Size-related premiums

by

Thiago de Oliveira Souza

Discussion Papers on Business and Economics
No. 3/2018

FURTHER INFORMATION
Department of Business and Economics
Faculty of Business and Social Sciences
University of Southern Denmark
Campusvej 55, DK-5230 Odense M
Denmark

E-mail: lho@sam.sdu.dk / http://www.sdu.dk/ivoe

Size-related premiums

THIAGO DE OLIVEIRA SOUZA*

ABSTRACT

This paper theoretically links the stock characteristics size and value to risks. The size premium arises – and spans the value premium – exclusively for portfolios formed in high *market price of risk* states. This is when the cross-sectional differences in risk premiums dominate the differences in expected cash flows connecting size and risk. Otherwise, value links better to the same risks, as it scales size by a proxy for expected cash flows. The hypothesis that value and size are (constant) risk proxies is formally rejected in the data, challenging the use of size-related portfolios as risk factors along with several strands of the literature based on this hypothesis.

JEL Classification: G11, G12, G14.
Keywords: Size premium, Value premium, Risk, Conditional, Portfolio sorts.

*Department of Business and Economics, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark. Email: tsouza@sam.sdu.dk. I would like to thank Jonathan Berk, Jylha Petri, Alexander Schandlbauer, Tim Kroencke, Stefano Giglio, Paulo Maio, Christian Flor, Linda Larsen, Francisco Gomes, the participants in the World Finance Conference 2016, Young Scholars Nordic Finance Workshop 2017, and Midwest Finance Association 2018.

I. Introduction

The size premium exists, is larger than, spans (without being spanned by), and shares a strong factor structure with the value premium, but only for portfolios formed in high *market price of risk* states. I explain why this happens and how the link between risk and these (portfolio level) characteristics varies over time. I formally test and reject the hypothesis of an otherwise constant link in favor of the framework that I present. This challenges the risk adjustment proposed by Fama and French (1996) because it reveals that the SMB and HML portfolios are not unconditional and stationary risk factors.[1] This also contradicts several strands of the literature in which any of these characteristics are (constant) risk proxies, such as the one starting with Zhang (2005).[2]

The explanation builds on the seminal ideas of Berk (1995), which I briefly recapitulate from a risk perspective:[3] Let $\zeta = (\zeta_t)$ be the unique stochastic discount factor following the Brownian motion in Eq. (1), where dz_{1t} is a one-dimensional standard Brownian motion, r^f is the risk-free rate, and λ_t is the market price of risk (both constant until time T). Let $D_{i,T}$ be an uncertain cash flow paid at time T, such that its time t expectation, $x_{i,t} = \mathrm{E}_t[D_{i,T}]$, follows the process in Eq. (2), implying the price in Eq. (3):

$$d\zeta_t = -\zeta_t[r^f\, dt + \lambda_t\, dz_t], \tag{1}$$

$$dx_{i,t} = x_{i,t}\left[0\, dt + \sigma_{i,t}\, dz_{1t}\right], \tag{2}$$

$$\implies P_{i,t} = \mathrm{E}_t[D_{i,T}]e^{-(r^f+\sigma_{i,t}\lambda_t)(T-t)}. \tag{3}$$

Eq. (3) summarizes the framework in Berk (1995): If the stocks had the same expected

[1]According to Fama and French (1997), one of the issues with using industry portfolios for risk adjustment purposes is precisely their non-stationary risk exposures.

[2]A short list with other examples includes Favilukis and Lin (2016), Piotroski and So (2012), Kapadia (2011), Chava and Purnanandam (2010), Campbell, Hilscher, and Szilagyi (2008), Yogo (2006), Griffin and Lemmon (2002), Gompers and Metrick (2001), or Lee, Shleifer, and Thaler (1991).

[3]Appendix I contains the theoretical framework in more details.

cash flows, there would be a perfect one-to-one mapping between high risk, $\sigma_{i,t}$, and low price, $P_{i,t}$. So the risk and the expected returns on the SMB portfolio, respectively,

$$\sigma_{smb,t} = \sigma_{small,t} - \sigma_{big,t}, \tag{4}$$

$$\mu_{smb,t} = \sigma_{smb,t}\lambda_t, \tag{5}$$

would both be very large because $\sigma_{small,t} >> \sigma_{big,t}$. But in reality the firms have different expected cash flows, and the size ranking does not seem to separate risky and safe stocks very well: The typical SMB portfolio has a low, often insignificant, risk premium.

The main new insight of the paper is the following: Extending the analysis in Berk (1995) reveals that a larger market price of risk, λ_t, increases the risk of the SMB portfolio (and the size premium) by strengthening the link between risk and size. For example, with $\lambda_t = 0$ in Eq. (3), the cash flows of all firms are discounted at the same risk-free rate, r^f. This induces a perfect one-to-one mapping between size and expected cash flows (**instead of risks**). Without a size-risk link in place, the small and the big stocks tend to be similarly risky, $\sigma_{small,t} \approx \sigma_{big,t}$ in Eq. (4), the SMB portfolio formed in this state has no risk, and the expected size premium is zero. On the other hand, a large enough λ_t in Eq. (3) guarantees that the prices of the risky stocks, with larger $\sigma_{i,t}$, are smaller than the prices of the safer stocks. This creates a perfect link between risk and size. Hence, the SMB portfolios formed in these states have the maximum risks and premiums.[4]

The same logic applies to the value premium: The price-to-book (PB) characteristic of firm i is given by

$$\frac{P_{i,t}}{BE_{i,t}} = \frac{\mathrm{E}_t[D_{i,T}]}{BE_{i,t}} e^{-(r^f+\sigma_{i,t}\lambda_t)(T-t)}, \tag{6}$$

where $BE_{i,t}$ is the firm's book equity (BE). As in Berk (1995), assume that the BE is a

[4] Appendix I.A contains a more detailed explanation of this result.

proxy for expected cash flows, $\mathrm{E}_t[D_{i,T}] \approx BE_{i,t}$, so that

$$\frac{P_{i,t}}{BE_{i,t}} \approx e^{-(r^f+\sigma_{i,t}\lambda_t)(T-t)}, \tag{7}$$

where the approximation depends on how well the BE proxies for expected cash flows. If the approximation is valid, the riskier stocks (with high $\sigma_{i,t}$) should have smaller PBs from Eq. (7), even when the market price of risk, λ_t, is small. This generates a risk-PB link even if the risk premiums, $\sigma_{i,t}\lambda_t$, are not as large as the ones needed to overcome differences in expected cash flows and establish the size-risk link in Eq. (3).[5] For some low market price of risk values, λ_t, the risk-PB link exists, but the risk-size link does not. So the HML portfolio captures risks that the SMB portfolio does not, and the value premium is larger than the size premium in these states. When the market price of risk is high, the link between size and risk is already established. Therefore, dividing size by BE, which is an imperfect proxy for cash flows, re-shuffles the ranking and reduces the risk of the HML portfolio compared to the SMB portfolio. The size premium is larger than the value premium in these states.

Finally, the value and the size premiums arise due to the same risks, as they are not intrinsically related to different risk dimensions.[6] In particular, there is no risk associated with the BE or with the PB characteristic itself. Hence, the two premiums should be highly correlated (and share a strong factor structure) when the market price of risk is high. However, when the market price of risk is low, the size ranking captures no risks and the correlation should be very low (with no factor structure).

The predictions above are exactly in line with the data. First, I show that both the value and especially the size premiums increase with the market price of risk proxy. Next, I confirm the remaining predictions by partitioning history into high and low market price

[5]See the explanation leading to Eq. (61) in the appendix for more details.

[6]Even in a multi-dimensional risk setting, in which $\sigma_{i,t}$ is a vector of exposures to orthogonal risk dimensions, we obtain conditions similar to Eq. (3) and Eq. (6), as detailed in Appendix I.C.

of risk states, using a backward-looking state classification (as explained in Section II.C) based on the significance of the size premium in each subsample.

