NBER WORKING PAPER SERIES

DURATION-BASED STOCK VALUATION

Jules H. van Binsbergen

Working Paper 27367
http://www.nber.org/papers/w27367

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2020

I thank Anat Admati, Jonathan Berk, Peter DeMarzo, Darrell Duffie, Vincent Glode, Hanno Lustig, Christian Opp, Anamaria Pieschacon, Josh Rauh, Myron Scholes and Ken Singleton for helpful comments and suggestions. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

Duration-Based Stock Valuation
Jules H. van Binsbergen
NBER Working Paper No. 27367
June 2020
JEL No. E2,E21,E4,G1,G11,G12,G15,G31,G32,G5,G52,O4

ABSTRACT

Interest rates across maturities have dropped to all-time low levels around the world. These unexpected shocks to discount rates have an important effect on the valuation of long duration assets. To quantify this effect, I construct a number of counterfactual fixed income portfolios that match the duration of the dividend strips of the aggregate stock market. I show that such fixed income portfolios have performed as well, if not better, than the U.S. stock market in the past five decades, while exhibiting similar (or higher) levels of volatility. Therefore, investors have received little to no compensation for taking long duration nominal dividend risk in the past half century. Further, if anything, stocks seem to have too little volatility (not excess volatility) compared to these fixed income counterfactuals. I discuss several explanations for these findings, including a secular decline in economic growth rates, dividends' potential to hedge against inflation, as well as the diversification of dividend risk across maturities. These results also have important implications for research on the cross-section of stock returns and capital structure.

Jules H. van Binsbergen
The Wharton School
University of Pennsylvania
3620 Locust Walk
Philadelphia, PA 19104
and NBER
julesv@wharton.upenn.edu

1. Introduction

In the past several decades both nominal and real 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 roughly doubled in price (on May 22nd 2020). Its yield-to-maturity has fallen below 1%, increasing its duration to 57 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 these unexpected downward shocks to 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 might have expected similar performance.[1] Answering this question is also important given that this secular decline in discount rates cannot continue much further due to a lower bound on nominal interest rates. These recent valuation windfalls are therefore unlikely to repeat themselves, making historical (average) realized returns less informative about future expected returns for all long duration assets. As such, one may wonder how useful these backward-looking data moments are to fit stationary theoretical asset pricing models.

To quantitatively evaluate the importance of the secular trends in interest rates on the measurement of realized returns, 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 S&P500 index, that investor had invested in a portfolio of U.S. 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.

Using data between 1996 and 2020, I show that this fixed income investor would have achieved at least similar and likely substantially better return performance than an investor

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

who invested in the stock market index. This indicates that over this sample period, investors have received no additional return compensation for taking long-duration dividend risk compared to nominally risk free government bonds. When expanding the sample to 1970-2020, I find similar results in that the realized average compensation for dividend risk is no more than about 1% per year and arguably much less. I discuss several explanations for this seeming stock market underperformance including (1) a secular decline in long-term expected dividend growth rates, (2) an increase in the dividend risk premium going forward (3) the inflation-hedging properties of long-term dividends, and (4) the diversification of dividend risk across maturities.

The results presented in this paper can also give some guidance to theoretical asset pricing models. The asset pricing literature has spent a considerable amount of time and effort studying four major puzzles. First, the average return on stocks has been 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 second puzzle), risk premia seem 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 of explaining the average return differentials between stocks in the cross-section.

Binsbergen, Brandt, and Koijen (2012), Binsbergen, Hueskes, Koijen, and Vrugt (2014) and Binsbergen and Koijen (2017) show that all these four puzzles occur when pricing short duration dividend claims with maturities up to 7 years. Using data starting in 1996, these authors show that (1) the average returns on short-term dividend strips are higher than their corresponding government bond strip, 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 relatively low, (4) the average returns on short-term strips are also higher than what the CAPM predicts. These results on short-term claims are 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. That is, if in the data dividend strips exhibit the same behavior as the index, but the models can only explain that behavior for the index (the long-duration claim), the key mechanism that resolves the puzzle may not be right.

