Duration-Based Stock Valuation: Reassessing Stock Market Performance and Volatility*

Jules H. van Binsbergen†

November 6, 2020

Abstract

Using a panel of international government bond data, I construct fixed income portfolios that match the duration of the dividend strips of the corresponding local aggregate stock market index. I find that these bond portfolios have performed as well as — if not better than — their stock counterparts in the past half century while exhibiting similar (or even higher) levels of volatility. These results provide a novel perspective on both the equity risk premium and excess volatility puzzles (bubbles). I present several potential explanations, and discuss further the implications for macroeconomics, monetary economics, asset pricing, and corporate finance.

*I thank Anat Admati, Jonathan Berk, Martijn Boons, Peter DeMarzo, Darrell Duffie, Vincent Glode, Joao Gomes, Andrei Goncalves, Bob Hodrick, Hanno Lustig, Max Miller, Christian Opp, Anamaria Pieschacon, Josh Rauh, Michael Roberts, Riccardo Sabbatucci, Andrea Tamoni, Myron Scholes, Ken Singleton, Jessica Wachter and seminar participants at Stanford GSB, University of Texas at Dallas and the Vienna Institute for Finance, Banking and Insurance for helpful comments and suggestions.

†Wharton, University of Pennsylvania, NBER and CEPR; julesv@wharton.upenn.edu.

1. Introduction

In the past several decades interest rates have been on a steady decline, reaching all-time low levels around the world. As an illustration, consider the 100-year Austrian government bond that was issued on September 20th 2017. This bond was issued at par at a 2.1% coupon rate with a duration of 42 years. Since that issuance, the bond has exhibited very large price volatility and its price has more than doubled (as of October 22nd 2020). Its yield-to-maturity has fallen below 1% increasing its duration to over 50 years. Given that interest rates of even these extended durations have dropped to levels this low, the valuation and realized returns of other long duration assets are likely affected as well. This raises the question what part of the increase in the prices of stocks, including the recent COVID-19 recovery, is driven by this downward path of long-term interest rates. After all, stocks have a comparable duration and cashflow-to-price ratio as the 100-year Austrian government bond described above, and so one may have expected similar performance.[1]

Answering this question is also important given potential concerns that central banks have been inflating asset values, including stocks, through accommodating monetary policy. As the corresponding decline in discount rates cannot continue much longer due to a lower bound on nominal interest rates, the associated valuation windfalls are unlikely to repeat themselves.[2] This makes historical (average) realized returns less informative about future expected returns for all long duration assets. One may thus also wonder how useful these backward-looking data moments are to fit stationary macro and asset pricing models.

To quantitatively evaluate the importance of interest rate dynamics for stock markets, I compute several counterfactual fixed income portfolios to address the following question. What would the returns of an investor have been, if instead of investing in the local stock market index (e.g. the S&P500 index), that investor had invested in a portfolio of government bond strips whose duration matches the dividend strips that make up the index? This investor would have avoided dividend risk altogether (in nominal terms) and would thus not have earned the risk premium (if any) associated with such risk.

Starting with U.S. data between 1996 and 2020, when inflation expectations were low and stable, I show that this fixed income investor would have achieved at least similar and likely substantially better return performance than an investor who invested in the stock

[1] Over the same sample period, the Austrian stock market has fallen by about 30%.

[2] While recently observed nominal interest rates suggest that the lower bound is not actually at zero, the existence of physical currency should enforce at least some bound on how negative nominal interest rates can go.

market index. Put differently, without the decline in interest rates, ceteris paribus (keeping growth rates and dividend risk premia unaffected), the stock market would have lost value.[3] I then show that these results also hold for Europe and Japan. We can conclude that over this sample period, global investors have received little to no additional return compensation for taking long-duration dividend risk compared to nominally risk free government bonds.

In addition to these surprising results for the first moment of returns, the results for second moments also seem puzzling. The return volatility of the duration-matched fixed income portfolios, which have fixed cash flows in nominal terms, is comparable to (if not higher than) that of the stock market indices, which have time-varying cash flows. This challenges the often-held belief that dividend risk premium variation is required to obtain the observed "excess" volatility in stock returns, a notion introduced by Shiller (1981). If anything, stock returns seem too little volatile, not excessively volatile, once compared to duration-matched bond counterfactuals. Dividend growth variation (expected and unexpected) as well as risk premium variation (if any) affect stocks but not bonds, and seem to have an offsetting effect on the stock price volatility induced by long-duration risk free rate variation.

When expanding the sample to 1970-2020, which includes the turbulent inflationary period of the late seventies and early eighties, I find similar results in that (1) the realized average compensation for dividend risk is no more than about 1% per year and arguably substantially less (negative), and (2) the volatility of the counterfactual fixed income portfolios is higher than that of the corresponding stock markets. I then discuss several explanations for this seeming stock market underperformance and low volatility including (1) a secular decline in long-term expected nominal and real growth rates (e.g. Gordon (2016)), (2) an increase in the dividend risk premium going forward (Farhi and Gourio (2018)), (3) the diversification of dividend risk across maturities (Barro, Nakamura, Steinsson, and Ursua (2011)), and (4) the inflation-hedging properties of long-term dividends. Given that (a) the results hold for the 1996-2020 period when inflation expectations and realizations were low and stable, and (b) the literature has argued that the equity risk premium has been declining (not increasing) in the past few decades (e.g. Lettau, Ludvigson, and Wachter (2004) and Jagannathan, McGrattan, and Scherbina (2000)), explanations (1) and (3) seem particularly interesting to explore further.

The results presented in this paper have several implications for theoretical macroeconomic and asset pricing models. The literature has spent a considerable amount of time and effort studying four major puzzles. First, the average return on stocks has been substantially

[3] Of course, in equilibrium, growth rates and interest rates are tied, as further discussed in Section 9.

higher than that of a short-duration fixed income instrument, the so-called equity premium puzzle (Hansen and Singleton (1983), Mehra and Prescott (1985)). Second, the variation in stock prices is larger than the variation in dividends, the so-called excess volatility puzzle (Shiller (1981)). Third (and related to the second puzzle), risk premia are believed to vary a lot over time leading to the predictability of excess returns (Campbell and Shiller (1988)). Fourth, the Capital Asset Pricing Model (Sharpe (1964)) seems to do a poor job explaining the average return differentials between stocks in the cross-section.

The recent literature on dividend strips argues that all these four puzzles occur when pricing short duration dividend claims with maturities up to 7 years.[4] Using data starting in 1996, this literature finds that (1) the average returns on short-term dividend strips are higher than their corresponding government bond strips, and higher than what existing models predicted (an equity premium type puzzle), (2) dividend strip prices are more volatile than their corresponding dividend realizations at maturity (excess volatility) and (3) the excess returns on dividend strips are predictable. Finally, given that the CAPM beta of short duration strips is low, that literature finds that (4) the average returns on short-term strips are higher than what the CAPM predicts. These results on short-term dividend claims were important particularly in the context of the theoretical advances that rely on the long duration nature of the aggregate dividend claim to explain the four puzzles above.[5] That is, if in the data dividend strips exhibit the same puzzling behavior as the stock index, but the models can only explain that behavior for the index (the long-duration claim), the key mechanism that resolves the puzzles may not be right.

This paper adds to the literature on the term structure of equity by showing that long duration dividend risk has received little to no compensation over the past half century and that duration-matched fixed income portfolios already exhibit similar (if not higher) volatility as the aggregate stock market. Taken all these results together, it could be fruitful to more closely examine the forces that drive the excess volatility/return predictability and risk premium puzzles for short duration equity claims, which are not affected by the secular trend and fluctuations in long-term interest rates. Duration-related explorations of the cross-section of stock returns, as further discussed in Section 9.2, are another important avenue to make progress on these questions.[6]

[4] See Brennan (1998), Binsbergen, Brandt, and Koijen (2012), Binsbergen, Hueskes, Koijen, and Vrugt (2014) and Binsbergen and Koijen (2017).

[5] See for instance Campbell and Cochrane (1999), Bansal and Yaron (2004) and Bansal, Kiku, and Yaron (2006).

[6] See Goncalves (2020) and Gormsen and Lazarus (2020) for recent contributions in this area.

The results presented in this paper also have implications for the literature on defined benefit pension plans. These plans have been in an underfunding crisis for the past decades around the world.[7] Defined benefit plans make long duration, often risk free, promises (liabilities) to pension holders. They have been trying to invest their way out of their underfunded status through high exposure to equities. This exposure was motivated by the idea that the return differential implied by the equity risk premium would help them make up for their asset shortfalls. The results in this paper show that this strategy has not worked for the past 50 years. After all, under the assumption that the duration of the risk free promises (bonds) is the same as the duration of the equity market, the return differential these pension plans were betting on is exactly the long duration dividend risk premium I compute in this paper. This compensation for dividend risk has failed to materialize in the past five decades.

The paper proceeds as follows. In Section 2 I lay out the theoretical foundation for the empirical measurement. I describe the data sources in Section 3 and explore a set of constant maturity zero coupon bonds of varying maturities as counterfactuals for the stock market in Section 4 focusing on the sample period 1996-2020 when inflation realizations and expectations were low and stable. I then explore bond portfolios whose portfolio weights are determined by the weights of dividend strips in the index in Section 5. I study the sample from 1970 through 2020 in Section 6 and explore international data from Europe and Japan in Section 7. I discuss excess volatility in Section 8 and present several potential explanations in Section 9. I discuss the importance for the cross-section of stock returns and the implications for capital structure research in Section 9.2. I verify the accuracy of the bond data against tradable bond portfolio data from Vanguard in Section 10. In that section I also show that the results are not particularly dependent on government bond portfolios and are also obtained using corporate debt portfolios. I conclude in Section 11 with a stock market outlook based on the Japanese experience of the last 25 years.

2. Duration Matching

Let $\mathcal{P}_{t,n}$ denote the present value at time t of the expected dividend paid out at time $t+n$. That is:

$$\mathcal{P}_{t,n} = \frac{E_t\left[D_{t+n}\right]}{\exp\left(n\left(y_{t,n}+\theta_{t,n}\right)\right)}, \tag{1}$$

[7] See for example Novy-Marx and Rauh (2011) and Binsbergen and Brandt (2015).

where $y_{t,n}$ denotes the continuously compounded risk free spot interest rate at time t for maturity n, and $\theta_{t,n}$ denotes the normalized (by n) dividend risk term premium for a dividend of that maturity (Binsbergen, Koijen, Hueskes and Vrugt (2015)). The stock index S_t is a portfolio that includes one unit of each dividend strip:

$$S_t = \sum_{n=1}^{\infty} \mathcal{P}_{t,n}. \tag{2}$$

Define $w_{t,n}$ as the weight that each strip has in this index portfolio:

$$w_{t,n} = \frac{\mathcal{P}_{t,n}}{S_t}. \tag{3}$$

The one-period return on a dividend strip with maturity n is given by:

$$R^d_{t+1,n} = \frac{\mathcal{P}_{t+1,n-1}}{\mathcal{P}_{t,n}} - 1, \text{for } n > 1, \tag{4}$$

$$R^d_{t+1,n} = \frac{D_{t+1}}{\mathcal{P}_{t,n}} - 1, \text{for } n = 1. \tag{5}$$

The one-period return on the n-year bond is given by:

$$R^b_{t+1,n} = \frac{\exp\left(-(n-1)y_{t+1,n-1}\right)}{\exp\left(-ny_{t,n}\right)} - 1, \tag{6}$$

for which the conditional expectation is given by:

$$\mu^b_{t,n} = E_t\left[R^b_{t+1,n}\right]. \tag{7}$$

The additional expected holding period return $\psi_{t,n}$ that an investor who takes dividend risk earns in excess of the return on the corresponding risk free bond is defined as:

$$\psi_{t,n} \equiv E_t\left[R^d_{t+1,n}\right] - \mu^b_{t,n}. \tag{8}$$

The premium $\psi_{t,n}$ is different from $\theta_{t,n}$, as the latter fits the usual yield-to-maturity definition, whereas $\psi_{t,n}$ is the one-period expected return over and above the one-period expected return on the maturity-matched risk free bond.