A strict classification assigns around 10% of the sample to the high market price of risk bin. The monthly size premium in the high and low market price of risk bins are, respectively, 1.3% (2.73 t statistics) and 0.1% (1.07). The HML coefficients in the spanning regressions in high and low market price of risk states are, respectively, 0.3 (4.14) and 0.03 (0.81). The unspanned returns on the SMB portfolio, after the additional inclusion of the market premium in the regression, are 1.1% (2.55) and −0.02% (−0.21), in the high and low market price of risk bins. The adjusted R^2 values of the spanning regressions are also substantially larger for the portfolios formed in high market price of risk states, indicating a much stronger factor structure in these periods: 0.131 and 0.000 for the portfolios formed in high and low market price of risk states, respectively, when I regress the SMB on the HML returns.

Relaxing the classification criteria to obtain a significant value premium in the high market price of risk bin results in around 22% of the sample in this bin. The value premiums are, respectively, 0.7% (2.11) and 0.3% (3.11) for the portfolios generated in the high and low market price of risk states. Even in this extended subsample, the value premium left unspanned by the size premium is an insignificant 0.3% (0.97) for the portfolios formed in the high market price of risk states. Including the market premium in the regression reduces this even further to 0.1% (0.52). Again, the factor structure is concentrated in the high market price of risk states.[7]

More conclusively, the hypotheses of size and value as (constant) risk proxies are rejected in formal empirical tests: One hypothesis is that there are orthogonal underlying risks constantly associated with size and value. Under this hypothesis, the slope coeffi-

[7] Lowering the threshold also induces a negative correlation between the size and the value premiums in the low market price of risk subsample that is otherwise absent, but it does not create a significant factor structure between the returns on the SMB and HML portfolios.

cients in the spanning regressions should be zero. The slope estimates are even biased towards zero due to measurement error. Nevertheless, the slopes are significantly positive for the portfolios formed in high market price of risk states as explained earlier, which clearly rejects the null hypothesis of orthogonal underlying risks.

A second hypothesis is simply that there are underlying risks constantly associated with size and value. Under this hypothesis, the return spread between the HML and SMB portfolios must increase with the market price of risk.[8] However, the positive spread for low market price of risk states, 0.29% (2.20 t statistics), turns negative for the portfolios formed in high market price of risk states, -0.88% (1.39). The difference of 1.18% (1.81) is significant against the one-sided alternative hypothesis that the spread must decrease. As a consequence, the hypothesis of constant underlying risks is rejected in favor of the framework that I present.

Related literature: This paper provides an explanation for the risks associated with the size-related characteristics in general and with the SMB and HML portfolios in particular. This discussion relates primarily to Fama and French (1996) and to the large group of theories in different strands of the literature that tentatively explain the size and the value premiums mentioned in the first paragraph.

The paper also relates to the literature on the cross-sectional risk information contained in stock characteristics. For example, Chordia, Goyal, and Shanken (2015), Davis, Fama, and French (2000), and Daniel and Titman (1997) try to disentangle risk exposures from characteristics at the stock level (but not at the portfolio level). Other empirical papers document characteristic-related premiums, especially the independent contribution of each characteristic in explaining the premiums and whether the characteristic is really

[8]In summary, the test is based on the fact that it is always possible to represent expected returns in terms of a single risk factor (Roll, 1977) and that the HML portfolio earns significantly higher returns than the SMB portfolio in low market price of risk states, which implies that the SMB portfolio cannot be always riskier than the HML portfolio.

associated with a given risk dimension. A short list of recent contributions includes Kogan and Papanikolaou (2013), Dittmar and Lundblad (2017), Ball, Gerakos, Linnainmaa, and Nikolaev (2017), Green, Hand, and Zhang (2017), and Souza (2017).

Finally, the framework in the paper is very close to the ideas in Berk (1995), but within a time-varying risk framework. There is a myriad of models based on different assumptions that generate time variation in the market price of risk. Examples of these models include Rietz (1988), Epstein and Zin (1989), Constantinides and Duffie (1996), Campbell and Cochrane (1999), Hansen and Sargent (2001), Bansal and Yaron (2004), Barro (2006), Piazzesi, Schneider, and Tuzel (2007), Brunnermeier (2009), Bansal, Kiku, and Yaron (2012), Shiller (2014), and Garleanu and Panageas (2015).

II. Data description and variables

II.A. The stock returns

The return data on U.S. stocks, described in detail in Fama and French (1993), are from Kenneth French's data library. The monthly returns from July of 1926 to December of 2016 correspond to the size premium (as the return on the SMB portfolio, R_{smb}), the value premium (as the return on the HML portfolio, R_{hml}), and the market premium (as the difference between the return on the market portfolio and the risk free rate, R_{mp}).

II.B. The market price of risk proxy

The market price of risk proxy data are from Amit Goyal's website. I use both the Dow Jones' book-to-market (BM) of Pontiff and Schall (1998) (DJBM, starting in 1921) and the net equity expansion of Boudoukh, Michaely, Richardson, and Roberts (2007) (NTIS, starting in 1927). The NTIS forecasts negative returns, so I change its sign to give it a market price of risk interpretation, too. These variables are the only ones that significantly predict the one-month, one-year (and also the five-year) equity premium in-sample during the full 1927–2013 period according to the updated results from Welch and Goyal (2008) (Tables 1, 2, and 3).[9]

I decompose the two variables according to their principal components to check whether it is possible to obtain a single market price of risk proxy, which makes the analysis easier to digest. Table I displays the results of the principal component analysis for the DJBM and NTIS, showing that each component is responsible for approximately half of the total variance. Both the DJBM and the NTIS load positively on the first component and have opposite loadings on the second component. Figure 1 plots the time series of the DJBM, NTIS, and their components, both with a monthly and with an annual frequency (which keeps their values fixed in each July, when the SMB and HML portfolios are formed).

Next, I show that only the first of the two components, PC1, is a valid market price of risk proxy that significantly forecasts the market premium in regressions of the form

$$R_{mp(h),t+1} = \beta_{h,1} PC1_t + \beta_{h,2} PC2_t + \varepsilon_{h,t+1}, \tag{8}$$

where $R_{mp(h),t+1}$ is the market premium either with a one-month, $mp(h = m)$, or with a one-year horizon, $mp(h = y)$, starting from month $t + 1$. Table II reports each estimated coefficient, $\beta_{h,i}$, corresponding to the first and second principal components in month t, $PC1_t$ or $PC2_t$.

Table III presents the descriptive statistics for the DJBM, NTIS, and PC1. Panel A displays these results under the assumption that the series follow a first order AR process. The three variables are very persistent, but the NTIS and the DJBM are mostly uncorrelated in levels and also in innovations. Panel B shows that the three series are stationary, which reflects the fact that they are mostly ratios that cannot increase or

[9] Similar results are available for a larger selection of variables upon request.

decrease forever.

II.B.1. The market price of risk proxy on each (portfolio formation) date

Based on the results above, I consider the value of PC1 as the (single) market price of risk proxy for each month in the rest of the analysis,

$$\Lambda_t = PC1_t. \tag{9}$$

Given that the SMB and HML portfolios are formed every July, based on the information available at the end of June, the market price of risk proxy on the portfolio formation date is simply the value of PC1 in the latest (beginning of) July,

$$\Lambda_{f(t)} = PC1_{f(t)}, \tag{10}$$

where $f(t)$ is the latest July before time t.

II.C. The backward-looking percentile

The conditional tests require the identification of the years in which the market price of risk is high or low, and this is only possible in relative terms. One (forward-looking) solution is to simply rank the years in the full sample based on the market price of risk proxy. This ranking is time consistent, but it includes a look-ahead bias because the changes in the variable can be unexpected.

Instead, I calculate the percentiles recursively, given the historical values of the proxies up until each point in time. Essentially, this assumes that the changes in the variables are unexpected. The main advantage of this process is exactly to remove any look-ahead bias, and this is how I assign each year to a "high" or "low" market price of risk bin in the next sections.[10]

The process has three steps. The first step is to recursively find the historical mean of each variable in year t, $\overline{z}_t$. For a series that starts in year t_0,

$$\overline{z}_t = \sum_{i=t_0}^{t} \frac{z_i}{t - t_0 + 1}, \tag{11}$$

where z_i is the value of the variable in year i.

