What Moves Equity Markets?

A Term Structure Decomposition for Stock Returns

Andrei S. Gonçalves*†

This Version: July, 2021

Link to Most Recent Version

Abstract

Several papers decompose stock returns into cash flow and discount rate news to study equity market fluctuations. This paper develops an alternative return decomposition based on the fact that equity movements originate from variation in the present values of dividends with different maturities. I find that roughly 60% of equity volatility comes from the present value of dividends with maturities beyond 20 years and that cash flow shocks drive volatility in short-term present values whereas discount rate news are responsible for volatility in long-term present values. I also provide three further empirical applications of this new equity term structure decomposition.

JEL Classification: E32; G10; G11; G12.

Keywords: Equity Volatility; Equity Term Structure; Equity Strips; Equity Duration; Campbell-Shiller Decomposition.

*Kenan-Flagler Business School, University of North Carolina at Chapel Hill, 4110 McColl Building, 300 Kenan Center Drive, Chapel Hill, NC 27599. E-mail: Andrei_Goncalves@kenan-flagler.unc.edu.

†I am grateful for the very helpful comments from Scott Cederburg (discussant), Jennifer Conrad, Mohit Desai, Kewei Hou, Gregory Leonard, René Stulz, and Lu Zhang, as well as seminar participants at the 2020 NFA. All remaining errors are my own. First draft: May 2019.

Introduction

Understanding the drivers of fluctuations in equity markets has been a longstanding pursuit in asset pricing. Traditionally, the literature has approached this task using a decomposition of equity volatility into news about cash flows and discount rates (e.g., Campbell and Shiller (1989), Campbell (1991), and Binsbergen and Koijen (2010)). However, decomposing returns into cash flow and discount rate components is not the only way to study movements in equity markets. Since equity prices can be expressed as the discounted value of future dividends, an alternative perspective is to view equity volatility as originating from variability in the present values of dividends with different maturities.

In this paper, I formalize this alternative perspective by developing a new log-linear decomposition of equity returns into returns associated with the present values of dividends with different maturities. This "equity term structure decomposition" allows me to assess the relative importance of short- and long-term dividend present values in explaining equity volatility. Hence, I also provide an empirical exercise that reveals that a substantial fraction of equity return volatility (roughly 60%) is driven by variation in the present value of very long-term dividends (defined as dividends with maturities beyond 20 years).

To start, I show that an equity contract can be viewed as a collection of equity strips, where "equity strips" are dividend present values obtained by discounting expected dividends by their maturity-matched equity expected returns. I then demonstrate that the log-linear return decomposition in Campbell (1991) can be used to express equity returns as a weighted average of returns on these equity strips, resulting in a new term structure decomposition for stock returns. I also show that this equity term structure decomposition is intrinsically connected to the risk premia term structure literature (see Binsbergen and Koijen (2017) for a review), and that relying on equity strips has important advantages relative to the traded dividend claims used in this literature for the purpose of decomposing equity volatility into its term structure components (my literature review elaborates on these advantages).

Empirically, I estimate a Vector Autoregressive (VAR) system to predict dividend growth and equity returns for the overall US equity market and use the VAR results to decompose equity volatility into movements in equity strips of different maturities. From this exercise, I provide three novel empirical facts about equity volatility.

First, I find that roughly 60% of equity return volatility comes from the present value of dividends with maturities beyond 20 years. This 60% benchmark is highly robust to empirical decisions related to the VAR specification and estimation, and indicates that we cannot fully understand volatility in equity markets without a deep economic understanding of the drivers of time variation in the present value of very long-term dividends. In this vein, I show that the importance of long-term dividends in explaining equity return volatility originates from a combination of the long average equity duration of the market (of roughly 30 years) and the upward sloping term structure of equity strip exposures to the overall equity market, with both factors playing a non-trivial role in the decomposition.

Second, I demonstrate that most of the return volatility associated with the present value of short-term dividends comes from cash flow shocks while discount rate news are mainly responsible for return volatility linked to long-term dividends. This result indicates that movements in the short and long ends of the equity term structure are driven by fundamentally different economic forces, which suggests the need for rich economic models to properly explain the determinants of equity volatility.

And third, I provide evidence that the relative importance of long-term dividend present values in explaining equity volatility varies strongly over time. I start by showing that there are extended periods in which long-term equity strips outperform short-term ones (and vice versa). These periods then translate into cycles in the fraction of equity value due to the present value of long-term dividends, with periods in which it is as low as 35% (e.g., early 1980s) and periods in which it reaches well beyond 50% (e.g., late 1990s). These results imply that equity duration varies over time, and thus I develop a conditional version of my term structure decomposition to account for time-varying equity duration. Using this conditional variance decomposition, I show that the fraction of equity volatility due to long-term dividends largely varies around its mean (of roughly 60%), with periods in which it is close to 50% and periods in which it reaches almost 65%.

While it is well understood that equity movements can be associated with long-term expectations, to the best of my knowledge, the aforementioned results represent the first quantification in the literature of the importance of short- versus long-term dividend present values in explaining equity market volatility. As such, the fact that around 60% of equity volatility is due to the present value of dividends with maturities beyond 20 years provides a benchmark to discipline our belief about the importance of long-term expectations in explaining equity volatility. Prior to this paper, such benchmark was not available, and thus one could easily believe that this number was, for example, 30% or 90%. To further demonstrate the importance of my equity term structure decomposition, I consider three applications of its use beyond decomposing equity volatility.

First, I contrast equity strip yields during the COVID crisis and the 2007-2009 financial crisis (with equity strip yields defined analogously to the equity yields of Binsbergen et al. (2013)). Note that an increase in equity strip yields tends to reflect a deterioration in macroeconomic prospects in the sense that it captures an increase in priced risks and/or a decline in expected growth. During March 2020, the spread between the 1-year and the 20-year equity strip yields grew to around 20% but it fully reverted to its typical value of around 0% by the end of 2020. While these figures suggest the COVID crisis had a non-trivial temporary effect on macroeconomic prospects, an analysis of equity strip yields during the 2007-2009 financial crisis reveals a much worse picture. For instance, the spread between the 1-year and the 20-year equity strip yields reached levels above 40% in late 2008 and equity strip yields did not stabilize until 2010. As such, from the perspective of changes in expected growth and priced risk, the COVID crisis has been relatively moderate and short-lived, specially in comparison to the 2007-2009 financial crisis.

Second, I explore how the return correlation between equity strips and bond portfolios has changed over time. Several papers show that the return correlation between equities and bonds has largely declined over the last decades, effectively becoming negative post 2000 (e.g., Campbell, Sunderam, and Viceira (2017)). I show that while this pattern holds for equity strips of different maturities, it is much weaker for short-term equity strips so that the overall pattern described in the literature is mostly a phenomenon associated with long-term equity strips. For instance, while the return correlation between the 50-year equity strip and a 10-year bond portfolio substantially declined from pre to post 2000 (from 0.50 to -0.50), the analogous correlation for the 1-year equity strip declined must less (from 0.10 to -0.30). Mechanisms attempting to explain the time variation in the equity-bond return correlation need to be consistent with this novel fact to be empirically credible.

And third, I show that the time variation in equity strip expected returns (which reflect investment opportunities) can be summarized by two factors capturing expected returns on short-term and long-term equity strips. I then use this information to provide an economically motivated restriction on an empirical application of the Intertemporal CAPM (ICAPM) of Merton (1973) that guards against the "fishing license" introduced by reduced-form ICAPM implementations (Fama (1991)). Specifically, I test the ability of the ICAPM to jointly price equity, Treasury bond, and corporate bond portfolios sorted on cash flow duration. My ICAPM implementation allows for the pricing of all state variables in the VAR system, but restricts their risk prices so that state variables are priced only in accordance to their link to the short- and long-term equity strip expected return factors. The results suggest that the ICAPM prices the testing assets well (and much better than the CAPM). Moreover, imposing the economic restrictions described is important because otherwise the model overfits the data and yields statistically insignificant risk prices given that the unrestricted ICAPM has too many (spurious) degrees of freedom.

In summary, I provide a novel equity term structure decomposition, with its empirical implementation suggesting that around 60% of the overall equity market volatility is driven by variability in the present value of very long-term dividends and that this fraction varies strongly over time. Moreover, I argue that these dividend present values, labeled equity strips, have other important applications in finance and showcase three of these applications. First, I show that my equity term structure decomposition can provide important information about macroeconomic prospects over different horizons. Second, I find that the time variation in the equity-bond return correlation drastically differ across equity strips with different maturities. And third, I use equity strips to discipline an ICAPM implementation to price the cross-section of equity and bond returns sorted on cash flow duration.

The main contribution of this paper is to develop a new term structure decomposition for equity returns. The proposed decomposition offers a new quantitative tool to an extensive literature that uses the traditional decomposition of equity returns into cash flow and discount rate news to explore different economic questions.[1]

The main empirical results in this paper relate more directly to the subset of this literature that quantifies the importance of discount rate and cash flow news in explaining equity volatility (see Koijen and Nieuwerburgh (2011) for a review of this part of the literature). Collectively, these papers show that both cash flow and discount rate news are important determinants of equity volatility, with discount rate news being relatively more important, specially over the last several decades. I add to this literature by quantifying the importance of the present value of short- and long-term dividends in explaining movements in equity markets and I find that both are important, with very long-term dividend present values being particularly relevant.

This paper is also connected to the recent literature on the equity term structure (see Binsbergen and Koijen (2017) for a review).[2] While this literature relies on tradable dividend claims to study the term structure of equity risk premia, I directly rewrite the return decomposition in Campbell (1991) to decompose equity returns into the effects coming from dividend present values with different maturities (which I label equity strips). I show that if the level (but not the slope) of the term structure of equity risk premia varies over time, then shocks to equity strips equal shocks to the tradable dividend claims studied in this literature.

[1]Some papers in this literature are Campbell and Shiller (1989), Campbell (1991), Cochrane (1992), Campbell and Ammer (1993), Campbell and Mei (1993), Campbell (1996), Vuolteenaho (2002), Campbell and Vuolteenaho (2004), Larrain and Yogo (2008), Campbell, Polk, and Vuolteenaho (2009), Chen (2009), Binsbergen and Koijen (2010), Cochrane (2011), Chen, Da, and Priestley (2012), Chen, Da, and Zhao (2013), Golez and Koudijs (2018, 2020), Cederburg (2019), and Weber (2021).

[2]See, for instance, Binsbergen, Brandt, and Koijen (2012), Binsbergen et al. (2013), Binsbergen and Koijen (2017), Cejnek and Randl (2016), Cejnek and Randl (2020), Gormsen (2021), Gonçalves (2021a), and Binsbergen (2021).

While tradable dividend claims have the advantage of being observable, they cannot be used to fully decompose equity return volatility because only the first few maturities of dividend claims are observable (typically up to seven years) and for a relatively short period (typically starting in the 2000s). In contrast, when paired with an econometric model, equity strips can be recovered for all maturities and over arbitrary sample periods, making it feasible to fully decompose equity return volatility into its term structure components.[3]

As I show, equity duration is one of the components that influences the term structure decomposition of equity volatility. Hence, my work is also linked to the equity duration literature.[4] My contribution to this literature is to show that while equity duration is an important component of whether equity volatility is driven by short- or long-term dividend present values, it is not the only component, with the term structure of equity strip exposures to the overall equity market playing an equally important role.

Finally, my equity term structure decomposition can be used to ask many questions that are linked to literatures unrelated to decomposing equity volatility. To showcase this aspect, my empirical applications demonstrate how to use my equity term structure decomposition to shed light on issues related to changes in macroeconomic prospects over different horizons, the time-varying equity-bond return correlation, and the ICAPM of Merton (1973).

The rest of this paper is organized as follows. Section 1 develops the new equity term structure decomposition, Section 2 details the empirical and econometric design used in this paper, Section 3 provides the main empirical results, Section 4 details three further empirical applications of my equity term structure decomposition, and Section 5 concludes. The Internet Appendix contains technical derivations, details about the econometric methodology as well as data sources/measurement, and supplementary empirical results.

[3] Some recent papers (e.g., Gonçalves (2021a), Giglio, Kelly, and Kozak (2020) and Andrews and Gonçalves (2020)) use an asset pricing model together with an econometric specification to recover the prices of tradable dividend claims at arbitrary maturities and sample periods. The advantage of equity strips is that they do not require an asset pricing model. The disadvantage is that they are not informative about the behavior of tradable dividend claims, but this disadvantage is irrelevant for the purpose of this paper as I do not study the term structure of equity risk premia.

[4] See, for example, Dechow, Sloan, and Soliman (2004), Da (2009), Weber (2018), Gonçalves (2021b), Chen and Li (2018), Golez and Koudijs (2020), Gormsen and Lazarus (2021), and Chen (2021).

1 A New Term Structure Decomposition for Stock Returns

This section presents a new term structure decomposition for equity returns. Subsection 1.1 defines equity strips, which are present values of single dividends, Subsection 1.2 shows how these equity strips can be used to decompose equity returns and volatility, and Subsection 1.3 provides a link between equity strips and the tradable dividend claims studied in the equity term structure literature (see Binsbergen and Koijen (2017)). Internet Appendix A provides all technical derivations.

1.1 Defining Equity Strips

Start from the definition of a gross equity return and isolate the equity price to get $P_t = (P_{t+1} + D_{t+1})/R_{t+1}$, which can be iterated forward to yield:

$$P_t = \sum_{h=1}^{\infty} \mathbb{E}_t \left[\left(\prod_{j=1}^{h} R_{t+j} \right)^{-1} D_{t+h} \right] \equiv \sum_{h=1}^{\infty} PV_t^{(h)} \tag{1}$$

where $PV^{(h)}$ represents the "present value" of the h-year dividend.

As the definition shows, $PV^{(h)}$ is obtained by discounting the h-year expected dividend using the h-year equity discount rate (instead of the h-year dividend discount rate). Consequently, $PV^{(h)}$ does not represent the price of a tradable claim. Nevertheless, $PV^{(h)}$ still represents the contribution of the dividends accruing in h years to the current equity value, and thus it can be used to strip the equity price into maturity specific components. As such, I refer to $PV^{(h)}$ as an equity strip present value. This terminology is analogous to the "dividend strip" terminology used in the literature to refer to tradable dividend claims (see Subsection 1.3 for details on the link between equity strips and dividend strips).

Using a log-linear approximation for $PV_t^{(h)}/D_t$, Equation 1 implies that equity strip present values depend on current dividends, expected dividend growth, and expected equity returns:[5]

[5] If r and Δd are conditionally homoskedastic and normally distributed, Equation 2 holds exactly after adding a Jensen's inequality constant to it. This constant does not affect any the term structure decomposi-

$$log(PV_t^{(h)}) = log(D_t) + \underbrace{\mathbb{E}_t\left[\sum_{j=1}^{h} \Delta d_{t+j}\right]}_{g_t^{(h)}} - \underbrace{\mathbb{E}_t\left[\sum_{j=1}^{h} r_{t+j}\right]}_{dr_t^{(h)}} \tag{2}$$

and combining this present value equation with the definition of equity strip log returns, $r_t^{(h)} = log(PV_t^{(h-1)}/PV_{t-1}^{(h)})$, yields:

$$r_t^{(h)} - \mathbb{E}_{t-1}[r_t^{(h)}] = \Delta d_t - \mathbb{E}_{t-1}[\Delta d_t] + \left(g_t^{(h-1)} - \mathbb{E}_{t-1}[g_t^{(h-1)}]\right) - \left(dr_t^{(h-1)} - \mathbb{E}_{t-1}[dr_t^{(h-1)}]\right)$$

or in more compact notation (with $\sim$ representing shocks to contemporaneous variables and N news about future information):

$$\widetilde{r}_t^{(h)} = \widetilde{\Delta d}_t + N_{g,t}^{(h-1)} - N_{dr,t}^{(h-1)} \tag{3}$$

Equation 3 shows that all equity strips depend on an identical dividend growth shock ($\widetilde{\Delta d}$), but are subject to different news about future dividend growth ($N_g^{(h)}$) and discount rates ($N_{dr}^{(h)}$). For instance, the annual return on a 1-year equity strip is only subject to dividend growth shocks. In contrast, the annual return on a 10-year equity strip is exposed not only to dividend growth shocks, but also to news about the remaining nine years dividend growth and discount rate. As a consequence, the heterogeneity in equity strip returns with different maturities originates from the heterogeneity in these two components.

1.2 Decomposing Stock Returns into Equity Strip Returns

Manipulating Equation 1, we can show that equity returns can be seen as a portfolio of equity strip returns. However, it is empirically easier to work with a log-linear version of this equation. Specifically, Equation 3 is the equity strip analogue of the stock return decomposition in Campbell (1991) and I explore this connection to demonstrate that log stock returns

tions provided in the text beyond adding a constant to them, which is irrelevant when decomposing equity return volatility.

are a weighted average of log equity strip returns.

Start by noting that Campbell (1991)'s log-linear stock return decomposition is given by:

$$\tilde{r}_t \; = \widetilde{\Delta d}_t \; + \; N_{g,t} \; - \; N_{dr,t} \tag{4}$$

where

$N_{g,t} = (\mathbb{E}_t - \mathbb{E}_{t-1}) \left[\sum_{h=1}^{\infty} \rho^h \cdot \Delta d_{t+h} \right]$ is cash flow news (i.e., news about expected growth)

$N_{dr,t} = (\mathbb{E}_t - \mathbb{E}_{t-1}) \left[\sum_{h=1}^{\infty} \rho^h \cdot r_{t+h} \right]$ is discount rate news (i.e., news about expected returns)

Then, Internet Appendix A shows how to rewrite Equation 4 to express stock returns as returns on a portfolio of equity strips:

$$\begin{aligned} \widetilde{r}_t \; &= \widetilde{\Delta d}_t \; + \sum_{h=1}^{\infty} w^{(h)} \cdot N_{g,t}^{(h-1)} \; - \sum_{h=1}^{\infty} w^{(h)} \cdot N_{dr,t}^{(h-1)} \\ &= \sum_{h=1}^{\infty} w^{(h)} \cdot \widetilde{r}_t^{(h)} \end{aligned} \tag{5}$$

where $w^{(h)} = \rho^{h-1} - \rho^h$ are weights that decrease in maturity and satisfy $\sum_{h=1}^{\infty} w^{(h)} = 1$ (with $\rho = 1/(1 + e^{\overline{dp}})$ capturing Campbell and Shiller (1989)'s log-linearization constant).[6]

Internet Appendix A also shows that $\mathbb{E}_{t-1}[r_t^{(h)}] = \mathbb{E}_{t-1}[r_t]$, and thus the term structure decomposition holds for raw returns as well:

$$r_t \; = \sum_{h=1}^{\infty} w^{(h)} \cdot r_t^{(h)} \tag{6}$$

Taking covariance with any arbitrary variable, x, on both sides of Equation 5 yields:[7]

$$Cov(\widetilde{r}_t, x_t) \; = \sum_{h=1}^{\infty} w^{(h)} \cdot Cov(\widetilde{r}_t^{(h)}, x_t) \tag{7}$$

[6] Note that the weights, $w^{(h)}$, depend only on ρ and h but not on t, and thus are constant over time. In Subsection 3.5, I also explore an alternative (conditional) log-linear approximation that allows ρ (and thus the weights) to vary over time.

[7] Given an asset pricing model, Equation 7 can also be applied to risk factors to provide a term structure decomposition for the equity premium. Gonçalves (2021a) explores this result to study the term structure of equity risk premia in the context of Campbell (1993)'s ICAPM.

and, if $x_t = \widetilde{r}_t$, Equation 7 simplifies to:

$$Var(\widetilde{r}_t) \ = \sum_{h=1}^{\infty} w^{(h)} \cdot Cov(\widetilde{r}_t^{(h)}, \widetilde{r}_t) \tag{8}$$

Therefore, stock return variance depends on the covariances between equity strips and stock returns. I explore this result in my empirical analysis to demonstrate that a large fraction of stock volatility is due to variation in the present values of dividends with maturities beyond 20 years. That is, $\Sigma_{h>20} w^{(h)} \cdot Cov(\widetilde{r}_t^{(h)}, \widetilde{r}_t)/Var(\widetilde{r}_t)$ is large (close to 60% in my baseline specification).

Equations 3 and 4 imply that equity return variance can also be decomposed into covariances with cash flow shocks/news and discount rate news:

$$Var(\widetilde{r}_t) \ = Cov(\widetilde{\Delta d}_t, \widetilde{r}_t) \ + \ Cov(N_{g,t}, \widetilde{r}_t) \ + \ Cov(-N_{dr,t}, \widetilde{r}_t) \tag{9}$$

$$Var(\widetilde{r}_t^{(h)}) \ = Cov(\widetilde{\Delta d}_t, \widetilde{r}_t^{(h)}) \ + \ Cov(N_{g,t}^{(h-1)}, \widetilde{r}_t^{(h)}) \ + \ Cov(-N_{dr,t}^{(h-1)}, \widetilde{r}_t^{(h)}) \tag{10}$$

Equations 9 is the traditional Campbell (1991) variance decomposition while Equation 10 is the analogous decomposition applied to equity strips with different maturities. In my empirical analysis, I also explore these two equations to show that short-term equity strip volatility is largely due to cash flow shocks/news while discount rate news are mainly responsible for the volatility of long-term equity strips.