The one-period return on the index is given by:

$$R^s_{t+1} = \frac{S_{t+1} + D_{t+1}}{S_t} - 1. \tag{9}$$

Because the return on the index is a weighted average of the returns on the strips, the expected return on holding the index for one period can be written as the weighted average of expected returns on the strips:

$$\mu_t^s = E_t\left[R_{t+1}^s\right] = \sum_{n=1}^{\infty} w_{t,n} E_t\left[R_{t+1,n}^d\right]. \tag{10}$$

Given these definitions, we can now define the main object of interest in this paper. It is the unconditional average returns that the index-implied portfolio of risky dividends of all maturities earn over and above their government bond counterparts:

$$\Psi_0 = E\left[E_t \sum_{n=1}^{\infty} w_{t,n}\psi_{t,n}\right] = E\left[\mu_t^s - E_t \sum_{n=1}^{\infty} w_{t,n}\mu_{t,n}^b\right]. \tag{11}$$

I estimate this quantity through:

$$\hat{\Psi}_0 \approx \frac{1}{T}\sum_{t=1}^{T}\left[\frac{S_{t+1}+D_{t+1}}{S_t} - \sum_{n=1}^{N} w_{t,n} R_{t+1,n}^b\right]. \tag{12}$$

There are only two inputs to this calculation that are not obvious from the data. First there is the weighting scheme $w_{t,n}$ which is not observable beyond the available dividend strip data. However, it is not hard to make informed choices on this weighting schemes for longer maturities given the available short-maturity data in combination with standard stock valuation formulas. To allay any concerns about this input I will conduct a large number of sensitivity analyses regarding these weights.

Second, there is the cutoff (CO) point N. Choosing N involves an important tradeoff. On the one hand, high quality term structure data is not available for the U.S. beyond 30 years of maturity and such yields are mainly based on extrapolation. On the other hand, a substantial fraction of a stock index's value is represented by strips beyond 30 years. One way to resolve this issue is to simply assign all the remaining weights beyond 30 years to this 30-year government bond strip. Alternatively, one could allow for positive weights beyond 30 years at the risk of using Nelson Siegel extrapolated term structure data. While I will conduct both analyses in this paper, to be conservative, I will use the former approach as the baseline analysis. As longer and longer maturity bonds are included in the calculation, the estimated long duration dividend risk premium $\hat{\Psi}_0$ becomes increasingly negative.

The usual stationarity assumption in asset pricing would imply that $\hat{\Psi}_0$ is informative about the expected value outside this estimation window. Later, I will discuss further the conditions under which this is true and under which this is likely not true. If the secular trend in risk free interest rates affects both stock returns and fixed-income instruments equally, the implied non-stationarities could at least partially cancel in the computation of $\hat{\Psi}_0$. However, this is potentially not true if there are also secular trends in long-term dividend growth rates and/or risk premia, as discussed in Section 9. One way to alleviate concerns related to the external validity of results for the U.S. and Europe is to study Japanese data spanning the period 1996-2020 as interest rates were already at very low levels in Japan in 1996. This analysis is presented in Section 7.

3. Data

Data on the S&P500 index are obtained from Global Financial Data. I use both the total return index and the price index. Monthly dividends are computed in the standard way by taking the difference between the monthly total return on the index and the monthly price appreciation of the index, multiplied by the lagged index level. To construct zero coupon bond strips, I use the updated term structure data provided by the Federal Reserve following the approach by Gurkaynak, Sack, and Wright (2006). I also use monthly return data for tradable long duration bond index funds from Vanguard as an additional source for long-term bond return data. I use these returns to verify (where possible) the high accuracy of the implied bond returns that follow from the yield data provided by Gurkaynak, Sack, and Wright (2006), as illustrated in Section 10.1. Data on zero coupon bond yields for Germany (Nelson-Siegel-Svenson parameters) and Japan are from the Bundesbank and the Bank of Japan, respectively. Dividend strip data uses the same data sources and procedures as in Binsbergen and Koijen (2017) and exchange traded data from Bloomberg since 2015.

4. Constant Duration Counterfactuals

To get a first sense of stock versus bond performance, I compare in this section the performance of the stock market to a set of constant maturity bond portfolios. To facilitate the comparison it is helpful to compute a reasonable range for the duration of the stock market, which I present in section 4.1. Section 4.2 then presents the comparison of returns. In the next section (5) I then compute strip-matched counterfactuals.

4.1. Duration of the Stock Market

The duration (Dur) of an asset is commonly computed as the ceteris paribus percentage change of the value of the asset for a 1 percentage point decrease in its discount rate. This is approximately equal to the weighted average time it takes for the asset to return the discounted cash flows to its owner:

$$\text{Dur} = \sum_{n=1}^{\infty} w_{t,n} n. \tag{13}$$

Equation 13 shows that data on the weighting scheme $w_{t,n}$ is sufficient to compute the duration of the stock market in each period.

Alternatively, we can use the static Gordon growth equation to compute the duration. This equation expresses the value of the stock market index S_t as a function of the dividend D_t, the expected return on the index μ^s and the expected growth rate g:

$$S_t = \frac{D_{t+1}}{\mu^s - g}. \tag{14}$$

If we take the natural logarithm of both sides of Equation 14 and compute the derivative with respect to μ^s, we obtain:

$$\text{Dur} \approx -\frac{\partial \ln S_t}{\partial \mu^s} = \frac{\partial \ln(\mu^s - g)}{\partial \mu^s} = \frac{1}{\mu^s - g} = \frac{S_t}{D_{t+1}}. \tag{15}$$

The equation shows that the duration equals the inverse of the dividend yield. In the U.S. in the past few decades this dividend yield has varied between 2% and 3% corresponding to a duration between 33.3 and 50 years. For a dividend yield of 6%, which briefly occurred several decades ago, the duration is 16.7 years. For the samples that I study, durations between 20 and 50 years seem reasonable. I will provide further empirical evidence to support this claim in Section 5.2.

One could wonder why the duration of the stock claim is so high, given that the duration of the physical capital stock of firms appears to be much lower. After all, particularly in recent years, the depreciation rate on capital has been substantial (e.g. IT equipment), lowering the overall duration of the capital stock.[8] The answer is twofold. First, owning equity on a firm entitles the owner not just to the current capital stock, but to all its future incarnations as well (growth options). Second, the duration of the whole firm is a weighted

[8] I thank Larry Summers for this suggestion.

average of the duration of corporate debt and equity. As such, for a given maturity of the assets of the firm (the left-hand side of the balance sheet) issuing short-term debt, increases the duration of the equity of the firm.

Another consideration is whether the present value formula should be based on cash dividends alone, or on total payout that includes issuances and repurchases. The present-value relationship for the S&P500 index should hold both on a per-share basis (i.e. a buy-and-hold investor who invests in the index and collects its dividends), as well as on a total equity value (not per share) basis that takes into account *both* issuances and repurchases. Repeating the calculations in this paper for the total equity value of the U.S. stock market that includes the net of repurchases and issuances is an interesting avenue for future research. That said, equity ownership stakes entitle the holder to a long stream of cash flows. Even if the owner wishes to sell this stake, the new owner will value the remaining stream of cash flows taking into account its duration. As such, the appropriate counterfactual bond portfolio will necessarily include longer duration claims that have performed substantially better than their short-duration counterparts. It seems unlikely that the counterfactual bond portfolio should have a duration shorter than 16 years or alternatively have weights corresponding to a payout yield of more than 6%, though future research in this direction may prove otherwise.

Maturity in years	FF	1	1.5	2	3	4	5	10	15	20	25	30	S&P500
Mean	0.0018	0.0023	0.0026	0.0029	0.0034	0.0039	0.0044	0.0062	0.0075	0.0086	0.0096	0.0109	0.0079
St. Dev.	0.0017	0.0026	0.0036	0.0048	0.0076	0.0104	0.0131	0.0260	0.0373	0.0469	0.0561	0.0676	0.0439
Mean log	0.0018	0.0023	0.0026	0.0028	0.0034	0.0038	0.0043	0.0059	0.0069	0.0075	0.0081	0.0086	0.0068

Table 1

Monthly Returns on Constant Maturity Zero Coupon Bonds. The second row in the table lists the average monthly bond returns ($\hat{\mu}_n^b$) on constant maturity zero coupon bond strategies for maturities ranging between 1 year and 30 years using monthly data between January 1996 and April 2020. The third row reports the monthly standard deviation. The second column lists of the table lists the corresponding statistics for the risk free return as used in the Fama French model and the last column lists those statistics for the S&P500 index.

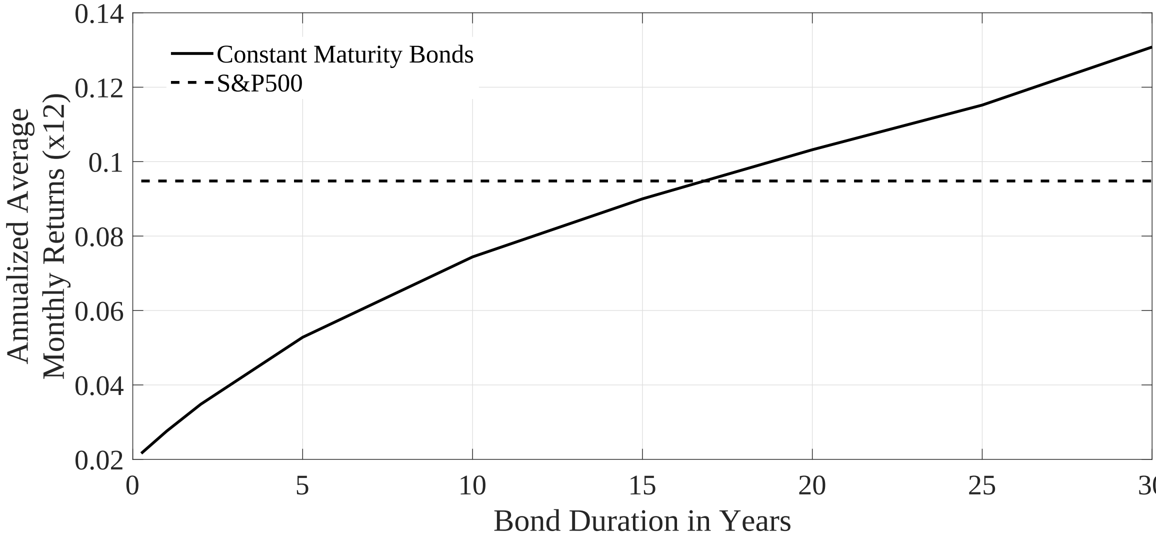

FIGURE I

Annualized Average Monthly Returns: Bonds vs Stocks The graph plots the annualized ($\times 12$) average monthly bond returns ($\hat{\mu}_n^b$) on constant maturity zero coupon bond strategies for maturities ranging between 1 year and 30 years using monthly data between January 1996 and April 2020. The maturity (duration) of the bonds is on the x-axis and the average return is on the y-axis. The dashed line represents the annualized average monthly return on the S&P500 index.

4.2. Empirical Results on Constant Duration Counterfactuals

To get a first sense of how well the index has performed relative to long duration bonds, I first present the average returns of constant maturity bond strategies of maturities varying between 1 and 30 years. The results are reported in Table 1 and plotted in Figures I and II. The second row in the table lists the average monthly bond returns ($\hat{\mu}_n^b$) on constant maturity zero coupon bond strategies for maturities ranging between 1 year and 30 years using monthly data between January 1996 and April 2020. The third row reports the monthly standard deviation. The second column of the table lists the corresponding statistics for the risk free return as used in the Fama French model (from Ken French's website) and the last column lists the corresponding statistics for the S&P500 index. The annualized mean returns are plotted in Figure I and the annualized volatilities are plotted in Figure II.