The second step is to calculate the difference between the value of z_t and its historical mean estimated until the previous year,

$$Dev_{z,t} = z_t - \overline{z}_{t-1}. \tag{12}$$

The final step is to calculate a percentile rank, $\Gamma_{z,t}$, based on how the time t deviation, $Dev_{z,t}$, compares with all previous deviations until time t,

$$\Gamma_{z,t} = \frac{\sum_{i=t_0}^{t} \left(I_{\{Dev_{z,i} < Dev_{z,t}\}} + 0.5 I_{\{Dev_{z,i} = Dev_{z,t}\}} \right)}{t - t_0 + 1}, \tag{13}$$

where $I_{\{.\}}$ is the indicator function. Intuitively, the ranking $\Gamma_{z,t}$ is high when the market price of risk proxy is large with respect to its historical (moving) average.

Fig. 2 and Fig. 3 show that the means of both the size and the value premiums increase for portfolios formed at higher percentiles of the DJBM, NTIS (both included for comparison), and PC1. Fig. 4 shows that the significance of the size premium for the portfolios formed in states in which the market price of risk proxy is above certain thresholds also tends to increase (until a certain point) with the thresholds, but Fig. 5 shows that the significance of the value premium actually tends to decrease with the thresholds. One explanation is that the significance of the value premium decreases

[10] A forward-looking classification gives similar results.

because the sample size reduces as the thresholds increase. Given that controlling for the market price of risk is not as crucial for creating a risk-value link as it is for creating a risk-size link, the significance of the value premium tends to simply decrease monotonically with the reduction in the sample size even if the mean of the premium increases, as we see in Fig. 3.

III. The empirical evidence

III.A. Size, value, and the market price of risk

The starting point of the model is that both the expected size and value premiums vary with the market price of risk, λ_t, in particular on the portfolio formation date, $\lambda_{f(t)}$, in line with Eq. (5). This section shows that, empirically, the expected premiums indeed increase with the market price of risk proxies, and this is especially true for the size premium. The evidence relies on running predictive regressions of the form

$$R_{i,t+1} = \alpha_i + \beta_{i,X} X_t + \varepsilon_{i,t+1}, \tag{14}$$

where $R_{i,t+1}$ is the time $t+1$ monthly return on the HML or the SMB portfolio, α_i is a constant intercept, $\beta_{i,X}$ is the constant slope on the time t value of the forecasting variable, X_t, and $\varepsilon_{i,t+1}$ is the error term.

The forecasting variable is the market price of risk proxy, the first principal component of DJBM and NTIS, PC1. It can be either its value in the latest month t, which gives $X_t = \Lambda_t$, or its value on the latest portfolio formation date for the SMB/HML portfolios in July, $X_t = \Lambda_{f(t)}$. Both are explained in Section II.B.1.

The left-hand portion of Table IV shows that, in the full sample, the proxy (strongly) predicts both the size and the value premium. Indeed, using the values on portfolio formation dates, $\Lambda_{f(t)}$, yields even better predictions than using the most recent month, Λ_t, in the full sample. This is in line with the hypothesis that the market price of risk on the portfolio formation date is fundamental to determine the risks (and therefore the premiums) associated with the SMB and HML portfolios. However, it is surprising to find this evidence in the data because the PC1 is a highly persistent variable, as we see in Table III, and does not change much within portfolio formation periods. Indeed, this specific result is not particularly robust, as we see on the right-hand side of Table IV: Removing the portfolio formed in 1932 from the sample also removes the evidence of superior performance for $\Lambda_{f(t)}$. Nevertheless, the market price of risk strongly forecasts the return on the SMB portfolio in this subsample, too. This is again in line with the hypothesis that the size premium is more sensitive to changes in the market price of risk than the value premium.

III.B. Partitioning history into high and low market prices of risk

In summary, Table IV confirms that there is a robust positive relation between the market price of risk and the returns on the SMB portfolio: The size premium should be large in the subsample in which the market price of risk is high, and vice versa. Hence, it is possible to split the data into high and low market prices of risk based on the magnitude and significance of the size premium.

III.B.1. The size premium threshold for the market price of risk, Λ^*_{MSE}

For the purpose of splitting the sample into two groups, assume that the size premium is constant when the market price of risk is above or below a certain threshold, instead of modeling it as function of the market price of risk, as in Eq. (5). This simplifies the estimation and reflects the fact that the size premium is in fact zero when the market price of risk is low, and positive otherwise.

Let the (observable) threshold for the market price of risk proxy, Λ^*, be a proxy for the actual threshold, λ^*, above which there is a link between size and risk, as explained in the introduction. The time t expectation of the size premium one period ahead is now

$$\mathrm{E}_t[R_{smb,t+1}] = \begin{cases} 0 & \Lambda_{f(t)} < \Lambda^* \\ \overline{R}_{smb,h} & \Lambda_{f(t)} \geq \Lambda^* \end{cases}, \tag{15}$$

where $\Lambda_{f(t)}$ is the time t market price of risk proxy (on the portfolio formation date), Λ^* is the threshold for this proxy, and $\overline{R}_{smb,h}$ is the mean return on the SMB portfolios that are formed in periods in which the market price of risk is above this threshold (high),

$$\overline{R}_{smb,h} = \frac{\sum_{t=1}^{T} \left(R_{smb,t} \; I_{\Lambda_{f(t)} \geq \Lambda^*} \right)}{\sum_{t=1}^{T} I_{\Lambda_{f(t)} \geq \Lambda^*}}, \tag{16}$$

where $I_{\{\}}$ is the indicator function and T is the sample size.

The mean squared error of this forecast (MSE), which is a function of the threshold, Λ^*, is

$$MSE(\Lambda^*) = \frac{1}{T} \sum_{t=1}^{T} \left(\mathrm{E}_{t-1}[R_{smb,t}] - R_{smb,t} \right)^2, \tag{17}$$

where $\mathrm{E}_{t-1}[R_{smb,t}]$ is given by Eq. (15). The natural choice for Λ^* is the one that minimizes the MSE in Eq. (17), with the additional restriction that the size premium must be significant in the high market price of risk subsample, denoted by

$$\begin{aligned} \Lambda^*_{MSE} \equiv \quad & \arg\min_{\Lambda^*} MSE(\Lambda^*) \\ & \qquad\qquad \text{r.t.} \quad t(\overline{R}_{smb,h}) \geq 2, \end{aligned} \tag{18}$$

where $t(\overline{R}_{smb,h})$ is the t statistics of the estimated mean in Eq. (16).

Figure 6 shows the MSE associated with different percentile thresholds, Λ^*, for the variables DJBM, NTIS (included again for comparison), and their principal component, PC1. The vertical dashed black lines appearing at very high percentiles for the DJBM and for the PC1 (but not for the NTIS) indicate the MSE thresholds for each variable.

The first two columns of Panel A in Table V present the size premium estimates conditioned on this threshold (with PC1 as the state variable). The results are in line with the theoretical predictions: The size premium is only significant for the portfolios formed in high market price of risk states, 1.3% per month (2.73 t statistics). The premium is small and insignificant for the portfolios formed in low market price of risk states, 0.1% (1.07).

III.B.2. The value premium threshold for the market price of risk, Λ^*_{HML}

The previous threshold, Λ^*_{MSE}, is restrictive: Only around 10% of the sample is assigned to the high market price of risk bin. On the positive side, this offers a precise identification of the years in which the size premium arises. The problem is that this subsample is too small to obtain a significant value premium. Panel B in Table V shows that the point estimates of the value premium are the same for the portfolios formed in high or low market price of risk states in this case: 0.4%. But the t statistics are 0.59 and 4.16, respectively.

For the conditional tests comparing the joint properties of the size and the value premiums it is possible to relax the threshold Λ^*_{MSE}, so that the value premium is also significant when the market price of risk is high under this new classification, $\Lambda_t \geq \Lambda^*_{HML}$. The only drawback is that some years of low market price of risk are misclassified.

