N\frac{1}{1+\frac{1}{r}} \tag{A.18}$$

$$P_1 = \frac{C/N}{r}(1+r) = (C/N)\left(1+\frac{1}{r}\right) \tag{A.19}$$

The price per share at $t = 0$ is unaffected since $P_0 = \frac{1}{1+r}\frac{C/(N-X)}{r} = \frac{C/N}{r}$. The duration of dividends is one year longer that in case 1, since shareholders get a perpetuity but starting only at $t = 2$. Since P_0 is unchanged relative to Case 1, the price elasticity with respect to r (accounting for effects of r via both discounting and dividends) is unchanged relative to Case 1. It is thus no longer the case that $\Psi^R = -\mathbb{D}$. The dividend duration overstates the overall interest rate risk because a higher interest rate improves cash flows from $t = 2$ onward: At a higher r, P_1 is lower, allowing the firm to buy back more shares with C at $t = 1$. This increases dividends per share from $t = 2$ onward.

These cases illustrate how discount rate effects on expected dividends emerge if a firm is expected to buy back shares and is expected to fund repurchases by lowering dividends (Case 3). By contrast, no discount rate effects on expected dividends emerge from repurchases if the firm is expected to fund repurchases with debt issuance (and the debt is rolled over perpetually) (Case 2). A third way to fund repurchases would be to use resources available inside the firm that would otherwise have been kept invested in the firm. This would work out much like Case 2, with no discount effects of repurchases on expected dividends. Expected share issuance will work with opposite effects of share repurchases, leading to discount rate effects on expected dividends if issuance proceeds are expected to be used to pay dividends, but no such effects if issuance proceeds are expected to be used to reduce debt.

C. Implementation of the Martin lower bound

We use option price data from OptionsMetrics to construct a time series of the Martin (2017) lower bound of the equity risk premium.[17] For each date-maturity observation, we estimate the equity risk premium by discretizing the right-hand side of Martin (2017)'s lower bound of the equity risk premium

$$E_t R_T - R_{f,t} \geq \frac{2}{S_t^2}\left[\int_0^{F_{t,T}} \text{put}_{t,T}(K)\, dK + \int_{F_{t,T}}^{\infty} \text{call}_{t,T}(K)\, dK\right] \tag{A.20}$$

where $\text{put}_{t,T}(K)$ and $\text{call}_{t,T}(K)$ are time t put and call option prices for strike price K and maturity T. The forward price $F_{t,T}$ is the unique solution K where $\text{put}_{t,T}(K) = \text{call}_{t,T}(K)$.

To clean the data and generate the risk-premium estimates we take the following steps. First we drop observations if the bid price or ask price is missing and if the best bid price is zero. We then calculate the mid price as the average of the best bid and best ask price and, for date-maturity-strike-type combinations where there are multiple mid-prices, we use the option with the highest open interest.[18]

We next keep date-maturity-strike observations where there is both a call and put mid-price,

[17] For the application to 2020, we have replicated the risk premium results using use option price data from CBOE intra-day traded price data. The results mirror those reported in the main results using OptionMetrics end of day quotes.

[18] Type refers to call or put option. The existence of multiple maturity-strike-type observations is common in the later years of the sample and is caused by the increase in the range of options issued on the exchange. For example, if a newly issued weekly option has the same expiry as an existing quarterly option that is expiring in a weeks time then we observe two mid-prices. Typically, shorter horizon options are the most liquid and have highest open interest (i.e. weekly options liquid than quarterly options)

and then select the option with the lowest mid-price. This step automatically deletes put options greater than $F_{t,T}$ and deletes call option prices less than $F_{t,T}$, as required for equation (A.20). Before implementing the discretizion of the integral, we then take two extra steps to ensure the accuracy of the discretizion approximation. First, we drop date-maturity observations where the number of unique strikes is less than 10. Second, we drop date-maturity observations where the difference between the maximum put strike and the minimum call strike is greater than 50 (100) for maturities less (greater) than 1-year. This step makes sure discretization is not too coarse in the most important range of the integral (where prices are at the highest levels). We allow a larger gap between strike prices for long-maturity options as these are typically issued with a lower range of strikes.

Equation (A.20) is then estimated, providing date-maturity equity risk premium estimates at the maturities that correspond to the expires of the options in issue. To generate constant maturity risk premium estimates across dates, we linearly interpolate risk premium estimates across maturities within dates. We also extrapolate to extend maturity. However, to avoid over extrapolation, we limit this extrapolation to a maximum of half a year greater than the longest maturity option available at that date.

D. Is the Martin lower bound a good measure of the equity risk premium?

We proxy the equity risk premium with the Martin lower bound, thus assuming the bound is tight. This section provides additional empirical evidence to support this assumption as well as theoretical results on how the *change* in the bound relates to the *change* in the true equity premium.

A. The tightness of the Martin lower bound

Martin (2017) documents an average lower bound over the 1996-2012 period of about 5%, close to the equity premium estimates obtained by Fama and French (2002) using average realized dividend (or earnings) growth rates as an estimate of ex-ante expected capital gains. Martin also tests whether the lower bound is a good predictor of the realized excess return, with small

intercepts. He estimates the relation

$$\frac{1}{T-t}\left(R_T - R_{f,t}\right) = a + b \times \frac{1}{T-t}\frac{\text{var}_t^* R_T}{R_{f,t}} + \epsilon_{i,t} \tag{A.21}$$

We extend the Martin (2017) empirical results by first re-estimating equation (A.21) over the 1996-June 2022 for the S&P500 index.[19] The estimated parameters are shown in Table III, Panel A. We cannot reject the null of $b = 1$, $a = 0$ for horizons $T - t = 1$, 2, 3, 6 and 12 months.