This paper adds to this literature by showing that the compensation for 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 equity premium puzzles for short duration 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 are another important avenue to make progress on these questions.[2]

These findings also have important implications for defined benefit pension plans around the world, which have been in an underfunding crisis for the past decades.[3] These defined benefit plans that have long duration, often risk free, promises (liabilities) to pension holders have been trying to gamble their way out of their underfunded status by investing in long duration equities. 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 promises 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, which has been zero or even negative in the past half century.

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. 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 in Section 7. I discuss several potential explanations in Section 8 and further discuss the implications for excess volatility and the importance of stock repurchases and tradable bond portfolio data. I discuss the importance for the cross-section of stock returns in Section 9 and conclude in Section 10 with a stock

[2]See Goncalves (2020) and Gormsen and Lazarus (2020) for recent contributions in this area.
[3]See for example Novy-Marx and Rauh (2011) and Binsbergen and Brandt (2015).

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:

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

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:

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

The additional expected holding period return $\psi_{t,n}$ that an investor earns over and above the return on the corresponding risk free bond by investing in risky dividends is then 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:

$$\mu^s_t = E_t\left[R^s_{t+1}\right] = \sum_{n=1}^{\infty} w_{t,n} E_t\left[R^d_{t+1,n}\right]. \tag{10}$$

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 risk free government bond counterparts:

$$\Psi_0 = E\left[E_t \sum_{n=1}^{\infty} w_{t,n}\psi_{t,n}\right] = E\left[\mu^s_t - E_t \sum_{n=1}^{\infty} w_{t,n}\mu^b_{t,n}\right]. \tag{11}$$

I estimate this quantity through:

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

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 given 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 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 Ψ_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, as discussed in Section 8.

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 bond index funds as an additional source of long-term bond data and to verify (where possible) the accuracy of the implied bond returns that follow from the yield data provided by Gurkaynak, Sack, and Wright (2006).

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.

4. 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. 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.

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).

The duration of the aggregate stock market plausibly lies somewhere between 10 and 50 years. 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 ($\hat{\mu}_n^b$).

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 dropping 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 shows a highly similar pattern. However, it should be noted that average simple returns are the preferred measure to estimate expected returns, not mean log returns.

5. Strip-Matched Counterfactuals

In this section I discuss several weighting schemes as defined in Equation 3. I discuss several 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 four main cases. As a first simple weight-generating model, consider the static Gordon growth model:

$$S_t = \frac{D_t(1+g)}{\mu^s - g} \tag{13}$$

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

$$\mathcal{P}_{t,n} = D_t \left(\frac{1+g}{1+\mu^s} \right)^n, \tag{14}$$

which implies a weight equal to:

$$w_{t,n} = (\mu^s - g) \frac{(1+g)^{n-1}}{(1+\mu^s)^n}. \tag{15}$$

As the length of the holding period return converges to 0 (that is, going from annual, to monthly to daily returns), this weighting scheme is only a function of $\mu^s - g$, 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). This is implies that I set the portfolio weight of this cutoff point to the sum of the remaining weights:

$$w_t^{CO} = \sum_{n=CO}^{\infty} w_{t,n} \tag{16}$$

The three graphs below plot the weighting schemes aggregated up to annual weights for $\mu^s - g = 0.06$ (in Figure I), $\mu^s - g = 0.03$ (in Figure II) and $\mu^s - g = 0.02$ (in Figure III) 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 16.

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 in the computations I present below. As such, the estimations I present will be conservative.