1.3 Relation Between Equity Strips and Dividend Strips

The equity term structure literature (see Binsbergen and Koijen (2017) for a review) uses tradable dividend claims (such as dividend futures) to decompose equity returns. Specifically, under the law of one price, there exists a Stochastic Discount Factor, M_t, such that equity prices can be decomposed into the prices of single dividend claims (called dividend strips):

$$P_t = \sum_{h=1}^{\infty} \mathbb{E}_t \left[M_{t \to t+h} \cdot D_{t+h} \right] \equiv \sum_{h=1}^{\infty} P_t^{(h)} \tag{11}$$

Defining dividend strip log returns in the usual way, $r_{d,t}^{(h)} = log(P_t^{(h-1)}/P_{t-1}^{(h)})$, we have:

$$log(P_t^{(h)}) = log(D_t) + \underbrace{\mathbb{E}_t\left[\sum_{j=1}^{h} \Delta d_{t+j}\right]}_{g_t^{(h)}} - \underbrace{\mathbb{E}_t\left[\sum_{j=1}^{h} r_{d,t+j}^{(h-j+1)}\right]}_{ddr_t^{(h)}} \tag{12}$$

and thus

$$\widetilde{r}_{d,t}^{(h)} \;=\widetilde{\Delta d}_t \;+\; N_{g,t}^{(h-1)} \;-\; N_{ddr,t}^{(h-1)} \tag{13}$$

which is analogous to the equity strip return decomposition in Equation 3. The key difference is that $N_{dr,t}^{(h)}$ reflects news about equity discount rates while $N_{ddr,t}^{(h)}$ reflects news about dividend discount rates.

If the level (but not the slope) of the term structure of risk premia varies over time, $\mathbb{E}_t[r_{d,t+1}^{(h)}] = \mathbb{E}_t[r_{t+1}] + \mathbb{E}[r_{d,t+1}^{(h)} - r_{t+1}]$, then $N_{ddr,t}^{(h-1)} = N_{dr,t}^{(h-1)}$ and (unexpected) returns on dividend strips are identical to (unexpected) returns on equity strips, $\widetilde{r}_{d,t}^{(h)} = \widetilde{r}_t^{(h)}$. As such, the equity return decomposition I study in this paper is closely connected to the equity term structure literature, which studies the relation between equity returns and $r_{d,t}^{(h)}$.

I view equity strips and dividend strips as complementary. The advantage of studying equity strips is that, given an econometric model, we can explore the equity term structure over a long sample period and across arbitrary maturities. In contrast, studying dividend strips limits the sample period (as dividend futures started trading in the 21st century) and maturity range (typically to the first seven years), so that we cannot empirically estimate a $Var(\widetilde{r}_t)$ term structure decomposition using dividend strips.[8] The advantage of studying dividend strips is that returns can be obtained directly from trading prices without an auxiliary econometric model.

[8] As explained in the introduction, some recent papers (e.g., Gonçalves (2021a), Giglio, Kelly, and Kozak (2020), and Andrews and Gonçalves (2020)) use an asset pricing model together with an econometric specification to recover the prices of tradable dividend claims at arbitrary maturities and sample periods. Relative to these papers, the advantage of equity strips is that they do not require an asset pricing model. The disadvantage is that they are not informative about the behavior of tradable dividend claims, but this disadvantage is irrelevant for the purpose of this paper as I do not study the term structure of equity risk premia.

2 Empirical Design

This section outlines the empirical implementation of the equity term structure decomposition introduced in the previous section. Internet Appendices B and C provide further details about the econometric design and data measurement/sources.

As explained in the previous section, decomposing equity returns into a portfolio of equity strip returns requires an econometric model to measure shocks to expected dividend growth and expected equity returns. Letting s_t be a state vector that includes Δd_t, my empirical analysis assumes $z_t = [r_{f,t} \quad xr_t \quad s_t]$ evolves as a Vector Autoregression system of order one, VAR(1), in which s_t contains the relevant predictive variables:[9]

$$\begin{aligned} z_t &= \Phi_0 \; + \; \Phi_1 \cdot z_{t-1} \; + \; \widetilde{z}_t \\ &= \Phi_0 \; + \; \Phi_{s,1} \cdot s_{t-1} \; + \; \widetilde{z}_t \end{aligned} \tag{14}$$

where $xr_t = r_t - r_{f,t}$ and $\widetilde{z}_t \overset{i.i.d}{\sim} N\left(0, \Sigma_z\right)$.

Then, letting I_Φ be an identity matrix of the same dimension as Φ_1 and $\mathbf{1}_x$ be a selector vector such that $\mathbf{1}_x^{'} z_t = x_t$, I define $B^{(h)} = (\Phi_1 - \Phi_1^{h+1})(\mathrm{I}_\Phi - \Phi_1)^{-1}$ and substitute it into Equation 3 to get

$$\begin{aligned} \tilde{r}_t^{(h)} &= \mathbf{1}_{\Delta d}^{'} \widetilde{z}_t \; + \; \mathbf{1}_{\Delta d}^{'} \cdot B^{(h-1)} \widetilde{z}_t \; - \; \mathbf{1}_r^{'} \cdot B^{(h-1)} \widetilde{z}_t \\ &= [\mathbf{1}_{\Delta d}^{'} + (\mathbf{1}_{\Delta d} - \mathbf{1}_r)^{'} B^{(h-1)}] \widetilde{z}_t \end{aligned} \tag{15}$$

so that all quantities used in my empirical analysis can be obtained directly from the VAR estimates. For instance, $Cov(\widetilde{r}_t^{(h)}, \widetilde{r}_t) = [\mathbf{1}_{\Delta d}^{'} + (\mathbf{1}_{\Delta d} - \mathbf{1}_r)^{'} B^{(h-1)}] \Sigma_z \mathbf{1}_r$.

I construct r_f from returns on the one-month Treasury bill (available in Kenneth French's data library). Moreover, I measure equity market returns and dividend growth based on a value-weighted portfolio containing all common stocks available in the CRSP dataset. The system also relies on other five predictive variables embedded into z_t (all measured

[9]Following Gonçalves (2021a), I do not include r_t in the set of state variables because the current equity return is not a significant predictor of future equity returns. This decision keeps the VAR more parsimonious and has basically no effect on my results.

in natural log units): dividend yield (dp), one year Treasury yield (ty), term spread (TS), credit spread (CS) and value spread (VS).[10] These are the same predictive variables used in Gonçalves (2021a) to study the equity term structure in the context of an ICAPM. Moreover, all of these variables have been explored in the prior literature as important predictors of dividend growth and equity returns.[11] To demonstrate the robustness of my term structure decomposition findings to the specific predictive variables used, Internet Appendix D provides results from specifications that (i) exclude predictive variables one at a time from s_t or (ii) set $s_t = dp_t$ so that only the dividend yield predicts future dividend growth and returns as in some papers in the literature (e.g., Cochrane (2011)).

The dividend measurement used for the dividend growth and dividend yield variables is based on the sum of annual dividends with no compounding to avoid introducing properties of returns into dividend growth (see Binsbergen and Koijen (2010) and Chen (2009)) and includes M&A paid in cash (as suggested by Allen and Michaely (2003)). Both aspects serve to make the dividend yield, which is an important state variable in my analysis, more stationary (consistent with Koijen and Nieuwerburgh (2011) and Sabbatucci (2015)).[12]

I estimate the VAR system by Ordinary Least Squares (OLS) equation by equation and add the Pope (1990) correction for the small sample bias in VAR systems with persistent predictors.[13]

[10]Dividend yield is the log of aggregate dividends over a normalized index price. The term spread is the difference between the 10-year and 1-year log Treasury yields. The credit spread is the difference between Moody's corporate BAA and AAA log yields. Following Campbell and Vuolteenaho (2004), the value spread is the difference between the log book-to-market ratios of the value and growth portfolios formed based on small stocks with an adjustment to account for within year movements in market equity.

[11]Chen and Zhao (2009) use lagged dividend growth as a dividend growth predictor. Several papers use the dividend yield as a predictor for both dividend growth and stock returns, with theoretical justification provided by the valuation identity of Campbell and Shiller (1989). The treasury yield (Fama (1981) and Fama and Schwert (1977)), term spread (Campbell (1987) and Fama and French (1989)), and credit spread (Keim and Stambaugh (1986)) are classical equity return predictors. Finally, Campbell and Vuolteenaho (2004), Campbell, Polk, and Vuolteenaho (2009), and Campbell et al. (2018) rely on the value spread as an important predictor of stock returns.

[12]Internet Appendix C further discusses these adjustments and Internet Appendix D demonstrates that the overall results are similar when I use a dividend measurement that does not account for M&A activity.

[13]Internet Appendix D provides results estimating the VAR system (i) without the small sample bias correction and (ii) using Projection Minimum Distance (Jordà and Kozicki (2011)), which is a generalization of OLS that targets predictability at multi-year horizons. The results obtained using these alternative

Flow variables (such as dividend growth and returns) are deflated using the CPI index. Moreover, I use monthly observations of annual flows to estimate the VAR system, which means that my flow observations overlap for eleven months.[14] Given the overlapping nature of some variables in the VAR system and the fact that quantities of interest are transformations of the VAR parameters, statistical inference relies on a bootstrap analysis, with a detailed description provided in Internet Appendix B.

The final dataset is a multivariate time series of monthly observations in which flow variables have annual measurement; this dataset extends from 12/1952 to 12/2019.[15] The starting date is selected to strike a balance between a long sample period and consistency in the behavior of the state variables used in the analysis. In particular, the analysis focuses on the post-war period and starts after the Fed-Treasury Accord of 1951 that restored independence to the Fed, affecting monetary policy. However, Internet Appendix D shows that my results are similar if I start the analysis in 1926 (i.e., long-term equity strips still explain more than 50% of the overall equity volatility).

3 Main Empirical Results

This section reports empirical results from applying the term structure decomposition developed in Section 1 to the aggregate U.S. equity market. Subsection 3.1 provides results for the VAR estimation, Subsection 3.2 decomposes stock return volatility into the effects of short- and long-term equity strips, Subsection 3.3 provides some comparative statics exercises on this equity term structure decomposition, Subsection 3.4 decomposes equity strip volatility into cash flow and discount rate news, and Subsection 3.5 explores time variation in the fraction of return volatility explained by long-term equity strips.

estimation methods are similar to the ones reported in the main text.

[14] The rationale for this approach is that annual dividend growth does not suffer from the seasonality issues that affect monthly and quarterly dividend growth.

[15] Data starting up to December 1951 is used in the OLS estimation given the annual VAR. However, $\widetilde{z}_t$ and Σ_z are based on shocks realized over the baseline sample period (12/1952 to 12/2019)

3.1 The VAR Estimation Results

Table 1 reports results for the estimation of the VAR system in Equation 14. Panel A provides the VAR coefficients while Panel B shows news terms at different horizons as linear functions of the s_t shocks. In Panel B, the discount rate news, N_{dr}, are split into interest rate news, N_{ir}, and equity premium news, N_{ep}. In both panels, coefficients are normalized to standard deviation units and bootstrap t-statistics are provided in parentheses.

The key message from Table 1 is that each state variable is significantly related to at least one news term, and thus all state variables are relevant for the equity term structure decomposition. One exception is that Δd shows no statistical relation to any of the news terms reported. However, Equation 15 shows that $\widetilde{\Delta d}$ needs to be in $\widetilde{z}$ to recover $\widetilde{r}^{(h)}$, and thus I keep $\widetilde{\Delta d}$ in $\widetilde{s}$.[16]

Table 2 shows the correlations among VAR shocks and news to cash flow growth (N_g) and discount rate components (N_{ir} and N_{ep}). The table also shows the correlations between these shocks/news and equity strip returns. The most important observation from this table is that equity returns are much more correlated with long-term equity strips than with short-term ones. For instance, $Cor(\widetilde{r}, \widetilde{r}^{(100)}) = 0.93$ while $Cor(\widetilde{r}, \widetilde{r}^{(1)}) = 0.20$. This result indicates that short- and long-term equity strips have fundamentally different sources of variation, which has important implications for empirical tests of asset pricing models (explored further in Subsection 4.3).

Overall, the VAR estimation results suggest that even though I consider a relatively parsimonious system with six variables, the implied equity strip returns display striking variation across horizons, which I explore in the rest of this paper.

3.2 Term Structure Decomposition of $Var(\widetilde{r})$: Baseline

Equation 8 demonstrates that aggregate equity variance can be decomposed into a weighted average of covariances between equity returns and equity strips re-

[16] In Internet Appendix D, I show that keeping $\widetilde{\Delta d}$ in $\widetilde{z}$ but not in $\widetilde{s}$ yields term structure decomposition results that are similar to the ones presented in the main text.

turns, $Var(\widetilde{r}) = \Sigma_{h=1}^{\infty} w^{(h)} \cdot Cov(\widetilde{r}^{(h)}, \widetilde{r})$. Figure 1 shows the term structures of $w^{(h)}$ and $\beta^{(h)} = Cov(\widetilde{r}^{(h)}, \widetilde{r})/Var(\widetilde{r})$ as well as the combination of the two, which reflects the equity variance decomposition.

Figure 1(a) displays the $w^{(h)}$ term structure, which starts close to 3.5% at the 1-year maturity and decreases to zero as we increase equity strip maturity. If the $\beta^{(h)}$ term structure was flat, the $w^{(h)}$ term structure in Figure 1(a) would fully summarize the $Var(\widetilde{r})$ term structure decomposition. However, Figure 1(b) shows that the $\beta^{(h)}$ term structure is strongly upward sloping, effectively counteracting the pattern observed in the $w^{(h)}$ term structure, and thus inducing a larger role for long-term equity strips in explaining equity volatility.

Figure 1(c) combines the weights with betas, $w^{(h)} \cdot \beta^{(h)}$, to decompose equity market variance, showing that the percentage of $Var(\widetilde{r})$ explained by equity strips of different maturities has a hump-shape, with equity strips of maturities between 10 and 20 years explaining slightly more than 2% of $Var(\widetilde{r})$ each while the 1-year and the very long-term equity strips are responsible for 0.5% of less of $Var(\widetilde{r})$ each.

Figure 1(d) displays the cumulative fraction of equity variance explained by equity strips and, strikingly, suggests that close to 60% of equity volatility comes from variation in the present value of very long-term dividends (i.e., dividends with maturities beyond 20 years). This result indicates that we cannot fully understand volatility in equity markets without a deep economic understanding of the drivers of time variation in the present value of very long-term dividends.

Moreover, the empirical result that around 60% of equity volatility is due to the present value of dividends with maturities beyond 20 years provides a benchmark to discipline our belief about the importance of long-term expectations in explaining equity volatility. Prior to this paper, such benchmark was not available, and thus one could easily believe that this number was, for example, 30% or 90%.

3.3 Term Structure Decomposition of $Var(\widetilde{r})$: Comparative Statics

To gain further insights into the $Var(\widetilde{r})$ term structure decomposition in this paper, I now explore comparative statics exercises focused on the two components that drive such a decomposition: equity duration and the term structure of equity strip βs.

(a) The Effect of Equity Duration

From Equation 8 (and from Figure 1), it is clear that the $w^{(h)}$ term structure is an important driver of the $Var(\widetilde{r})$ term structure decomposition. As such, a comparative statics exercise that changes the $w^{(h)}$ term structure can be informative. To design such an exercise, note that equity duration is given by (see Gonçalves (2021b) for more on equity duration)[17]

$$Dur \ = 1/(1-\rho) \ = \sum_{h=1}^{\infty} h \cdot w^{(h)} \tag{16}$$

Consequently, equity duration determines whether the weights in $w^{(h)}$ concentrate in low or high hs, which directly affects the $Var(\widetilde{r})$ decomposition. In particular, a longer equity duration leads to a higher importance of long-term equity strips in explaining the overall equity variance (and vice versa).

From the average dividend yield in my sample, $Dur = 30.8$ years. In Figure 2, I explore comparative statics exercises that change the $w^{(h)}$ term structure to be consistent with counterfactually low and high equity duration values given by $Dur = 20.8$ years and $Dur = 40.8$ years, respectively.[18]

[17] To derive Equation 16, note that the Campbell and Shiller (1989) equity valuation identity can be written as

$$log(P_t) \ = constant \ + \ log(D_t) \ + \ \sum_{j=1}^{\infty} \rho^{h-1} \cdot \mathbb{E}_t \left[\Delta d_{t+j} \right] \ - \ \sum_{j=1}^{\infty} \rho^{h-1} \cdot \mathbb{E}_t \left[r_{t+j} \right]$$

so that equity duration is given by

$$Dur \ = \ - \frac{\partial log(P_t)}{\partial \mathbb{E}_t[r]} \ = \ 1/(1-\rho) \ = \ (1-\rho) \cdot \sum_{h=1}^{\infty} h \cdot \rho^{h-1} \ = \ \sum_{h=1}^{\infty} h \cdot w^{(h)}$$

[18] Specifically, given a counterfactual duration value, Dur_c, we have $\rho_c = 1 - 1/Dur_c$, which implies the

Figures 2(a) and 2(b) show that if equity duration was counterfactually low (i.e., $Dur = 20.8$ years), then the $w^{(h)}$ term structure would concentrate more in short horizons and this would lead to a lower fraction of $Var(\widetilde{r})$ being explained by long-term equity strips (roughly 50% in contrast to the 60% in the baseline results). Similarly, Figures 2(c) and 2(d) show that if equity duration was counterfactually high (i.e., $Dur = 40.8$ years), then the $w^{(h)}$ term structure would concentrate more in long horizons and this would lead to a higher fraction of $Var(\widetilde{r})$ being explained by long-term equity strips (slightly more than 65% in contrast to the 60% in the baseline results).

Despite equity duration having a qualitatively intuitive effect on the $Var(\widetilde{r})$ term structure decomposition, its quantitative effect is somewhat modest. That is, even with large changes (of 10 years) on the overall market equity duration, the fraction of $Var(\widetilde{r})$ explained by long-term equity strips still remains relatively close to the baseline value of 60%.

(b) The Effect of the Equity β Slope

From Equation 8 (and from Figure 1), the other component that drives the $Var(\widetilde{r})$ term structure decomposition is the term structure of equity βs. My estimated term structure of equity betas in Figure 1(b) is based on the estimated VAR, except for the 1-year equity strip since $\beta^{(1)} = Cov(\widetilde{r}^{(1)}, \widetilde{r})/Var(\widetilde{r}) = Cov(\widetilde{\Delta d}, \widetilde{r})/Var(\widetilde{r})$ does not depend on the VAR dynamics.[19] As such, I design comparative statics exercises by constructing counterfactual term structures of equity βs that keep $\beta^{(1)}$ fixed and satisfy $\Sigma_{h=1}^{\infty} w^{(h)} \cdot \beta^{(h)} = 1$, but that have different slopes.

Specifically, I start by fitting a Nelson and Siegel (1987) term structure model to the VAR-implied βs while requiring the model to perfectly fit $\beta^{(1)}$ and $\Sigma_{h=1}^{\infty} w^{(h)} \cdot \beta^{(h)} = 1$. I then shift the slope parameter up and down to create counterfactual term structures of equity βs

counterfactual weights $w_c^{(h)} = \rho_c^{h-1} - \rho_c^h$. I use this result to obtain weights that are consistent with the counterfactual values of $Dur_c = 20.8$ years and $Dur_c = 40.8$ years.

[19] Strictly speaking, the shocks still depend on the VAR dynamics. However, using Δd and r instead of their shocks ($\widetilde{\Delta d}$ and $\widetilde{r}$) results in $\beta^{(1)} = Cov(\Delta d, r)/Var(r) = 0.09$, which is very similar to the $\beta^{(1)} = Cov(\widetilde{\Delta d}, \widetilde{r})/Var(\widetilde{r}) = 0.16$ value obtained from the VAR.

with low and high equity β slopes while adjusting the other parameters to continue to fit $\beta^{(1)}$ and $\Sigma_{h=1}^{\infty} w^{(h)} \cdot \beta^{(h)} = 1$. Further details are provided in Internet Appendix B.

Figure 3(a) shows the original term structure of equity βs (black dotted line) together with the counterfactual term structure of equity βs (red solid line) with a lower equity β slope, which effectively implies lower βs for short-term equity strips and higher βs for long-term equity strips. Figure 3(b) combines the estimated $w^{(h)}$ term structure from Figure 2(a) with this counterfactual term structure of equity βs, which leads to a higher fraction of $Var(\widetilde{r})$ being explained by long-term equity strips (roughly 70% in contrast to the 60% in the baseline results).

Similarly, Figure 3(c) shows the original term structure of equity βs (black dotted line) together with the counterfactual term structure of equity βs (red solid line) with a higher equity β slope, which effectively implies higher βs for short-term equity strips and lower βs for long-term equity strips. Figure 3(d) combines the baseline $w^{(h)}$ term structure with this counterfactual term structure of equity βs, which leads to a lower fraction of $Var(\widetilde{r})$ being explained by long-term equity strips (roughly 50% in contrast to the 60% in the baseline results).

Just as with the equity duration results, the term structure of equity βs has a qualitatively intuitive effect on the $Var(\widetilde{r})$ term structure decomposition, but its quantitative effect is somewhat modest. That is, even with substantial changes in the term structure of equity βs, the fraction of $Var(\widetilde{r})$ explained by long-term equity strips always remains relatively close to the baseline value of 60%.

3.4 Cash Flow vs Discount Rate Decomposition of $Var(\widetilde{r}^{(h)})$

Given that short- and long-term equity strips largely differ in terms of their contribution to equity volatility, I now study the drivers of equity strip volatility at different maturities. To this end, Table 3 and Figure 4 decompose each equity strip variance (as well as the overall equity variance) into cash flow and discount rate components.

Consistent with the literature, realized dividend growth shocks, cash flow news, and especially discount rate news are important drivers of equity volatility, with the three components explaining respectively 16.1%, 31.4%, and 53.6% of equity return variability.

Long-term equity strips tend to display similar decompositions to the one observed for equity returns in the sense that discount rate news explain a substantial portion of return volatility.[20] In contrast, short-term equity strips are mostly driven by the cash flow components. For instance, 100% of the volatility in the 1-year equity strip comes from dividend growth and even at 7-year maturity discount rate news explain less than 50% of the variability in equity strip returns. This result is consistent with the argument in Golez and Koudijs (2020) that shifts in the overall equity market duration affect the relative importance of discount rate and cash flow news in explaining equity volatility.