We re-estimate equation A.21 over the 1996-2022 as shown in Table III. Over this longer sample, we find that β is higher than one for most horizons, though not significantly so for most horizons. The intercept is close to zero and insignificant across all horizons.

We find that b is greater than one in all estimations, though not significantly. Our decomposition relies on *changes* in equity risk premia. The b estimates above one imply that the true risk premium change exceeds that of the change in the lower bound. It is possible, however, that realized excess returns exceeded expected returns over this particular time period, more so in times of stress (high values of the risk-neutral variance). Fama and French (2002) argue that realized returns exceeded expected returns even over a sample as long as 1951-2000. Cieslak, Morse, and Vissing-Jorgensen (2019) argue that over the post-1994 period, unexpectedly accommodating monetary policy has contributed to much of the realized excess return on the stock market. If the unexpected positive component of realized returns is sufficiently correlated with risk-neutral variance, then an estimated b above one may not imply that changes in the lower bound are smaller than the true changes in the equity risk premium for a given horizon.

Given the lack of conclusive empirical evidence on whether $b = 1$ or $b > 1$ it is relevant to ask what theory says about how the *change* in the bound relates to the *change* in the true equity risk premium.

[19] Martin's defines a variable $SVIX^2_{t\to T} = \frac{1}{T-t}\text{var}_t^*\left(\frac{R_T}{R_{f,t}}\right)$ and his regressor is thus expressed as $R_{f,t}SVIX^2_{t\to T}$.

B. Theoretical results for the Martin lower bound in changes

Suppose an underlying state variable s_t changes and that s_t is signed such that $\frac{\partial\left[\frac{1}{R_{f,t}}\text{var}_t^* R_T\right]}{\partial s_t} > 0$. Then

$$\begin{aligned}\frac{\partial\left[E_t R_T - R_{f,t}\right]}{\partial s_t} &= \frac{\partial\left[\frac{1}{R_{f,t}}\text{var}_t^* R_T\right]}{\partial s_t} - \frac{\partial cov_t\left(M_T R_T, R_T\right)}{\partial s_t} \\ &\geq \frac{\partial\left[\frac{1}{R_{f,t}}\text{var}_t^* R_T\right]}{\partial s_t} \text{ iff } \frac{\partial cov_t\left(M_T R_T, R_T\right)}{\partial s_t} \leq 0\end{aligned}$$

It follows that the change in the lower bound is, on average, equal to the true change in the risk premium if the regression coefficient b in (A.21) equals one. If instead $b > 1$ that would suggest that the regressor is positively correlated with the omitted variable $-cov_t\left(M_T R_T, R_T\right)$ implying that $\frac{\partial cov_t(M_T R_T, R_T)}{\partial s_t} < 0$ and thus that the true change in the risk premium is larger than the change in the lower bound.

To assess theoretically whether $b > 1$ is likely, assume conditional log-normality as follows:

$$\begin{aligned}M_T &= e^{-r_{f,t}+\sigma_{M,t}Z_{M,T}-\frac{1}{2}\sigma_{M,t}^2} \\ R_T &= e^{\mu_{R,t}+\sigma_{R,t}Z_{R,T}-\frac{1}{2}\sigma_{R,t}^2}\end{aligned}$$

where $Z_{M,t}$ and $Z_{R,t}$ are (potentially correlated) standard normal random variables. Martin (2017) shows that in the log-normal case, the NCC holds iff the conditional Sharpe ratio exceeds the conditional standard deviation:

$$cov_t\left(M_T R_T, R_T\right) \leq 0 \text{ iff } e^{r_{f,t}+\sigma_{R,t}^2} \leq e^{\mu_{R,t}} \iff \sigma_{R,t} \leq \frac{\mu_{R,t} - r_{f,t}}{\sigma_{R,t}}$$

The following result states conditions that allow us to relate the true change in the risk premium to the change in the lower bound.

Result 5 (**True change vs. lower bound change in equity risk premium**). *Suppose an underlying state variable s_t changes such that $\frac{\partial\left[\frac{1}{R_{f,t}}var_t^* R_T\right]}{\partial s_t} > 0$ and $\frac{\partial \sigma_{R,t}}{\partial s_t} \geq 0$. The true change in the equity risk premium is at least as large as the change in the lower bound iff $\frac{\partial cov_t(M_T R_T, R_T)}{\partial s_t} \leq 0$. Under log-normality, it is sufficient for $\frac{\partial cov_t(M_T R_T, R_T)}{\partial s_t} \leq 0$ that*
(1) The NCC holds: $cov_t\left(M_T R_T, R_T\right) \leq 0 \iff \frac{\mu_{R,t} - r_{f,t}}{\sigma_{R,t}} \geq \sigma_{R,t}$, and

(2) $\frac{\partial}{\partial s_t}\left[\frac{\mu_{R,t}-r_{f,t}}{\sigma_{R,t}}\right] \geq \frac{\partial \sigma_{R,t}}{\partial s_t}$.