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

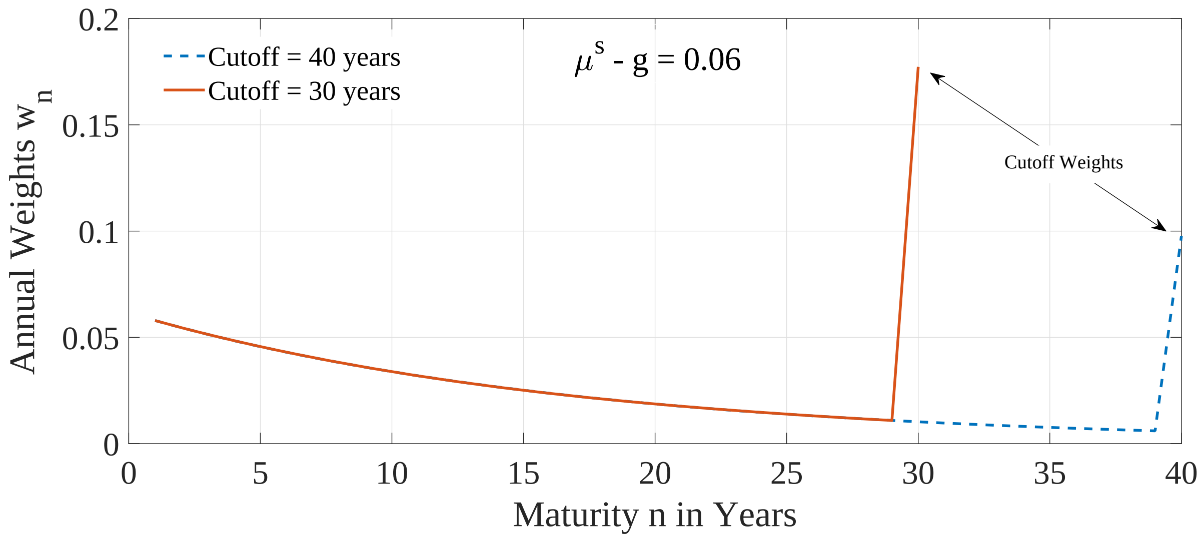

FIGURE I

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.

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 (i.e $\mu^s - g$) set to lower values, the longer duration bonds receive more weight in the calculation, thereby increasing the average return of the counterfactual portfolio and lowering the implied realized long-term dividend premium. When the cutoff is set to 360 months and $\mu^s - g$ (i.e. the long-term dividend yield) is set to 0.06, the estimated compensation that investors have received for dividend risk equals 1.36% (in line with 10-15 year constant maturity bond counterfactuals). For all the other counterfactuals, the duration is larger leading to

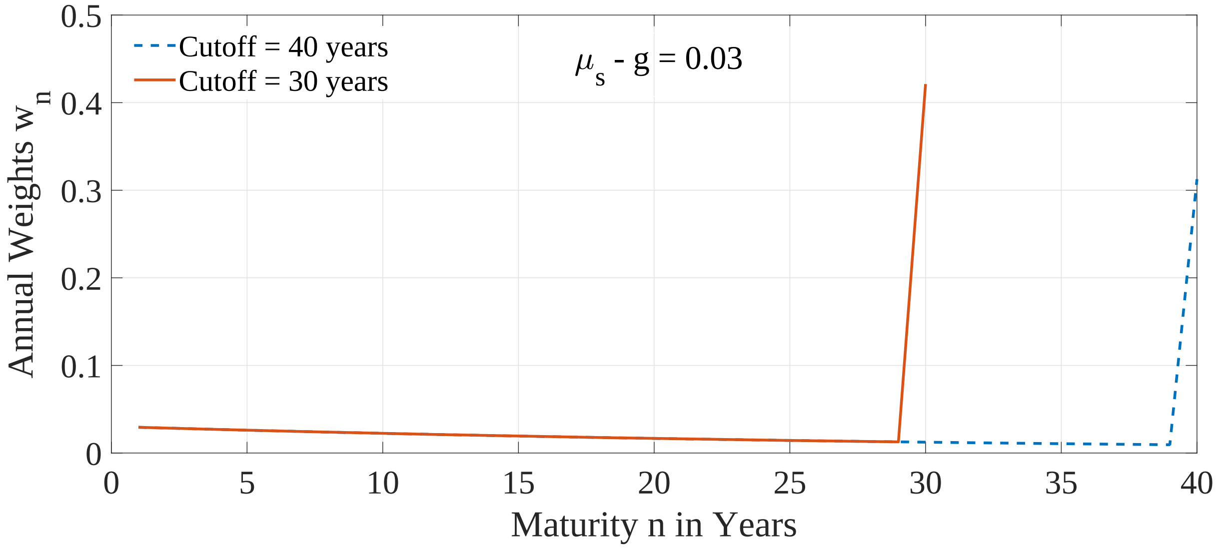

FIGURE II

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.

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 V and VI are very similar to those of counterfactuals III and IV. 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 IV and VI.

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. The results are plotted in Figure IV.