Let us define the equivalent to Eq. (16) for the value premium,

$$\overline{R}_{hml,h} = \frac{\sum_{t=1}^{T}\left(R_{hml,t}\ I_{\Lambda_{f(t)}\geq\Lambda^*}\right)}{\sum_{t=1}^{T} I_{\Lambda_{f(t)}\geq\Lambda^*}}, \tag{19}$$

where $R_{hml,t}$ is the time t return on the HML portfolio and $\overline{R}_{hml,h}$ is the mean return on the HML portfolios that are formed in periods in which the market price of risk is above this threshold. This second threshold is the highest value for the market price of risk proxy that generates mean value and size premiums two standard errors above zero,

$$\begin{aligned} \Lambda^*_{HML} \equiv \quad & \underset{\Lambda^*}{\arg\max} = \Lambda^* \\ & \text{r.t.} \quad t(\overline{R}_{smb,h}) \geq 2 \\ & \qquad\quad t(\overline{R}_{hml,h}) \geq 2, \end{aligned} \tag{20}$$

where $t(\overline{R}_{i,h})$ is the t statistics of the estimated means in Eq. (16) and Eq. (19). In Figure 6, these are the vertical dashed light gray lines that appear at lower percentiles for the DJBM and for the PC1 compared to Λ^*_{MSE} (for the NTIS the two thresholds are the same).

The first two columns of Panel C in Table V display a significant value premium in both subsamples based on this threshold (for the variable PC1). The mean returns are, respectively, 0.3% (3.11) and 0.7% (2.11) for the portfolios formed in low or high market price of risk states according to this threshold. The lower Λ^*_{HML} threshold increases the proportion of years in the high market price of risk group to a little over 22%. This corresponds to an increase of more than 125% compared to the previous classification.

III.C. Conditional spanning tests: Value is redundant when size links to risk

Grouping high and low market price of risk periods based on $\Lambda^* = \Lambda^*_{HML}$ or $\Lambda^* = \Lambda^*_{MSE}$, allows us to test whether the premium linked to the value characteristic can be spanned by the premium linked to the size characteristic conditioned on the states in which the portfolios are formed. I test this by separately running, for the portfolios formed in high market price of risk states, $\Lambda_{f(t)} > \Lambda^*$, or, for the portfolios formed in low market price of risk states, $\Lambda_{f(t)} < \Lambda^*$, the spanning regression:

$$R_{hml,t} = \alpha_{hml} + s_{hml} R_{smb,t} + \varepsilon_{hml,t}, \tag{21}$$

where $R_{hml,t}$ is the time t return on the HML portfolio, α_{hml} is the constant intercept, s_{hml} is the constant coefficient on $R_{smb,t}$, which is the time t return on the SMB portfolio, and $\varepsilon_{hml,t}$ is the error term. According to the model, the coefficients for these regressions are

$$\alpha_{hml} \begin{cases} > 0 & \Lambda_{f(t)} < \Lambda^* \\ = 0 & \Lambda_{f(t)} \geq \Lambda^* \end{cases} \quad \text{and} \quad s_{hml} \begin{cases} = 0 & \Lambda_{f(t)} < \Lambda^* \\ > 0 & \Lambda_{f(t)} \geq \Lambda^* \end{cases}. \tag{22}$$

For portfolios formed in high market price of risk states, $\Lambda_{f(t)} \geq \Lambda^*$, there is no premium associated with the value characteristic that is not also associated with the size characteristic because there is already a size-risk link in place. This explains why α_{hml} is zero. In addition, the premiums related to both characteristics arise in response to the same risks, implying a positive s_{hml}. Otherwise, for the portfolios formed in low market price of risk states, $\Lambda_{f(t)} < \Lambda^*$, the size ranking captures no priced risks, so the value premium should be completely unspanned by and uncorrelated with the size premium.

The exposure to market risk can also affect the correlation between the returns on the size and value portfolios. Including the market premium in Eq. (21) gives

$$R_{hml,t} = \alpha_{hml} + \beta_{hml} R_{mp,t} + s_{hml} R_{smb,t} + \varepsilon_{hml,t}, \tag{23}$$

where β_{hml} is the constant coefficient on the market premium, $R_{mp,t}$. The predicted coefficients are the same as those presented in Eq. (22).

III.C.1. Results for the size premium threshold, Λ^*_{MSE}

The estimates in Panel B in Table V, based on Λ^*_{MSE}, strongly support the theoretical predictions regarding every coefficient in Eq. (22). However, the results involving the constant intercept, α_{hml}, are less informative because the value premium is already statistically insignificant in the (small) subsample in which the market price of risk is high: 0.4%, (0.59 t statistics).

For the portfolios formed in high market price of risk states (columns $\geq \Lambda^*_{MSE}$), there is a strong factor structure between the returns on the SMB and HML portfolios: The adjusted R^2 values are 0.131 and 0.335 for the formulations in Eq. (21) and Eq. (23), respectively. On the other hand, the size premium explains almost none of the variation in the value premium for the portfolios formed in low market price of risk states, with adjusted R^2 values of 0.000 and 0.003, respectively (columns $< \Lambda^*_{MSE}$). The s_{hml} coefficients also reveal how the market price of risk on the portfolio formation dates influences the risk profiles of the two portfolios: The coefficients reflecting the correlation between the two premiums are significantly large and positive in the former case, 0.5 (4.14) and 0.3 (2.35), respectively, considering Eq. (21) and Eq. (23), compared to essentially zero in the latter: 0.03 (0.81) and 0.004 (0.13).

III.C.2. Results for the value premium threshold, Λ^*_{HML}

Relaxing the market price of risk threshold by using $\Lambda^*_{HML} < \Lambda^*_{MSE}$ from Eq. (20) results in a significant value premium in both high and low market price of risk subsamples, as explained in Section III.B.2 and shown in the first two columns of Panel C in Table V. The drawback is that the restrictions in Eq. (22) are less likely to hold in this extended sample because the separation between the portfolios formed in high and low market price of risk states is less sharp.

Nevertheless, the results in Panel C in Table V are largely in line with the theoretical predictions in Eq. (22). Most importantly, the size premium, alone or in conjunction with the market premium, spans the value premium for the portfolios formed in the high market price of risk states: The intercepts in columns 4 and 6 are, respectively, 0.3% (0.97) and 0.1% (0.52) for the spanning tests excluding or including the market premium. Otherwise, the intercepts are significantly positive for the portfolios that are formed in low market price of risk states.

Again, there is a strong factor structure exclusively among the returns on the SMB and HML portfolios that are formed in high market price of risk states. The adjusted R^2 values are much lower for the portfolios formed in the low market price of risk states: 0.010 compared to 0.169 (based on Eq. (21)), and 0.009 compared to 0.310 (based on Eq. (23)).

The only minor difference with respect to the predicted coefficients in Eq. (22) is the negative correlation between the size and the value premiums in low market price of risk states, given by the coefficient s_{hml} in columns 3 and 5 of Panel C, Table V. As mentioned, these correlations translate to an almost non-existent factor structure, and they are zero conditioned on Λ^*_{MSE}. For the portfolios formed in high market price of risk states, the correlation between the returns on the SMB and HML portfolios is even stronger in Panel C (based on Λ^*_{HML}) than in Panel B (based on Λ^*_{MSE}): The s_{hml} coefficients in columns 4 and 6 are 0.5% (7.13) and 0.4% (4.71), not including and including the market premium in the spanning regression, respectively.