The table and the figures show that both the average return and the volatility of the stock index equals that of a constant maturity zero coupon bond strategy with a duration of about 17 years. Constant maturity bond strategies with higher durations than 17 years have higher returns and volatilities than the stock market. Therefore, compared to these

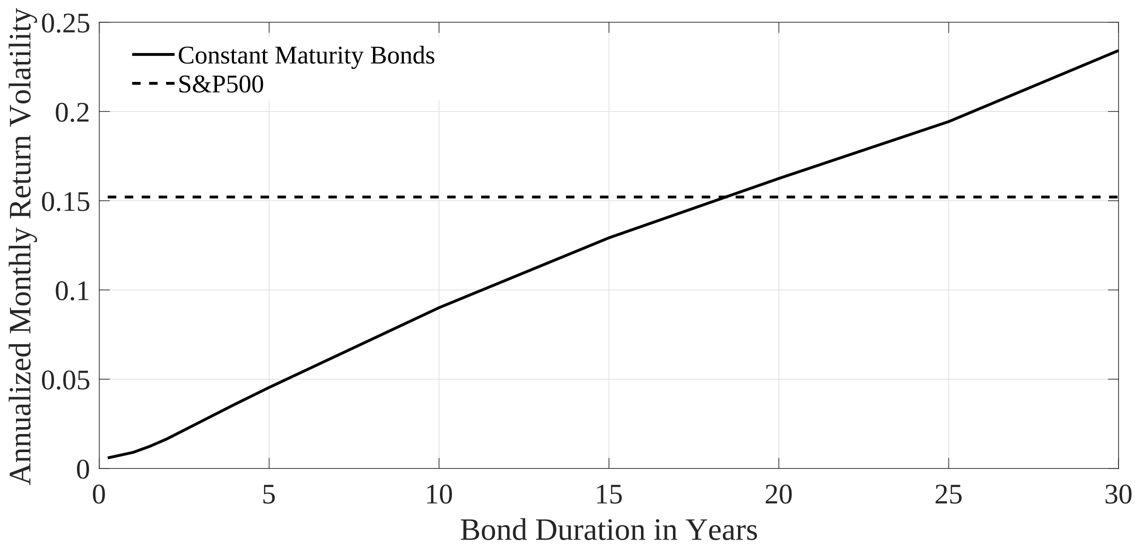

FIGURE II

Annualized Volatility Monthly Returns: Bonds vs Stocks The graph plots the annualized ($\times\sqrt{12}$) volatility of monthly bond returns on constant maturity zero coupon bond strategies for maturities ranging between 1 year and 30 years using monthly data between January 1996 and April 2020. The maturity (duration) of the bonds is on the x-axis and the volatility is on the y-axis. The dashed line represents the annualized monthly return volatility on the S&P500 index over the sample.

constant maturity zero coupon bond strategies, over this sample period, stocks only have puzzlingly high returns and volatilities (the equity premium and excess volatility puzzles) if one believes that the duration of the stock market is substantially less than 16 years. As argued in the previous subsection, a stock market duration of less than 16 years corresponds to a dividend yield larger than 6%, which was not observed over this sample period.

In Table 2 I compute the difference between the average monthly returns on the S&P500 index ($\hat{\mu}^s$) and those of the corresponding long-term bond portfolios ($\hat{\mu}^b_n$). The table shows that the average annual realized compensation for dividend risk between 1996 and 2020 is low. If we take the 10-year government portfolio as the counterfactual, the estimated return differential is +2% per year. If we use the 15-year constant maturity bond as the counterfactual the average return differential shrinks to +47b.p. per year. For longer maturity counterfactuals, the difference turns negative and drops to -3.55% per year for the 30-year duration counterfactual bond portfolio. None of these return differentials are statistically significant. For completeness, the last row shows the annualized difference in log returns, which exhibits a similar pattern. However, it should be noted that average simple returns are the preferred measure to estimate expected returns, not mean log returns.

Duration in years	10	15	20	25	30
$\hat{\mu}^s - \hat{\mu}^b$	0.0017	0.0004	-0.0007	-0.0017	-0.0030
t-stat on difference	0.51	0.11	-0.16	-0.37	-0.56
$12(\hat{\mu}^s - \hat{\mu}^b)$	0.0200	0.0047	-0.0079	-0.0204	-0.0355
Annualized difference in mean log returns	0.0110	-0.0005	-0.0085	-0.0153	-0.0216

Table 2

Monthly Return Differences between the S&P500 and Constant Maturity Zero Coupon Bonds. The the second row in the table lists the difference between the monthly returns on the S&P500 index and the monthly returns on constant maturity zero coupon bonds ($\hat{\mu}_n^b$) for maturities ranging between 10 year and 30 years using monthly data between January 1996 and April 2020. The third row reports the t-statistic on the difference. For ease of interpretation, the third row reports the annualized difference (by multiplying by 12). The last row reports the annualized difference in the means of the monthly log returns (instead of simply returns).

5. Strip-Matched Counterfactuals

In this section I discuss several weighting schemes as defined in Equation 3. In particular, I compute several Gordon-growth-model-implied weighting schemes and compare them to the available dividend strip data.

5.1. Weighting Schemes

The previous section contained several comparisons of stock returns with constant maturity bond returns of varying maturities. In this section, I construct several simple strip replicating fixed income portfolios. I focus on eight main cases. As a first simple weight-generating model, consider the static Gordon growth model:

$$S_t = \frac{D_t(1+g)}{\mu^s - g} \tag{16}$$

where g is the dividends' growth rate. The n-th strip value is given by:

$$\mathcal{P}_{t,n} = D_t \left(\frac{1+g}{1+\mu^s} \right)^n, \tag{17}$$

which implies a weight equal to:

$$w_{t,n} = (\mu^s - g)\,\frac{(1+g)^{n-1}}{(1+\mu^s)^n}. \tag{18}$$

As the length of the holding period return converges to 0 (i.e., going from annual, to monthly to daily returns), this weighting scheme is only a function of $\mu^s - g$ (the difference), and not of μ^s and g separately. As I employ a monthly holding period return, this property effectively holds already. A second important choice is the point of the cutoff. I consider two cutoff levels, one at $CO = 360$ months (30 years) and one at $CO = 480$ months (40 years). I set the portfolio weight of this cutoff point to the sum of the remaining weights:

$$w_t^{CO} = \sum_{n=CO}^{\infty} w_{t,n} \tag{19}$$

The three graphs below plot the weighting schemes aggregated up to annual weights for $\mu^s - g = 0.06$ (in Figure III), $\mu^s - g = 0.03$ (in Figure IV) and $\mu^s - g = 0.02$ (in Figure V) for these two cutoff levels. The available data on dividend strips (when available) produces an average weighting scheme between ($\mu^s - g = 0.02$) and ($\mu^s - g = 0.03$) as further explored later. The first weighting scheme ($\mu^s - g = 0.06$) applies too much weight to the short duration assets relative to the available dividend data. The final points on each curve represent the cutoff weights described in Equation 19.

The graphs show important differences between the curves. The duration of the low dividend yield scenario ($\mu^s - g = 0.02$) is substantially higher than that of the high dividend yield scenario ($\mu^s - g = 0.06$), though the duration of both is capped through the imposition of the cutoff at 30 or 40 years. Even higher cutoff points (and thus durations) could certainly be justified from a stock pricing perspective, though I am limited by the available bond term structure data. It is important to keep in mind though that because of these cutoff points, the estimates I present will be conservative in that the true dividend risk premium is arguably even lower than what my estimates suggest.

The combination of two dividend yield levels and two cutoffs provides 6 counterfactual scenarios, labeled I-VI. In addition, I explore a time-varying weighting scheme that uses the real-time dividend yield on the index as the value for $\mu^s - g$. That is, in each month, I take the ratio of the sum of the past twelve monthly dividends and divide them by the index level. I then generate in each month Gordon growth model-implied weights that are consistent with that months dividend yield and apply those portfolio weights to the next

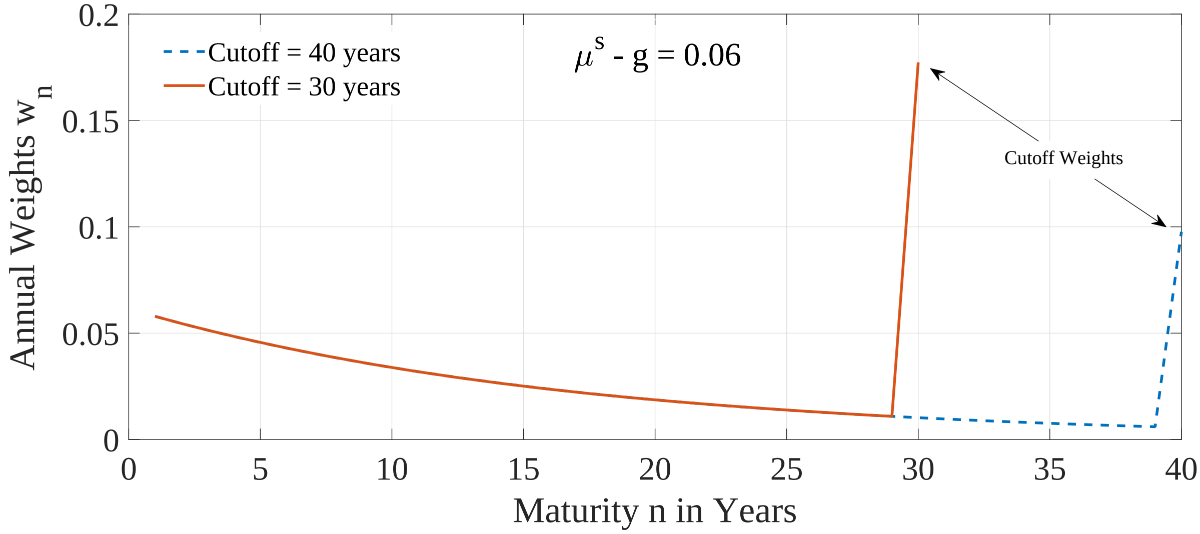

FIGURE III

Strip Weights by Year: High Dividend Yield The graph plots the weights implied by the Gordon growth formula for $\mu^s - g = 0.06$. The values for μ^s and g used are 0.12 and 0.06 (both scaled by 12 to arrive at monthly numbers), though up to a first order approximation, the weights are only dependent on the difference between the two. Two cutoffs are considered: 30 years and 40 years. That is, the weight at 30 years (40 years) equals the sum of all weights with maturity 30 (40) or higher.