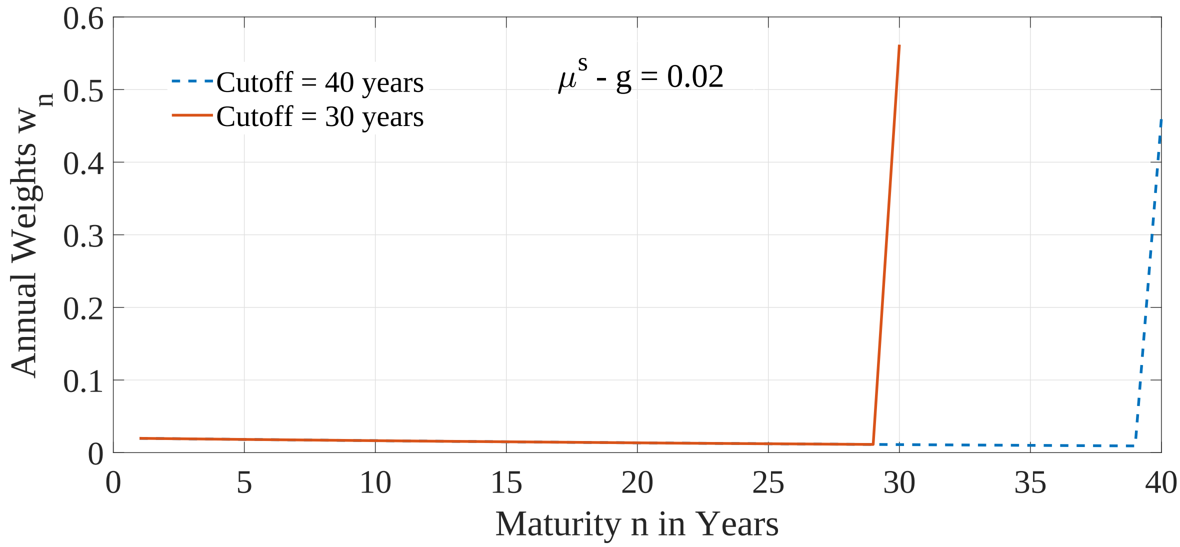

FIGURE III

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.

The figures shows that for the available 15 years of annual 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 would lead to upper bounds 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.

6. Expanded Sample: 1970-2020

I now double the sample from 25 years to 50 years.[4]. 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 recent sample. The average returns on the S&P500 index are roughly

[4]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.

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
$\sum_{n=1}^{\infty} w_{t,n}\mu_{t,n}^b$	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
Annualized diff in mean log rets	0.0063	0.0011	-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 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

equivalent to those of constant maturity bonds between 15 and 20 years of maturity.