Overall, the results indicate that movements in the short and long ends of the equity term structure are driven by fundamentally different economic forces. While short-term equity strips are mostly affected by cash flows components, long-term equity strips are largely driven by discount rate news. Such a result clearly demonstrates the importance of treating short- and long-term equity strips separately when thinking about equity volatility.

3.5 Time Variation in the Importance of Long-term Equity Strips

The fact that short- and long-term equity strips are driven by different economic forces raises the possibility for extended periods in which long-term equity strips outperform short-term equity strips (and vice versa). If so, these periods likely translate into cycles in the relative importance of the present value of long-term dividends to equity prices. Figure 5 explores this aspect.

Figure 5(a) plots the time variation in the 5-year performance of long-term equity strips relative to short-term equity strips.[21] It is clear from the graph that there are extended periods of good (and bad) performance of long-term equity strips relative to short-term

[20] One difference, however, is that realized dividend growth, $\widetilde{\Delta d}$, explains effectively 0% of the variability in long-term equity strips whereas it still has a non-trivial explanatory power (of 16.1%) for equity returns.

[21] In mathematical terms, this equity strip performance spread is given by the 5-year moving average of $r_t^{(h>20)} - r_t^{(h\leq 20)} \equiv (1/\Sigma_{h>20} w^{(h)}) \cdot \Sigma_{h>20} w^{(h)} \cdot r_t^{(h)} - (1/\Sigma_{h\leq 20} w^{(h)}) \cdot \Sigma_{h\leq 20} w^{(h)} \cdot r_t^{(h)}$.

ones. For instance, long-term equity strips perform much better than short-term ones in the late 1990s, which indicates the tech boom was largely due to an increase in the value of long-term dividends, consistent with the idea that the tech boom was related to growth companies.

Figure 5(b) uses Equation 2 to display the time variation in the fraction of the equity market value due to dividends beyond 20 years, $\Sigma_{h=21}^{\infty} PV_t^{(h)}/P_t$. This fraction varies strongly over time, with periods in which it is as low as 35% (e.g., early 1980s) and periods in which it reaches well beyond 50% (late 1990s). Clearly, variation in the relative performance of long-term equity strips translates into a time-varying importance of the present value of long-term dividends to equity prices.

These results indicate that equity duration varies over time, which should induce variation in the $w^{(h)}$ term structure and lead to changes in the fraction of equity volatility explained by long-term equity strips. Since (as shown in Equation 16), the Campbell and Shiller (1989) log-linear approximation that I have relied on up to this point assumes a constant equity duration, I also develop a conditional version of this log-linear approximation in which the Taylor series expansion point varies over time as expected dividend yield varies. This approximation implies the conditional equity term structure decomposition given by

$$\widetilde{r}_{t+1} = \sum_{h=1}^{\infty} w_t^{(h)} \cdot \widetilde{r}_{t+1}^{(h)} \tag{17}$$

which can be combined with the VAR to provide a conditional variance decomposition of equity returns given by

$$Var_t(\widetilde{r}_{t+1}) = \sum_{h=1}^{\infty} w_t^{(h)} \cdot Cov(\widetilde{r}_{t+1}^{(h)}, \widetilde{r}_{t+1}) \tag{18}$$

with all relevant details about this conditional equity term structure decomposition (including the expression for $w_t^{(h)}$) provided in Internet Appendix A.3.

Figure 6 relies on Equation 18 to display the time series of the fraction of $Var_t(\widetilde{r}_{t+1})$ explained by equity strips with maturities beyond 20 years (i.e., long-term equity strips), which varies as the $w_t^{(h)}$ term structure moves over time. On average, the fraction is around 59%, which is very close to the result obtained using my unconditional term structure decomposition (see Figure 1(d)). Moreover, the fraction of equity volatility explained by long-term equity strips varies over time and follows a pattern very similar to what we observe for the fraction of equity value due to the present value of long-term dividends (in Figure 5(b)). In particular, the relative importance of long-term equity strips in explaining equity volatility has been 1 to 5 percentage points above its unconditional average over the last twenty years, which is a consequence of the longer equity duration in the market over the last couple of decades.

Overall, the results in this subsection indicate that the fraction of equity value and equity volatility attributable to the present value of long-term dividends varies strongly over time. This finding is important because it suggests that there are times in which long-term equity strips are more relevant to understanding equity price and volatility dynamics and times in which they are less relevant.

4 Three Further Empirical Applications

The previous sections use the concept of equity strips to provide a term structure decomposition of equity variance. In this section, I provide three other applications of the term structure decomposition developed, which are meant to illustrate that equity strips can be used to answer important research questions beyond applications related to decomposing equity volatility. Subsection 4.1 studies the behavior of equity strips during different economic environments, including the recent COVID crisis, Subsection 4.2 explores the economic insights that the decomposition can provide for the time-varying equity-bond return correlation, and Subsection 4.3 relies on equity strips to discipline an intertemporal CAPM when pricing the cross-section of risk premia.

4.1 Equity Strip Yields during the COVID Crisis

A growing literature is exploring the macroeconomic consequences of the recent COVID crisis (see Gormsen and Koijen (2020) for an example in the equity term structure literature). This subsection shows that, from the perspective of changes in expected growth and priced risk, the COVID crisis has been relatively moderate and short-lived, specially in comparison to the 2007-2009 financial crisis.

To start, define an h-year equity strip yield analogously to how Binsbergen et al. (2013) define equity yields:

$$\begin{aligned} ey_t^{(h)} &= \frac{1}{h} \cdot log\left(D_t/PV_t^{(h)}\right) \\ &= \underbrace{\mathbb{E}_t\left[\frac{1}{h} \cdot \sum_{j=1}^{h} r_{t+j}\right]}_{\overline{dr}_t^{(h)}} - \underbrace{\mathbb{E}_t\left[\frac{1}{h} \cdot \sum_{j=1}^{h} \Delta d_{t+j}\right]}_{\overline{g}_t^{(h)}} \end{aligned} \tag{19}$$

where $\overline{dr}_t^{(h)}$ and $\overline{g}_t^{(h)}$ reflect the average annual discount rate and growth rate over the subsequent h years.

Adverse macroeconomic events tend to induce increases in discount rates through priced risks and declines in expected growth, with both channels inducing increases in equity yields. As such, changes in the term structure of equity yields can be informative about changes in macroeconomic prospects over different horizons. To analyse this issue, I combine the VAR estimates from Section 3.1 (which rely on data until 2019) with daily values for the s_t vector to obtain out-of-sample estimates for equity strip yields daily over the year 2020.[22]

Figure 7(a) displays equity strip yields with maturities of $h = 1$, 5, 10, and 20 years during 2020. Before March, equity strip yields were fairly close to 0% with a small upward sloping term structure. The equity strip yield curve inverted in the beginning of March, with the 1-year equity strip yield reaching 20% in late March while the 10-year and 20-year equity strip yields barely changed. Such inversion suggests that short-term (and potentially mid-term)

[22]Internet Appendix C describes in details the construction of the daily values for s_t.

macroeconomic prospects deteriorated while long-term macroeconomic prospects remained stable. However, this situation was relatively short-lived. By the end of August, equity strip yields were all close to zero and by December the equity strip yield curve looked similar to how it was in January.

For comparison, Figure 7(b) shows the same term structure of equity strip yields from August 2006 to the April 2010, which covers the 2007-2009 financial crisis. As it is clear from the graph, the inversion of the equity strip yield curve was much more severe, with the 1-year equity strip yield reaching beyond 50%. Moreover, the effect on macroeconomic prospects was much longer lived, with the equity strip yield curve not stabilizing until 2010.

In summary, the findings in this subsection show that while the COVID crisis had a non-trivial effect on macroeconomic prospects (as captured by expected dividend growth and priced risk), its effect was moderate and short-lived in comparison to the effect of the 2007-2009 financial crisis.

4.2 Time Variation in the Equity-Bond Return Correlation

Several recent papers study the time variation in the equity-bond return correlation (e.g., Campbell, Sunderam, and Viceira (2017)). This correlation is important from an investment perspective since it affects the optimal allocation to equities and bonds, and it is also relevant when evaluating the ability of theoretical models to line up with important moments of the data. As Chapter 9.4 in Campbell (2018) makes it clear, the equity-bond return correlation has largely declined over the last decades, effectively becoming negative post 2000. This subsection shows that while this pattern holds for equity strips of different maturities, it is much weaker for short maturities so that the overall pattern described in the literature is mostly a phenomenon associated with long-term equity strips.

Figure 8(a) displays the time series of $Cor(r^{(1)}, R_b^{(10)})$ and $Cor(r^{(50)}, R_b^{(10)})$, which reflect the correlations of a short-term and a long-term equity strip with a portfolio of Treasury bonds with maturities up to ten years. The correlations are calculated on a rolling window of 20 years so that the rolling window in the last sample year starts in 2000, which is the year in which the equity-bond correlation started to change sign (see Fig. 9.3 in Campbell (2018)).

The main observation from Figure 8(a) is that both $Cor(r^{(1)}, R_b^{(10)})$ and $Cor(r^{(50)}, R_b^{(10)})$ declined and became negative by the end of the sample (which reflects the correlation over the period from 2000 to 2019). However, the $Cor(r^{(50)}, R_b^{(10)})$ time variation is extremely high, with a movement from around 0.50 correlation to roughly -0.50 correlation. In contrast, the movement in $Cor(r^{(1)}, R_b^{(10)})$ is more modest, with a movement from close to 0.10 correlation to around -0.30 correlation.

Figure 8(b) shows that the pattern observed in Figure 8(a) is quite general. Specifically, each line shows the term structure of the change in $Cor(r^{(h)}, R_b^{(k)})$ (for different k values) from the 2000-2019 period relative to the period that ends in 1999. As can be seen from the figure, it is generally the case that the change was negative, but it is smaller in magnitude for short-term equity strips.

In summary, the findings in this subsection show that while the decline in the equity-bond return correlation is quite general, it is quantitatively much weaker for short-term equity strips in comparison to long-term equity strips. Mechanisms attempting to explain the time variation in the equity-bond return correlation need to be consistent with this novel fact to be empirically credible.

4.3 Disciplining Intertemporal Asset Pricing Models

The seminal work of Merton (1973) introduced the Intertemporal CAPM (or simply ICAPM), which generalizes the CAPM to account for the hedging demand arising from variation in investment opportunities. In equilibrium, this hedging demand induces the marginal investor to price state variables that reflect investment opportunities. However, the framework in Merton (1973) does not impose strong economic restrictions on the relevant state variables and their risk prices, and thus empirical work has explored many state variables in a somewhat unconstrained fashion. This subsection shows how equity strips can be used to provide economically motivated restrictions on empirical implementations of the ICAPM.

To put things in perspective, a reduced form implementation of the ICAPM could be specified using a Stochastic Discount Factor (SDF) given by:

$$M_t \ = exp\left\{a \ - \ \lambda_m \cdot \widetilde{r}_t \ - \ \lambda^{'}\widetilde{s}_t\right\} \tag{20}$$

which, in the context of my empirical analysis, adds six degrees of freedom relative to the CAPM to match the cross-section of risk premia.

This level of flexibility in empirical tests is clearly not desirable from the perspective of guarding against the "fishing license" introduced by the ICAPM (see Fama (1991)). One approach to deal with this problem is to fully specify the economic environment (tied to a specific utility function) and how investment opportunities vary over time. This approach is taken by Campbell (1993), Campbell et al. (2018), Campbell and Vuolteenaho (2004), and Gonçalves (2021a) among others.

For instance, the ICAPM of Campbell (1993) implies:

$$M_t \ = exp\left\{a \ - \ \gamma \cdot \widetilde{r}_t \ - \ (\gamma - 1) \cdot N_{dr,t}\right\} \tag{21}$$

where γ represents relative risk aversion and $N_{dr,t} = (\mathbb{E}_t - \mathbb{E}_{t-1})[\sum_{h=1}^{\infty} \delta^h \cdot r_{t+h}] = b'_{dr}\widetilde{s}_t$ captures discount rate news, with b_{dr} fully characterized by δ and the VAR parameters.

In this version of the ICAPM, the only degree of freedom added relative to the CAPM is in the δ parameter, which is constant of approximation directly linked to the investor's time discount factor.[23] Some papers (e.g., Campbell and Vuolteenaho (2004)) further restrict δ to equal the coefficient in Campbell and Shiller (1989)'s log-linear stock return approximation, ρ, which effectively keeps γ as the only free parameter to estimate ICAPM risk prices.

While the discipline imposed by Campbell (1993) and related specifications of the ICAPM is very effective in guarding against the ICAPM fishing license, it requires tight economic assumptions on preferences, which leaves the empirical tests more subject to specification issues. A natural goal then is to design an empirical specification that represents a compromise between the generality of Equation 20 and the economic discipline of Equation 21.

[23]Specifically, δ represents a coefficient in the log-linear approximation to the budget constraint and it equals the time discount factor if the intertemporal elasticity of substitution equals one.

I argue that equity strips provide a direct solution to this problem. The argument relies on the assumption even though N_{dr} may not fully summarize investment opportunities as the model in Campbell (1993) suggests, the full term structure of equity strip expected returns does since it captures investment opportunities at different horizons. For instance, if two factors fully summarize $N_{dr}^{(h)}$ for all h, then an ICAPM in which only these two factors are priced should hold. Such a specification would decrease the degrees of freedom (beyond the market risk price) in Equation 20 from six to two and guard against the spurious pricing of state variables that are not connected to investment opportunities.

Figure 9 displays the R^2 from regressing each $N_{dr}^{(h)}$ onto $N_{dr}^{(1)}$ and N_{dr} (separately and jointly). The general message is that $N_{dr}^{(1)}$ misses long-term variation in expected returns and N_{dr} misses short-term variation in expected returns, but jointly $N_{dr}^{(1)}$ and N_{dr} properly describe the full the term structure of equity strip expected returns (i.e., the term structure of discount rate news).

Motivated by these results, I consider an ICAPM of the form:

$$\begin{aligned} M_t &= exp\left\{a \;-\; \lambda_m \cdot \widetilde{r}_t \;-\; \theta \cdot N_{dr,t}^{(1)} \;-\; \varphi \cdot N_{dr,t}\right\} \\ &= exp\left\{a \;-\; \lambda_m \cdot \widetilde{r}_t \;-\; \lambda' \widetilde{s}_t\right\} \end{aligned} \tag{22}$$

which implies the system of linear constraints

$$\lambda \;= \theta \cdot b_{dr}^{(1)} \;+\; \varphi \cdot b_{dr} \tag{23}$$

where $b_{dr}^{(1)} = \mathbf{1}_{r,s}\Phi_1$ and $b_{dr} = \rho \cdot \mathbf{1}_{r,s}\Phi_1(\mathrm{I}_\Phi - \rho \cdot \Phi_1)^{-1}$, with Φ_1 reflecting the VAR matrix in Equation 14, I_Φ capturing an identity matrix of the same dimension as Φ_1, and $\mathbf{1}_{r,s}$ representing a selector matrix such that $\mathbf{1}_{r,s}\mathrm{X}$ returns a column vector containing the X row associated with r and the X columns associated with s.

To estimate the SDF characterized by Equations 22 and 23, I note that the model implies

$$\mathbb{E}\left[M_t \cdot (R_{j,t} - R_{i,t})\right] \;= 0 \tag{24}$$

for any long-short return $R_{j,t} - R_{i,t}$. I then obtain a set of long-short testing portfolios that focus on the overall differences across equities, Treasury bonds, and corporate bonds as well as on cash flow duration differences within each of these asset classes.[24] Specifically, I consider (i) the equity market portfolio in excess of the risk-free rate, (ii) the equity duration deciles 2 to 10 from Gonçalves (2021b) in excess of his equity duration decile 1, (iii) the seven Fama Bond Portfolios in CRSP containing Treasury bonds with maturities up to 1, 2, 3, 4, 5, 10, and 30 years (all measured in excess of the risk-free rate), and (iv) eight Barclays long-term and mid-term corporate bond portfolios with Moody's ratings of Aaa, Aa, A, or Baa (with returns obtained from Datastream and measured in excess of the risk-free rate). As a risk-free rate proxy, I use returns on the one-month Treasury bill available on Kenneth French's data library.[25]

Finally, I estimate λ by applying Generalized Method of Moments (GMM) on the moment Equation 24 using the testing portfolios aforementioned with an identity weighting matrix while imposing the linear constraint system in Equation 23 as well as the constraint that the model must perfectly price the equity market excess return.[26] Standard errors and t-statistics are obtained from a bootstrap exercise analogous to the one used for inference on the VAR parameters. Further estimation details are provided in Internet Appendix B.

Table 4 reports the estimated λ_m and λ (normalized by the volatility of their respective risk factor shocks) with their t-statistics in parentheses as well as the implied $\gamma = \lambda_m$ value (with no normalization) with its standard error in brackets. With risk neutral pricing, a zero risk price is assigned to all risk factors, and thus the risk price portion of the first column is

[24]Since the ICAPM restriction I rely on emphasizes the importance of expected returns on short- and long-term equity strips, it is natural to consider a cross-section of testing portfolios that reflects assets with different discount rate sensitivities (i.e., that vary in cash flow duration).

[25]Since the VAR system yields monthly observations of annual s_t shocks, I use monthly observations of (overlapping) annual returns when constructing the testing assets. The sample period for the ICAPM estimation is from June 1974 to July 2018, with the initial date restricted by the availability of equity duration portfolios as well as corporate bond portfolios and the end date restricted by the availability of corporate bond portfolios.

[26]In a model in which all risk factors are tradable, estimating risk prices by perfectly pricing the tradable risk factors themselves yields an efficient and robust GMM (see Chabi-Yo, Gonçalves, and Loudis (2020)). Since the market portfolio is the only tradable risk factor in the ICAPM considered here, I accordingly require the model to perfectly price the market portfolio.

empty.

In the CAPM, only the market factor is priced. A one standard deviation movement in the market factor implies a (statistically significant) $\lambda_m = 0.42$ movement in the log SDF. Moreover, the implied CAPM risk aversion is relatively small ($\gamma = 2.62$).

In the ICAPM restricted by Equation 23, the market factor has the much larger (and also statistically significant) normalized risk price of $\lambda_m = 2.30$.[27] Moreover, several state variables are priced in the ICAPM. The dividend yield is the most strongly priced state variable with $\lambda_{dp} = 2.16$ ($t_{stat} = 2.04$) followed by the term spread with $\lambda_{TS} = 0.83$ ($t_{stat} = 2.31$), the value spread with $\lambda_{VS} = -0.72$ ($t_{stat} = -2.16$), and the Treasury yield with $\lambda_{ty} = -0.50$ ($t_{stat} = -2.13$). The other two state variables, CS and Δd, have risk prices that are statistically insignificant. These results suggest that an increase in dp or TS (ceteris paribus) is associated with an improvement in investment opportunities while an increase in VS or ty (ceteris paribus) is associated with a deterioration of investment opportunities. The ICAPM implied risk aversion point estimate is $\gamma = 14.5$, which is relatively high, although the large standard error of 5.7 implies we cannot reject that risk aversion is at more moderate levels satisfying $\gamma < 10$.

For comparison, I also consider a version of the ICAPM in which λ_m is fixed to match the value in the restricted ICAPM, but λ is unrestricted (i.e., I ignore the restriction in Equation 23). This model adds six degrees of freedom relative to the CAPM in contrast to the restricted ICAPM, which only adds two degrees of freedom relative to the CAPM. While the risk price signs in the unrestricted ICAPM are consistent with the risk price signs in the restricted ICAPM, their magnitudes substantially change and the large model flexibility adds considerable statistical instability so that none of the state variables has a statistically significant risk price. These results illustrate the importance of restricting risk prices using the economically motivated restrictions in Equation 23.

[27] One may (wrongly) perceive the normalized $\lambda_m = 2.30$ as unrealistically high. However, it is important to point out that, in the context of the ICAPM, λ_m reflects the marginal effect of market risk when investment opportunities are held fixed. In equilibrium, market returns are highly correlated with investment opportunities. For example, $Cor(\widetilde{r}_m, \widetilde{dp}) = -0.65$ and $\lambda_{dp} = 2.62$ so that movements in r_m are counteracted by opposing movements in dp in terms of its effect on the SDF.

Table 4 also reports pricing errors for duration spread portfolios that buy the longest maturity portfolio and sell the shortest maturity portfolio in each asset class.[28] With risk neutral pricing, pricing errors equal risk premia so that the first column provides the risk premium on each spread portfolio. As the second column shows, the CAPM substantially reduces the pricing errors on the corporate bond duration spread portfolios, but has little (or even an adverse) effect on the pricing errors associated with the duration spread portfolios constructed from equities or Treasury bonds. In contrast, the restricted ICAPM largely reduces pricing errors with almost all of them becoming statistically insignificant. Two duration spread portfolios of corporate bonds have statistically significant ICAPM pricing errors, but the signs are the opposite of their risk premia signs, suggesting that, if anything, the model overcorrects these risk premia. The pricing error results from the unrestricted ICAPM are largely similar to the ones from the restricted ICAPM, suggesting that the extra flexibility in the unrestricted ICAPM simply leads to overfitting of risk premia.

Figure 10(a) provides a scatter plot with risk premia in the x-axis and the (restricted) ICAPM pricing errors in the y-axis for all testing assets used in the estimation of risk prices. An ineffective model would result in all points scattered around the 45 degree line (i.e., risk premia equal pricing errors). An effective model can still result in non-zero pricing errors due to sampling noise, but the pricing errors should scatter around the 0% horizon axis so that they are not predicted by the risk premia. This pattern is exactly what we observe with the restricted ICAPM. That is, pricing errors across and within asset classes scatter around the 0% horizontal line, indicating the ICAPM reasonably captures the risk premia of the testing assets studied.