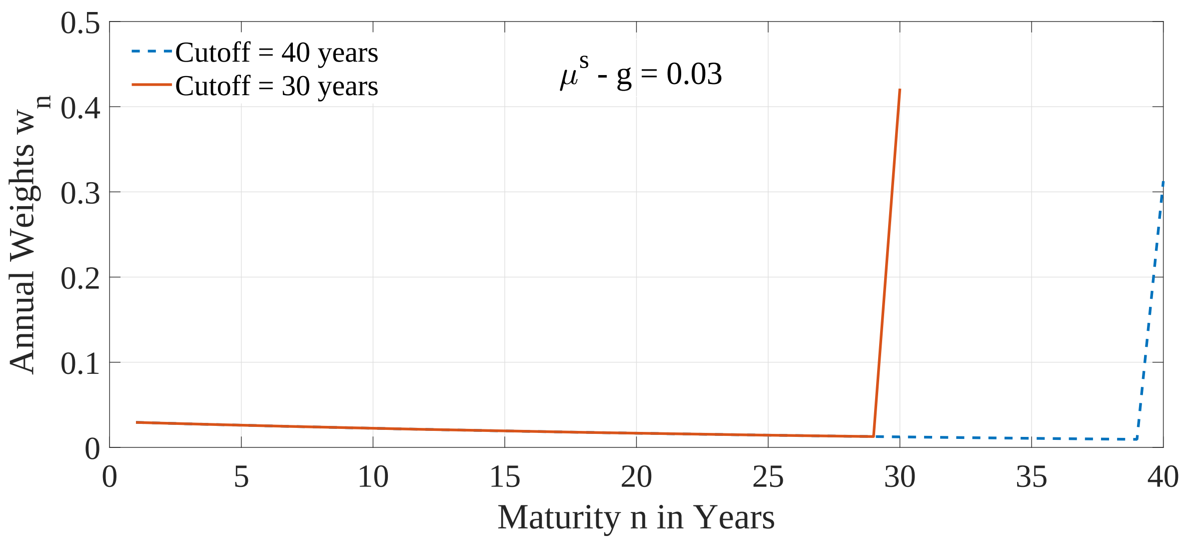

FIGURE IV

Strip Weights by Year: Medium Dividend Yield The graph plots the weights implied by the Gordon growth formula for $\mu^s - g = 0.03$. The values for μ^s and g used are 0.09 and 0.06 (both scaled by 12 to arrive at monthly numbers), though up to a first order approximation, the weights are only dependent on the difference between the two. Two cutoffs are considered: 30 years and 40 years. That is, the weight at 30 years (40 years) equals the sum of all weights with maturity 30 (40) or higher.

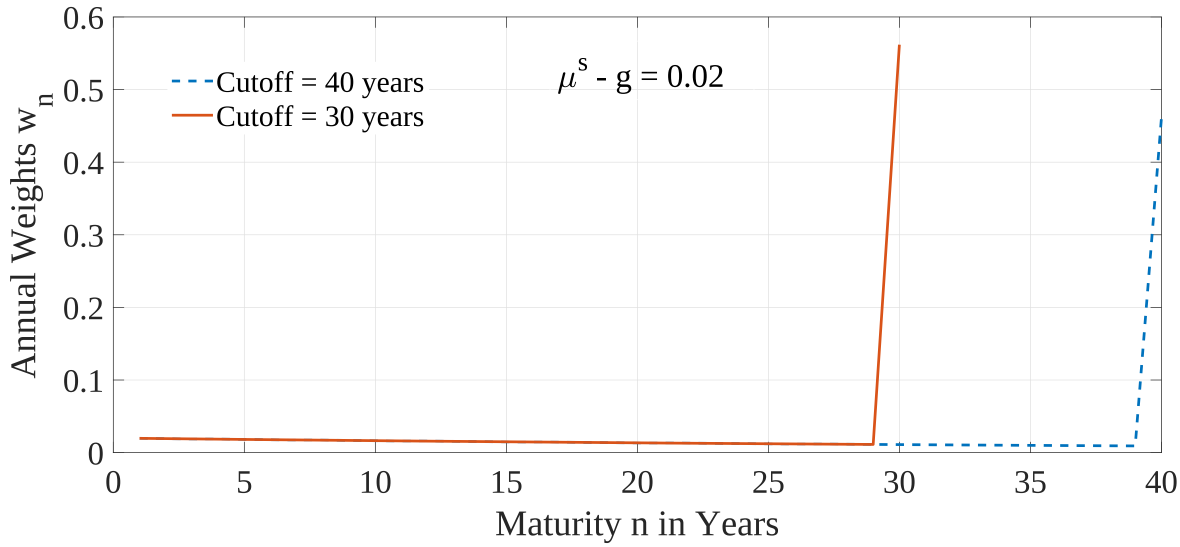

FIGURE V

Strip Weights by Year: Low Dividend Yield The graph plots the weights implied by the Gordon growth formula for $\mu^s - g = 0.02$. Two cutoffs are considered: 30 years and 40 years. That is, the weight at 30 years (40 years) equals the sum of all weights with maturity 30 (40) or higher.

month's returns. As before, I use two cutoff levels: 30 years (Counterfactual VII) and 40 years (Counterfactual VIII).

The results are presented in Table 3. The first row reports the model inputs to the Gordon growth formula to obtain the strip weights. The second row reports the cutoff month, which either happens after 30 years or after 40 years. The third row reports the monthly average returns of the counterfactual bond portfolio. The last row reports the implied realized annualized long-term dividend premium ($12\hat{\Psi}_0$) as defined in Equation 12.

Perhaps unsurprisingly, the results are in line with the constant maturity bond portfolios formed in the previous section. As the cutoff is picked later and the dividend yield ($\mu^s - g$) set to lower values, the longer duration bonds receive more weight in the calculation, thereby increasing the average return of the counterfactual bond portfolio and lowering the implied realized long-term dividend premium. When $\mu^s - g$ (i.e., the long-term dividend yield) is set to 0.06 and the cutoff is set to 360 months, the estimated compensation that investors have received for dividend risk equals 1.4% (in line with 10-15 year constant maturity bond counterfactuals). For all the other counterfactuals, the larger duration of the counterfactual bond portfolio implies an annual implied dividend premium that is less than a percent and often even negative. Because the dividend yield was on average about 2% during the 1996-2020 time period, the results of counterfactuals VII and VIII are very similar to those of counterfactuals V and VI. This suggests that the time variation in the dividend yield over this sample period, and the corresponding time-varying weights this induces, does not have much effect on the results. The standard deviation of the monthly index returns over this sample period equals 4.39% which is somewhat higher than those generated by counterfactuals I and II, somewhat lower than those generated by counterfactuals III and IV and quite a bit lower than those generated by counterfactuals V-VIII.

Counterfactual	I	II	III	IV	V	VI	VII	VIII
$\mu^s - g$	0.06	0.06	0.03	0.03	0.02	0.02	D_t/S_t	D_t/S_t
Cutoff month	360	480	360	480	360	480	360	480
$\sum w_{t,n}\mu^b_{t,n}$	0.0068	0.0073	0.0083	0.0099	0.0090	0.0113	0.0090	0.0114
Std. Dev.	0.0316	0.0348	0.0444	0.0558	0.0508	0.0691	0.0507	0.0697
$12\hat{\Psi}_0$ (Annual)	0.0135	0.0066	-0.0048	-0.0234	-0.0131	-0.0410	-0.0135	-0.0417
Annual diff in mean log rets	0.0063	0.0014	-0.0062	-0.0175	-0.0109	-0.0252	-0.0113	-0.0254

Table 3

Strip-Replicating Portfolios. The table reports the average monthly returns on the strip-replicating bond portfolios using a variety of different weighting schemes using data between January 1996 and April 2020. Columns 2 through 7 use a constant Gordon growth model to generate the weights, whereas the last two columns use the real-time dividend yield on the S&P500 to construct Gordon growth model weights. The first row reports the model inputs to the Gordon growth formula to obtain the strip weights. The second row reports the cutoff month, which either happens after 30 years or after 40 years. The third row reports the average of the monthly returns of the counterfactual bond portfolio, and the fourth row reports its standard deviation. The second-to-last row reports the implied realized annualized long-term dividend premium ($12\hat{\Psi}_0$) as defined in Equation 12 and the last row reports the difference in the annualized log return between the index return and the counterfactual bond portfolio.

5.2. Weighting Schemes: Dividend Strip Data

In the previous section I have explored 8 different counterfactuals based on model-implied weighting schemes. In this section I compare those weighting schemes with the available dividend strip data. Dividend strips are now directly traded on futures markets. A long position in a dividend futures contract implies that in exchange for a known payment due in n years from now, one receives the dividends paid on the underlying index over the year leading up to the settlement. Because the prices are listed as futures and not as spot prices, I first convert them to spot prices using the usual no-arbitrage relationship:

$$\mathcal{P}_{t,n} = \exp\left(-ny_{t,n}\right) F_{t,n}. \tag{20}$$

I then compute for each time t (annual) the cumulative weights of the first N years of dividend spot prices:

$$\sum_{n=1}^{N} w_n = \frac{\mathcal{P}_{t,1} + \ldots + \mathcal{P}_{t,N}}{S_t}. \tag{21}$$

I then compute the minimum, the maximum, and the average of these cumulative weights and plot them in Figure VI. As a comparison, I also plot the implied cumulative weights of the Gordon growth model for values of $\mu^s - g$ equal to 0.02, 0.03 and 0.06.

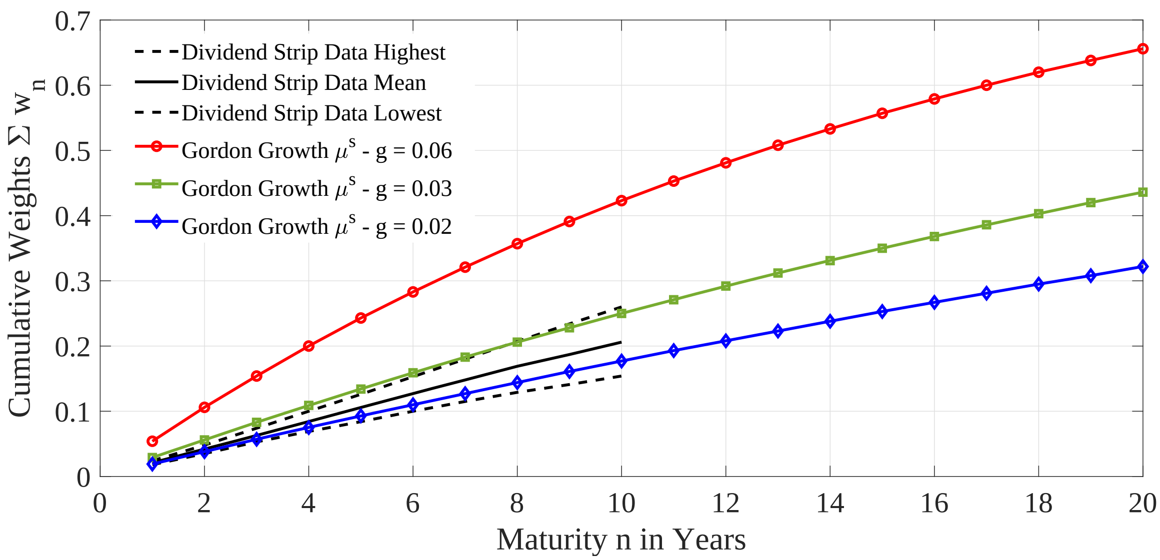

FIGURE VI

Cumulative Strip Weights by Maturity: Data vs Models The graph plots the cumulative weights $\sum_{n=1}^{N} w_n$, where N is on the x-axis, implied by the Gordon growth formula for $\mu^s - g = 0.02, 0.03$, and 0.06 and compares them to the available annual dividend strip data between December 2004 and December 2019.

The figure shows that for the available 16 years of annual dividend strip data (Dec 2004-Dec 2019), Gordon growth weights for a value of $\mu^s - g$ between 0.02 and 0.03 correspond to the average of the data. Gordon growth weights for $\mu^s - g = 0.03$ are close to the upper bound of the data and thus lead to an upper bound on the realized dividend premium. Recall further that setting the cutoff point at 360 or 480 months already lowers the duration of the counterfactual portfolios relative to the actual stock market. Gordon growth weights for $\mu^s - g = 0.06$ give too much weight to the early maturities. To be conservative in my estimates, I have still included these as possible counterfactuals in the computations.

6. Expanded Sample: 1970-2020

I now double the sample from 25 years to 50 years, which includes the high inflationary period between 1979 and 1983.[9] In Table 4 I report the long sample (1970-2020) results corresponding to Table 1. The patterns are very similar to those presented for the 1996-2020 sample. The average returns on the S&P500 index are roughly equivalent to those of constant maturity zero coupon bonds with a 17 year duration.