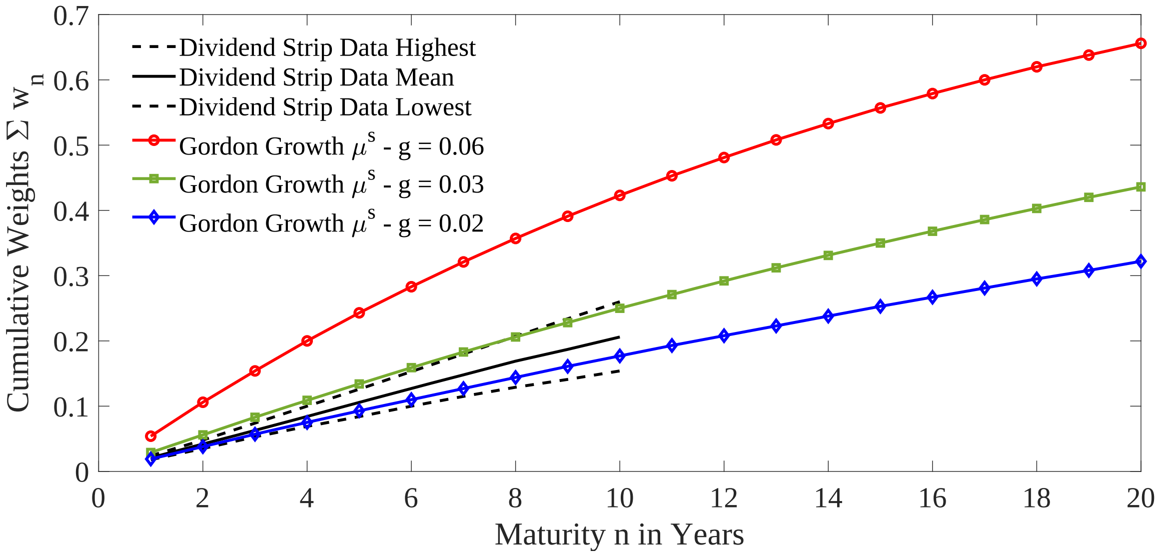

FIGURE IV

Cumulative Strip Weights by Maturity: Data vs Models The graph plots the cumulative weights $\sum_1 N w_n$, 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 2004 and 2020.

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.

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, the average return differential shrinks. Note that in this case there is a difference between mean log returns and mean simple, as the long-duration bond portfolios are more volatile in this expanded sample, and the difference between mean simple returns and mean log returns is $\frac{1}{2}\sigma^2$ under normality.

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 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 ($\hat{\mu}^b_n$) 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).

Finally, in V I plot 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. The graph once again confirms the insights from the previous analyses: bonds have done as well, if not better than the market. Interestingly, the graph reveals that there is a negative low-frequency correlation between the stock and replicating bond portfolio. This is consistent with long-term expected growth rates being positively correlated with interest rates, and risk premia being negatively correlated.

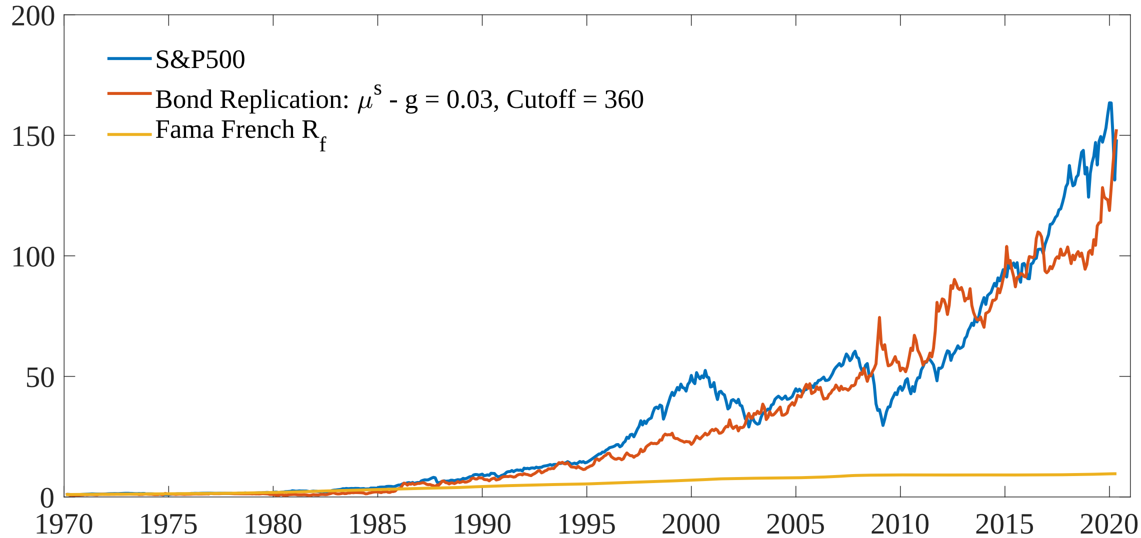

FIGURE V

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.

7. International Evidence

In this section I repeat the analysis using European data between January 1996 and April 2020 using data from Wright (2011), which has German bond maturities going out to 15 years. I use returns on the Eurostoxx index as the stock market proxy. The results for the constant maturity bond portfolio are reported in Table 6. All returns are in local currency. The results are even starker than for the U.S. in the sense that even the 10-year constant maturity zero coupon bonds have outperformed the index over this sample period.

In Table 7 I repeat the analysis from Table 2 using German data between 1996 and 2020. Once again, the results are similar (if not stronger) as those for the U.S. Due to the limited duration availability, I include the 5-year duration in the computation below, which is not included in the U.S. analysis.