To gain further insights into the ICAPM's mechanism, Figures 10(b) to 10(d) consider a linear approximation to the SDF around one so that

$$\mathbb{E}[R_i - R_j] \; = \gamma \cdot Cov(R_i - R_j, \widetilde{r}) \; + \; \theta \cdot Cov(R_i - R_j, N_{dr}^{(1)}) \; + \; \varphi \cdot Cov(R_i - R_j, N_{dr}) \quad (25)$$

which yields a decomposition of the risk premia on the different testing portfolios into market

[28]Pricing errors are obtained from $\alpha_j = \mathbb{E}[R_{j,t} - R_{i,t}] - Cov(M_t/\mathbb{E}[M_t]\,, R_{j,t} - R_{i,t})$, with the M_t estimated from each model.

risk, $\gamma \cdot Cov(R_i - R_j, \widetilde{r})$, short-term discount rate risk, $\theta \cdot Cov(R_i - R_j, N_{dr}^{(1)})$, and long-term discount rate risk, $\varphi \cdot Cov(R_i - R_j, N_{dr})$.

Figure 10(b) shows that market risk leads to risk premia that increase in equity duration. In contrast, exposures to short-term discount rate risk and specially long-term discount rate risk induce risk premia that decline in equity duration. Overall, the discount rate risk components dominate the market risk component, and thus the ICAPM-implied risk premia decrease with equity duration, matching the data.

Figure 10(c) demonstrates that market risk by itself generates an upward sloping bond term structure, but its slope is not as strong as in the data (and it would be quantitatively small with the CAPM $\lambda_m = 0.42$ value replacing the ICAPM $\lambda_m = 2.30$ value). Exposure to short-term discount rate risk further contributes to the upward sloping bond term structure and exposure to long-term discount rate risk also does but with an extra concavity effect coming from a declining pattern beyond the five-year maturity bond portfolio.

Figure 10(d) shows that market risk is the major driver of the upward sloping corporate bond term structure. Discount rate risk mostly serves to adjust the overall level of risk premia, with long-term discount rate risk also having a small (and negative) effect on the corporate bond term structure slope.

In summary, the findings in this subsection demonstrate that equity strips can be used to impose economically motivated restrictions on the risk prices of state variables capturing investment opportunities in empirical applications of the ICAPM. Such restrictions are crucial to statistically identify risk prices on relevant state variables and guard against the fishing license that is inherently present in unrestricted implementations of the ICAPM.

5 Conclusion

In this paper, I develop a novel term structure decomposition for stock returns and explore some of its implications for the aggregate US equity market. I find that roughly 60% of equity volatility comes from variation in the present value of dividends (i.e., equity strips) with maturities beyond 20 years and that this fraction varies strongly over time. Moreover, I show that variation in short- and long-term equity strips are driven by different economic forces (cash flows vs discount rates). I also provide three other empirical applications of my equity term structure decomposition that demonstrate that equity strips allow researchers to explore a rich set of questions beyond providing a way to decompose equity return volatility.

The equity term structure decomposition developed in this paper provides a new quantitative tool to an extensive literature that explores different economic questions by decomposing equity returns into cash flow and discount rate news. Therefore, analogously to Campbell and Shiller (1989) and Campbell (1991), this paper opens the door to many new questions that connect equity return dynamics to asset pricing. For instance, do firms and countries differ in their equity response to news about short- and long-term dividends? Are structural models that match equity volatility consistent with the source of such volatility? These are examples of important questions to be addressed by future research.

References

Allen, F. and R. Michaely (2003). "Payout Policy". In: *Handbook of the Economics of Finance*. Ed. by G. M. Constantinides, M. Harris, and R. M. Stulz. Vol. 1. A. Elsevier Science. Chap. 7, pp. 337–429.

Andrews, S. and A. Gonçalves (2020). "The Bond, Equity, and Real Estate Term Structures". Working Paper.

Binsbergen, J. H. v. (2021). "Duration-Based Stock Valuation: Reassessing Stock Market Performance and Volatility". Working Paper.

Binsbergen, J. H. v., M. W. Brandt, and R. S. J. Koijen (2012). "On the Timing and Pricing of Dividends". In: *American Economic Review* 102.4, pp. 1596–1618.

Binsbergen, J. H. v., W. Hueskes, R. S. J. Koijen, and E. Vrugt (2013). "Equity Yields". In: *Journal of Financial Economics* 110, pp. 503–519.

Binsbergen, J. H. v. and R. S. J. Koijen (2010). "Predictive Regressions: A Present-Value Approach". In: *Journal of Finance* 65.4, pp. 1439–1471.

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

Campbell, J. Y. (1987). "Stock Returns and the Term Structure". In: *Journal of Financial Economics* 18, pp. 373–399.

Campbell, J. Y. (1991). "A Variance Decomposition for Stock Returns". In: *The Economic Journal* 101.405, pp. 157–179.

Campbell, J. Y. (1993). "Intertemporal Asset Pricing without Comsumption Data". In: *American Economic Review* 83.3, pp. 487–512.

Campbell, J. Y. (1996). "Understanding Risk and Return". In: *Journal of Political Economy* 104.2, pp. 298–345.

Campbell, J. Y. (2018). "Financial Decisions and Markets: A Course in Asset Pricing". In: Princeton University Press. Chap. 9.4, The Term Structure of Risky Assets, pp. 290–295.

Campbell, J. Y. and J. Ammer (1993). "What Moves the Stock and Bond Markets? A Variance Decomposition for Long-Term Asset Returns". In: *Journal of Finance* 48.1, pp. 3–37.

Campbell, J. Y., S. Giglio, C. Polk, and R. Turley (2018). "An Intertemporal CAPM with Stochastic Volatility". In: *Journal of Financial Economics* 128.2, pp. 207–233.

Campbell, J. Y. and J. Mei (1993). "Where do Betas Come From? Asset Price Dynamics and the Sources of Systematic Risk". In: *Review of Financial Studies* 6.3, pp. 567–592.

Campbell, J. Y., C. Polk, and T. Vuolteenaho (2009). "Growth or Glamour? Fundamentals and Systematic Risk in Stock Returns". In: *Review of Financial Studies* 23.1, pp. 305–344.

Campbell, J. Y. and R. J. Shiller (1989). "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors". In: *Review of Financial Studies* 1.3, pp. 195–228.

Campbell, J. Y., A. Sunderam, and L. M. Viceira (2017). "Inflation Bets or Deflation Hedges? The Changing Risks of Nominal Bonds". In: *Critical Finance Review* 6.2, pp. 263–301.

Campbell, J. Y. and T. Vuolteenaho (2004). "Bad Beta, Good Beta". In: *American Economic Review* 94.5, pp. 1249–1275.

Cederburg, S. (2019). "Pricing intertemporal risk when investment opportunities are unobservable". In: *Journal of Financial and Quantitative Analysis* 54.4, pp. 1759–1789.

Cejnek, G. and O. Randl (2016). "Risk and return of short-duration equity investments". In: *Journal of Empirical Finance* 36, pp. 181–198.

Cejnek, G. and O. Randl (2020). "Dividend Risk Premia". In: *Journal of Financial and Quantitative Analysis* 55.4, pp. 1199–1242.

Chabi-Yo, F., A. S. Gonçalves, and J. Loudis (2020). "An Intertemporal Risk Factor Model". Working Paper.

Chen, L. (2009). "On the reversal of return and dividend growth predictability: A tale of two periods". In: *Journal of Financial Economics* 92, pp. 128–151.

Chen, L., Z. Da, and R. Priestley (2012). "Dividend Smoothing and Predictability". In: *Management Science* 58.10, pp. 1834–1853.

Chen, L., Z. Da, and X. Zhao (2013). "What Drives Stock Price Movements?" In: *Review of Financial Studies* 26.4, pp. 841–876.

Chen, L. and X. Zhao (2009). "Return Decomposition". In: *Review of Financial Studies* 22.12, pp. 5213–5249.

Chen, S. and T. Li (2018). "A Unified Duration-based explanation of the value, profitability, and investment anomalies". Working Paper.

Chen, Z. (2021). "Inferring Stock Durations Around FOMC Surprises: Estimates and Implications". In: *Journal of Financial and Quantitative Analysis*. Forthcoming.

Cochrane, J. H. (1992). "Explaining the Variance of Price-Dividend Ratios". In: *Review of Financial Studies* 5.2, pp. 243–280.

Cochrane, J. H. (2011). "Presidential Address: Discount Rates". In: *Journal of Finance* 66.4, pp. 1047–1108.

Da, Z. (2009). "Cash Flow, Consumption Risk, and the Cross-Section of Stock Returns". In: *Journal of Finance* 64.2, pp. 923–956.

Dechow, P. M., R. G. Sloan, and M. T. Soliman (2004). “Implied Equity Duration: A New Measure of Equity Risk”. In: *Review of Accounting Studies* 9, pp. 197–228.

Fama, E. F. (1981). “Stock Returns, Real Activity, Inflation, and Money”. In: *American Economic Review* 71.4, pp. 545–565.

Fama, E. F. (1991). “Efficient Capital Markets: II”. In: *Journal of Finance* 46.5, pp. 1575–1617.

Fama, E. F. and K. R. French (1989). “Business conditions and expected returns on stocks and bonds”. In: *Journal of Financial Economics* 25.1, pp. 23–49.

Fama, E. F. and G. W. Schwert (1977). “Asset Returns and Inflation”. In: *Journal of Financial Economics* 5, pp. 115–146.

Giglio, S., B. Kelly, and S. Kozak (2020). “Equity Term Structures without Dividend Strips Data”. Working Paper.

Golez, B. and P. Koudijs (2018). “Four Centuries of Return Predictability”. In: *Journal of Financial Economics* 127.2, pp. 248–263.

Golez, B. and P. Koudijs (2020). “Equity duration and predictability”. Working Paper.

Gonçalves, A. S. (2021a). “Reinvestment Risk and the Equity Term Structure”. In: *Journal of Finance* Forthcoming.

Gonçalves, A. S. (2021b). “The Short Duration Premium”. In: *Journal of Financial Economics* Forthcoming.

Gormsen, N. J. and E. Lazarus (2021). “Duration-Driven Returns”. Working Paper.

Gormsen, N. J. (2021). “Time Variation of the Equity Term Structure”. In: *Journal of Finance* Forthcoming.

Gormsen, N. J. and R. S. J. Koijen (2020). “Coronavirus: Impact on Stock Prices and Growth Expectations”. In: *Review of Asset Pricing Studies* 10, pp. 574–597.

Jordà, Ò. and S. Kozicki (2011). “Estimation and Inference by the Method of Projection Minimum Distance: An Application to the New Keynesian Hybrid Phillips Curve”. In: *International Economic Review* 52.2, pp. 461–487.

Keim, D. B. and R. F. Stambaugh (1986). "Predicting Returns in the Stock and Bond Markets". In: *Journal of Financial Economics* 17, pp. 357–390.

Koijen, R. S. J. and S. V. Nieuwerburgh (2011). "Predictability of Returns and Cash Flows". In: *Annual Review of Financial Economics* 3, pp. 467–491.

Larrain, B. and M. Yogo (2008). "Does firm value move too much to be justified by subsequent changes in cash flow?" In: *Journal of Financial Economics* 87.1, pp. 200–226.

Merton, R. C. (1973). "An Intertemporal Capital Asset Pricing Model". In: *Econometrica* 41.5, pp. 867–887.

Nelson, C. R. and A. F. Siegel (1987). "Parsimonious Modelling of Yield Curves". In: *Journal of Business* 60.4, pp. 473–489.

Pope, A. L. (1990). "Biases of Estimators in Multivariate Non-Gaussian Autoregressions". In: *Journal of Time Series Analysis* 11.3, pp. 249–258.

Sabbatucci, R. (2015). "Are Dividends and Stock Returns Predictable? New Evidence Using M&A Cash Flows". Working Paper.

Vuolteenaho, T. (2002). "What Drives Firm-Level Stock Returns?" In: *Journal of Finance* 57.1, pp. 233–264.

Weber, M. (2018). "Cash Flow Duration and the Term Structure of Equity Returns". In: *Journal of Financial Economics* 128.3, pp. 486–503.

Weber, R. (2021). "Institutional Investors, Households and the Time-Variation in Expected Stock Returns". Working Paper.

Table 1

VAR Estimation Results (1952-2019)

The table reports the (1952-2019) estimates from the VAR system in Equation 14. The s_t state vector includes the dividend yield (dp), dividend growth (Δd), 1-year Treasury yield (ty), Treasury yield term spread (TS), yield spread between BAA and AAA bonds (CS), and book-to-market spread between small value and small growth firms (VS). Panel A provides the VAR coefficients while Panel B shows news terms at different horizons as linear functions of the s_t shocks. In Panel B, the discount rate news, N_{dr}, are split into interest rate news, N_{ir}, and equity premium news, N_{ep}. In both panels, coefficients are normalized to standard deviation units and bootstrap t-statistics are provided in parentheses. Section 2 provides further details on s_t and the VAR estimation.

PANEL A - VAR Coefficients in Standard Deviation Units (with t_{stat})

$s_t =$	dp_t	Δd_t	ty_t	TS_t	CS_t	VS_t
dp_{t+1}	0.84	0.02	0.12	-0.02	-0.14	0.00
	(8.94)	(0.39)	(0.91)	(-0.28)	(-2.46)	(-0.01)
Δd_{t+1}	-0.17	0.05	0.06	0.11	-0.21	-0.14
	(-1.32)	(1.00)	(0.33)	(1.26)	(-3.03)	(-2.54)
ty_{t+1}	0.00	-0.02	0.94	0.04	-0.05	-0.05
	(0.03)	(-0.37)	(5.46)	(0.52)	(-0.79)	(-1.08)
TS_{t+1}	-0.03	-0.10	-0.11	0.45	0.27	0.03
	(-0.25)	(-1.66)	(-0.61)	(4.30)	(3.29)	(0.47)
CS_{t+1}	-0.01	0.16	0.28	-0.04	0.50	0.06
	(-0.09)	(1.51)	(1.12)	(-0.21)	(3.61)	(0.64)
VS_{t+1}	-0.06	0.01	-0.26	-0.22	0.10	0.48
	(-0.34)	(0.09)	(-1.12)	(-1.37)	(0.73)	(4.12)

PANEL B - News Coefficients in Standard Deviation Units (with t_{stat})

$s_t =$	dp_t	Δd_t	ty_t	TS_t	CS_t	VS_t
$N_{ir,t}^{(1)}$	0.16	0.32	0.97	0.54	-0.41	0.60
	(0.28)	(0.79)	(1.43)	(0.78)	(-0.76)	(1.24)
$N_{ep,t}^{(1)}$	0.42	-0.04	-0.42	0.32	0.30	-0.46
	(1.29)	(-0.10)	(-1.05)	(0.73)	(0.83)	(-1.48)
$N_{g,t}^{(1)}$	-0.39	0.18	0.12	0.33	-0.68	-0.48
	(-1.32)	(1.00)	(0.33)	(1.26)	(-3.03)	(-2.54)
$N_{ir,t}^{(5)}$	0.14	0.00	1.16	0.35	-0.24	0.21
	(0.25)	(-0.01)	(1.65)	(0.78)	(-0.73)	(0.83)
$N_{ep,t}^{(5)}$	0.65	0.03	-0.46	0.25	0.19	-0.27
	(2.15)	(0.23)	(-1.34)	(1.19)	(1.07)	(-1.85)
$N_{g,t}^{(5)}$	-0.63	-0.06	-0.02	0.47	-0.30	-0.43
	(-1.43)	(-0.43)	(-0.04)	(1.68)	(-1.37)	(-2.51)
$N_{ir,t}$	0.15	-0.06	1.11	0.26	-0.18	-0.05
	(0.45)	(-0.58)	(2.31)	(1.06)	(-1.06)	(-0.39)
$N_{ep,t}$	0.78	0.02	-0.56	0.16	0.03	-0.20
	(2.42)	(0.21)	(-1.45)	(0.79)	(0.21)	(-1.62)
$N_{g,t}$	-0.77	0.00	-0.25	0.32	-0.04	-0.27
	(-1.62)	(0.02)	(-0.44)	(1.10)	(-0.20)	(-1.66)

Table 2
Correlations Between Returns, News, and State Vector Shocks

The table reports (1952-2019) correlations among (shocks to) state variables, equity and equity strip returns, and news about interest rates (N_{ir}), equity premium (N_{ep}), and dividend growth (N_g). The state variables are the dividend yield (dp), dividend growth (Δd), 1-year Treasury yield (ty), Treasury yield term spread (TS), yield spread between BAA and AAA bonds (CS), and book-to-market spread between small value and small growth firms (VS). Section 2 provides all relevant empirical details.

	r	N_{ir}	N_{ep}	N_g	dp	Δd	ty	TS	CS	VS
r	1.00									
N_{ir}	-0.11	1.00								
N_{ep}	-0.54	-0.43	1.00							
N_g	0.45	-0.58	-0.21	1.00						
dp	-0.65	0.24	0.71	-0.82	1.00					
Δd	0.16	0.20	0.38	-0.61	0.65	1.00				
ty	-0.05	0.97	-0.52	-0.61	0.19	0.20	1.00			
TS	-0.07	-0.69	0.48	0.61	-0.15	-0.27	-0.80	1.00		
CS	-0.46	-0.37	0.41	-0.16	0.28	-0.10	-0.27	0.29	1.00	
VS	0.21	-0.22	-0.27	-0.06	-0.21	-0.06	-0.12	0.08	0.22	1.00
	$r^{(1)}$	$r^{(3)}$	$r^{(5)}$	$r^{(7)}$	$r^{(10)}$	$r^{(20)}$	$r^{(30)}$	$r^{(40)}$	$r^{(50)}$	$r^{(100)}$
r	0.20	0.62	0.81	0.91	0.98	0.98	0.97	0.95	0.95	0.93
N_{ir}	0.18	0.39	0.30	0.18	0.04	-0.20	-0.30	-0.35	-0.38	-0.41
N_{ep}	0.37	-0.20	-0.41	-0.50	-0.57	-0.56	-0.52	-0.49	-0.48	-0.45
N_g	-0.58	-0.26	0.02	0.20	0.36	0.56	0.61	0.63	0.64	0.65
dp	0.62	0.09	-0.22	-0.42	-0.58	-0.74	-0.76	-0.77	-0.77	-0.76
Δd	0.99	0.74	0.52	0.38	0.22	0.02	-0.02	-0.04	-0.05	-0.05
ty	0.17	0.37	0.30	0.20	0.08	-0.14	-0.23	-0.28	-0.31	-0.34
TS	-0.26	-0.45	-0.38	-0.29	-0.18	0.01	0.08	0.12	0.14	0.17
CS	-0.12	-0.65	-0.73	-0.68	-0.59	-0.40	-0.34	-0.31	-0.29	-0.27
VS	-0.06	-0.10	-0.06	0.02	0.12	0.24	0.28	0.30	0.31	0.31

Table 3

Equity Strip Volatility: Cash Flow vs Discount Rate Shocks

The table reports (1952-2019) volatilities in equity strip unexpected returns $\widetilde{r}^{(h)}$, discount rate news, $N_{dr}^{(h)}$, cash flow news, $N_{g}^{(h)}$, and realized dividend growth shocks, $\widetilde{\Delta d}$. The table also reports (in parentheses) the fraction of equity strip return volatility due to $N_{dr}^{(h)}$, $N_{g}^{(h)}$, and $\widetilde{\Delta d}$ based on Equations 9 and 10. Section 2 provides all relevant empirical details.

		$\widetilde{r}^{(h)}$	$N_{dr}^{(h-1)}$	$N_{g}^{(h-1)}$	$\widetilde{\Delta d}$
Equity	σ	15.0%	12.4%	10.4%	15.0%
	$(\%\sigma_r^2)$		(53.6%)	(31.4%)	(16.1%)
h=1	σ	15.0%	0.0%	0.0%	15.0%
	$(\%\sigma_r^2)$		(0.0%)	(0.0%)	(100.0%)
h=3	σ	15.3%	7.3%	6.3%	15.0%
	$(\%\sigma_r^2)$		(17.1%)	(9.1%)	(73.8%)
h=5	σ	14.6%	10.0%	8.1%	15.0%
	$(\%\sigma_r^2)$		(33.9%)	(11.6%)	(54.5%)
h=7	σ	14.3%	11.3%	8.8%	15.0%
	$(\%\sigma_r^2)$		(44.3%)	(15.7%)	(40.1%)
h=10	σ	14.7%	12.4%	9.6%	15.0%
	$(\%\sigma_r^2)$		(53.1%)	(24.0%)	(23.0%)
h=20	σ	16.8%	14.0%	11.5%	15.0%
	$(\%\sigma_r^2)$		(60.1%)	(38.3%)	(1.6%)
h=30	σ	17.8%	14.5%	12.4%	15.0%
	$(\%\sigma_r^2)$		(59.6%)	(43.0%)	(-2.6%)
h=40	σ	18.3%	14.6%	12.8%	15.0%
	$(\%\sigma_r^2)$		(58.7%)	(45.2%)	(-3.9%)
h=50	σ	18.5%	14.6%	13.1%	15.0%
	$(\%\sigma_r^2)$		(58.0%)	(46.4%)	(-4.4%)
h=100	σ	18.9%	14.7%	13.4%	15.0%
	$(\%\sigma_r^2)$		(57.0%)	(47.8%)	(-4.9%)

Table 4
ICAPM Risk Prices and Term Spread Pricing Errors

The upper portion of the table reports the estimated λ_m and λ (normalized by the volatility of their respective risk factor shocks) as well as the implied relative risk aversion (i.e., $\gamma = \lambda_m$ with no normalization). The lower portion of the table reports pricing errors for duration spread portfolios that buy the longest maturity portfolio and sell the shortest maturity portfolio in each asset class. With risk neutral pricing, a zero risk price is assigned to all risk factors and pricing errors equal risk premia. As such, in the first column, the risk price portion is empty and the pricing errors provide the risk premium on each spread portfolio. t-statistics are reported in parentheses and standard errors in brackets. Section 4.3 provides all relevant empirical details.