III.D. Conditional spanning tests: Size is never redundant when it is linked to risk

It is possible that the premium related to the size characteristic is also spanned by the one related to the value characteristic. This can be tested with a spanning regression equivalent to Eq. (21)

$$R_{smb,t} = \alpha_{smb} + h_{smb}R_{hml,t} + \varepsilon_{smb,t}. \tag{24}$$

Or, controlling for market risk, with coefficients equivalent to Eq. (23),

$$R_{smb,t} = \alpha_{smb} + \beta_{smb}R_{mp,t} + h_{smb}R_{hml,t} + \varepsilon_{smb,t}. \tag{25}$$

Within the same framework, the predicted coefficients are

$$\alpha_{smb} \begin{cases} = 0 & \Lambda_{f(t)} < \lambda^* \\ \geq 0 & \Lambda_{f(t)} \geq \lambda^* \end{cases} \quad \text{and} \quad h_{smb} \begin{cases} = 0 & \Lambda_{f(t)} < \lambda^* \\ > 0 & \Lambda_{f(t)} \geq \lambda^* \end{cases}. \tag{26}$$

The difference with respect to Eq. (22) is that the size premium can be unspanned by the value premium as long as (i) the portfolios are formed in periods in which the market price of risk is high and (ii) the BE is not a perfect proxy for expected cash flows. The intercept, α_{smb}, should be strictly positive in this case. Otherwise, for the portfolios formed in low market price of risk states, the intercept should be zero because the mean size premium is zero.

The results in Panel A in Table V, based on Λ^*_{MSE}, confirm the predictions in Eq. (26). Apart from the previous analysis about the factor structure and slope coefficients, the size premium is not spanned by the value premium when the market price of risk is high: The intercept, α_{smb}, is a significant 1.2% (2.69) based on the HML return alone, and 1.1% (2.55) with the market premium included in the spanning regression. When the market price of risk is low, the excess return on the SMB portfolio is zero. In summary, this is consistent with the hypothesis that it is the size portion of the value characteristic (market equity) that carries all risk information, and the BE is simply an imperfect proxy for expected cash flows. Hence, size is never redundant for the portfolios formed in high market price of risk states.

III.E. Formally rejecting the (constant) risk proxy hypotheses

Although the evidence in the previous sections supports the time-varying risk framework in the paper, it does not formally reject the alternative hypotheses of characteristics as risk proxies. I construct such formal empirical tests in this section.

Assume that there are constant underlying risks related to the SMB and HML portfolios. Let SMB^* and HML^*, respectively, be the portfolios that perfectly mimic these risks. In terms of the parameters in Eq. (1) and Eq. (2), the discrete time version of the returns on these portfolios is given by, respectively,

$$R_{smb^*,t} = \sigma_{smb^*}\lambda_{smb^*} + \sigma_{smb^*}\varepsilon_{smb^*,t}, \tag{27}$$

$$R_{hml^*,t} = \sigma_{hml^*}\lambda_{hml^*} + \sigma_{hml^*}\varepsilon_{hml^*,t}, \tag{28}$$

where λ_i is the average market price of risks i, σ_i is the exposure of portfolio i to these risks, and $\varepsilon_{i,t}$ is a (priced) shock on the premiums associated with these risks.

A linear regression of R_{smb^*} on R_{hml^*},

$$R_{smb^*,t} = \alpha + \beta R_{hml^*,t} + \varepsilon_t, \tag{29}$$

where ε_t is an error term, would have slope

$$\beta = \frac{\sigma_{smb^*}}{\sigma_{hml^*}}\frac{\text{cov}(\varepsilon_{smb^*,t}, \varepsilon_{hml^*,t})}{\text{Var}(\varepsilon_{hml^*,t})}, \tag{30}$$

and intercept

$$\alpha = \mathrm{E}[R_{smb^*,t}] - \beta\,\mathrm{E}[R_{hml^*,t}] = \sigma_{smb^*}\lambda_{smb^*} - \beta(\sigma_{hml^*}\lambda_{hml^*}). \tag{31}$$

The problem is that the portfolios SMB^* and HML^* are not observable, so the coefficients in Eq. (30) and Eq. (31) cannot be directly estimated.

III.E.1. The question within a measurement error framework

Let size and value be proxies for the true underlying and unobservable risks in the mimicking portfolios, SMB^* and HML^*, such that the returns on the SMB and HML portfolios contain the error terms u_t and v_t:

$$R_{smb,t} = R_{smb^*,t} + u_t, \tag{32}$$

$$R_{hml,t} = R_{hml^*,t} + v_t. \tag{33}$$

Substituting Eq. (32) and Eq. (33) in Eq. (29), gives

$$R_{smb,t} = \alpha + \beta(R_{hml,t}) + w_t, \tag{34}$$

where the error term $w_t = -\beta v_t + \varepsilon_t + u_t$ is correlated with $R_{hml,t}$, rendering the least squares estimates of the coefficients inconsistent. In particular, the estimate of β ($\hat{\beta}$) is biased towards zero, and the estimate of α ($\hat{\alpha}$) is biased in the opposite direction.

III.E.2. Rejecting the orthogonal risk proxy hypothesis

Under the hypothesis that value and size proxy for orthogonal risks, Eq. (30) implies that β is zero because

$$\mathrm{cov}(\varepsilon_{smb^*,t}, \varepsilon_{hml^*,t}) = 0. \tag{35}$$

The estimated $\hat{\beta}$ from Eq. (34) is in fact biased towards zero, which adds an extra hurdle to obtaining a significant value but implies that the null hypothesis of orthogonal risks,

$$H_0\colon \beta = 0, \tag{36}$$

is still rejectable if the evidence is strong enough.

The alternative hypothesis is that the risk of the portfolios varies according to the framework that I present. Within this framework, detailed in Section III.C, the SMB portfolios formed in low market price of risk states capture no risks, and the model also predicts that β in Eq. (34) is zero. However, the SMB and the HML portfolios formed in high market price of risk states capture the same risks under the alternative: Conditioned on portfolios formed in high market price of risk states, the alternative hypothesis is that beta is larger than zero:

$$H_a\colon \beta \geq 0 \mid \lambda_{f(t)} \geq \lambda^*. \tag{37}$$

The evidence presented in Section III.C shows that the null hypothesis in Eq. (36) is inconsistent with the data. More specifically, the fourth column of Panel A in Table V shows a positive $\hat{\beta}$ (0.3, 4.14 t statistics), rejecting the null hypothesis of orthogonal risks in favor of the alternative of positively correlated risks exclusively for the portfolios formed in high market price of risk states.

III.E.3. Rejecting the constant risk proxy hypothesis in general

The returns on the mimicking portfolios in Eq. (27) and Eq. (28) can also be expressed in terms of exposures to the single risk factor on the mean variance frontier that prices every asset in the economy (Roll, 1977), such that

$$\lambda_{mv} \equiv \lambda_{smb^*} = \lambda_{hml^*}, \tag{38}$$

$$\varepsilon_{mv,t} \equiv \varepsilon_{smb^*,t} = \varepsilon_{hml^*,t}. \tag{39}$$

Under the hypothesis that these mimicking portfolios have constant risk exposures, Eq. (30) and Eq. (31) imply

$$\beta = \frac{\sigma_{smb^*}}{\sigma_{hml^*}}, \tag{40}$$

$$\alpha = 0. \tag{41}$$

The two portfolios should span each other perfectly, which is rejected by the significant intercept in the fourth column of Panel A in Table V, for example. The problem is that the biases in estimates $\hat{\beta}$ and $\hat{\alpha}$ from Eq. (34) imply that this test cannot provide clean evidence against the null hypothesis because $\hat{\alpha}$ is overestimated.