In addition to log-normality, assume CRRA utility,

$$M_T \;=\; \beta\left(\frac{C_T}{C_t}\right)^{-\gamma} = e^{\ln\beta-\gamma\ln(C_T/C_t)}$$

with $\ln(C_T/C_t) \sim N(\mu_{c,t}, \sigma^2_{c,t})$ conditional on information known at t. Define the consumption beta relative to the market as $\beta^C_t = \frac{cov_t(\ln R_T, \ln(C_T/C_t))}{\sigma^2_{R,t}}$. The following result emerges.

Result 6 (**True change vs. lower bound change in equity risk premium**).
In the log-normal CRRA case,

$$\frac{\mu_{R,t}-r_{f,t}}{\sigma_{R,t}} = \gamma\beta^C_t\sigma_{R,t}. \tag{A.22}$$

so the NCC holds if $\gamma\beta^C_t \geq 1$. *Furthermore,*

$$\frac{\partial}{\partial s_t}\left[\frac{\mu_{R,t}-r_{f,t}}{\sigma_{R,t}}\right] = \gamma\beta^C_t\frac{\partial\sigma_{R,t}}{\partial s_t} + \gamma\frac{\partial\beta^C_t}{\partial s_t}\sigma_{R,t}$$

so it is sufficient for

$$\frac{\partial}{\partial s_t}\left[\frac{\mu_{R,t}-r_{f,t}}{\sigma_{R,t}}\right] \geq \frac{\partial\sigma_{R,t}}{\partial s_t}$$

that $\gamma\beta^C_t \geq 1$ *and* $\frac{\partial\beta^C_t}{\partial s_t} \geq 0$.

Therefore, the same condition that ensures the NCC holds, $\gamma\beta^C_t \geq 1$, also helps to ensure that the true change in the risk premium is larger than the change in the lower bound. Martin (2017) argues that the NCC is very likely to hold in the log-normal case since the Sharpe ratio based on realized returns has substantially exceeded the realized standard deviation. The additional condition, $\frac{\partial\beta^C_t}{\partial s_t} \geq 0$ holds if β^C_t is constant. This will be the case for an investor who is fully invested in the market, since then $\beta^C_t = 1$. It will also (approximately) be the case for an investor who is not fully invested in the market as long as the the investor has a roughly constant portfolio weight in the market and the covariance between the market and non-market risky assets is roughly constant over time.

Overall, theoretical considerations suggest that to the extent that the Martin lower bound is not exact, the most likely direction of any bias is that the true changes in the equity risk premium are larger than the changes in the Martin lower bounds.

E. Dividend strip weights and elasticities: Liquidity and robustness

Our dividend strip weights and elasticities are based on dividend futures prices. These are a relatively new product and liquidity may be an issue. Our dividend futures prices are trade prices from Bloomberg. Bloomberg does not provide bid-ask spreads for dividend futures. However, using a proprietary data source, Bansal et al. (2021) document that over their sample period from January 2010 to February 2017, bid-ask spreads on S&P500 dividend futures averaged around 2% for each maturity (out to 5 years in their data), with a time series standard deviation around 1% for a given maturity. This is large enough to have an economically large effect on the monthly returns generated by these contracts. Because price elasticities are based on the *level* of dividend futures prices and not on dividend futures price changes (to calculate dividend strip returns), the economic impact of bid-ask spreads on price elasticities is far lower than the impact on returns. To take an example, consider the first year's dividend strip weight, which is 1.80% of the total market on January 2nd, 2020. Even if we consider a bid-ask spread of 4% (the sample mean plus two sample standard deviations), in the extreme the dividend strip weight would be 1.764% using bid prices and 1.836% using ask prices. To illustrate the point, the top panel of Figure A.1 shows how the elasticities of the stock price with respect to expected dividends (left hand side) and with respect to expected returns (right hand side) vary with bid-ask spreads of 4%. The impact on elasticities is very small across all maturities.[20]

Beyond bid-ask spreads, another important issue in our context is that limited open interest in dividend futures may lead to a risk that these are priced by particular investors, rather than a broad set of investors relevant for the overall stock market. Across contracts from year 1 to year 10, open interest has increased over time, with the average daily open interest across 2020 was 60% higher than the average daily open interest across 2017. However, across maturities, open interest in 2020 (annual average) declines from \$802M at maturity 1, to \$351M at maturity 3, \$77M at maturity 5, \$17M at maturity 8, and less than \$10M at maturity 9 and 10.

Dividend weights beyond the maximum maturity of observed dividend futures are estimated using Result 4 in our methodology, with our baseline implementation using the maximum observed dividend maturity $N_D = 10$ years. We therefore assess the robustness of our results to shortening

[20] Gormsen and Koijen (2020) provide updated liquidity analysis on the dividend futures market in 2020, with a focus on the euro-area. The bid-ask spreads observed in this period are slightly lower than those measured in the Bansal et al. (2021) sample.

the maximum dividend futures maturity used. The bottom panels of Figure A.1 show how the elasticity of the stock price with respect to expected dividends (left hand side) and with respect to expected returns (right hand side) change when Result 4 is implemented with either the 8-year or 5-year dividend future selected as the maximum maturity. Price elasticities are broadly consistent across specifications. As with the impact of bid-ask spreads, we therefore find that the sensitivity of results to adjustments in the maximum maturity of dividends futures is limited.