			ICAPM	
	Risk Neutral	**CAPM**	**Restricted**	**Unrestricted**
γ		2.62	14.50	14.50
		[1.2]	[5.7]	[5.7]
λ_m		0.42	2.30	2.30
		(3.18)	(2.62)	(2.62)
λ_{dp}			2.16	1.11
			(2.05)	(0.47)
$\lambda_{\Delta d}$			0.00	0.64
			(0.37)	(0.27)
λ_{ty}			-0.50	-0.31
			(-2.13)	(-0.19)
λ_{TS}			0.83	0.61
			(2.31)	(0.32)
λ_{CS}			0.14	1.36
			(0.45)	(1.12)
λ_{VS}			-0.72	-0.93
			(-2.16)	(-0.86)
$\alpha_{EDur}^{(L-S)}$	-8.2%	-11.3%	-2.0%	-2.7%
	(-3.70)	(-5.16)	(-0.83)	(-2.08)
$\alpha_{TBond}^{(L-S)}$	3.4%	3.1%	0.9%	-0.6%
	(2.72)	(2.05)	(0.69)	(-0.42)
$\alpha_{AAA}^{(L-S)}$	1.4%	0.4%	-1.2%	-1.3%
	(1.83)	(0.43)	(-1.89)	(-1.42)
$\alpha_{AA}^{(L-S)}$	1.6%	1.1%	0.1%	-0.5%
	(2.39)	(1.48)	(0.10)	(-0.76)
$\alpha_{A}^{(L-S)}$	1.3%	0.5%	-1.2%	-0.8%
	(2.14)	(0.87)	(-2.22)	(-1.98)
$\alpha_{BAA}^{(L-S)}$	1.5%	0.4%	-0.8%	-0.2%
	(2.54)	(0.80)	(-1.27)	(-0.41)

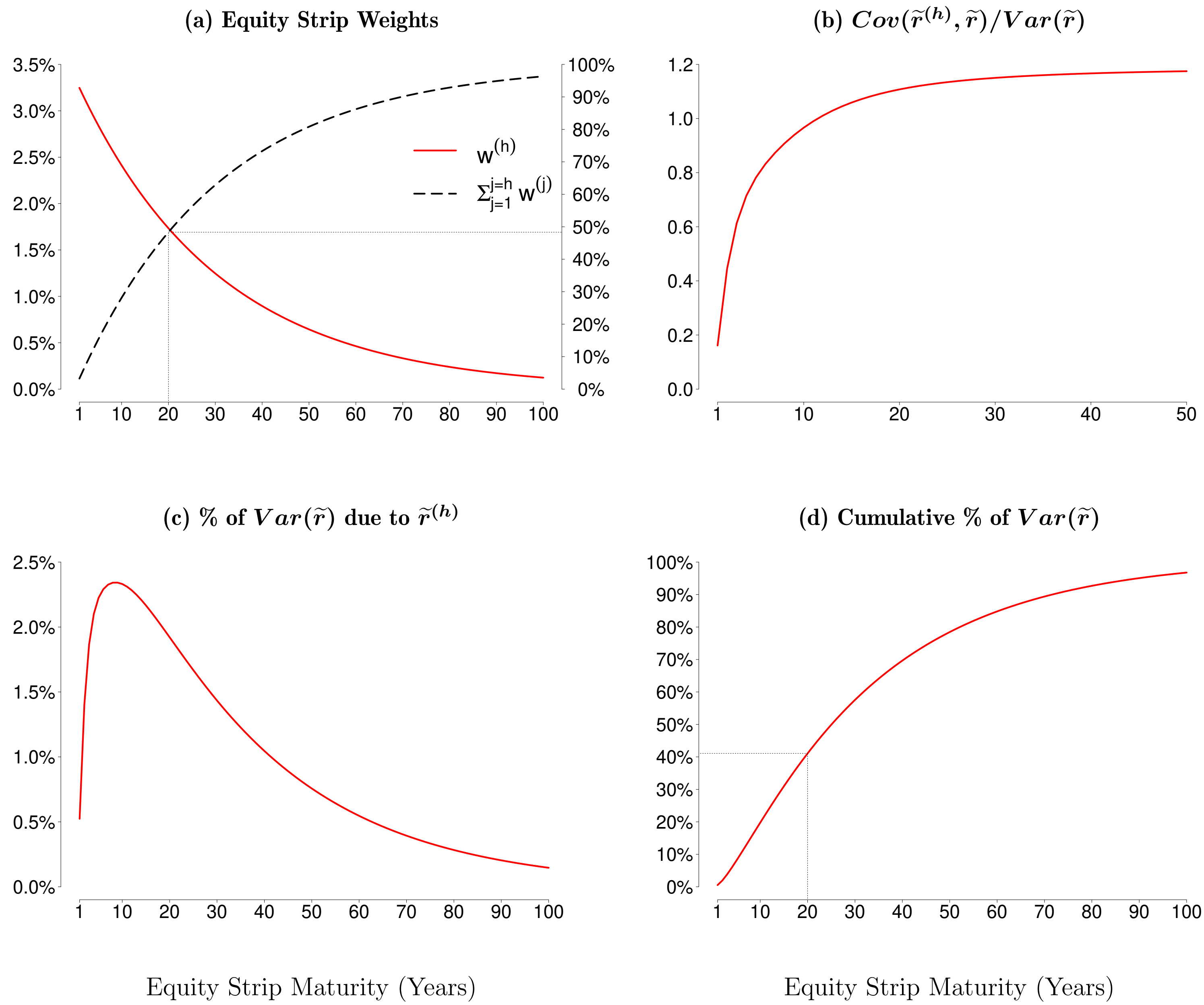


Figure 1
Decomposing $Var(\widetilde{r})$ into its Term Structure Components

The graphs collectively report the term structure decomposition of equity variance for the overall US equity market in Equation 8. Graph (a) provides the term structure of equity strip weights, $w^{(h)}$, and Graph (b) the term structure of equity betas, $\beta^{(h)} = Cov(\widetilde{r}^{(h)}, \widetilde{r})/Var(\widetilde{r})$. Graphs (c) and (d) report the fraction (raw and cumulative) of equity variance due to equity strips at different maturities, effectively combining the $w^{(h)}$ and $\beta^{(h)}$ term structures. Section 1 describes the equity term structure decomposition and Section 2 details its empirical implementation.

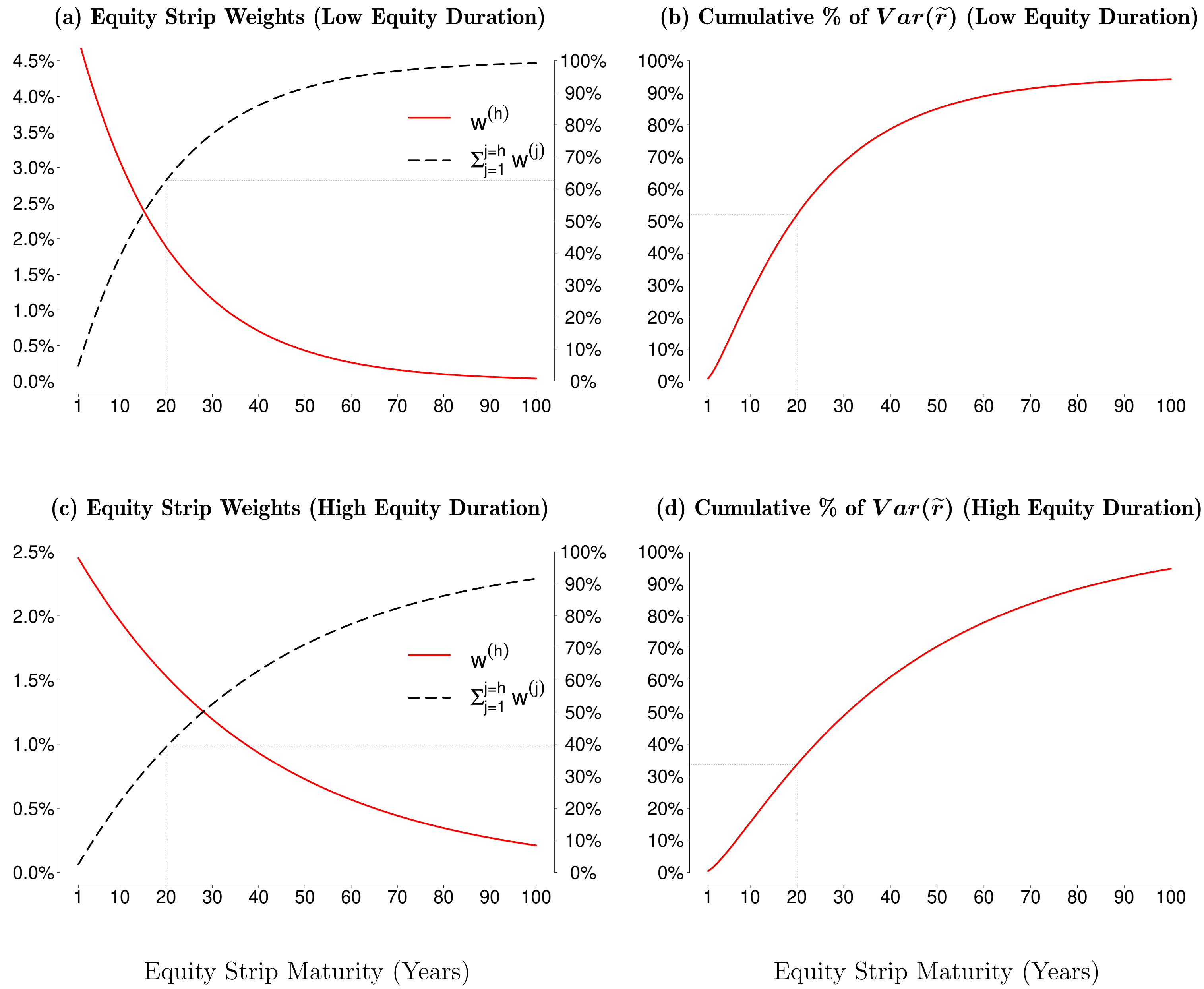


Figure 2
Decomposing $Var(\widetilde{r})$ into its Term Structure Components: the Effect of Equity Duration

The graphs collectively report the term structure decompositions of equity variance for the overall US equity market in Equation 8 after adjusting equity strip weights, $w^{(h)}$, to be consistent with counterfactually low or high equity duration values given by $Dur = 20.8$ years or $Dur = 40.8$ years (see Section 3.3 for details). Graphs (a) and (c) provide the term structures of counterfactual $w^{(h)}$ and Graphs (b) and (d) report the term structures of the cumulative fraction of equity variance due to equity strips at different maturities. Section 1 describes the equity term structure decomposition and Section 2 details its empirical implementation.

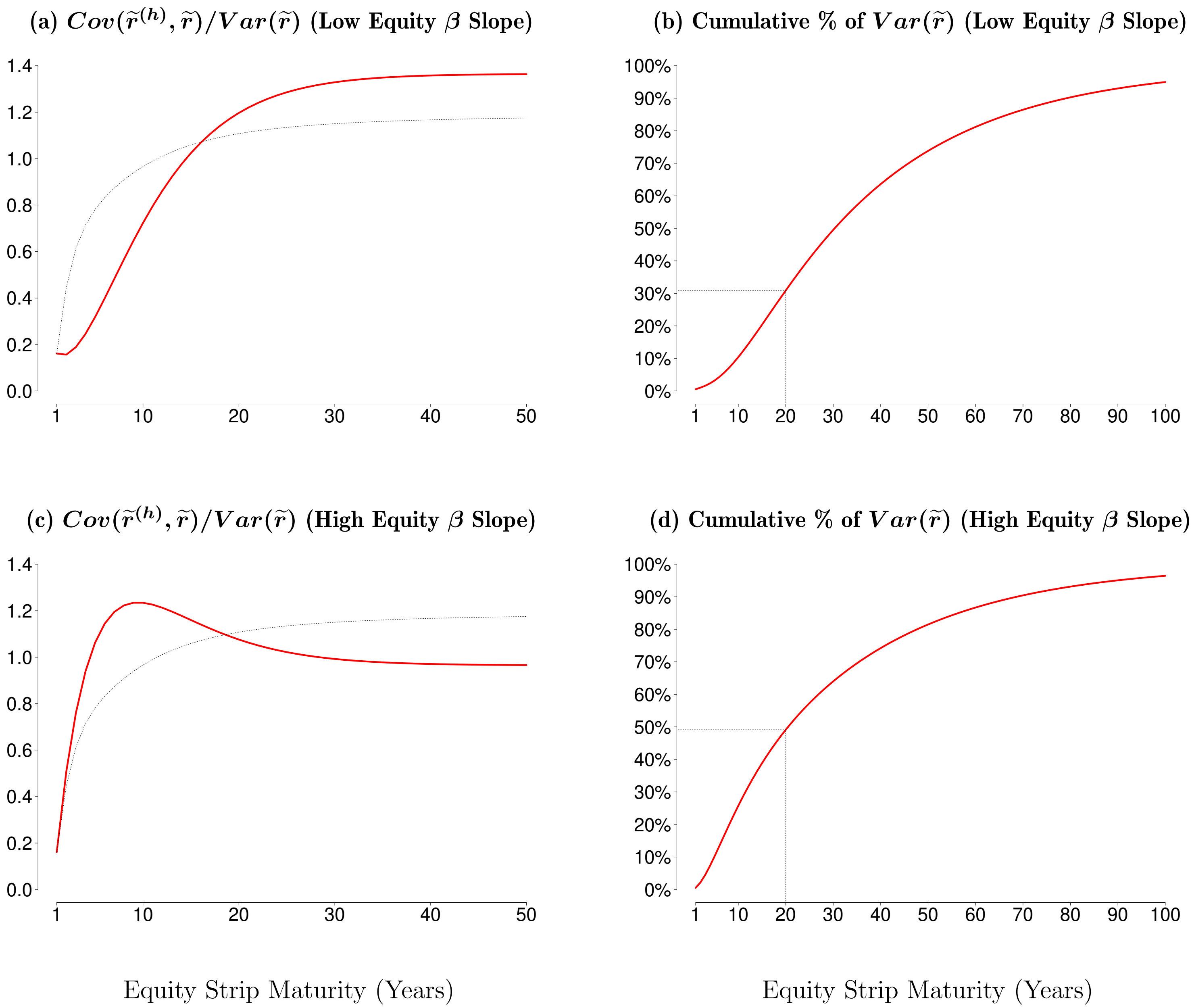


Figure 3
Decomposing $Var(\widetilde{r})$ into its Term Structure Components: the Effect of the Equity β Slope

The graphs collectively report the term structure decompositions of equity variance for the overall US equity market in Equation 8 after adjusting the term structure of equity betas, $\beta^{(h)} = Cov(\widetilde{r}^{(h)}, \widetilde{r})/Var(\widetilde{r})$, to be consistent with counterfactually low or high equity beta slopes (see Section 3.3 for details). Graphs (a) and (c) provide the term structures of counterfactual $\beta^{(h)}$ in the red solid line with the baseline $\beta^{(h)}$ term structure in the dotted black line and Graphs (b) and (d) report the term structures of the cumulative fraction of equity variance due to equity strips at different maturities. Section 1 describes the equity term structure decomposition and Section 2 details its empirical implementation.

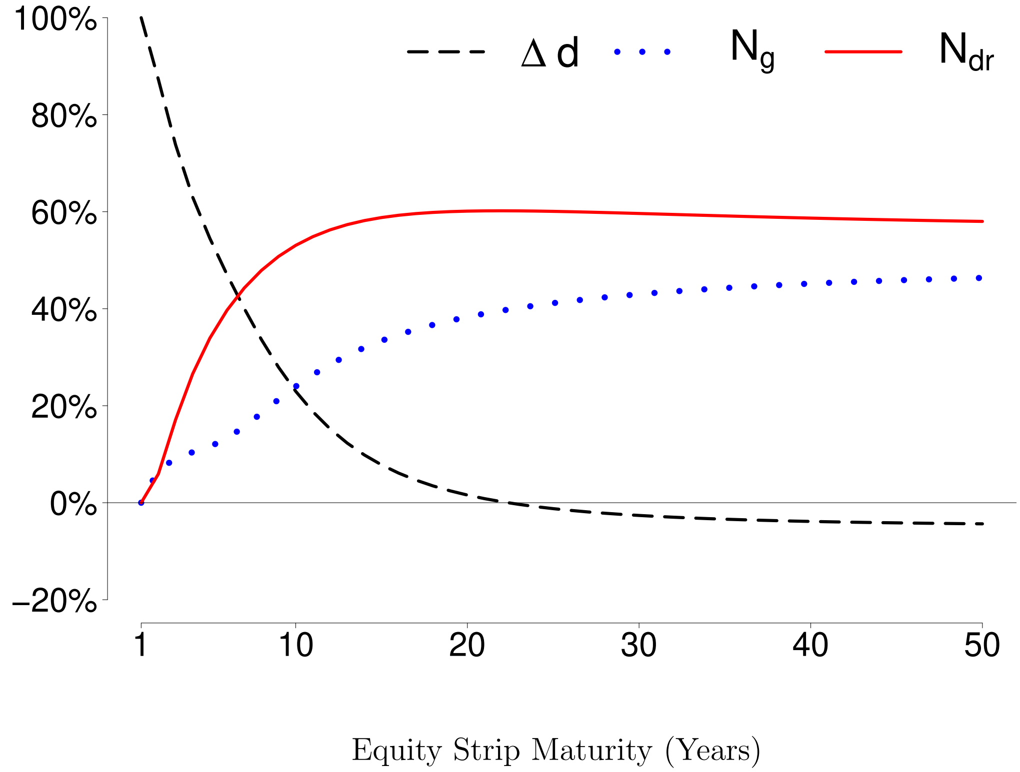
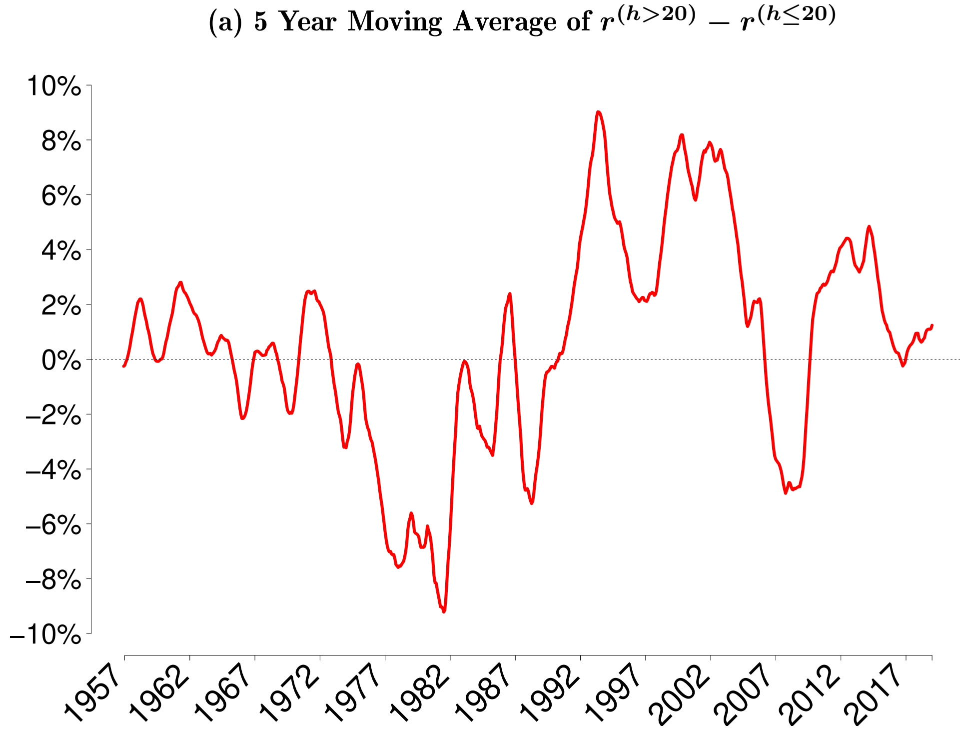


Figure 4
Decomposing Equity Strip Volatility into Cash Flow and Discount Rate Effects

The graph reports the fraction of the variance of equity strip unexpected returns due to discount rate news, N_{dr}, cash flow news, N_g, and realized dividend growth shocks, $\widetilde{\Delta d}$, based on Equation 10. Table 3 reports the underlying values in the graph, Section 1 describes the equity term structure decomposition, and Section 2 details its empirical implementation.