Alternatively, consider the expected return spread between the HML and SMB portfolios, obtained from Eq. (27), Eq. (28), Eq. (32), and Eq. (33), conditioned on the market price of risk,

$$\mathrm{E}[SP_{t+1} \mid \lambda_{mv,t}] \equiv \mathrm{E}[R_{hml,t+1} - R_{smb,t+1} \mid \lambda_{mv,t}] = (\sigma_{hml^*} - \sigma_{smb^*})\lambda_{mv,t}. \tag{42}$$

The spread is significantly positive for the portfolios formed in low market price of risk states if and only if $\sigma_{hml^*} > \sigma_{smb^*}$. In this case, the spread must increase with the market price of risk under the null hypothesis of constant risk exposures:

$$H_0\colon\ \mathrm{E}[SP_{t+1} \mid \lambda_{mv,t} < \lambda^*] > 0 \implies \mathrm{E}[SP_{t+1} \mid \lambda_{mv,t} \geq \lambda^*] > \mathrm{E}[SP_{t+1} \mid \lambda_{mv,t} < \lambda^*]. \tag{43}$$

On the other hand, the link between size and risk arises exclusively for the portfolios formed in high market price of risk states within the framework that I present. Under this alternative hypothesis, the spread must be positive for the portfolios formed in low market price of risk states, but it must decrease with the market price of risk:

$$H_a\colon\ \mathrm{E}[SP_{t+1} \mid \lambda_{mv,t} \geq \lambda^*] < \mathrm{E}[SP_{t+1} \mid \lambda_{mv,t} < \lambda^*]. \tag{44}$$

Based on the market price of risk proxy threshold, Λ^*_{MSE}, Table VI shows that the hypothesis of a negative spread for the low market price of risk states is strongly rejected in the data: The mean spread for the portfolios formed in these states is 0.29% per month (2.20 t statistics), as predicted by the model that I present. This is also consistent with the null hypothesis in case $\sigma_{hml^*} > \sigma_{smb^*}$ and in case the spread increases with the market price of risk. However, this second part of the null hypothesis is rejected in favor of the alternative hypothesis of a decreasing spread: The mean spread for the portfolios formed in high market price of risk states is -0.88% per month (1.39 t statistics), and the difference of 1.18% (1.81 one-sided t statistics) implies rejection of the null hypothesis of constant risk exposures.

IV. Summary

This paper changes our understanding of the value and the size premiums in several ways. First, we learn about the existence of a very explicit, "first principles", connection between size and risk. This adds to Berk (1995) by showing that the size-related regularities appear in the data because of our (wrong) models of risk, not because investors do not care about risk. Second, it shows that the hypothesis of a direct link between risk and the BE or BM characteristics of the firms, which is predicted by long strands of the literature, such as the one following Zhang (2005), for example, is inconsistent with the data. Finally, the paper reveals several issues with the risk adjustment in Fama and French (1996): A smaller problem is multicollinearity, given that the HML is redundant in the periods in which there are risks associated with the SMB factor. A bigger problem is that the SMB and HML factors are not constant and stationary risk proxies: The unconditional "loadings" on these factors and their respective intercepts are not accurate indicatives of risk exposures and excess returns. The time-varying risk profiles of the SMB and HML portfolios resemble exactly those of the industry portfolios criticized by Fama and French (1997) as being inappropriate for risk adjustments. This is also a problem for the 25 portfolios double-sorted by size and value as test assets.

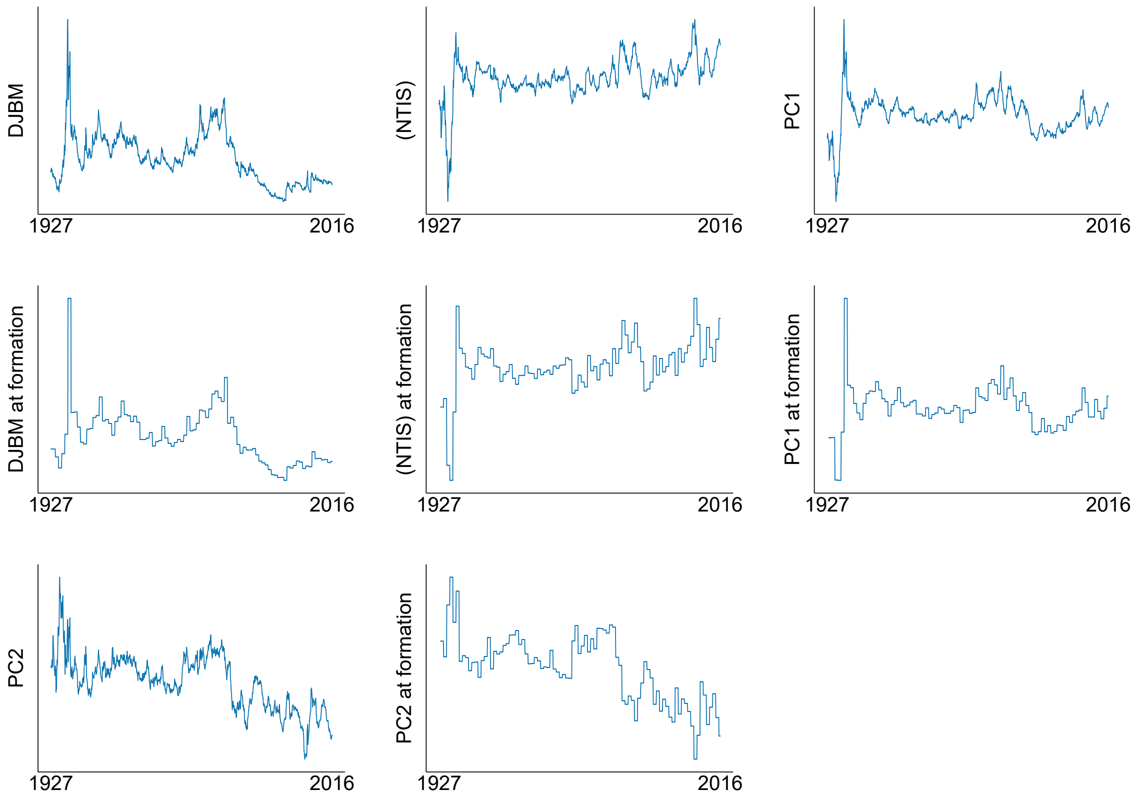

Figure 1. The market price of risk proxies in time series from 1927 to 2016. The panels plot the time series of the BM of the Dow Jones stocks (DJBM), (the negative of) the net equity expansion (NTIS), and the two factors corresponding to their two principal components, PC1 and PC2 at a monthly frequency. The values on the formation date are for the end of June in each year (fixed for 12 months).

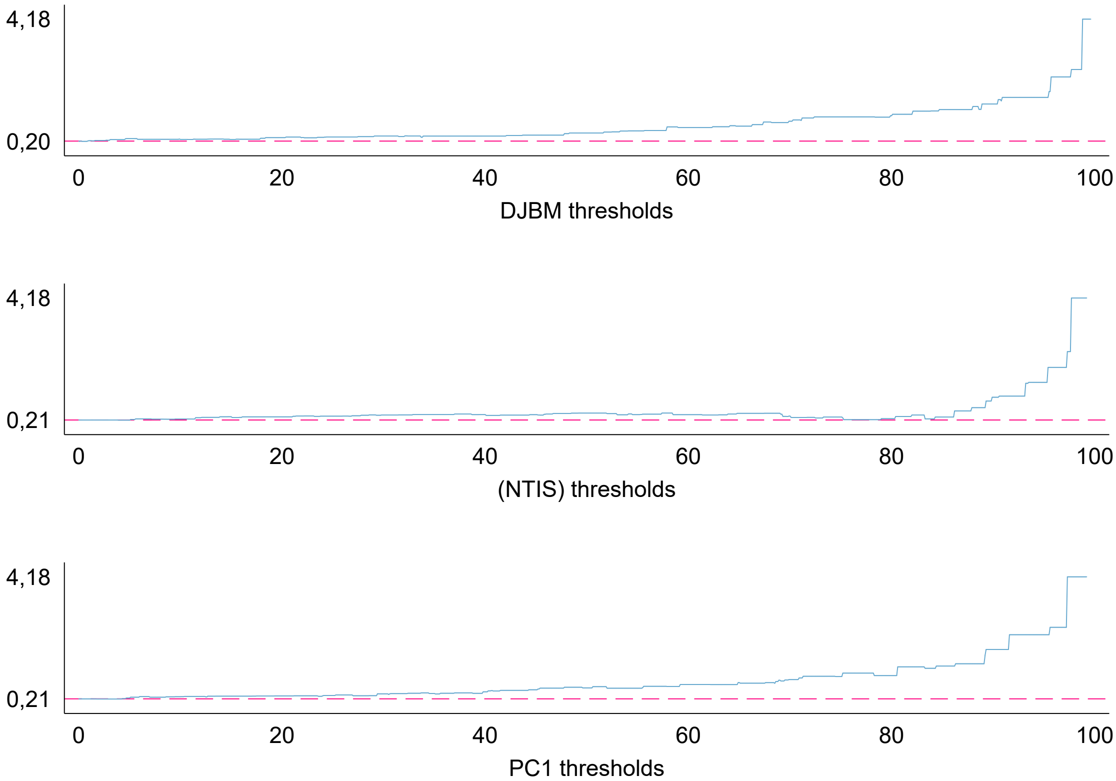

Figure 2. Mean size premium for the portfolios formed in states in which the market price of risk proxy is above certain thresholds in 1926–2016. The panels plot 1) the unconditional mean size premium, given by the return on the SMB portfolio, as the dashed pink horizontal line and 2) the mean size premium above different backward-looking thresholds for the market price of risk proxies, Λ^*. The market price of risk proxies are the BM of the Dow Jones stocks, DJBM, (the negative of) the net equity expansion, (NTIS), and the factor corresponding to their first principal component, PC1.