F. An observables approach based on the Campbell-Shiller log-linearization

The Campbell-Shiller (CS) log-linearization allows one to express the log price as a function of expected log dividends and expected log returns. This allows for a decomposition of realized log returns (and realized returns) based on changes to expected log dividends and to expected log returns (and thus log real yields and log equity risk premia when one breaks out expected log returns into components). After stating this well-known decomposition, we make two observations. First, that dividend futures can be used to guide the choice of ρ in the CS log-linearization. Second, that observable data on expected log returns can be used to implement the CS decomposition in practice. Therefore, while the two approaches provide fundamentally different decompositions (since one uses inputs in level and the other inputs in logs), our idea of implementing decomposition based on observables is feasible with either decomposition. Both are approximations, ours relying on the accuracy of Assumption 1 and the CS log-linearization relying on the accuracy of the log-linearization.

A. Capital gain and return decomposition using the CS log-linearization

The CS log-linearization starts from $r_t = \ln R_t$, $p_t = \ln P_t$, and $d_t = \ln D_t$, and the expression for the log-return

$$\begin{aligned} r_{t+1} &= \ln\left(P_{t+1} + D_{t+1}\right) - \ln P_t \\ &= \ln\left(P_{t+1}\left[1 + \frac{D_{t+1}}{P_{t+1}}\right]\right) - \ln P_t \\ &= p_{t+1} - p_t + \ln\left(1 + \exp\left(d_{t+1} - p_{t+1}\right)\right) \end{aligned} \tag{A.23}$$

This is log-linearized using a first-order Taylor approximation of $\ln(1+\exp(d_{t+1}-p_{t+1}))$ around a particular value of $d_{t+1}-p_{t+1}$, denoted by $\overline{d-p}$:

$$\ln(1+\exp(d_{t+1}-p_{t+1})) = k+(1-\rho)[d_{t+1}-p_{t+1}]$$

with

$$\rho = \frac{1}{1+\exp(\overline{d-p})}, \qquad k = -\ln\rho-(1-\rho)\ln\left(\frac{1}{\rho}-1\right) \tag{A.24}$$

Substituting back into (A.23) implies

$$r_{t+1} \approx p_{t+1}-p_t+k+(1-\rho)[d_{t+1}-p_{t+1}] \Longrightarrow p_t \approx \rho p_{t+1}+k+(1-\rho)d_{t+1}-r_{t+1} \tag{A.25}$$

Iterating forward, and assuming the no bubble condition that $\lim_{j\to\infty}\rho^j p_{t+j}=0$, and taking expectations,

$$p_t \approx \frac{k}{1-\rho}+\sum_{j\geq 0}\rho^j(1-\rho)E_t(d_{t+1+j})-\sum_{j\geq 0}\rho^j E_t(r_{t+1+j}). \tag{A.26}$$

It follows that the log of the capital gain from t to $t+1$ can be expressed as follows (the CS equivalent of our Result 1).[21]

$$p_{t+1}-p_t \approx \sum_{j\geq 0}\rho^j(1-\rho)[E_{t+1}(d_{t+2+j})-E_t(d_{t+1+j})]-\sum_{j\geq 0}\rho^j[E_{t+1}(r_{t+2+j})-E_t(r_{t+1+j})] \tag{A.28}$$

As for returns, focusing on the CS equivalent of our Result 2(a), the one-period log return on the stock market, $ln(R_{t+1})=r_{t+1}$, can be decomposed as $r_{t+1}\approx p_{t+1}-p_t+k+(1-\rho)[d_{t+1}-p_{t+1}]$ with $p_{t+1}-p_t$ from equation (A.28).[22] From (A.28), the capital gain resulting from a change to the generic expected log return k periods out is thus characterized by

[21] Similar to the G-factors in our approach, the expectations-updating terms in (A.28) have both "roll" and "news" components. This can be seen from rewriting as follows, using d as an example:

$$E_{t+1}(d_{t+2+j})-E_t(d_{t+1+j}) = [E_t(d_{t+2+j})-E_t(d_{t+1+j})]+[E_{t+1}(d_{t+2+j})-E_t(d_{t+2+j})] \tag{A.27}$$

Taking the expectation in (A.28), the expected log capital gain in the CS approach will capture the roll effects while the unexpected log capital gain will capture the news effects.

[22] Since (A.28) is in logs, the CS log-linearization implies a multiplicative decomposition of the level of the gross return:

$$R_{t+1} \approx exp\left(\sum_{j\geq 0}\rho^j(1-\rho)[E_{t+1}(d_{t+2+j})-E_t(d_{t+1+j})]\right) exp\left(-\sum_{j\geq 0}\rho^j[E_{t+1}(r_{t+2+j})-E_t(r_{t+1+j})]\right) exp(k+(1-\rho)[d_{t+1}-p_{t+1}]) \tag{A.29}$$

$$\psi_k^{R,CS} = \frac{p_{t+1} - p_t}{E_{t+1}\left(r_{t+k+1}\right) - E_t\left(r_{t+k}\right)} = -\rho^{k-1} \tag{A.30}$$

The capital gain resulting from a change to the generic expected log dividend k period out is characterized by

$$\psi_k^{D,CS} = \frac{p_{t+1} - p_t}{E_{t+1}\left(d_{t+k+1}\right) - E_t\left(d_{t+k}\right)} = \rho^{k-1}\left(1-\rho\right) \tag{A.31}$$

B. Implementing the CS decomposition based on observables, guided by dividend futures

It is well-known in the literature that the accuracy of the CS approximation depends on the variance of the dividend yield around its chosen value in the log-linearization parameter. While the CS linearization is accurate on average, Gao and Martin (2021) show the approximation error is largest when the dividend price ratio is far below its long-run mean. Figure A.3 plots the dividend yield since 1872, and shows that over our sample period of observable data (2004-2022), the dividend yield has been below its long-run mean, with the issue particularly acute over 2020, one of our main application periods.