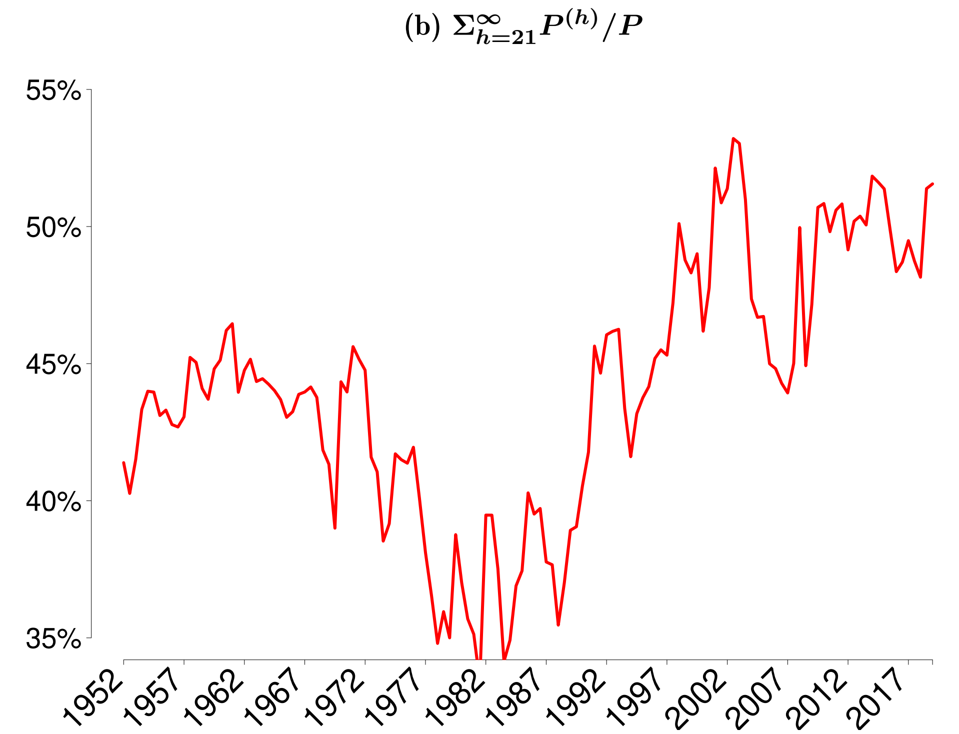


Figure 5

Relative Performance of Short-term versus Long-term Equity Strips

Graph (a) plots the 5-year return spread $r_t^{(h>20)} - r_t^{(h\leq20)} \equiv (\frac{1}{\Sigma_{h>20}w^{(h)}}) \cdot \Sigma_{h>20}w^{(h)} \cdot r_t^{(h)} - (\frac{1}{\Sigma_{h\leq20}w^{(h)}}) \cdot \Sigma_{h\leq20}w^{(h)} \cdot r_t^{(h)}$, which reflects the 5-year moving average performance of long-term equity strips relative to short-term equity strips. Graph (b) plots the fraction of the equity market value due to dividends beyond 20 years, $\Sigma_{h=21}^{\infty} PV_t^{(h)}/P_t$. Section 1 describes the equity term structure decomposition and Section 2 details its empirical implementation.

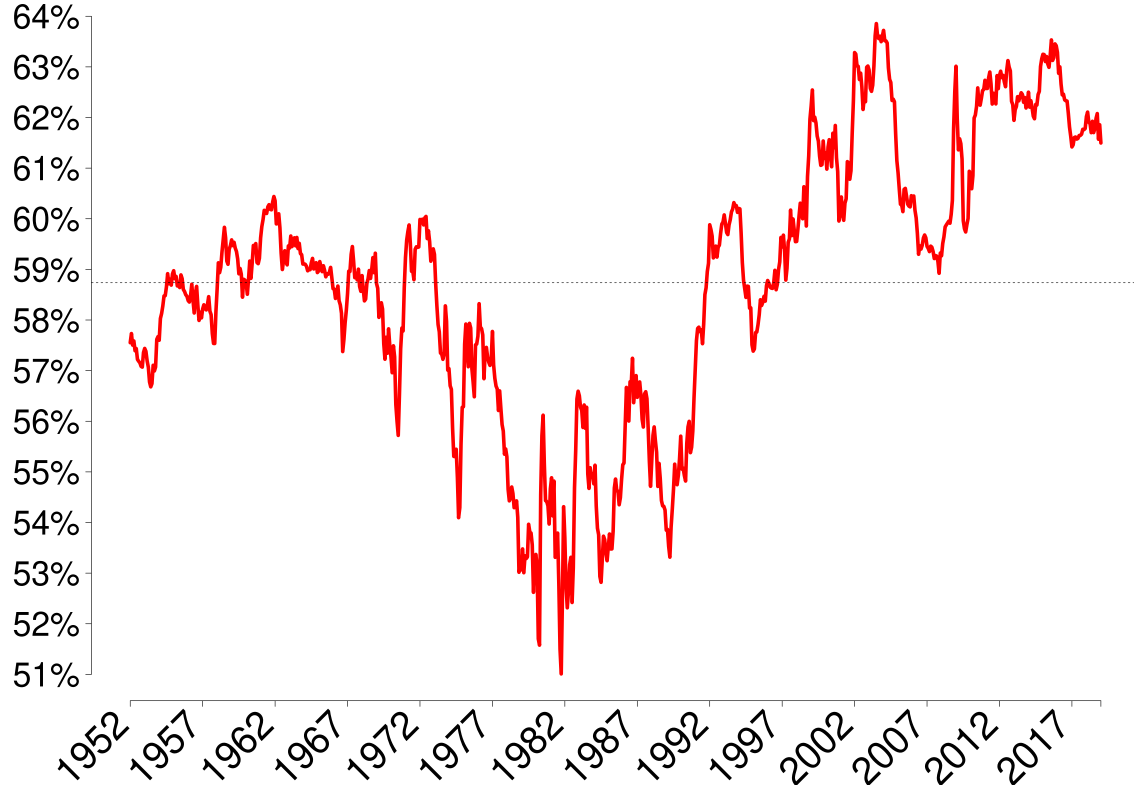


Figure 6
Fraction of $Var_t(\widetilde{r})$ Explained by Equity Strips with $h > 20$

The graph reports the fraction of equity variance explained by equity strips with maturities beyond twenty years. The conditional term structure decomposition of equity variance used for this graph is described in Subsection 3.5 (see Equation 18) and the VAR specification/estimation necessary to implement the decomposition empirically is described in Section 2.

(a) COVID Crisis

(b) 2007-2009 Financial Crisis

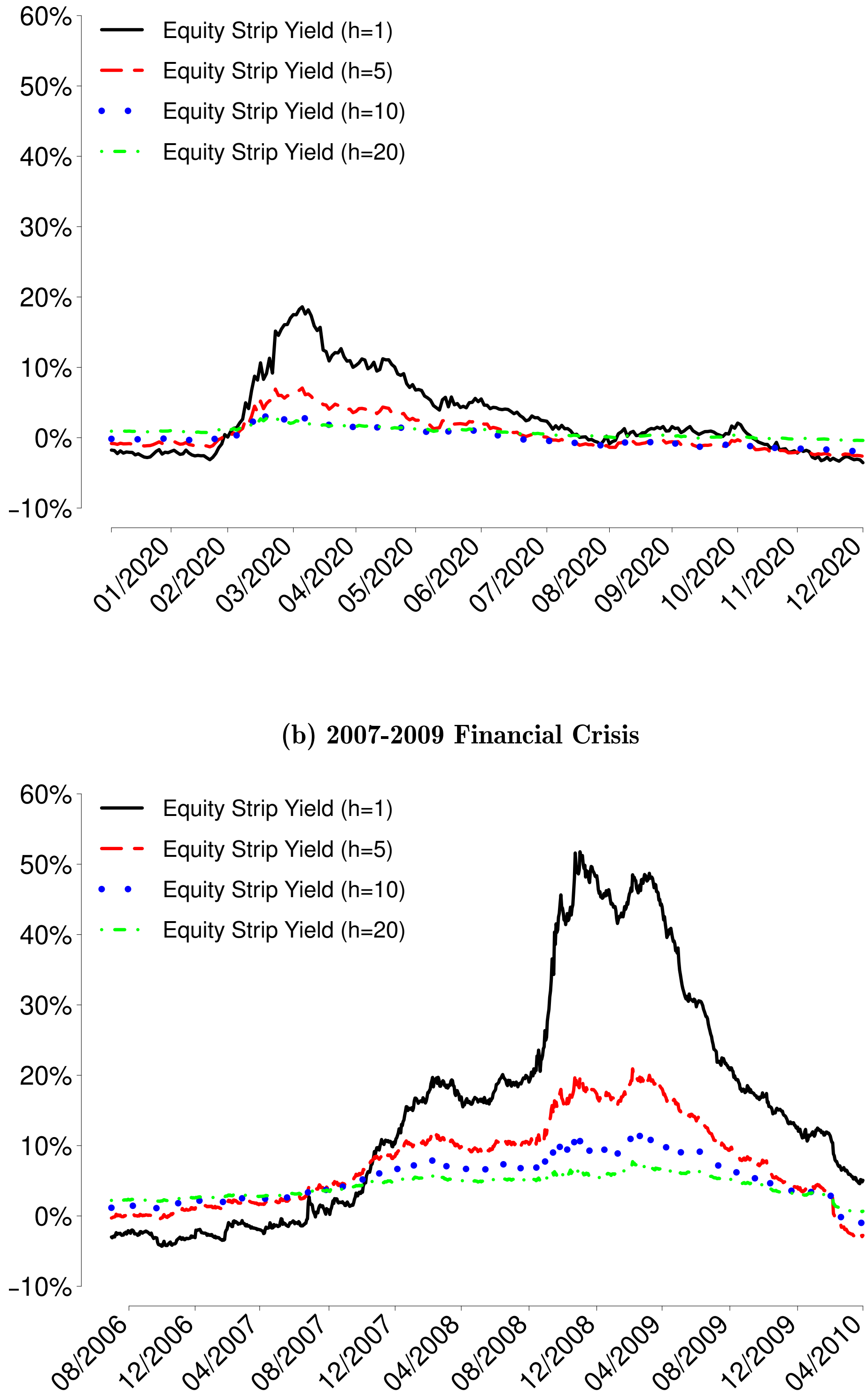
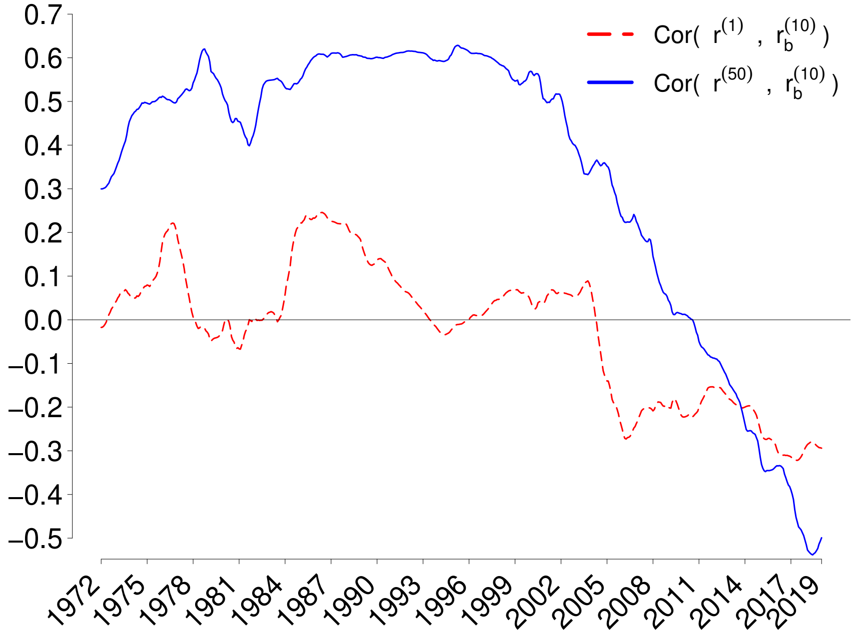


Figure 7
Equity Strip Yields During the COVID Crisis and the 2007-2009 Financial Crisis

The graphs display daily equity strip yields, $ey_t^{(h)} = (1/h) \cdot log(D_t/PV_t^{(h)})$, for the maturities of h=1, 5, 10, and 20 years. Graph (a) focuses on the year 2020, which covers the COVID crisis, and Graph (b) focuses on the period from August 2006 to the April 2010, which covers the 2007-2009 financial crisis. Section 1 defines and describes equity strips and Section 2 details their empirical estimation.

(a) Equity-Bond Return Correlation over Time



(b) Change in the Equity-Bond Return Correlation from $t < 2000$ to $t \geq 2000$

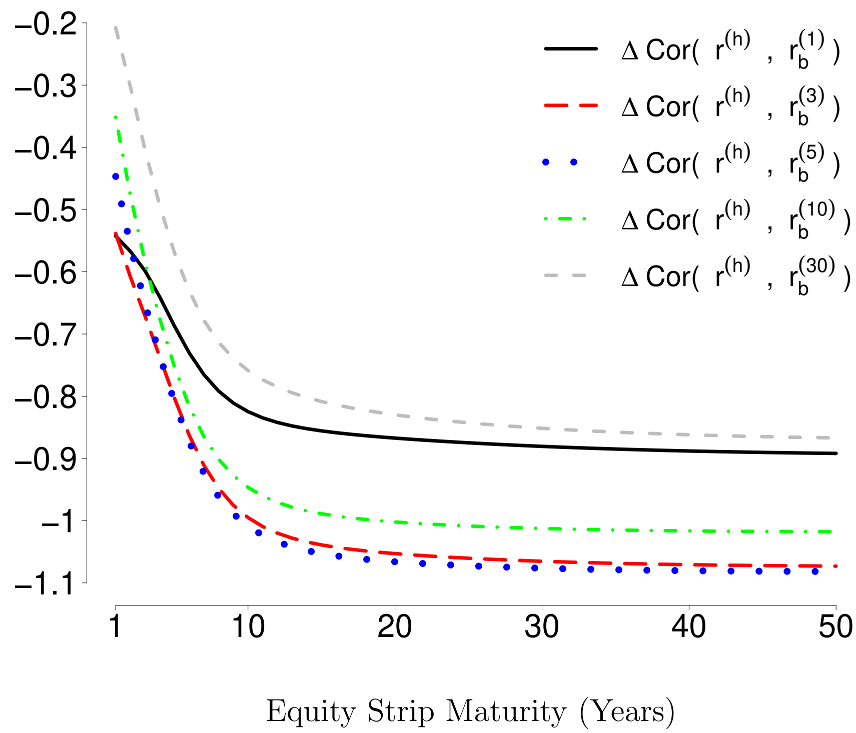
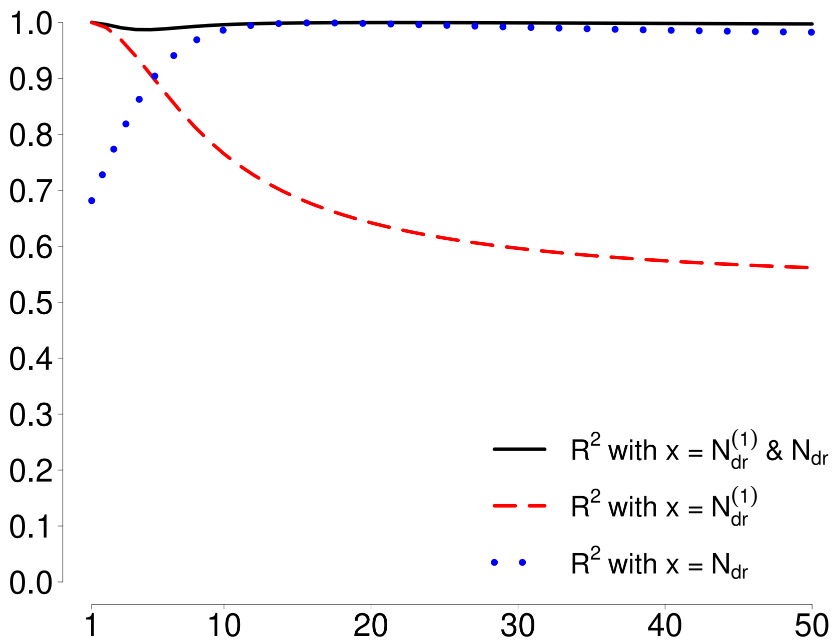


Figure 8
Time Variation in the Equity-Bond Return Correlation

Graph (a) plots the time series of the 20-year moving average correlation between the 1-year and 50-year equity strips ($r^{(1)}$ and $r^{(50)}$) and a portfolio of Treasury bonds with maturities up to ten years ($r_b^{(10)}$). Graph (b) plots the change in the correlation of equity strips (for all hs) with different bond portfolios from the period ending in 1999 to the period starting in 2000. The bond portfolios are the same ones used in the ICAPM estimation of Table 4 and all bond returns are measured in logs to match the equity strip log returns. Section 1 defines and describes equity strips and Section 2 details their empirical estimation.



Equity Strip Maturity (Years)

Figure 9

R^2 from Projecting $N_{dr}^{(h)}$ onto $N_{dr}^{(1)}$ and N_{dr}

The graph displays the R^2 of regressions of each $N_{dr}^{(h)}$ onto $N_{dr}^{(1)}$ and N_{dr} (separately and jointly), with $N_{dr}^{(h)}$ reflecting the h-year discount rate news (defined in Equation 3) and N_{dr} capturing discount rate news for the overall equity market (defined in Equation 4). Section 1 describes the equity term structure decomposition and Section 2 details its empirical implementation.

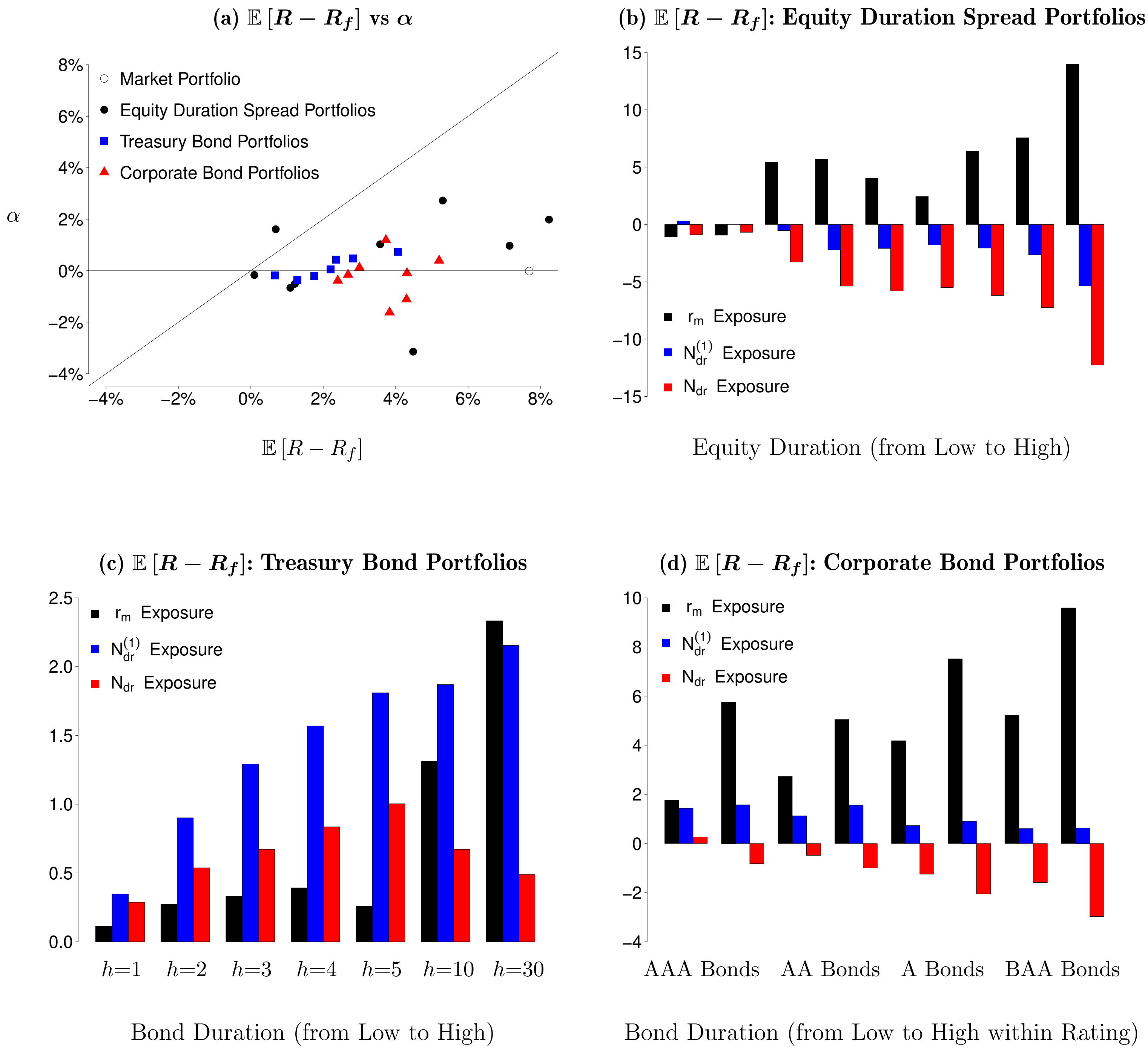


Figure 10
ICAPM Risk Premia and Pricing Errors

The graphs collectively report results related to ICAPM risk premia and pricing errors. Graph (a) provides a scatter plot with average risk premia in the x-axis and the (restricted) ICAPM pricing errors in the y-axis for all testing assets used in the estimation of risk prices in Table 4. Graphs (b) to (d) provide a decomposition of the ICAPM risk premia on the different testing portfolios into market risk, $\gamma \cdot Cov(R_i - R_j, \widetilde{r})$, short-term discount rate risk, $\theta \cdot Cov(R_i - R_j, N_{dr}^{(1)})$, and long-term discount rate risk, $\varphi \cdot Cov(R_i - R_j, N_{dr})$. Section 4.3 provides all relevant details on the ICAPM and its empirical implementation.

Internet Appendix

"What Moves Equity Markets? A Term Structure Decomposition for Stock Returns"

By Andrei S. Gonçalves

This Internet Appendix is organized as follows. Section A contains technical details/derivations required to support the results in the paper. Section B outlines the econometric methodology used for the main results provided. Section C details data sources and measurement for the analysis and Section D describes some further results that supplement the main findings in the paper.

A Technical Derivations

This section outlines the technical details/derivations that support the equity term structure decomposition developed in this paper. As in the main text, we define $\widetilde{x}_{t+1} = x_{t+1} - \mathbb{E}_t[x_{t+1}]$ for arbitrary x_t variable.