Maturity in years	1	1.5	2	3	4	5	10	15	Eurostoxx
Mean	0.0023	0.0026	0.0029	0.0035	0.0041	0.0047	0.0068	0.0083	0.0055
St. Dev.	0.0022	0.0032	0.0043	0.0066	0.0087	0.0107	0.0200	0.0293	0.0473
Mean log	0.0018	0.0020	0.0022	0.0027	0.0031	0.0035	0.0051	0.0060	0.0040

Table 6

Monthly Returns on Constant Maturity Zero Coupon Bonds: Germany. 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 Eurostoxx index.

Duration in years	5	10	15
$\hat{\mu}^s - \hat{\mu}^b$	0.0018	-0.0005	-0.0023
t-stat on difference	0.59	-0.15	-0.66
$12(\hat{\mu}^s - \hat{\mu}^b)$	0.0210	-0.0057	-0.0271
Annualized difference in mean log returns	0.0056	-0.0128	-0.0243

Table 7

Monthly Return Differences between the Eurostoxx and Constant Maturity German Zero Coupon Bonds. The the second row in the table lists the difference between the monthly returns on the Eurostoxx 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).

8. Potential Explanations and Discussion

In this section I discuss four potential explanations for the findings as well as several remaining issues and potential concerns.

8.1. 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. This would suggest an important (inverse) link between the size of the inflation risk premium and the dividend risk premium.

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.[5]

Third, there is the possibility that while investors were expecting a large return compensation for long-duration dividend risk ex ante, they ended up not receiving it ex post for the past five decades. 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 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, as I will further explore in the conclusion. Future dividend growth realizations will then be low.

Finally, there is the possibility that stocks have not performed well due to a secular increase in long-term future 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, returns can only be high of future growth remains high. After all, the Gordon growth formula states the expected returns is the sum of the dividend yield and the expected

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

growth rate.

8.2. Excess Volatility

One of the most surprising features of short-term dividend strips is that they seem excessively volatile relative to their subsequent realizations.[6] 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 variation deepens the Shiller (1981) excess volatility puzzle.

In fact, generating excess volatility for long duration claims is relatively straightforward, as small changes in discount rates correspond to large changes in prices. After all, bond prices are clearly much more volatile than their subsequent fixed 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. 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 move opposite to interest rates, and growth expectations in lock-step with interest rates (as suggested by V, they both have a tempering effect on the volatility of the equity claim compared to those of the bond claim.

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 most 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. This logic also implies that if long-term dividend risk premia are on average low (and potentially even zero), they have less potential of generating excess volatility, though obviously the duration still acts as a lever.

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

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 expectations.

8.3. Stock Repurchases

One question that could come to mind is whether stock repurchases, which have gained in popularity in recent decades, could affect the computations in this paper. The present-value relationship 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 issuances and repurchases. Repeating the calculations here for the total equity value of the U.S. stock market that includes issuances and repurchases is an interesting avenue to take. 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 would have a duration shorter than 10 years or alternatively have weights corresponding to a dividend yield of more than 6%, though future research in this direction may prove otherwise.

8.4. Tradable Indices

The 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 VII plots the historical effective duration of this fund between June 2002 and April 2020.[7]

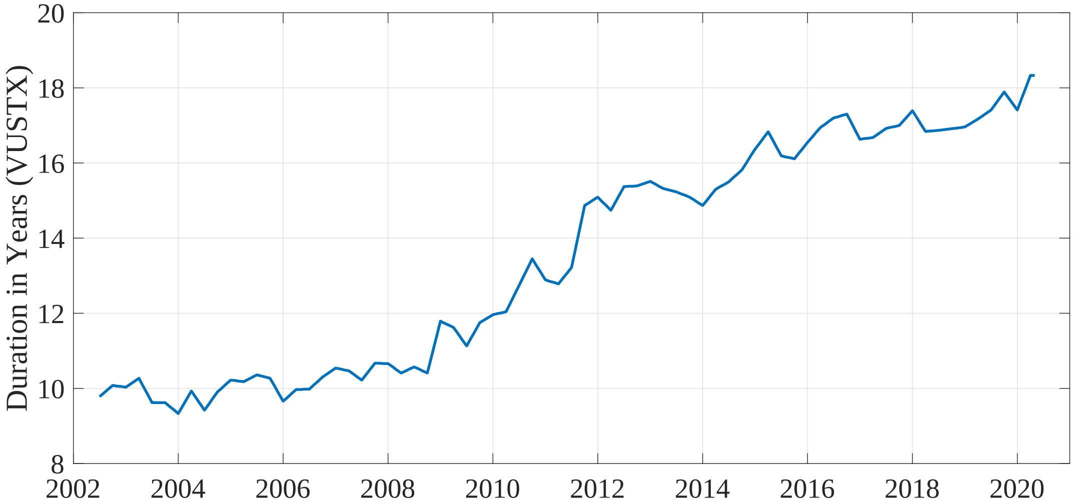

FIGURE VI

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

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 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.