Conceptually, the most accurate value of the log-linearization parameter uses the dividend yield investors expect going forward, i.e., a value that is close to the average expected ρ across future periods. If investors view the D/P ratio as a random walk, then the latest observed value will be a good choice. If they think the D/P ratio will revert to its historical mean, then the historical average may be better. We suggest a third possible choice: The value of the D/P ratio implied by the longest observed dividend futures contract (year 10).[23]

Assume constant expected growth and expected returns beyond year 10. Then, from year 10 on, $P = D\frac{G}{R-G}$ and thus $D/P = \frac{R-G}{G}$. From Result 4, the market's perception of $\frac{R-G}{G}$ can be calculated as $P_t^{(10)}/L_t$. Investors' perceived long-run dividend yield can therefore be inferred directly from dividend futures prices as $P_t^{(10)}/L_t$. Figure A.3 bottom panel provides a time series plot of the three choices for D/P ratios. It is clear that investors' expected long-run D/P ratio based on dividend futures is substantially below the historical average up to the start of the sample and closer to the latest D/P ratio, with notable exceptions in 2008 and 2009 following large changes to the latest D/P ratio (suggesting these were perceived as somewhat transitory). Given their forward-looking nature, using dividend futures to guide the choice for ρ is likely to be

[23] An alternative would be to allow ρ to vary by horizon and estimate a statistical model for its evolution based on historical data. This approach is used in Gonçalves (2021c) via a VAR framework.

most accurate.

Using data as of the start of year 2020, Figure A.4 illustrates how log capital gain sensitivities vary with the choice of the log-linearization parameters. The left figure shows the sensitivity of the log capital gain to future expected log dividends, and the right figure shows the sensitivity of the log capital gain to future expected log returns. The latter is quite sensitive to the chosen value of ρ.

The result from our preferred approach (using ρ's based on dividend futures prices) are shown in red. The dividend futures imply a certain amount of mean-reversion in the D/P ratio, as demonstrated by the dividend future implied log-price sensitivities falling somewhere between the two other approaches. The largest differences between the various approaches are apparent in the log capital gain sensitivity with respect to expected returns, especially at longer maturities. For example, the sensitivity with respect to the expected return in year 30 is -0.60 when basing ρ_{t+j} on the January 2, 2020, D/P ratio, -0.29 when ρ_{t+j} is based on the historical average D/P ratio, and -0.53 when ρ_{t+j} is based on dividend future implied D/P ratios. To the extent that long-dated expected log returns vary, the different assumptions will therefore lead to different stock return decomposition results when using a CS-based approach. The large gap between approach 1 and 2 for determining values for ρ emerges because dividend yields have fallen a long way below their historical averages in recent decades (see Figure A.3 top panel). This makes the duration of the stock market higher today than it has been previously.

In terms of observable inputs for the CS approach, one would decompose expected log returns, $E_t r_{t+k}$, as the sum of the forward log real risk-free rate for year t+k and the expected log excess returns (exploiting $r = ln(R) = ln\left(R^f \times \frac{R}{R^f}\right) = r^f_{t+k} + \left(r_{t+k} - r^f_{t+k}\right)$). Gao and Martin (2021) provide a lower bound on expected log excess returns (which one could again assume is tight). As for the log risk-free rate, it is just the log of the forward real risk-free rate. To save space, we do not provide application results based on the dividend-futures guided CS log-linearization. As in our decomposition of returns, data is only available out to various maximum horizons when implementing the CS decomposition, so the empirical implementation of the CS-based approach will again have a long-term news component that captures news past these horizons.