In Table 5 I repeat the analysis from Table 2 for the long sample. Once again a similar pattern emerges. As we choose larger duration bond portfolios as the counterfactual, the average return differential shrinks. Note that in this case there is a difference between mean log returns and mean simple returns. The long-duration bond portfolios are substantially more volatile in this expanded sample, particularly during the 1979-1983 period, and the difference between mean simple returns and mean log returns is $\frac{1}{2}\sigma^2$, assuming normality.

[9]The full sample of bond returns provided by Gurkaynak, Sack, and Wright (2006) goes back a few more years but has some outliers in the data that I wish to avoid as they have an outsized influence on the volatility estimates.

Maturity in years	FF	1	1.5	2	3	4	5	10	15	20	25	30	S&P500
Mean	0.0038	0.0046	0.0049	0.0051	0.0055	0.0059	0.0062	0.0075	0.0086	0.0096	0.0113	0.0147	0.0093
St. Dev.	0.0028	0.0054	0.0073	0.0090	0.0123	0.0153	0.0182	0.0327	0.0478	0.0646	0.0874	0.1216	0.0440
Mean log	0.0038	0.0046	0.0048	0.0050	0.0054	0.0057	0.0060	0.0070	0.0074	0.0076	0.0076	0.0075	0.0083

Table 4

Monthly Returns on Constant Maturity Zero Coupon Bonds. The second row in the table lists the average monthly bond returns ($\hat{\mu}_n^b$) on constant maturity zero coupon bond strategies for maturities ranging between 1 year and 30 years using monthly data between January 1970 and April 2020. The third row reports the monthly standard deviation. The second column lists of the table lists the corresponding statistics for the risk free return as used in the Fama French model and the last column lists those statistics for the S&P500 index.

Duration in years	10	15	20	25	30
$\hat{\mu}^s - \hat{\mu}^b$	0.0018	0.0007	-0.0003	-0.0020	-0.0054
t-stat on difference	0.83	0.30	-0.12	-0.55	-1.07
$12(\hat{\mu}^s - \hat{\mu}^b)$	0.0211	0.0088	-0.0041	-0.0245	-0.0648
Annualized difference in mean log returns	0.0156	0.0104	0.0085	0.0082	0.0090

Table 5

Monthly Return Differences between the S&P500 and Constant Maturity Zero Coupon Bonds. The second row in the table lists the difference between the monthly returns on the S&P500 index and the monthly returns on constant maturity zero coupon bonds ($\hat{\mu}_n^b$) for maturities ranging between 10 year and 30 years using monthly data between January 1970 and April 2020. The third row reports the t-statistic on the difference. For ease of interpretation, the fourth row reports the annualized difference (by multiplying by 12). The last row reports the annualized difference in the means of the monthly log returns (instead of simply returns).

The long-sample results corresponding to Table 3 are presented in Table 6. As before, the first row reports the model inputs to the Gordon growth formula to obtain the strip weights. The second row reports the cutoff month, which either happens after 30 years or after 40 years. The third row reports the monthly average returns of the counterfactual bond portfolio. The last row reports the implied realized annualized long-term dividend premium ($12\hat{\Psi}_0$) as defined in Equation 12.

As the average dividend yield of the S&P500 over the 1970-2020 sample was 2.8% (compared to 1.9% over the 1996-2020 sample), the results for counterfactuals VII and VIII (the time-varying weighting schemes based on the current dividend yield) are closer to counterfactuals III and IV instead of V and VI. The results from counterfactuals V and VI are therefore less relevant for this sample period. We can conclude that, as before, the results are largely in line with the constant maturity bond portfolios: bonds have generally outperformed stocks with similar or even higher volatility.

Finally, in VII I plot the cumulative performance between January 1970 and April 2020 of the S&P500 index and compare it to a fixed income counterfactual where the portfolio weights are based on the Gordon growth formula for $\mu^s - g = 0.03$, using a cutoff point of 360 months (counterfactual III). Interestingly, the graph reveals that there seems to be a low or even negative low-frequency correlation between the stock market and the replicating bond portfolio. This is consistent with long-term expected growth rates being positively correlated with interest rates, and/or dividend risk premia being negatively correlated, as further discussed in Section 8.

Counterfactual	I	II	III	IV	V	VI	VII	VIII
$\mu^s - g$	0.06	0.06	0.03	0.03	0.02	0.02	D_t/S_t	D_t/S_t
Cutoff month	360	480	360	480	360	480	360	480
$\sum w_{t,n}\mu^b_{t,n}$	0.0089	0.0109	0.0108	0.0169	0.0121	0.0212	0.0109	0.0158
Std. Dev.	0.0473	0.0580	0.0716	0.1112	0.0868	0.1508	0.0644	0.0938
$12\hat{\Psi}_0$ (Annual)	0.0053	-0.0109	-0.0184	-0.0911	-0.0332	-0.1431	-0.0192	-0.0789
Annual diff in mean log rets	0.0066	-0.0116	-0.0006	-0.0378	-0.0018	-0.0442	-0.0069	-0.0420

Table 6

Strip-Replicating Portfolios. The table reports the average monthly returns on the strip-replicating bond portfolios using a variety of different weighting schemes using data between January 1970 and April 2020. Columns 2 through 7 use a constant Gordon growth model to generate the weights, whereas the last two columns use the real-time dividend yield on the S&P500 to construct Gordon growth model weights. The first row reports the model inputs to the Gordon growth formula to obtain the strip weights. The second row reports the cutoff month, which either happens after 30 years or after 40 years. The third row reports the average of the monthly returns of the counterfactual bond portfolio, and the fourth row reports its standard deviation. The second-to-last row reports the implied realized annualized long-term dividend premium ($12\hat{\Psi}_0$) as defined in Equation 12 and the last row reports the difference in the annualized log return between the index return and the counterfactual bond portfolio.

7. International Evidence

In this section I repeat the computations from the previous section for Europe and Japan.

7.1. Europe

For European government bonds, I use data between October 1972 (the start date of the Bundesbank yield curve data) and April 2020. I once again study two samples: 1996-2020 and the full sample 1972-2020. I use returns on the Eurostoxx index as the stock market proxy. All returns are computed in local currency. Further, the average dividend yield on the Eurostoxx index over these two sample periods is 0.024 (short sample) and 0.031 (long sample) corresponding to durations of 32-42 years. The results for the constant maturity bond portfolio are reported in Table 7 and are even starker than for the United States. Even the 15-year constant maturity zero coupon bonds have outperformed the index over this sample period.

In the top panel of Table 8 I repeat the analysis from Table 2 using data between 1996 and

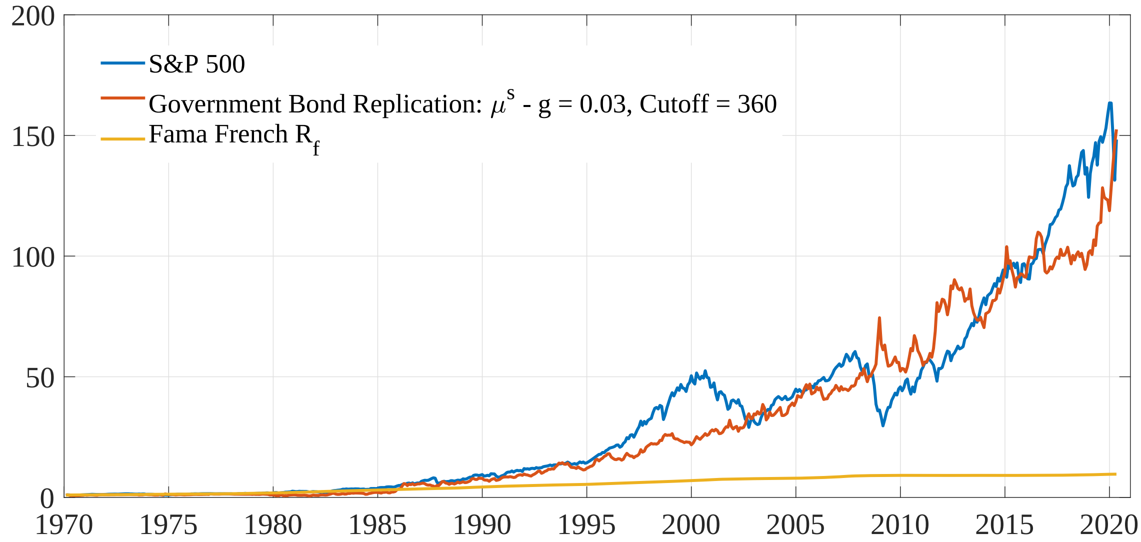

FIGURE VII

Cumulative Return Performance of the S&P500 Index Relative to Fixed Income Matched Portfolio The graph plots the cumulative performance between January 1970 and April 2020 of the S&P500 index and compares it to a fixed income counterfactual where the portfolio weights are based on the Gordon growth formula for $\mu^s - g = 0.03$, using a cutoff point of 360 months.

Mat. in years	1	2	5	10	15	20	25	30	Eurostoxx
Sample: 1996-2020									
Mean	0.0017	0.0022	0.0038	0.0060	0.0078	0.0094	0.0110	0.0129	0.0060
St. Dev.	0.0023	0.0041	0.0099	0.0190	0.0288	0.0391	0.0500	0.0625	0.0525
Mean log	0.0017	0.0022	0.0037	0.0058	0.0072	0.0085	0.0096	0.0108	0.0046
Sample: 1972-2020									
Mean	0.0038	0.0043	0.0055	0.0069	0.0083	0.0106	-	-	0.0079
St. Dev.	0.0043	0.0064	0.0126	0.0236	0.0417	0.0719	-	-	0.0484
Mean log	0.0037	0.0043	0.0054	0.0066	0.0074	0.0081	-	-	0.0067

Table 7

Monthly Returns on Constant Maturity Zero Coupon Bonds: Germany. The second row of the first panel lists the average monthly bond returns ($\hat{\mu}_n^b$) on constant maturity zero coupon bond strategies for maturities ranging between 1 year and 30 years using monthly data between January 1996 and April 2020. The third row reports the monthly standard deviation. The last two columns list those statistics for the Eurostoxx index. The second panel repeats all statistics but now for the longer sample 1972-2020.

2020 and using the Eurostoxx as the stock market proxy. Once again, the results are similar (if not stronger) as those for the United States. The bottom panel repeats the analysis for the long sample (1972-2020), once again with similar results.

Duration in years	10	15	20	30
Sample: 1996-2020				
$\hat{\mu}^s - \hat{\mu}^b$	0.0001	-0.0016	-0.0032	-0.0067
t-stat on difference	0.0188	-0.4424	-0.7993	-1.3659
$12(\hat{\mu}^s - \hat{\mu}^b)$	0.0008	-0.0196	-0.0385	-0.0800
Annualized difference in mean log returns	-0.0167	-0.0341	-0.0488	-0.0769
Sample: 1972-2020				
$\hat{\mu}^s - \hat{\mu}^b$	0.0010	-0.0004	-0.0027	-
t-stat on difference	0.4634	-0.1401	-0.7551	-
$12(\hat{\mu}^s - \hat{\mu}^b)$	0.0123	-0.0043	-0.0319	-
Annualized difference in mean log returns	0.0013	-0.0082	-0.0164	-

Table 8

Monthly Return Differences between the Eurostoxx and Constant Maturity German Zero Coupon Bonds. The second row in the table lists the average difference between the monthly returns on the Eurostoxx 50 index ($\hat{\mu}^s$) and the monthly returns on constant maturity zero coupon bonds ($\hat{\mu}^b_n$) for maturities ranging between 5 and 20 years using monthly data between January 1996 and April 2020. The third row reports the t-statistic on the difference. For ease of interpretation, the third row reports the annualized difference (by multiplying by 12). The last row reports the annualized difference in the means of the monthly log returns (instead of simply returns).