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

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.

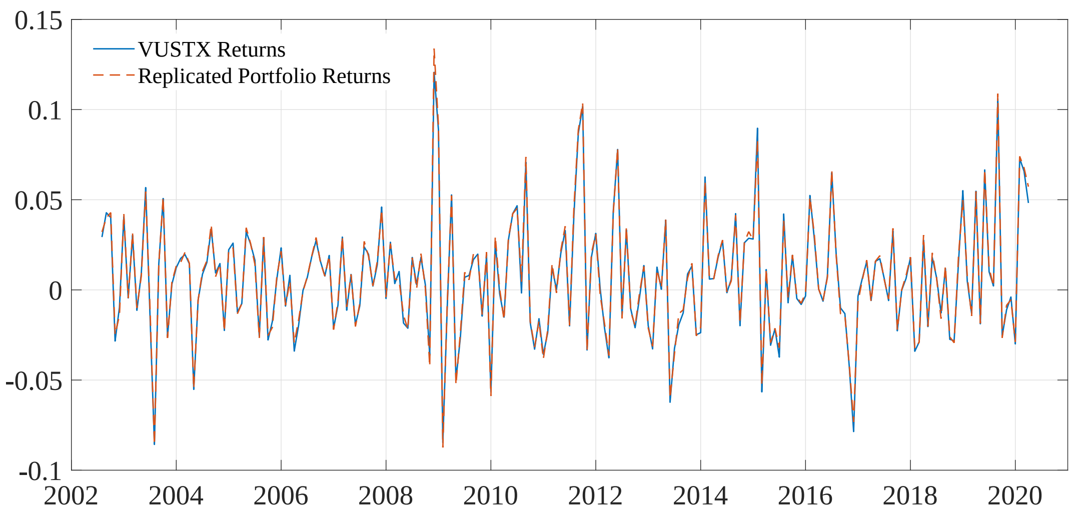

FIGURE VII

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).

9. Implications for the Cross-Section of Stock Returns and Real Estate

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. 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). Interest rate and inflation risk (or lack thereof)? This seems particularly pressing given that the CAPM is not known to price government bonds (of all maturities) well.

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.[8]

10. 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 the stock market has exhibited little to no outperformance over these fixed income counterfactuals implying that investors have not received any compensation for taking long duration dividend risk. 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 such 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 divi-

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

dend growth rates over these decades. While the expectations hypothesis for interest rates makes it somewhat easier to observe unexpected downward shocks to the risk free discount rates, observing unexpected shocks to long-term expected dividend growth rates is more challenging as long-run expected growth measures are not in ample supply.

One may wonder whether the results in this paper have external validity going forward. Perhaps the Japanese experience 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 in value. Furthermore, long-term bond yields have decreased even further. This suggests that the American and European results presented in this paper could in fact repeat themselves. The Japanese data is also informative about the type of stock and bond returns investors should expect in the U.S. and Europe for the next 25 years. It seems increasingly likely 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 major 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 such that return moments can be meaningfully interpreted in stationary model environments. Alternatively, it could be helpful to explicitly model secular trends 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.[9] 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 previously ignored insights.

[9] See also recent work by Lettau and Wachter (2007) and Lettau and Wachter (2010) in this area.

References

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. 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.

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.

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.

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.

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.

MEHRA, R., AND E. C. PRESCOTT (1985): "The Equity Premium Puzzle," *Journal of Monetary Economics*, 15, 145–161.

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

Wright, J. (2011): "Term Premia and Inflation Uncertainty: Empirical Evidence from an International Panel Dataset," *American Economic Review*, 101(4), 1514–1534.