G. Appendix Figures

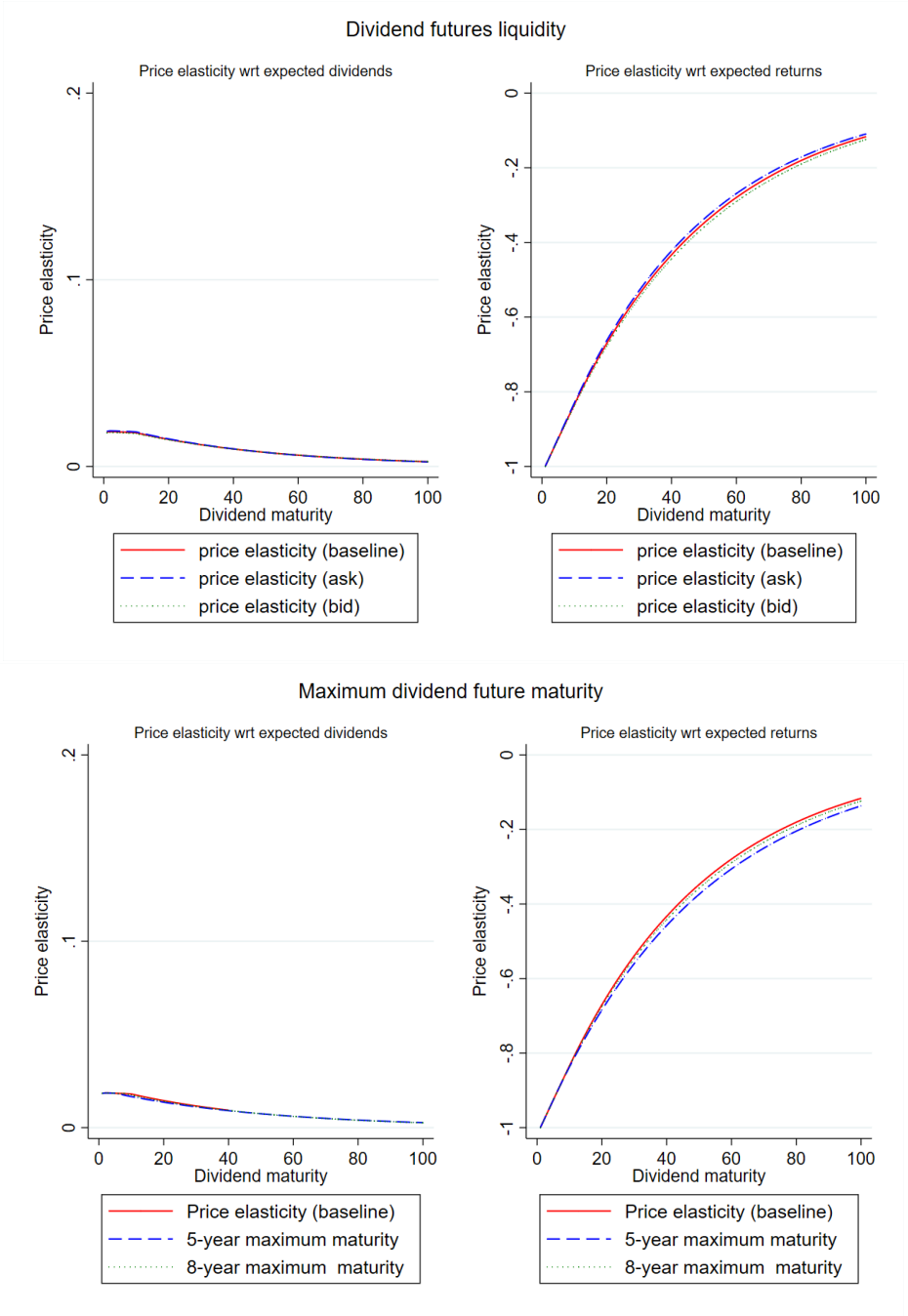


Figure A.1. Price elasticities and dividend futures liquidity.

This figure shows how the price elasticity of the stock market with respect to expected returns changes with different measures of dividend strips, and how these variations in elasticities impact on the estimation in yield curve news over 2020. The elasticities are measured on the first day of 2020. The different measures of dividend strips are adding / subtracting a bid-ask spread (top panels), reducing the maximum maturity of dividend strip used (middle panels) and using different log-linearization parameter choices (bottom panel).