7.2. Japan

Next, I repeat the analysis for Japan. The results for the constant maturity bond portfolio are reported in Table 9. All returns are once again in local currency. The results on average returns are similar to (if not stronger than) those of the United States. This is particularly interesting for the 1996-2020 subsample. After all, in Japan, bond yields of all maturities already reached very low levels in 1996, comparable to the bond yields observed today in the United States and Europe. This illustrates that such low levels of bond yields by no means guarantee that stocks will outperform bonds over the next quarter century. Perhaps unsurprisingly, the volatility generated by risk free rate variation is smaller compared to the U.S. and Europe over this sample period.

In the top panel of Table 10 I repeat the analysis from Table 2 using data between 1996 and 2020 and using the Topix index as the Japanese stock market proxy. Once again, the results are similar (if not stronger) as those for the United States. The bottom panel reports results for the longer sample (1985-2020), with stronger results.

Mat. in years	1	2	3	4	5	10	15	Topix	Nikei225
Sample: 1996-2020									
Mean	0.0002	0.0004	0.0007	0.0011	0.0014	0.0031	0.0043	0.0023	0.0028
St. Dev.	0.0007	0.0017	0.0030	0.0044	0.0059	0.0133	0.0205	0.0504	0.0553
Mean log	0.0002	0.0004	0.0007	0.0011	0.0014	0.0030	0.0041	0.0010	0.0011
Sample: 1985-2020									
Mean	0.0014	0.0017	0.0021	0.0026	0.0030	0.0046	0.0059	0.0037	0.0038
St. Dev.	0.0025	0.0043	0.0063	0.0083	0.0102	0.0191	0.0288	0.0548	0.0590
Mean log	0.0014	0.0017	0.0021	0.0025	0.0029	0.0044	0.0054	0.0023	0.0023

Table 9

Monthly Returns on Constant Maturity Zero Coupon Bonds: Japan. The second row of the first panel lists the average monthly bond returns ($\hat{\mu}_n^b$) on constant maturity zero coupon bond strategies for maturities ranging between 1 year and 30 years using monthly data between January 1996 and April 2020. The third row reports the monthly standard deviation. The last two columns list those statistics for the Topix index and the Nikkei 225. The second panel repeats all statistics but now for the longer sample 1985-2020.

Duration in years	10	15
Sample: 1996-2020		
$\hat{\mu}^s - \hat{\mu}^b$	-0.0008	-0.0020
t-stat on difference	-0.2410	-0.5700
$12(\hat{\mu}^s - \hat{\mu}^b)$	-0.0093	-0.0235
Annualized difference in mean log returns	-0.0248	-0.0495
Sample: 1985-2020		
$\hat{\mu}^s - \hat{\mu}^b$	-0.0009	-0.0022
t-stat on difference	-0.3324	-0.7278
$12(\hat{\mu}^s - \hat{\mu}^b)$	-0.0114	-0.0267
Annualized difference in mean log returns	-0.0272	-0.0527

Table 10

Monthly Return Differences between the Nikkei 225 and Constant Maturity German Zero Coupon Bonds. The second row in the table lists the average difference between the monthly returns on the Nikkei 225 index ($\hat{\mu}^s$) and the monthly returns on constant maturity zero coupon bonds ($\hat{\mu}_n^b$) for maturities ranging between 10 and 15 years using monthly data between January 1996 and April 2020. The third row reports the t-statistic on the difference. For ease of interpretation, the third row reports the annualized difference (by multiplying by 12). The last row reports the annualized difference in the means of the monthly log returns (instead of simply returns).

8. Excess Volatility and Variance Decomposition

One of the more surprising features of short-term dividend strip prices is that they seem excessively volatile relative to their subsequent realizations.[10] As such, dividend strips, while relatively safe investments as measured by their CAPM beta, are risky investments as measured by their volatility. An important open question is where this excess strip price volatility is coming from, as this short duration price variation deepens the Shiller (1981) excess volatility puzzle.

In fact, as this paper shows, generating large amounts of volatility for long duration claims is relatively straightforward, as small changes in discount rates correspond to large changes in prices. Government bond prices are clearly much more volatile than their subsequent predetermined coupon and principal payments. In fact, without additional dividend risk premium variation, the counterfactual bond portfolios I construct already have similar (if not higher) volatility profiles as those of the index, despite the cash flows being fixed. This questions the common notion that excess volatility is a puzzle for equities. If anything, the stock market claim is too little volatile compared to what we should expect based on the fixed-income calculations I present here. If dividend risk premia and interest rates are negatively correlated, and growth expectations and long-term interest rates have a positive correlation (as suggested by Figure VII), then variation in both the dividend risk premium (if any) and dividend growth (expected and unexpected) have a tempering effect on the volatility of the equity claim compared to those of the bond claim. This also puts into context the common belief that stock markets are particularly prone to bubble-type episodes. At least compared to long-term bonds, there seems to be little evidence for excessively volatile or bubble-like stock price movements. Alternatively, it is of course possible that long-term bonds are even more exposed to bubbles than stock markets are. Evidence on potential excess volatility in bond markets is provided by Shiller (1979), Giglio and Kelly (2018) and Lustig, Brooks, and Katz (2019).

Based on the duration-matched counterfactual (CF) fixed income portfolios we can decompose stock returns in excess of the short-duration fixed income instrument into two components:

$$R^S_{t+1} - R^b_{t+1,1} = \underbrace{R^S_{t+1} - R^b_{t+1,CF}}_{\text{Dividend Risk Premium}} + \underbrace{R^b_{t+1,CF} - R^b_{t+1,1}}_{\text{Term Premium}}, \tag{22}$$

[10] See Binsbergen, Brandt, and Koijen (2012), Binsbergen, Hueskes, Koijen, and Vrugt (2014) and Binsbergen and Koijen (2017).

which leads to the following variance decomposition:

$$\begin{aligned}\operatorname{var}\left(R^S_{t+1}-R^b_{t+1,1}\right) &= \operatorname{var}\left(R^S_{t+1}-R^b_{t+1,CF}\right)+\operatorname{var}\left(R^b_{t+1,CF}-R^b_{t+1,1}\right)\\ &+ 2\operatorname{cov}\left(R^S_{t+1}-R^b_{t+1,CF},R^b_{t+1,CF}-R^b_{t+1,1}\right).\end{aligned}$$

	1970 - 2020		1996 - 2020	
$\operatorname{var}\left(R^S_{t+1}-R^b_{t+1,1}\right)$	0.0020	100%	0.0019	100%
Decomposition				
$\operatorname{var}\left(R^S_{t+1}-R^b_{t+1,CF}\right)$	0.0062	317%	0.0049	254%
$\operatorname{var}\left(R^b_{t+1,CF}-R^b_{t+1,1}\right)$	0.0052	264%	0.0020	102%
$2\operatorname{cov}\left(R^S_{t+1}-R^b_{t+1,1}\right),$	-0.0094	-482%	-0.0049	-256%

Table 11

Variance Decomposition U.S. Excess Stock Returns. The table provides a variance decomposition of U.S. stock returns in excess of the short duration fixed income instrument (in this case the risk free rate used in the Fama French factors) into the contribution of the realized term premium and the realized dividend risk premium. The long-duration fixed income counterfactual (CF) uses Gordon growth weights corresponding to $\mu^s - g = 0.03$ and a cutoff of 360 months. In columns 2 and 4 the table reports the raw variance and covariance numbers for both the full sample as well as the 1996-2020 subsample. Columns 3 and 5 report the numbers scaled by the variance of $R^S_{t+1} - R^b_{t+1,1}$, so that the decomposition terms sum up to 100%.

Table 11 summarizes this variance decomposition, where the long-duration fixed income counterfactual uses Gordon growth weights corresponding to $\mu^s - g = 0.03$ and a cutoff of 360 months. The table shows that the realized term premium and realized dividend risk premium are strongly negatively correlated. Both in the full sample as in the 1996-2020 sample, this correlation is -0.8. This suggests that shocks to long-term risk free discount rates are either positively correlated with innovations in long term growth rates, or negatively correlated with long-duration dividend risk premia (or both) as further explored in the next section.[11]

As a final comment, one important unexplored constraint on risk premia's ability to generate excess volatility imposed by most models is that they need to have a substantial positive mean. After all, if risk premia are substantially volatile, but also need to always remain positive (as suggested by many macro finance models), they need to be large on average. Under these model constraints, for dividend strip excess volatility to be driven by risk premium variation, the average risk premium on dividend strips has to be large.

[11]Note that over the full sample stock returns and the long-term bond returns are somewhat positively correlated (0.15) whereas over the subsample 1996-2020 they are negatively correlated -0.25. See Campbell, Pflueger, and Viceira (2020) for the economic drivers of this time-varying correlation.

This logic also implies that if long-duration dividend risk premia are on average low (and potentially even zero), they have less potential of generating or tempering volatility, though obviously the duration still acts as a lever. There are two alternative views that challenge the "risk premia always need to be positive" constraint. First, if dividends are a hedge against other types of economic risks, they can in fact have a negative risk premium. Secondly, it is possible that excess volatility in dividend strip and stock prices is truly excessive and unrelated to risk premia variation. Rather, this volatility is a reflection of mismeasured or overextrapolated expectations.[12]

9. Potential Explanations and Discussion

In this section I discuss four potential explanations for the findings as well as further implications.

9.1. Four Potential Explanations

In addition to the recent literature that has proposed theoretical foundations for a downward sloping equity term structure (see Binsbergen and Koijen (2017) for a review), there are at least four other explanations to consider.

First, there is the possibility that dividends are less risky than nominal government bonds as the latter offer fixed nominal payments that are not protected against inflation. Long-term dividends can be increased with inflation. While this explanation could be important for the first quarter century of the sample (1970-1995), it seems less important for the second half of the sample (1996-2020) when inflation expectations were low and stable.[13]

Second, there is the possibility that while short-term dividends are exposed to large disaster type risks, such as the government-imposed dividend cuts in the financial crisis and COVID epidemic, long-term dividends will mean revert to trend levels and are therefore less risky.[14] This could also explain why short-term dividends, that do not have enough time

[12]For recent work in this area on the aggregate stock market see Bordalo, Gennaioli, LaPorta, and Shleifer (2020) and the references therein. For a recent exploration of such mismeasured expectations for real investment, see Binsbergen and Opp (2013).

[13]See also Katz, Lustig, and Nielsen (2015) for a discussion on the degree to which stocks are real assets.

[14]See Barro, Nakamura, Steinsson, and Ursua (2011) for a model that features such dynamics, as well as Cejnek, Randl, and Zechner (2020) who aptly relate dividend smoothing policies to the downside risk of dividends.

to recover after a disaster, command a risk premium over their corresponding zero coupon bond.

Third, there is the possibility that while investors were expecting at least some return compensation for long-duration dividend risk, they ended up not receiving it ex post. Because interest rates and (expected) growth are tied in equilibrium, the series of unexpected downward shocks to interest rates was accompanied by a series of unexpected downward shocks to long-term growth rates. This could explain why the stock market has not performed too well relative to duration-matched fixed income portfolios whose fixed cash flows are shielded from such growth shocks. This could imply that the U.S. and Europe are now also stuck in a Japan-type scenario of long-term low growth and low interest rates, implying that future dividend growth realizations will be low.

More formally, if we assume (1) log-normal consumption growth with average growth rate g and variance of σ_c^2, and (2) power utility preferences over aggregate consumption with risk aversion coefficient γ and subjective discount factor β:

$$U = E_0 \left[\sum_{t=1}^{\infty} \beta^t \frac{C_t^{1-\gamma}}{1-\gamma} \right], \tag{23}$$

the steady state interest rate is tied to long-term growth as follows:

$$y = -\ln(\beta) + \gamma g - \frac{1}{2}\gamma^2 \sigma_c^2. \tag{24}$$

Comparing steady states, there is a one-to-one relation between growth (g) and the interest rate (y) when the risk aversion coefficient (or the inverse of the Intertemporal Elasticity of Substitution) is 1. In that case the effect of lower interest rates and lower growth cancel when computing realized returns. For a risk aversion close to, but different than 1, the two effects partially cancel and, depending on the sign of the deviation, can contribute to either a positive or negative correlation between bonds and stocks. This level of risk aversion also implies that the precautionary savings effect (the last term in Equation 24) is small. In most models, such a low risk aversion coefficient will imply low risk premia, though given the findings in this paper, that may be less of a concern when matching the model to the data of the past 50 years. It seems more of a concern when matching data from the earlier part of the twentieth century. However, during that period investors were less insured and were forced to hold less diversified portfolios, as the mutual fund sector was small, potentially leading to larger risk premia.