A.1 Equity Strip Prices and Returns (Equations 2 and 3)

Here, I derive equity strip present values and returns (Equations 2 and 3).

Equation 1 defines equity strip present values, PV. Under the assumption that r and Δd are conditionally homoskedastic and normally distributed, PVs are given by:

$$\begin{aligned} PV_t^{(h)} &= \mathbb{E}_t\left[D_{t+h} \cdot \left(\Pi_{j=1}^{h} R_{e,t+j}\right)^{-1}\right] \\ &= D_t \cdot \mathbb{E}_t\left[exp\left(\Sigma_{j=1}^{h} \Delta d_{t+j} - r_{e,t+j}\right)\right] \\ &= D_t \cdot exp\left(\mathbb{E}_t[\Sigma_{j=1}^{h} \Delta d_{t+j} - r_{e,t+j}] + 0.5 \cdot Var[\Sigma_{j=1}^{h} \Delta d_{t+j} - r_{e,t+j}]\right) \end{aligned} \tag{IA.1}$$

which yields Equation 3 (ignoring constants) after taking log on both sides of the equation.

This Equation can be derived as an approximation even if r and Δd are not conditionally homoskedastic and/or normally distributed. Specifically:

$$\begin{aligned} e^{log(PV_t^{(h)}/D_t)} &= \mathbb{E}_t\left[exp\left(\Sigma_{j=1}^{h} \Delta d_{t+j} - r_{t+j}\right)\right] \\ &\approx \mathbb{E}_t\left\{e^{log(PV_t^{(h)}/D_t)} + e^{log(PV_t^{(h)}/D_t)} \cdot [\Sigma_{j=1}^{h}(\Delta d_{t+j} - r_{e,t+j}) - log(PV_t^{(h)}/D_t)]\right\} \\ &\Downarrow \\ log(PV_t^{(h)}/D_t) &= \mathbb{E}_t[\Sigma_{j=1}^{h}(\Delta d_{t+j} - r_{t+j})] \end{aligned} \tag{IA.2}$$

where the approximation is based on a Taylor expansion of $\Sigma_{j=1}^{h} \Delta d_{t+j} - r_{t+j}$ around $log(P_t^{(h)}/D_t)$.

Now, letting $R_t^{(h)} = PV_t^{(h-1)}/PV_{t-1}^{(h)}$, we have:

$$\begin{aligned} r_t^{(h)} &= log(PV_t^{(h-1)}) \ - \ log(PV_{t-1}^{(h)}) \\ &= log(D_t) \ + \ \mathbb{E}_t[\Sigma_{j=1}^{h-1}(\Delta d_{t+j} - r_{t+j})] \ - \ log(PV_{t-1}^{(h)}) \\ &\Downarrow \\ \widetilde{r}_t^{(h)} &= \widetilde{\Delta d}_t \ + \ (\mathbb{E}_t - \mathbb{E}_{t-1})[\Sigma_{j=1}^{h-1}(\Delta d_{t+j} - r_{t+j})] \end{aligned} \tag{IA.3}$$

and Equation 3 follows by definition.

A.2 Equity Term Structure Decompositions (Equations 5, 6, and 8)

I now derive the term structure decompositions of equity returns and volatility in Equations 5, 6, and 8.

Comparing Equations 4 (derived in Campbell (1991)) and 3 (derived above), we have that the $\widetilde{\Delta d}_t$ component cancels out in Equation 5. Thus, showing that

$$\lim_{H\to\infty}\left(\sum_{h=1}^{H}\rho^h \cdot A_h\right) = \lim_{H\to\infty}\left(\sum_{h=1}^{H} w^{(h+1)} \cdot \sum_{j=1}^{h} A_j\right) \tag{IA.4}$$

holds for a generic variable A is sufficient to prove equation 5.

Start by reorganizing the right side of this equation as:

$$\sum_{h=1}^{H} w^{(h+1)} \cdot \sum_{j=1}^{h} A_j = A_1 \cdot \left(\sum_{h=1}^{H} w^{(h+1)}\right) + A_2 \cdot \left(\sum_{h=2}^{H} w^{(h+1)}\right) + ... + A_H \cdot w^{(h+1)} \tag{IA.5}$$

Now, use $w^{(h)} = \rho^{h-1} - \rho^h$ and the formula for the finite sum of a geometric progression to get:

$$\sum_{h=1}^{H} w^{(h+1)} \cdot \sum_{j=1}^{h} A_j = \sum_{h=1}^{H} \left(\frac{\rho^h - \rho^{h+1} + \rho^{H+1} - \rho^H}{1-\rho}\right) \cdot A_h \tag{IA.6}$$

As H goes to infinity, the $\rho^{H+1} - \rho^H$ term vanishes since $\rho < 1$ and, thus, all we need to show is that $\frac{\rho^h - \rho^{h+1}}{1-\rho} = \rho^h$, which can be demonstrated by multiplying both sides of this equation by $1 - \rho$. Hence, we have our result and Equation 5 holds.

To obtain Equation 6, simply note that $\mathbb{E}_t[r_{t+1}^{(h)}] = \mathbb{E}_t[pd_{t+1}^{(h-1)}] - pd_t^{(h)} + \mathbb{E}_t[\Delta d_{t+1}] = \mathbb{E}_t[r_{t+1}]$ so that we can add $\mathbb{E}_t[r_{t+1}]$ on both sides of Equation 5 to get Equation 6.

Equation 8 follows directly from taking covariance with respect to r_t on both sides of Equation 6.

A.3 Conditional Equity Term Structure Decompositions

In Subsection 3.5 in the main text, I also explore a conditional equity term structure decomposition. This subsection derives this conditional decomposition.

The return definition, $R_{t+1} = (P_{t+1} + D_{t+1})/P_t$, implies

$$dp_t = r_{t+1} \; - \; \Delta d_{t+1} \; - \; log(e^{-dp_{t+1}} + 1) \tag{IA.7}$$

Now, consider the sequence of first order Taylor approximations given by

$$log(e^{-dp_{t+h}} + 1) \; \approx k_{t-1}^{(h)} \; - \; \rho_{t-1}^{(h)} \cdot dp_{t+h} \tag{IA.8}$$

where

$$k_{t-1}^{(h)} = log\left(e^{-\mathbb{E}_{t-1}[dp_{t+h}]} + 1\right) \; + \; \rho_{t-1}^{(h)} \cdot \mathbb{E}_{t-1}[dp_{t+h}] \tag{IA.9}$$

and

$$\rho_{t-1}^{(h)} = e^{-\mathbb{E}_{t-1}[dp_{t+h}]} / \left(e^{-\mathbb{E}_{t-1}[dp_{t+h}]} + 1\right) \tag{IA.10}$$

Equation IA.8 reflects a first order Taylor expansion around $dp_{t+h} = \mathbb{E}_{t-1}[dp_{t+h}]$, which is a generalization of Campbell and Shiller (1989) approximation in which the conditional expansion point is more accurate than in Campbell and Shiller (1989). As h increases, the accuracy of the approximation decreases and it approaches the accuracy of Campbell and Shiller (1989) approximation as $h \to \infty$ since $\mathbb{E}_{t-1}[dp_{t+h}] \overset{h \to \infty}{\to} \mathbb{E}[dp]$ given that dp_t is stationary.

Substituting Equation IA.8 into Equation IA.7, we have:

$$dp_t \;\; = \; - \; k_{t-1}^{(1)} \; + \; r_{t+1} \; - \; \Delta d_{t+1} \; + \; \rho_{t-1}^{(1)} \cdot dp_{t+1} \tag{IA.11}$$

and more generally

$$dp_{t+h} \;=\; -\,k_{t-1}^{(1+h)} \;+\; r_{t+1+h} \;-\; \Delta d_{t+1+h} \;+\; \rho_{t-1}^{(1+h)} \cdot dp_{t+1+h} \tag{IA.12}$$

Then, substituting Equation IA.12 recursively in Equation IA.11, we have:

$$\begin{aligned}
dp_t \;&=\; -\,k_{t-1}^{(1)} + r_{t+1} - \Delta d_{t+1} + \rho_{t-1}^{(1)} \cdot dp_{t+1} \\
&=\; -\,k_{t-1}^{(1)} + r_{t+1} - \Delta d_{t+1} + \rho_{t-1}^{(1)} \cdot [-k_{t-1}^{(2)} + r_{t+2} - \Delta d_{t+2} + \rho_{t-1}^{(2)} \cdot dp_{t+2}] \\
&=\; -\,[k_{t-1}^{(1)} + \rho_{t-1}^{(1)} \cdot k_{t-1}^{(2)}] + [r_{t+1} + \rho_{t-1}^{(1)} \cdot r_{t+2}] - [\Delta d_{t+1} + \rho_{t-1}^{(1)} \cdot \Delta d_{t+2}] + \rho_{t-1}^{(1)} \cdot \rho_{t-1}^{(2)} \cdot dp_{t+2} \\
&\;\vdots \\
\rho_{t-1}^{(0)} \cdot dp_t \;&=\; \sum_{h=1}^{\infty} \left(\prod_{\tau=0}^{h-1} \rho_{t-1}^{(\tau)} \right) \cdot r_{t+h} - \sum_{h=1}^{\infty} \left(\prod_{\tau=0}^{h-1} \rho_{t-1}^{(\tau)} \right) \cdot \Delta d_{t+h} - \sum_{h=1}^{\infty} \left(\prod_{\tau=0}^{h-1} \rho_{t-1}^{(\tau)} \right) \cdot k_t^{(h)} \\
&\quad + \; \lim_{h\to\infty} \left(\prod_{\tau=0}^{h} \rho_{t-1}^{(\tau)} \right) \cdot dp_{t+h} \\
&\Downarrow
\end{aligned}$$

$$\rho_{t-1}^{(0)} \cdot dp_t \;=\; \sum_{h=1}^{\infty} \phi_{t-1}^{(h)} \cdot \mathbb{E}_t[r_{t+h}] \;-\; \sum_{h=1}^{\infty} \phi_{t-1}^{(h)} \cdot \mathbb{E}_t[\Delta d_{t+h}] \;-\; \sum_{h=1}^{\infty} \phi_{t-1}^{(h)} \cdot k_{t-1}^{(h)} \tag{IA.13}$$

where the last equality takes conditional expectation on both sides of the equation, defines $\phi_{t-1}^{(h)} = \prod_{\tau=0}^{h-1} \rho_{t-1}^{(\tau)}$, and imposes the transversality condition $\lim_{h\to\infty} \phi_{t-1}^{(h+1)} \cdot \mathbb{E}_t[dp_{t+h}] = 0$.

Then, subtracting from Equation IA.13 its $t-1$ conditional expectation, we have:

$$\rho_{t-1}^{(0)} \cdot \widetilde{dp}_t \;=\; \sum_{h=1}^{\infty} \phi_{t-1}^{(h)} \cdot \widetilde{\mathbb{E}}_t[r_{t+h}] \;-\; \sum_{h=1}^{\infty} \phi_{t-1}^{(h)} \cdot \widetilde{\mathbb{E}}_t[\Delta d_{t+h}] \tag{IA.14}$$

Finally, using the Taylor approximation in Equation IA.8 with $h = 0$, we have

$$dp_{t-1} \;=\; -\,k_{t-1}^{(0)} \;+\; r_t \;-\; \Delta d_t \;+\; \rho_{t-1}^{(0)} \cdot dp_t \tag{IA.15}$$

which implies a conditional version of the Campbell (1991) decomposition:

$$\begin{aligned}
\widetilde{r}_t \;&=\; \widetilde{\Delta d}_t \;-\; \rho_{t-1}^{(0)} \cdot \widetilde{dp}_t \\
&=\; \widetilde{\Delta d}_t \;+\; \sum_{h=1}^{\infty} \phi_{t-1}^{(h)} \cdot \widetilde{\mathbb{E}}_t[\Delta d_{t+h}] \;-\; \sum_{h=1}^{\infty} \phi_{t-1}^{(h)} \cdot \widetilde{\mathbb{E}}_t[r_{t+h}]
\end{aligned} \tag{IA.16}$$

To convert Equation IA.16 into a conditional equity term structure decomposition, I define the conditional weights $w_{t-1}^{(h)} \equiv \phi_{t-1}^{(h-1)} - \phi_{t-1}^{(h)}$ with $\phi_{t-1}^{(0)} \equiv 1$, which is analogous to the $w^{(h)} = \rho^{h-1} - \rho^{h}$ expression in my unconditional term structure decomposition. I then calculate $\sum_{h=1}^{\infty} w_{t-1}^{(h)}$ and $\sum_{h=1}^{\infty} w_{t-1}^{(h)} \cdot \widetilde{r}_t^{(h)}$ for every observation in my sample and verify that $\sum_{h=1}^{\infty} w_{t-1}^{(h)} = 1$ and $\sum_{h=1}^{\infty} w_{t-1}^{(h)} \cdot \widetilde{r}_t^{(h)} = \widetilde{r}_t$ with $\widetilde{r}_t$ calculated from Equation IA.16.

Given the paragraph above, we have the conditional equity term structure decomposition

$$\widetilde{r}_{t+1} = \sum_{h=1}^{\infty} w_t^{(h)} \cdot \widetilde{r}_{t+1}^{(h)} \tag{IA.17}$$

which can be used to obtain an unconditional variance decomposition given by

$$Var(\widetilde{r}_{t+1}) = \sum_{h=1}^{\infty} Cov(w_t^{(h)} \cdot \widetilde{r}_{t+1}^{(h)}, \widetilde{r}_{t+1}) \tag{IA.18}$$

as well as a conditional variance decomposition given by

$$Var_t(\widetilde{r}_{t+1}) = \sum_{h=1}^{\infty} w_t^{(h)} \cdot Cov_t(\widetilde{r}_{t+1}^{(h)}, \widetilde{r}_{t+1}) \tag{IA.19}$$

When applying Equation IA.19 to study the conditional variance decomposition in my empirical analysis, I use the fact that Σ_z is constant in the VAR, which implies conditional covariances are constant and allows me to simplify Equation IA.19 to

$$Var_t(\widetilde{r}_{t+1}) = \sum_{h=1}^{\infty} w_t^{(h)} \cdot Cov(\widetilde{r}_{t+1}^{(h)}, \widetilde{r}_{t+1}) \tag{IA.20}$$

Equation IA.20 incorporates the effect of time-varying equity duration into the conditional variance decomposition, but not the effect of time-varying covariances, which would require modeling a time varying Σ_z and is outside of the scope of this paper.

B Econometric Details

This section provides methodological details about the econometric analysis in the paper.

B.1 VAR Estimation and Inference

This subsection details the estimation of the VAR system in Equation 14 (reproduced below for convenience):

$$\begin{aligned} z_t &= \Phi_0 + \Phi_1 \cdot z_{t-1} + \widetilde{z}_t \\ &= \Phi_0 + \Phi_{s,1} \cdot s_{t-1} + \widetilde{z}_t \end{aligned} \tag{IA.21}$$

where $\widetilde{z}_t \overset{i.i.d}{\sim} N(0, \Sigma_z)$, $z_t = [r_{f,t} \quad xr_t \quad s_t]$, and $\Phi_{s,1}$ represents the non-zero portion of Φ_1 (i.e., the coefficients on the s_t variables).

(a) Baseline Estimation

For the main text (i.e., my baseline specification), I estimate the VAR system in two steps. In the first step, I estimate each VAR equation by Ordinary Least Squares (OLS). The data for the left side of the equation starts in Dec-1952 while the data for the right side of the equation starts in Dec-1951. VAR systems are potentially subject to a small sample bias due to persistent predictors (Stambaugh (1999)). To adjust the VAR for this bias, I use the method in Pope (1990) and treat as independent observations only the total number of year (as opposed to the number of months), which is a conservative adjustment for the fact that I use monthly observations with flow variables that overlap for 11 months given their annual measurement.

Note that the Campbell and Shiller (1989) approximation implies $\mathbb{E}_t[r_{t+1}] = k + \mathbb{E}_t[\Delta d_{t+1}] - \rho \cdot \mathbb{E}_t[dp_{t+1}] + dp_t$, which imposes a restriction on the VAR coefficients. Since dp_t is included in s_t, these restrictions are automatically imposed (as approximations) in my OLS estimation, with no adjustment needed.

(b) Alternative Estimations

For Internet Appendix Section D, I also estimate the VAR using two alternative approaches. The first is OLS without the bias adjustment in Pope (1990). The second is a method called Projection Minimum Distance (PMD), which was developed in Jordà and Kozicki (2011) as a robust (to mispecification) way to estimate VARMA models and builds on the local projections methodology in Jordà (2005), which is often used as a robust (to mispecification) way to estimate impulse response functions. The idea is to obtain long-run regressions (up to 10 years in my case) of each variable on s_t and then solve for the VAR $\Phi_{s,1}$ matrix that best reflect those projections. The Internet Appendix of Gonçalves (2021) provides details on how to estimate VAR models by PMD in closed-form.

(c) Inference

Statistical inference on the VAR parameters and quantities derived from them is done through bootstrap analysis. I use a residual bootstrap approach since fully non-parametric bootstrap is unlikely to properly capture the long-term dependence in the VAR variables (z_t). Specifically, I use the estimated residuals, $\widetilde{z}_t$, obtained from the VAR estimation as the relevant multivariate time-series to draw from. I fix the z_t for the period before Dec-1952 and (independently) draw 5-year blocks from the monthly observations of $\widetilde{z}_t$ (which are annual residuals) to create the full simulated z_t series. Once the simulated dataset is created, I estimate all relevant quantities in the simulated sample. I use 10,000 simulations. To calculate the t-statistic for a generic quantity θ, I use $t^{\theta}_{stat} = \widehat{\theta}/\widehat{\sigma}(\widehat{\theta})$ with $\widehat{\sigma}(\widehat{\theta})$ obtained from the standard deviation of $\widehat{\theta}$ across simulations (with both tails of the distribution trimmed at 1% to obtain a consistent estimator of $\sigma(\widehat{\theta})$ as suggested in Hansen (2019)).

B.2 $\beta^{(h)}$ Term Structure Comparative Statics

In Subsection 3.3, I design a comparative statics exercise by creating counterfactual term structures of equity βs that keep $\beta^{(1)}$ fixed and satisfy $\Sigma_{h=1}^{\infty} w^{(h)} \cdot \beta^{(h)} = 1$, but that have different slopes. I start by fitting a Nelson and Siegel (1987) term structure model to the

VAR-implied βs while requiring the model to perfectly fit $\beta^{(1)}$ and $\Sigma_{h=1}^{\infty} w^{(h)} \cdot \beta^{(h)} = 1$. I then shift the slope parameter up or down to create counterfactual term structures of equity βs with low or high equity β slope while adjusting the other parameters accordingly to continue to fit $\beta^{(1)}$ and $\Sigma_{h=1}^{\infty} w^{(h)} \cdot \beta^{(h)} = 1$. Below, I explain all the details of this process.

To fit a Nelson and Siegel (1987) smooth function to $\beta^{(h)}$, I estimate the parameters in

$$\begin{aligned} \beta^{(h)} &= \widehat{\beta}^{(h)} + \epsilon^{(h)} \\ &= b_0 \ + \ b_1 \cdot exp(-(h-1)/\tau) \ + \ b_2 \cdot ((h-1)/\tau) \cdot exp(-(h-1)/\tau) \ + \ \epsilon^{(h)} \\ &= b_0 \ + \ b_1 \cdot L_1(h,\tau) \ + \ b_2 \cdot L_2(h,\tau) \ + \ \epsilon^{(h)}, \end{aligned} \tag{IA.22}$$

where the last equality defines the L_1 and L_2 functions and the measurement error, $\epsilon^{(h)}$, is assumed to be independent of h.

The constraint to have $\widehat{\beta}^{(1)} = \beta^{(1)}$ implies

$$b_0 = \beta^{(1)} - b_1 \tag{IA.23}$$

and thus Equation IA.22 can be simplified to:

$$\begin{aligned} \beta^{(h)} - \beta^{(1)} &= b_1 \cdot [L_1(h,\tau) - 1] \ + \ b_2 \cdot L_2(h,\tau) \ + \ \epsilon^{(h)} \\ &= b_1 \cdot H_1(h,\tau) \ + \ b_2 \cdot L_2(h,\tau) \ + \ \epsilon^{(h)} \end{aligned} \tag{IA.24}$$

where $L_1^*(h,\tau) = L_1(h,\tau) - 1$.

Similarly, the constraint to satisfy $\Sigma_{h=1}^{\infty} w^{(h)} \cdot \widehat{\beta}^{(h)} = 1$ implies

$$b_2 = \frac{1 - \beta^{(1)} - b_1 \cdot H_1^{\infty}(\tau)}{L_2^{\infty}(\tau)} \tag{IA.25}$$

and thus Equation IA.24 can be further simplified to:

$$\beta^{(h)} - \beta^{(1)} = b_1 \cdot H_1(h,\tau) + \frac{1 - \beta^{(1)} - b_1 \cdot H_1^{\infty}(\tau)}{L_2^{\infty}(\tau)} \cdot L_2(h,\tau) + \epsilon^{(h)}$$

$$\Downarrow$$

$$\underbrace{\beta^{(h)} - \beta^{(1)} - \frac{L_2(h,\tau)}{L_2^{\infty}(\tau)} \cdot (1 - \beta^{(1)})}_{y_t^{(h)}} = b_1 \cdot \underbrace{\left[H_1(h,\tau) - \frac{L_2(h,\tau)}{L_2^{\infty}(\tau)} \cdot H_1^{\infty}(\tau)\right]}_{x_t^{(h)}} + \epsilon^{(h)} \qquad \text{(IA.26)}$$

where $H_1^{\infty}(\tau) \approx \Sigma_{h=1}^{1000} H_1(h,\tau)$ and $L_2^{\infty}(\tau) \approx \Sigma_{h=1}^{1000} L_2(h,\tau)$.