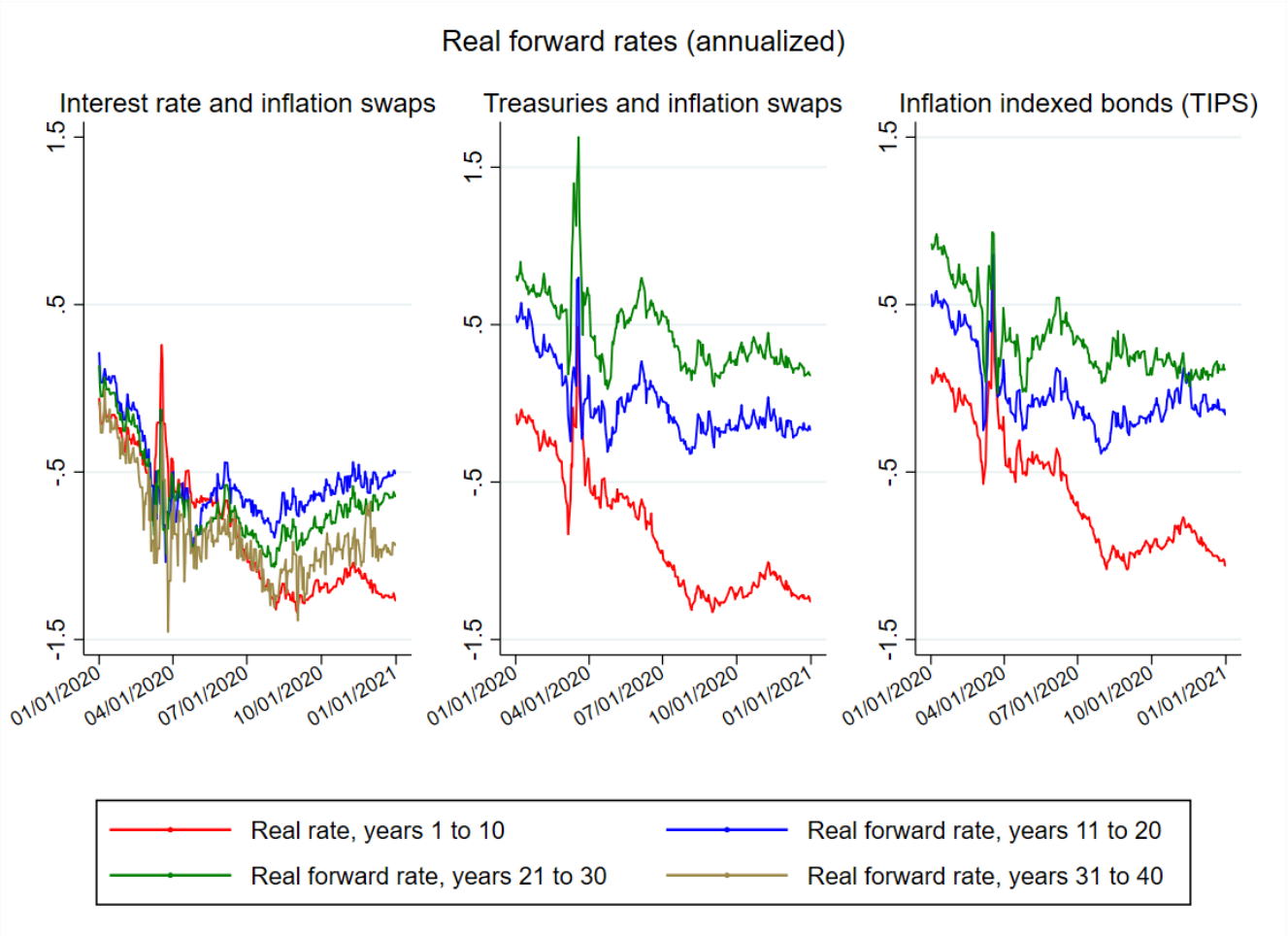


Figure A.2. Observed forward real yields

This figure shows real forward rates over 10 year periods as observed from swaps, Treasuries and swaps, and TIPS respectively.

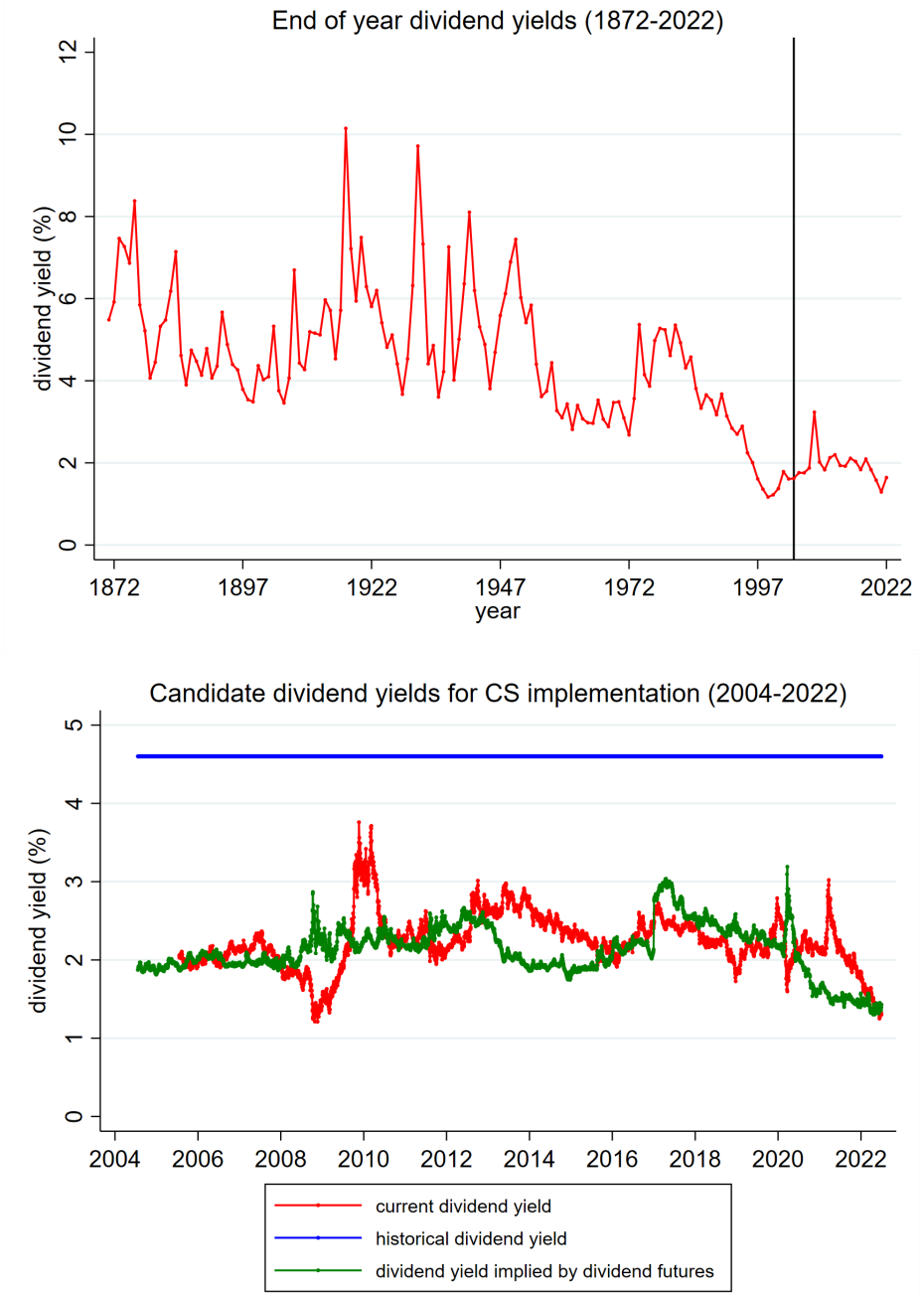


Figure A.3. Dividend yield analysis.