Finally, there is the possibility that stocks have not performed well due to a secular increase in long-term future dividend risk premia (Farhi and Gourio (2018)), which occurs when long-term risk free discount rates are negatively correlated with long duration dividend risk premia. This increasing path of risk premia going forward has suppressed the value of the equity claim relative to the fixed income claim. In that case, future excess return realizations should be high. Given that dividend yields are currently low, the Gordon growth formula prescribes that expected returns can only be high of future growth remains high. After all, that formula states that the expected return is the sum of the dividend yield and the expected growth rate.

9.2. Further Implications

The results presented above also have potentially important implications for the cross-section of stock returns. To the extent that different stocks have different durations of cash flows, the valuation of those stocks will be differentially affected by the secular decline in long-term interest rates. For example, if value stocks have shorter duration cash flows than growth stocks, it may be less surprising that the latter have outperformed the former in recent years, though growth expectations may also have developed differently for these two sets of firs. As argued in the introduction, the valuation windfalls for long duration assets are not likely to repeat themselves given the lower bound on interest rates. Furthermore, if assets that are exposed to cash flow risks have not outperformed their fixed cash flow counterparts, this raises the question of what risk premium the CAPM is exactly supposed to capture when estimating the slope of the Security Market Line (SML). Is it interest rate risk, or inflation risk (or lack thereof)? This seems particularly pressing given that the version of the CAPM that uses the stock market as a proxy for the wealth portfolio is not known to price government bonds (of all maturities) well.[15]

Future research could construct counterfactual fixed income portfolios at the stock level to evaluate duration-matched outperformance in the cross-section of stock returns. This would better separate the differential returns that investors receive for investing in risky earnings/dividends as opposed to those that result from interest rate changes. A similar argument can be held for real estate assets that may differ in duration in important ways.[16]

Finally, the results are important for corporate finance studies. First, the secular down-

[15]See also Lustig, Verdelhan, and Nieuwerburgh (2013).

[16]See Giglio, Maggiori, and Stroebel (2014) for an exploration of the term structure of discount rates in real estate markets.

ward trend in interest rates in the past 50 years implies that the market value of corporate debt (bonds) is higher than the book value. The results in this paper suggest that for firms that rely on long term debt, this difference can be substantial. As such the amount of market leverage for those firms is higher than what conventional measurement (which uses the book value of debt as a proxy for the market value) would imply. Second, when studying capital structure, the common assumption is that as the leverage of the firm increases, ceteris paribus, the required return on equity increases as the CAPM beta scales up the risk premium. The results in this paper suggest that the measured (realized) risk premium effects from leverage increases are likely small or even negative in the past 50 years.

10. Tradable Portfolios and Corporate Bonds

In this section I explore two robustness analysis. One that compares the returns implied by the Gurkaynak, Sack, and Wright (2006) curves with those of a tradable long-term government bond index portfolio provided by Vanguard, and one that compares realized long-term government bond returns with a tradable corporate bond portfolio of the same maturity.

10.1. Tradable Indices

The U.S. government bond portfolio returns presented so far are based on bootstrapped zero curves (Gurkaynak, Sack, and Wright (2006)). One could therefore be concerned that the returns are affected by these curve fitting methods. One straightforward way to address this concern is to simply compare the implied returns from these yield curves to those of traded bond funds. For example, the Vanguard Long-Term Treasury Fund Investor Shares (VUSTX) holds a diversified portfolio of long duration government bonds. The fund advertises that the average maturity of the bonds varies between 15 and 30 years. The duration of the fund varies between 10 and 20 years. Figure IX plots the historical effective duration of this fund between June 2002 and April 2020.[17]

I construct a set of long-term bond portfolios of varying maturities and price them using the bootstrapped zero curves of Gurkaynak, Sack, and Wright (2006). Because of the downward trend in interest rates, bonds that were issued at par are soon trading at a premium. I use coupon rates that are 2% above the prevailing 10-year constant maturity yield. I then

[17] I thank John Ameriks for generously sharing this data with me.

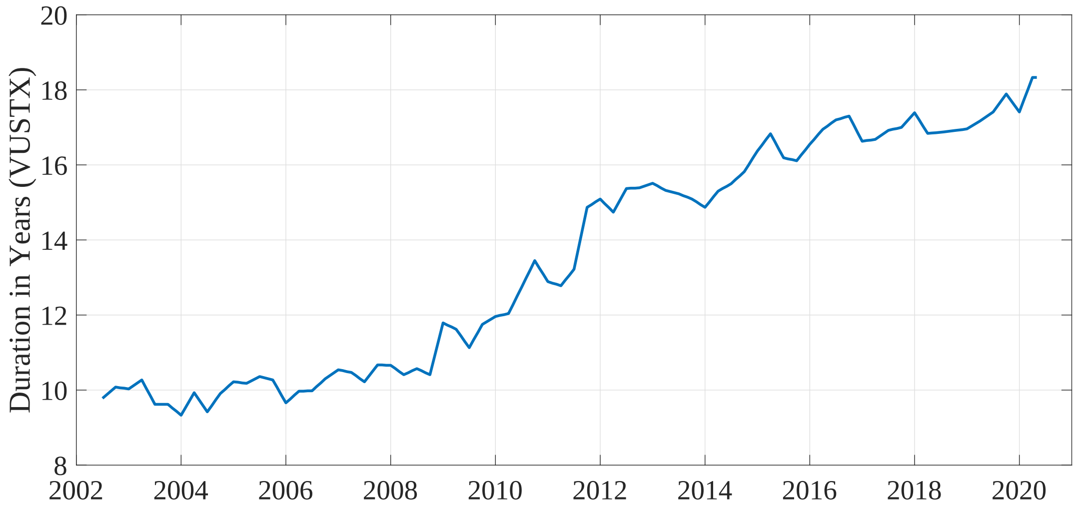

FIGURE VIII

Vanguard Long-Term Government Bond Index Fund The graph plots the effective duration of the Vanguard Long-Term Government Bond Index Fund (VUSTX).

compute monthly bond returns on these bonds and form a portfolios of the two closest bonds in duration and take a weighted average of their returns to match the effective duration of the VUSTX fund. The graph below compares the returns of this replicated portfolio to those of the VUSTX fund. The graph shows that the returns are highly similar with an almost identical standard deviation of monthly returns that equals 3.29% for the replicated portfolio and 3.26% for the actual monthly returns of VUSTX to investors. The correlation between the two monthly return series is 0.998. The returns on the replicated portfolio are 4.8b.p. per month higher than the VUSTX returns, which corresponds to about 58 basis points per year. The annual fees on the investor class VUSTX fund is 28b.p. in 2002 and drops to 20b.p. by the end of the sample. So this explains a little under half the difference. This leaves about 3b.p. per month for replication errors in my approach (the cross-sectional variation in bond maturities in the VUSTX fund is a bit larger than in my replication, somewhat suppressing its performance) as well as trading costs. Overall, we can conclude that the implied returns of the zero curves provided by Gurkaynak, Sack, and Wright (2006) over this sample period lead to fairly accurate representations of the actual trading data.

10.2. Corporate Bonds

In the results presented above, I have compared strip-matched government bonds with stocks. In this section, I explore whether similar results can be obtained using long-term

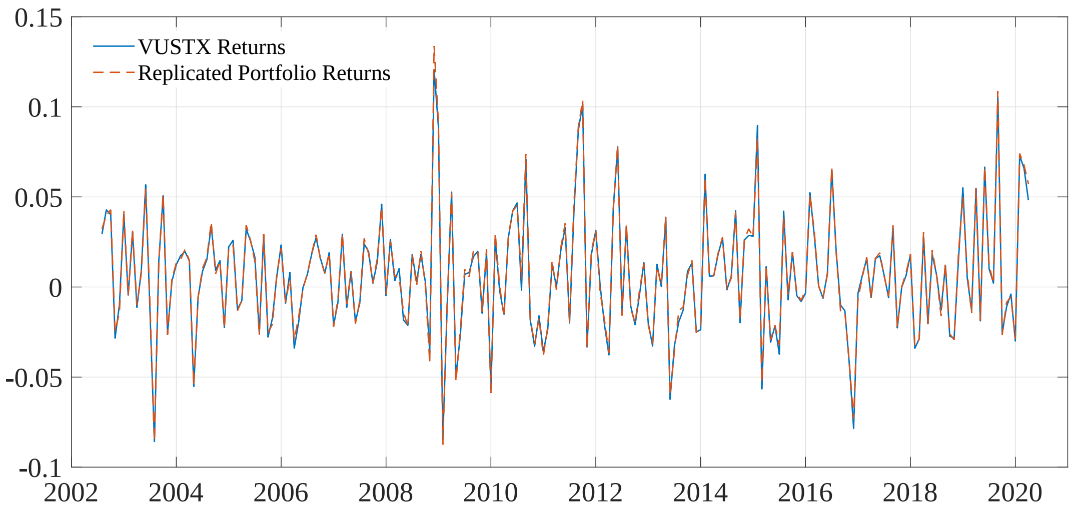

FIGURE IX

Vanguard Long-Term Government Bond Returns (VUSTX) vs Replicated Returns The graph plots the monthly returns to investors on Vanguard's long-term government bond index fund (VUSTX) and plots it against replicated returns on duration-matched portfolios based on the zero curves provided by Gurkaynak, Sack and Wright (2006).

corporate bond returns instead of government bond returns. This also helps address concerns related to the potential specialness or mispricing of government debt, for example due to convenience yields associated with the money-like features of government debt.[18]

To explore this question I study the returns on the Barclays Bloomberg long duration corporate bond index. This index is designed to measure the value-weighted performance of U.S. corporate bonds that have a maturity of greater than or equal to 10 years. The duration of this portfolio, as provided by Bloomberg, is plotted in Figure X. Following the same procedure as in the previous subsection (i.e. the Vanguard fund), I then construct government bonds that best match the duration of this portfolio and compute their returns using the zero coupon yield curves from Gurkaynak, Sack and Wright (2006). The cumulative return of the Bloomberg Barclays portfolio as well as that of the replicating bond portfolio is plotted in Figure XI. The graph shows that government bonds have done about as well as corporate bonds over this sample period for these long duration bonds, suggesting that the comparison with equity returns would have led to similar conclusions if 20 to 30-year duration corporate bond returns had been available. Finally, the results suggest that there is not much evidence for a corporate credit risk premium over this sample period either,

[18]See Krishnamurthy and Vissing-Jorgensen (2012), Nagel (2016), Lustig, Jiang, Nieuwerburgh, and Xiaolan (2019) and Binsbergen, Diamond, and Grotteria (2019) for recent contributions in this area.