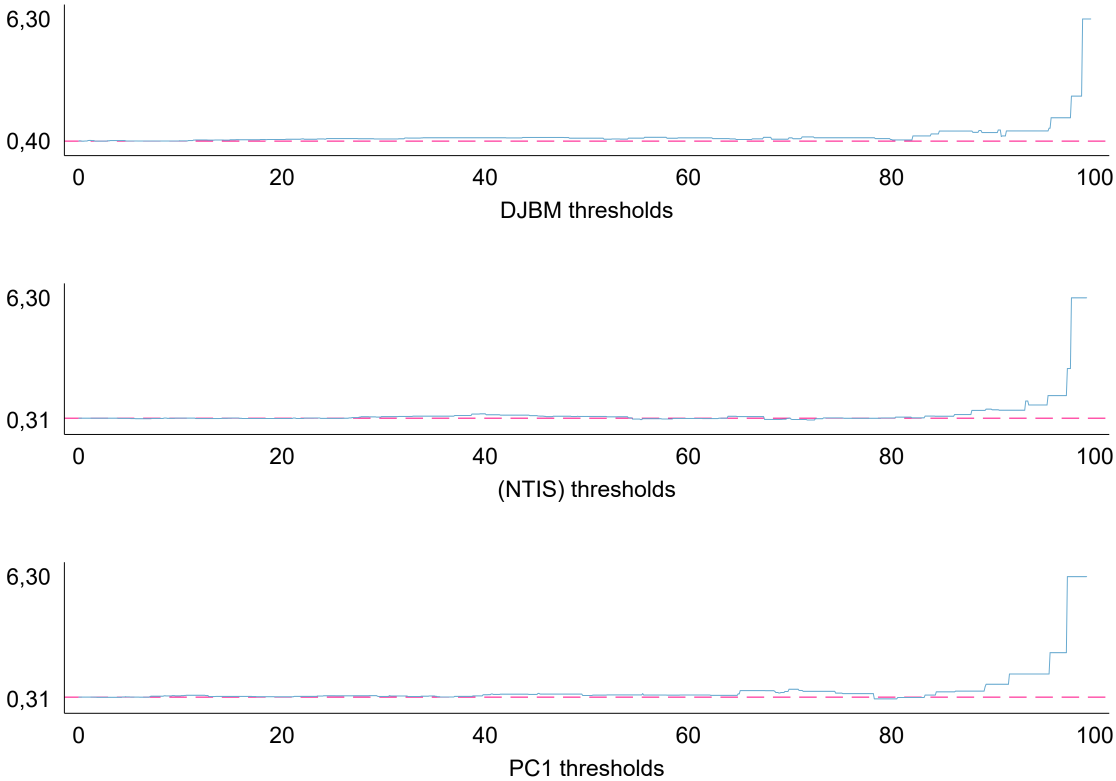

Figure 3. Mean value premium for the portfolios formed in states in which the market price of risk proxy is above certain thresholds in 1926–2016. The panels plot 1) the unconditional mean value premium, given by the return on the HML portfolio, as the dashed pink horizontal line and 2) the mean size premium above different backward-looking thresholds for the market price of risk proxies, Λ^*. The market price of risk proxies are the BM of the Dow Jones stocks, DJBM, (the negative of) the net equity expansion, (NTIS), and the factor corresponding to their first principal component, PC1.

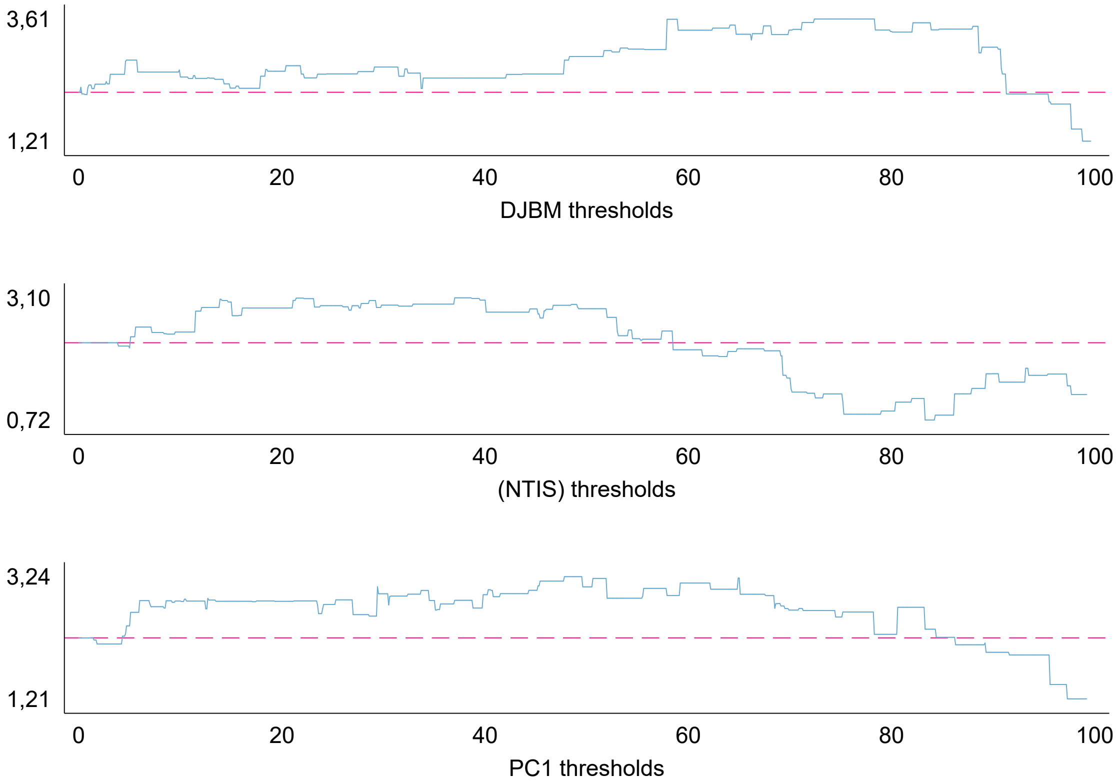

Figure 4. Significance of the size premium for the portfolios formed in states in which the market price of risk proxy is above certain thresholds in 1926–2016. The panels plot the t statistics of the mean size premium in Eq. (16) for 1) the values above different backward-looking percentiles; or 2) the unconditional estimation (horizontal dashed pink line). The market price of risk proxies are the BM of the Dow Jones stocks, DJBM, (the negative of) the net equity expansion, (NTIS), and the factor corresponding to their first principal component, PC1.