The top panel of this figure shows the dividend yield of the US Stock market at the end of each year from 1872 to 2022 using data from Shiller's website. The black vertical line indicated the start of our main sample of observable data in 2004. For the sample 2004-2022, at a daily frequency the bottom panel shows the current dividend yield, the historical average of the dividend yield from 1872-2004 (4.6%), and the dividend yield implied by dividend futures (using Result 4).

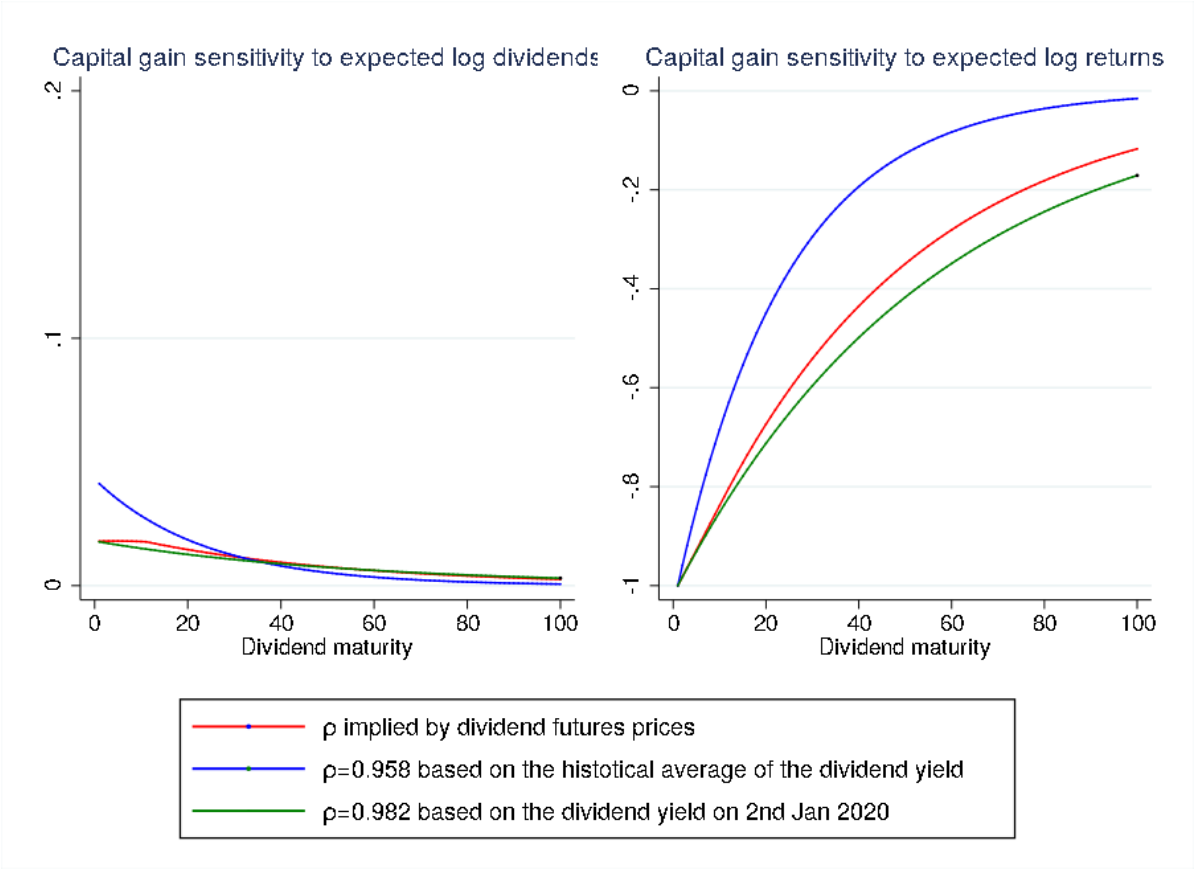


Figure A.4. Log price sensitivities using Campbell-Shiller log-linearization.

This figure shows the sensitivity of the log stock price at t with respect to expected log dividends in year $t+k$ (left panel) and the sensitivity of log stock price at t with respect to the expected log return in year $t+k$ (right panel). In each panel, we plot the sensitivity for three different choices of the future dividend yields, $\overline{D_{t+k}/P_{t+k}}$, from which to compute the Campbell-Shiller log-linearization parameter $\rho_{t+k} = \frac{1}{1+\overline{D_{t+k}/P_{t+k}}}$. For the first sensitivity, we use the historical average of the dividend yield ($\overline{D/P} = 4.2\%$, $\rho = 0.958$) for all future dividend yields. For the second sensitivity, we use the dividend yield at the start of our main sample on January 2nd 2020 ($\overline{D/P} = 1.8\%$, $\rho = 0.982$) for all future dividend yields. For the third sensitivity, we use the long-run dividend yield implied by dividend futures prices (see Result 4) for dividend yields beyond year $t+10$, and linearly interpolate from the current dividend yield to the long-run dividend yield for the implied dividend yields from year $t+1$ to year $t+10$.