I estimate τ and b_1 by Non-linear Least Squares (NLS) with objective function $\Sigma_{h=1}^{h=100}[\epsilon^{(h)}]^2$ and then recover b_0 and b_2 from Equations IA.23 and IA.25. To avoid potential issues with multidimensional nonlinear optimization, I simplify the estimation of (τ, b_1) by developing an algorithm that relies on a set of OLS estimates and a unidimensional search to find the NLS estimate. Specifically, I note that, given τ, the objective function can be minimized by estimating b_1 using OLS. As such, I create a τ grid (with $\tau = 0.1, 0.2, ..., 100$) and estimate b_1 by OLS for each grid point, selecting at the end the (τ, b_1) pair at the grid point with the lowest objective function.

The process above provides a Nelson and Siegel (1987) smooth fit to the $\beta^{(h)}$ term structure. To obtain the counterfactually low and high slope $\beta^{(h)}$ term structures in Subsection 3.3, I change b_1 to $b1 - 0.2$ and $b_1 + 0.2$ and recover the corresponding b_0 and b_2 values from Equations IA.23 and IA.25.

B.3 ICAPM Estimation

The ICAPM Euler condition is given by

$$\mathbb{E}\left[exp\left\{a - \lambda_m \cdot \widetilde{r}_t - \theta \cdot N_{dr,t}^{(1)} - \varphi \cdot N_{dr,t}\right\} \cdot xR\right] = 0 \qquad \text{(IA.27)}$$

where xR reflects the vector of excess returns used to estimate the model.

I estimate λ_m, θ, and φ by applying Generalized Method of Moments (GMM) on the moment Equation IA.27 while imposing the constraint that the model must perfectly price the equity market excess return and the normalization that $\mathbb{E}[M \cdot e^{r_f}] = 1$. Specifically, I choose θ and φ to minimize $e'e$ in

$$\frac{1}{T} \cdot \sum_{t=1}^{T} exp\left\{a \; - \; \lambda_m \cdot \widetilde{r}_t \; - \; \theta \cdot N_{dr,t}^{(1)} \; - \; \varphi \cdot N_{dr,t}\right\} \cdot xR_{o,t} \; = e \tag{IA.28}$$

while solving for a and λ_m in each step of the optimization to satisfy

$$\frac{1}{T} \cdot \sum_{t=1}^{T} exp\left\{a \; - \; \lambda_m \cdot \widetilde{r}_t \; - \; \theta \cdot N_{dr,t}^{(1)} \; - \; \varphi \cdot N_{dr,t}\right\} \cdot xR_{e,t} \; = 0 \tag{IA.29}$$

and

$$\frac{1}{T} \cdot \sum_{t=1}^{T} exp\left\{a \; - \; \lambda_m \cdot \widetilde{r}_t \; - \; \theta \cdot N_{dr,t}^{(1)} \; - \; \varphi \cdot N_{dr,t}\right\} \cdot e^{r_{f,t}} \; = 1 \tag{IA.30}$$

where xR_e reflects the excess return on the equity market portfolio and xR_o reflects the excess returns on the other testing assets.

Once θ and φ are estimated, I recover the λ parameter vector from the system of linear constraints

$$\lambda \; = \theta \cdot b_{dr}^{(1)} \; + \; \varphi \cdot b_{dr} \tag{IA.31}$$

where $b_{dr}^{(1)} = \mathbf{1}_{r,s}\Phi_1$ and $b_{dr} = \rho \cdot \mathbf{1}_{r,s}\Phi_1(\mathrm{I}_\Phi - \rho \cdot \Phi_1)^{-1}$, with Φ_1 reflecting the VAR matrix in Equation IA.21, I_Φ capturing an identity matrix of the same dimension as Φ_1, and $\mathbf{1}_{r,s}$ representing a selector matrix such that $\mathbf{1}_{r,s}\mathrm{X}$ returns a column vector containing the X row associated with r and the X columns associated with s.

Model implied risk premia and pricing errors are obtained from

$$\widehat{\mathbb{E}}[xR] \; = \; - \; \frac{Cov(xR_t, M_t)}{\mathbb{E}\left[M_t\right]} \tag{IA.32}$$

and

$$\widehat{\alpha} \; = \mathbb{E}[xR] - \widehat{\mathbb{E}}[xR] \tag{IA.33}$$

where $M_t = exp\left\{\widehat{a} \; - \; \widehat{\lambda}_m \cdot \widetilde{r}_t \; - \; \widehat{\theta} \cdot N_{dr,t}^{(1)} \; - \; \widehat{\varphi} \cdot N_{dr,t}\right\}$.

Statistical inference on risk prices and pricing errors is done through a bootstrap analysis (similar to other quantities in the paper). Specifically, I draw 5-year blocks from the monthly observations of $\widetilde{r}_t$, $N_{dr,t}^{(1)}$, and $N_{dr,t}$ jointly with xR and r_f and estimate, in each simulation, all risk prices and pricing errors as described above. I use 10,000 simulations and calculate the t-statistic for a generic quantity θ using $t^{\theta}_{stat} = \widehat{\theta}/\widehat{\sigma}(\widehat{\theta})$ with $\widehat{\sigma}(\widehat{\theta})$ obtained from the standard deviation of $\widehat{\theta}$ across simulations (with both tails of the distribution trimmed at 1% to obtain a consistent estimator of $\sigma(\widehat{\theta})$ as suggested in Hansen (2019)).

C Data Sources and Measurement

This section details the data sources and measurement for variables used in the VAR estimation described in Section 2. The final dataset is a multivariate time series of monthly observations in which flow variables (such as returns and dividend growth) have annual measurement. Since the left side of the predictive regressions starts in Dec-1952, the measurement of the VAR variables start in Dec-1951 given the annual VAR.

C.1 Returns, Dividend Growth (Δd), and Dividend Yield (dp)

Annual r_f is obtained by compounding monthly returns on the one-month Treasury bill (available in Kenneth French's data library) over overlapping 12-month periods.[29]

Equity returns (r_e) and dividend growth (Δd) are based on a value-weighted portfolio containing all common stocks available in the CRSP data set and their measurement accounts for delistings and mergers and acquisitions (M&A) paid in cash. I do not use the CRSP value-weighted index because it includes all issues listed on NYSE, NASDAQ, and AMEX with, on average, 5.3% of the market capitalization in the index referring to noncommon stock issues (see Sabbatucci (2015)). Moreover, accounting for delistings and M&A activity requires a "bottom-up" approach.

I start by adjusting returns for delistings. For each firm for which I can identify a delisting (delisting code available and different from 100), I adjust the (ex- and cum-dividend) return for the month in which the distribution of proceeds took place by assigning the delisting return to that month. If no delisting return is available, I base the delisting return on the findings in Shumway (1997) and assign to the delisting month a return of -30% if the delisting was for cause (delisting code between 400 and 599) and of 0% otherwise. I assign a 0% return to all months between delisting and distribution when there is a temporal gap between the two events.

With ex- and cum-dividend returns accounting for delistings, I construct returns based on

[29] See http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

a value-weighted equity portfolio. I start by selecting all common shares (share codes 10 and 11) listed on NYSE, NASDAQ, or AMEX (exchange code 1, 2, and 3) and then calculate value-weighted cum- and ex-dividend monthly returns ($R^{cum}_{m,t}$ and $R^{ex}_{m,t}$).

Since my dividend measurement accounts for M&A paid in cash (as suggested in Allen and Michaely (2003)), I also construct a monthly "M&A yield" ($M\&Ay = M\&A_t/P_{t-1}$) at the aggregate level. Specifically, each month I sum all proceeds from distributions that can be classified as originating from an M&A paid in cash (distribution code between 3000 and 3400) across all firms that have lagged market equity available, and I divide this value by the sum of the lagged market equity for these firms.

To get dividends that incorporate M&A activity, I first adjust aggregate ex-dividend monthly returns by $\widehat{R}^{ex}_{m,t} = R^{ex}_{m,t} - M\&Ay$ and calculate a normalized aggregate price series, $\widehat{P}_t$, by cumulating $\widehat{R}^{ex}_{m,t}$. I then calculate dividends from cum- and ex-dividend returns as is standard in the literature (see Koijen and Nieuwerburgh (2011)), but relying on the adjusted ex-dividend return so that $\widehat{D}_{m,t} = \left(R^{cum}_{m,t} - \widehat{R}^{ex}_{m,t}\right) \cdot \widehat{P}_{t-1}$.[30]

The monthly series of annual dividends ($\widehat{D}_t$) is based on the sum of the monthly dividends ($\widehat{D}_{m,t}$) over the respective period. I sum the dividend as opposed to reinvesting them into the stock market to avoid introducing properties of returns into dividend growth (see Binsbergen and Koijen (2010)).

Dividend growth is given by $\Delta d = log(\widehat{D}_t/\widehat{D}_{t-12})$ and dividend yield by $dp = log(\hat{D}_t/\widehat{P}_t)$. To get annual returns that are consistent with the assumption of no dividend reinvestment, I use $r_{e,t} = log((\widehat{P}_t + \widehat{D}_t)/\widehat{P}_{t-12})$ as opposed to compounding $R^{cum}_{m,t}$. Finally, I subtract annual (log) inflation from r_f, r_e, and Δd using the CPI index to get real quantities.

The Internet Appendix of Gonçalves (2021) shows that including M&A activity in the

[30] It is important to note that the somewhat natural approach of calculating M&A based on $\widehat{D}_{m,t} = (R^{cum}_{m,t} - R^{ex}_{m,t}) \cdot P_{t-1} + M\&Ay \cdot P_{t-1}$, where P_t is constructed by cumulating $R^{ex}_{m,t}$, is incorrect as it produces price and dividend series that are inconsistent with the cum return provided: $R^{cum}_{m,t} \neq (P_t + \widehat{D}_{m,t})/P_{t-1}$. The method I use ensures that $R^{cum}_{m,t} = (\widehat{P}_t + \widehat{D}_{m,t})/\widehat{P}_{t-1}$, which is important because accounting for M&A activity in dividend payments should not affect the cum-dividend return delivered by equities. It simply affects the split between how much of that return comes from dividends and price appreciation.

dividend measurement changes the dynamics of Δd and dp and helps alleviate nonstationarity concerns (with these results being consistent with the findings in Sabbatucci (2015)). All other adjustments (e.g., related to delistings and the aggregation of firm-level data) have negligible effects.

C.2 Predictive Variables ($s_t = [dp \quad \Delta d \quad ty \quad TS \quad CS \quad VS]$)

Sources and measurement for dividend yield (dp) and dividend growth (Δd) are detailed above. The Treasury yield (ty) is the one-year log Treasury yield and comes from Global Financial Data until May 1952 and from the CRSP Fama-Bliss discount bond file after that. The term spread (TS) is the difference between the 10-year log Treasury yield and ty, where the former comes from Global Financial Data until March 1953 and from the Federal Reserve of St. Louis website thereafter. The credit spread (CS) is the difference between Moody's corporate BAA and AAA log yields with both coming from the Federal Reserve of St. Louis website. The value spread is the difference between the log book-to-market ratios of the value and growth portfolios formed based on small stocks and adjusting for within-year movements in market equity. The data come from Kenneth French's data library and measurement follows Campbell and Vuolteenaho (2004).

In the COVID crisis analysis of Section 4.1, I also construct these predictive variables daily and combine them with VAR estimates to recover daily equity strip yields. In the case of ty, TS, and CS, the Federal Reserve of St. Louis website provides daily values for the relevant variables. In the case of VS, I use the same construction of the monthly VS, except that I update the log book-to-market ratios of the value and growth portfolios formed based on small stocks using daily return for the respective portfolios (also available on Kenneth French's data library). The case of dp and Δd is more involved to make sure the daily values match the monthly values at the end of each month. For the price measure (used in dp), I first obtain daily ex-dividend returns following a procedure analogous to the one used to obtain my monthly $\widehat{R}^{ex}_{m,t}$ and then obtain daily prices by deflating the end of month price used in my monthly dp measure by these daily ex-dividend returns over the relevant period. For the dividend measure (used in dp and Δd), I first obtain daily dividends following a procedure analogous to the one used to obtain my monthly $\widehat{D}_{m,t}$ and then calculate the fraction of the total monthly dividend due to the given day by dividing the daily dividend by the sum of daily dividends over the month. I then multiply this fraction by my $\widehat{D}_{m,t}$ measure, which yields adjusted daily dividends that are consistent with my monthly dividends. Finally, each day I sum these adjusted daily dividends over the past year. For example, on 04/22/2020, I sum the daily dividends paid from all days that come after 04/22/2019 up to 04/22/2020.

D Supplementary Empirical Results

This section provides the main result in the paper (i.e., the term structure decomposition of equity variance) under different empirical specifications. Figure IA.1 summarizes these robustness findings.

First, Figures IA.1(a) to IA.1(f) consider results under different specifications for the state vector, s_t. Figure IA.1(a) uses $s_t = dp_t$, which is the simplest state vector one can create and is often used in the literature as the benchmark univariate forecasting regression for returns and dividend growth (e.g., Cochrane (2011)). Figure IA.1(b) to IA.1(f) consider alternative specifications that take our benchmark $s_t = [dp_t \ \Delta d_t \ ty_t \ TS_t \ CS_t \ VS_t]$ vector and drop one state variable at a time, with the exception of dp_t the Campbell and Shiller (1989) approximation implies $\mathbb{E}_t[r_{t+1}] = k + \mathbb{E}_t[\Delta d_{t+1}] - \rho \cdot \mathbb{E}_t[dp_{t+1}] + dp_t$, and thus state variables have to spam dp_t for a proper predictive system. For all these cases, the fraction of equity volatility explained by the present value of long-term dividends (i.e., dividends beyond twenty years) is fairly close to the 60% benchmark in the main text.

Second, Figures IA.1(g) and IA.1(h) consider different specifications for dividends. Figure IA.1(g) measures dividends without incorporating M&A paid in cash while Figure IA.1(h) obtains dividend growth directly from Campbell and Shiller (1989) approximation as $\Delta d_{t+1} = -k + r_{t+1} + \rho \cdot dp_{t+1} - dp_t$. In both cases, the fraction of equity variance explained by long-term dividends remains high. In FigureIA.1(h) it is close to the 60% benchmark in the main text and in Figure IA.1(g) it is even higher (close to 70%).

Third, Figures IA.1(i) and IA.1(j) consider alternative sample periods. Specifically, Figure IA.1(i) shows that the fraction of equity volatility explained by long-term dividends is slightly higher than the 60% benchmark when we exclude the great recession. Figure IA.1(j) shows that this fraction is lower (but still higher than 50%) if we start our sample period in 1926. While a longer sample period has advantages, it can limit the application of a VAR with fixed parameters, especially given that s_t contains interest rate variables that are likely affected by the Fed-Treasury Accord of 1951 that restored independence to the Fed, affecting monetary policy.

Finally, Figures IA.1(k) and IA.1(l) consider alternative estimation methods for the VAR (described in Internet Appendix Subsection B.1). As it is clear from the figures, the VAR estimation method has almost no impact on the fraction of volatility explained by long-term dividend present values.

References for Internet Appendix

Allen, F. and R. Michaely (2003). "Payout Policy". In: *Handbook of the Economics of Finance*. Ed. by G. M. Constantinides, M. Harris, and R. M. Stulz. Vol. 1. A. Elsevier Science. Chap. 7, pp. 337–429.

Binsbergen, J. H. v. and R. S. J. Koijen (2010). "Predictive Regressions: A Present-Value Approach". In: *Journal of Finance* 65.4, pp. 1439–1471.

Campbell, J. Y. (1991). "A Variance Decomposition for Stock Returns". In: *The Economic Journal* 101.405, pp. 157–179.

Campbell, J. Y. and R. J. Shiller (1989). "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors". In: *Review of Financial Studies* 1.3, pp. 195–228.

Campbell, J. Y. and T. Vuolteenaho (2004). "Bad Beta, Good Beta". In: *American Economic Review* 94.5, pp. 1249–1275.

Cochrane, J. H. (2011). "Presidential Address: Discount Rates". In: *Journal of Finance* 66.4, pp. 1047–1108.

Gonçalves, A. S. (2021). "Reinvestment Risk and the Equity Term Structure". In: *Journal of Finance* Forthcoming.

Hansen, B. (2019). "Econometrics," in: Unpublished manuscript.

Jordà, Ò. (2005). "Estimation and Inference of Impulse Responses by Local Projections". In: *American Economic Review* 95.1, pp. 161–182.

Jordà, Ò. and S. Kozicki (2011). "Estimation and Inference by the Method of Projection Minimum Distance: An Application to the New Keynesian Hybrid Phillips Curve". In: *International Economic Review* 52.2, pp. 461–487.

Koijen, R. S. J. and S. V. Nieuwerburgh (2011). "Predictability of Returns and Cash Flows". In: *Annual Review of Financial Economics* 3, pp. 467–491.

Nelson, C. R. and A. F. Siegel (1987). "Parsimonious Modelling of Yield Curves". In: *Journal of Business* 60.4, pp. 473–489.

Pope, A. L. (1990). "Biases of Estimators in Multivariate Non-Gaussian Autoregressions". In: *Journal of Time Series Analysis* 11.3, pp. 249–258.

Sabbatucci, R. (2015). "Are Dividends and Stock Returns Predictable? New Evidence Using M&A Cash Flows". Working Paper.

Shumway, T. (1997). "The Delisting Bias in CRSP Data". In: *Journal of Finance* 52.1, pp. 327–340.

Stambaugh, R. F. (1999). "Predictive regressions". In: *Journal of Financial Economics* 54, pp. 375–421.

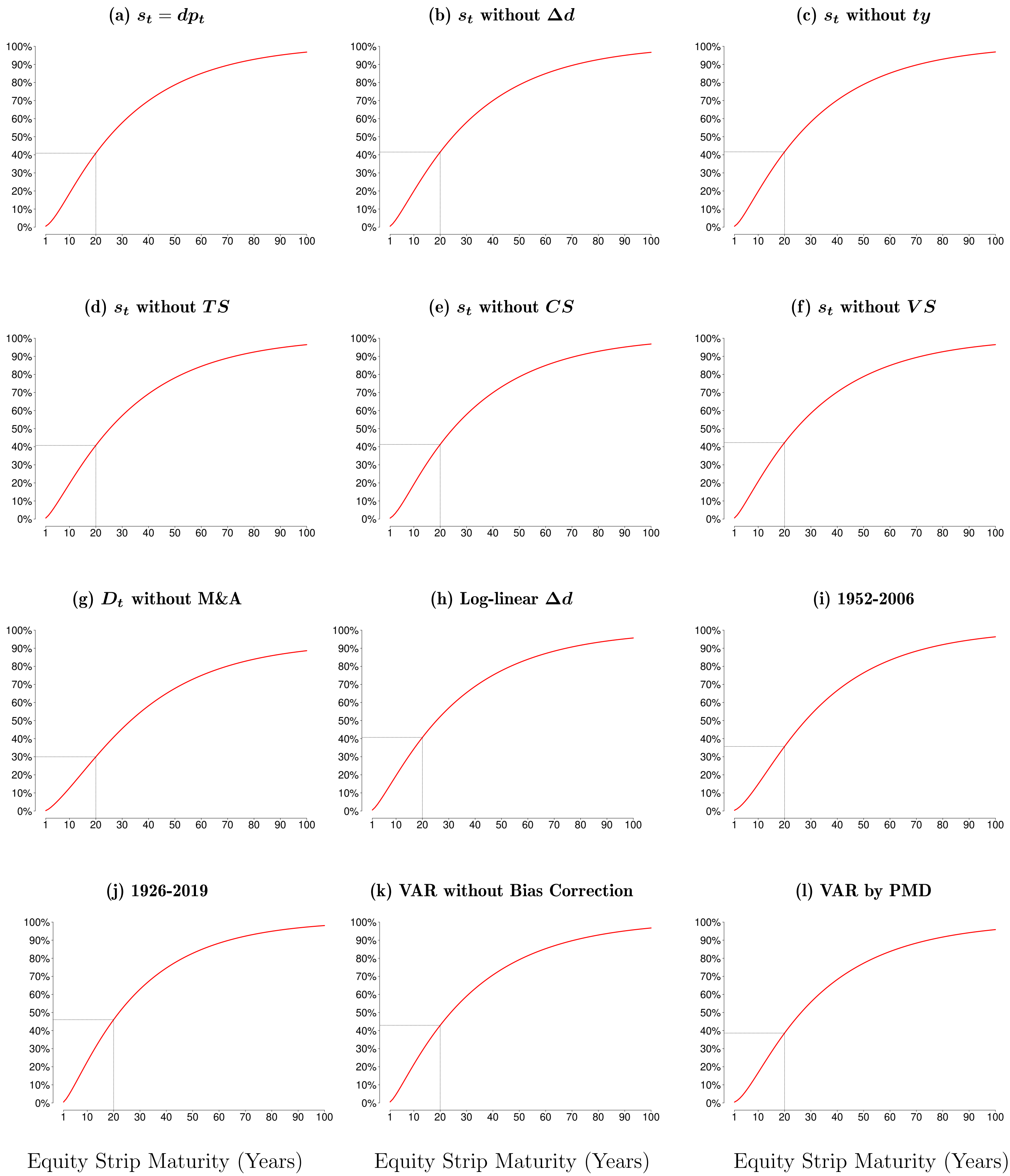


Figure IA.1
Cumulative % of Equity Variance Explained by Equity Strips: Robustness Checks

The graphs report the cumulative fraction of equity variance explained by equity strips with maturity from 1 to 100 years. Each graph reports the results after a given modification to the baseline specification. Further details are provided in Internet Appendix Section D.