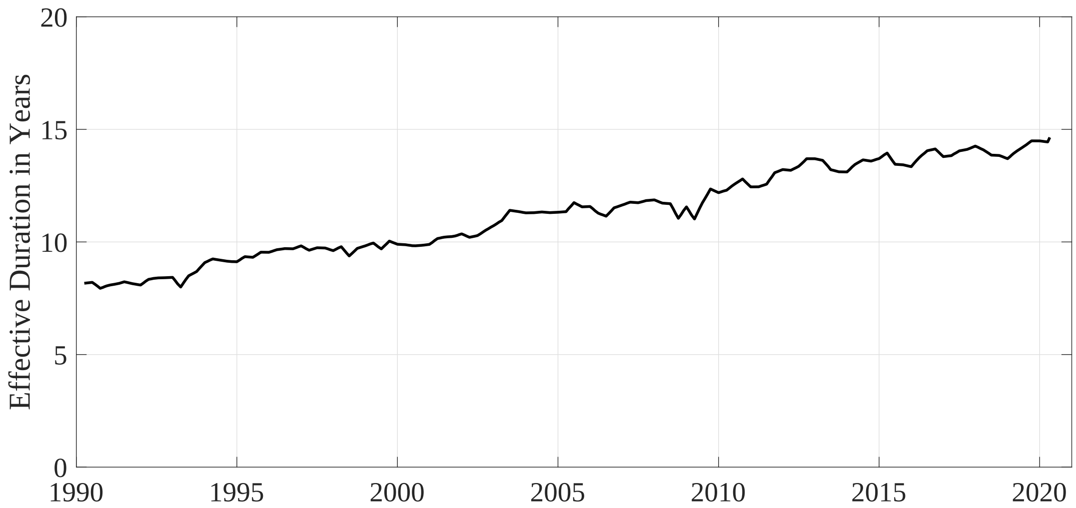

FIGURE X

Barclays Bloomberg Long Duration Corporate Bond Index: Duration The graph plots the effective duration of the Barclays Bloomberg long duration corporate bond index between 1990 and 2020.

consistent with the idea that all risk premia over this sample period have been low.

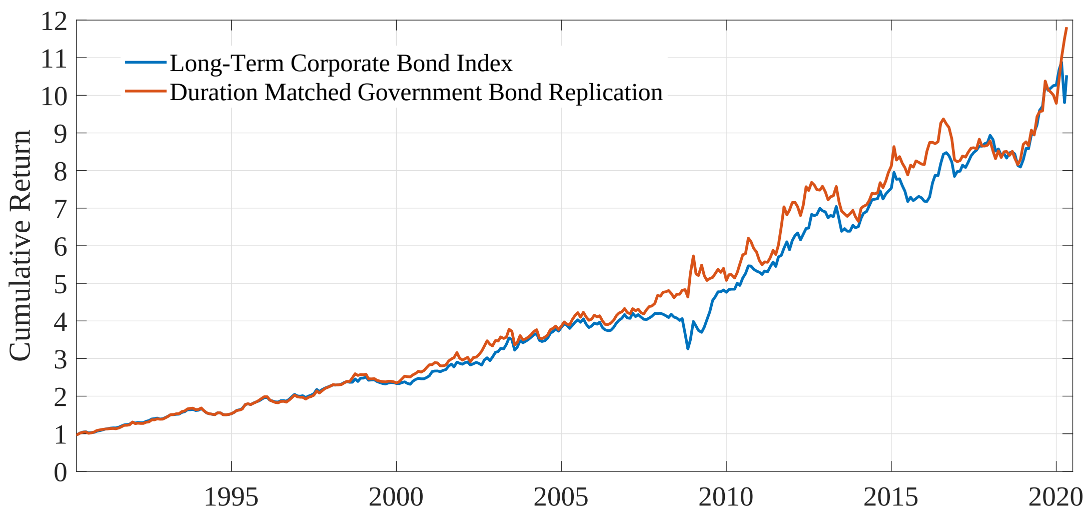

FIGURE XI

Barclays Bloomberg Long Duration Corporate Bond Index: Cumulative Performance The graph plots the performance of the Barclays Bloomberg long duration corporate bond index between 1990 and 2020 and compares it with the performance of a duration-matched U.S. government bond portfolio.

11. Conclusion

In this paper I have constructed a set of plausible counterfactual fixed income portfolios that span the plausible range for the duration profile of the aggregate stock market. I find that over the past five decades stock market indices have exhibited no outperformance over these fixed income counterfactuals with comparable volatility profiles. This implies that investors have not received much compensation for taking long duration dividend risk. This result holds across several world regions. Given that this result is obtained under conservative assumptions regarding the duration of the stock market, the realized dividend risk premium may very well be substantially negative (i.e. multiple percentage points) over this sample period. One could argue that this simply means that the equity premium puzzle has resolved itself.

However, the fact that investors have not received compensation for long duration dividend risk does not necessarily mean that investors were not expecting to receive at least some compensation. It could mean that stocks had poor long-term performance compared to their fixed income counterparts as a consequence of a secular decline in long-term expected dividend growth rates (and/or secular increase in long-term risk premia) over these decades. While the expectations hypothesis for interest rates provides us with some information regarding the unexpected downward shocks to the risk free discount rates, observing unexpected shocks to long-term (30-50 year) expected growth rates is more challenging, as very long-run expected growth measures are not in ample supply. Arguably, the stock market itself is the best predictor of long-term growth, providing a potentially bleak outlook.

Even though the performance of stock and bond markets in the past 50 years are important in their own right, one may wonder whether the results on means and volatilities presented in this paper have further external validity (in addition to the international evidence provided), particularly going forward. Perhaps the Japanese results presented above can give some guidance here. Interest rates already reached very low levels in Japan in 1996 for all maturities. In the 25 years since, the Nikkei 225 has not increased much in value. Furthermore, long-term bond yields have decreased even further, leading once again to bond outperformance over stocks. This suggests that the American and European results presented in this paper could in fact repeat themselves. The Japanese data is therefore also potentially informative about the level of stock and bond returns investors should expect in the U.S. and Europe for the next 25 years. It seems at least possible that expected returns on both stocks and bonds have reached all-time low levels. After all, the Gordon growth math, where expected returns equal growth g plus dividend yield, should still hold. If both the dividend yield as well as growth (and inflation) expectations are very low, there is little room for either nominal or real expected returns (and risk premia) going forward. This has important implications for retirement savings, as it means that workers should save a substantially higher percentage of their annual incomes to achieve an acceptable living standard in retirement.

To conclude, it seems important to adjust asset pricing moments for secular trends in interest rates (and potentially growth rates) such that they can be meaningfully interpreted in stationary environments. Alternatively, it could be helpful to explicitly model secular trends, investors' perceptions of them, and their underlying causes, in asset pricing theories. Also, a further decomposition of asset valuations into the effects of the term structure of interest rates and the term structure of dividend risk premiums (or other risk premiums) seems an important avenue for future research.[19] More generally, I would argue that organizing and comparing available assets by maturity instead of within traditional asset class categorizations (i.e. stocks, bonds, real estate and commodities) can provide important and interesting insights that have previously been ignored.

[19] See also recent work by Lettau and Wachter (2007) and Lettau and Wachter (2010) for a joint treatment of the equity and bond term structure.

References

Bansal, R., D. Kiku, and A. Yaron (2006): "Long Run Risks: Estimation and Inference," Unpublished paper, Duke University and the Wharton School.

Bansal, R., and A. Yaron (2004): "Risks for the Long-Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, 59(4), 1481–1509.

Barro, R., E. Nakamura, J. Steinsson, and J. Ursua (2011): "Crises and Recoveries in an Empirical Model of Consumption Disasters," Working paper.

Binsbergen, J. H. v., and M. W. Brandt (2015): "Asset Liability Management," *Handbook of Fixed Income*.

Binsbergen, J. H. v., M. W. Brandt, and R. S. Koijen (2012): "On the Timing and Pricing of Dividends," *American Economic Review*, 102, 1596–1618.

Binsbergen, J. H. v., W. Diamond, and M. Grotteria (2019): "Risk Free Interest Rates," Working Paper, University of Pennsylvania.

Binsbergen, J. H. v., W. Hueskes, R. S. Koijen, and E. B. Vrugt (2014): "Equity Yields," *Journal of Financial Economics*, 110(3), 503–519.

Binsbergen, J. H. v., and R. S. Koijen (2017): "The Term Structure of Returns: Facts and Theory," *Journal of Financial Economics*, 124, 1–21.

Binsbergen, J. H. v., and C. Opp (2013): "Real Anomalies," *Journal of Finance*, 74, 1659–1706.

Bordalo, P., N. Gennaioli, R. LaPorta, and A. Shleifer (2020): "Expectations of Fundamentals and Stock Market Puzzles," Working Paper, Harvard University.

Brennan, M. J. (1998): "Stripping the S&P 500 Index," *Financial Analyst Journal*, 54(1), 12–22.

Campbell, J., C. Pflueger, and L. Viceira (2020): "Macroeconomic Drivers of Bond and Equity Risks," *Journal of Political Economy*, 128(8), 3148–3185.

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, 205–251.

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, 195–227.

Cejnek, G., O. Randl, and J. Zechner (2020): "The Covid-19 Pandemic and Corporate Dividend Policy," Working Paper.

Farhi, E., and F. Gourio (2018): "Accounting for macro-finance trends," *Brookings Papers on Economic Activity*, pp. 147–223.

Giglio, S., and B. Kelly (2018): "Excess Volatility: Beyond Discount Rates," *The Quarterly Journal of Economics*, 133(1), 71–127.

Giglio, S., M. Maggiori, and J. Stroebel (2014): "Very Long-run Discount Rates," *Quarterly Journal of Economics, forthcoming.*

Goncalves, A. (2020): "The Short Duration Premium," Working Paper.

Gordon, R. J. (2016): *The Rise and Fall of American Growth.* Princeton University Press, Princeton, NJ, 1 edn.

Gormsen, N., and E. Lazarus (2020): "Duration Driven Returns," Working Paper.

Gurkaynak, R., B. Sack, and J. Wright (2006): "The U.S. Treasury Yield Curve: 1961 to the Present," Working Paper.

Hansen, L. P., and K. Singleton (1983): "Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns," *Journal of Political Economy*, 91, 249–65.

Jagannathan, R., E. McGrattan, and A. Scherbina (2000): "The Declining U.S. Equity Premium," *Federal Reserve Bank of Minneapolis Quarterly Review*, 24(4), 3–19.

Katz, M., H. N. Lustig, and L. N. Nielsen (2015): "Are Stocks Real Assets?," Working Paper, AQR Capital Management, LLC.

Krishnamurthy, A., and A. Vissing-Jorgensen (2012): "The Aggregate Demand for Treasury Debt," *Journal of Political Economy*, 120(2), 233–267.

Lettau, M., S. C. Ludvigson, and J. A. Wachter (2004): "The Declining Equity Premium: What Role Does Macroeconomic Risk Play?," NBER Working Paper No. 10270.

Lettau, M., and J. A. Wachter (2007): "Why is Long-Horizon Equity Less Risky? A Duration-Based Explanation of the Value Premium," *Journal of Finance*, 62, 55–92.

——— (2010): "The Term Structures of Equity and Interest Rates," *Journal of Financial Economics,* forthcoming.

Lustig, H., J. Brooks, and M. Katz (2019): "Post-FOMC Announcement Drift in U.S. Bond Markets," Working Paper, Stanford GSB.

Lustig, H., Z. Jiang, S. V. Nieuwerburgh, and M. Z. Xiaolan (2019): "US Government Debt Valuation Puzzle," Working Paper, Stanford GSB.

Lustig, H., A. Verdelhan, and S. V. Nieuwerburgh (2013): "The Wealth-Consumption Ratio," *Review of Asset Pricing Studies*, 1, 38–94.

Mehra, R., and E. C. Prescott (1985): "The Equity Premium Puzzle," *Journal of Monetary Economics*, 15, 145–161.

Nagel, S. (2016): "The Liquidity Premium of Near-Money Assets," *Quarterly Journal of Economics*, 131, 1927.

Novy-Marx, R., and J. Rauh (2011): "Public Pension Promises: How Big Are They and What Are They Worth?," *Journal of Finance*, 66, 1207–1245.

Sharpe, W. (1964): "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk," *Journal of Finance*, 19, 425–444.

Shiller, R. (1979): "The Volatility of Long-Term Interest Rates and Expectations Model of the Term Structure," *Journal of Political Economy*, 87(6), 1190–1219.

Shiller, R. J. (1981): "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?," *American Economic Review*, 71, 421–436.