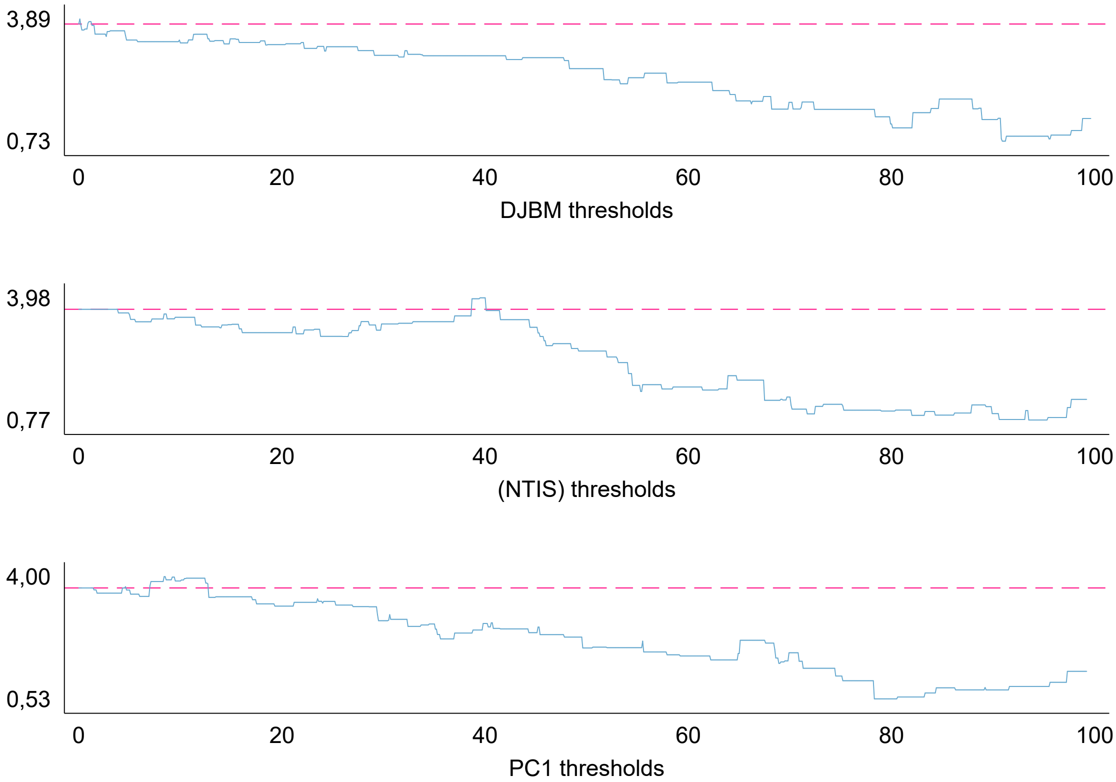

Figure 5. Significance of the value premium for the portfolios formed in states in which the market price of risk proxy is above certain thresholds 1926–2016. The panels plot the t statistics of the mean value premium in Eq. (19) for 1) the values above different backward-looking percentiles; or 2) the unconditional estimation (horizontal dashed pink line). The market price of risk proxies are the BM of the Dow Jones stocks, DJBM, (the negative of) the net equity expansion, (NTIS), and the factor corresponding to their first principal component, PC1.

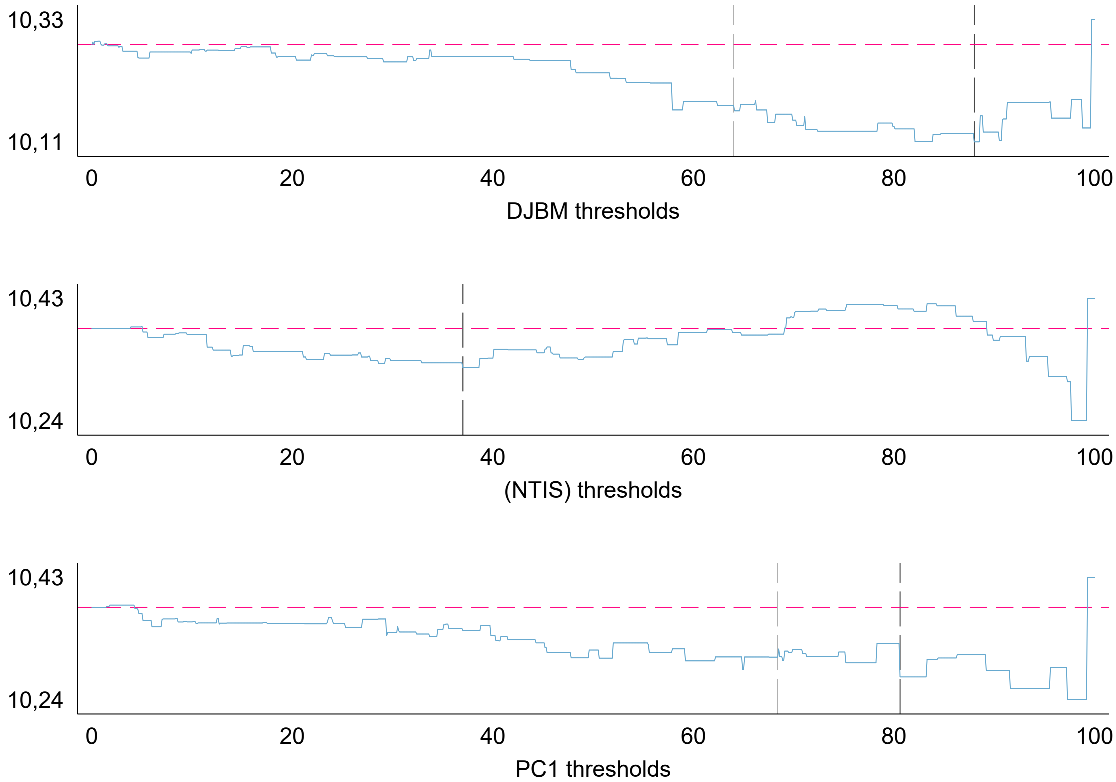

Figure 6. Conditional and unconditional mean squared errors from the prediction of the size premium in 1926–2016. The panels plot the MSE in Eq. (17) related to either 1) the forecast conditioned on different backward-looking thresholds for the market price or risk proxies, Λ^*; or 2) the unconditional forecast (horizontal dashed pink line). The vertical dashed black line indicates the percentile at which the MSE is minimum given that the size premium in the high market price of risk states is significant. The vertical dashed gray line indicates the maximum percentile at which both the size and the value premiums are significant in the high market price of risk states (if different from the previous point). The market price of risk proxies are the BM of the Dow Jones stocks, DJBM, (the negative of) the net equity expansion, (NTIS), and the factor corresponding to their first principal component, PC1.

Table I Results of Principal Component Analysis for monthly values of the DJBM and the NTIS from December 1926 to December 2016. The variables are the BM of the Dow Jones stocks, DJBM, and the negative of the net equity expansion, NTIS. The analysis uses the sample correlation matrix.

PCA – Eigenvectors

Variable	Comp. 1	Comp. 2
NTIS	0.71	0.71
DJBM	0.71	-0.71

PCA – Eigenvalues

Component	Eigenvalue	Difference	Proportion
Comp. 1	1.03	0.07	0.52
Comp. 2	0.97		0.48

Table II Monthly predictive regressions of the market premium based on the two principal components of the DJBM and the NTIS, PC1 and PC2, between July 1926 and December 2016. The regressions are of the form $R_{MP(h),t+1} = \beta_{h,1} PC1_t + \beta_{h,2} PC2_t + \varepsilon_{h,t+1}$, where $R_{MP(h),t+1}$ is the market premium either with a one-month, $MP(h = m)$, or with a one-year horizon, $MP(h = y)$, starting from month $t + 1$. The table reports each estimated coefficient $\beta_{h,i}$ corresponding to the first and second principal components in month t, $PC1_t$ or $PC2_t$.

	MP(m)	MP(y)
PC1	0.6^{***}	7.9^{***}
	(3.72)	(4.23)
PC2	0.04	0.2
	(0.23)	(0.11)
Constant	0.6^{***}	7.7^{***}
	(3.98)	(4.69)
Observations	1080	1069
Adjusted R^2	0.011	

t statistics in parentheses

Newey-West t-stat. (12 lags) for MP(y)

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table III Descriptive statistics and unit root tests for the market price of risk proxies in 1926–2016. Panel A displays the mean, standard error, first order autocorrelation, and correlations at a monthly frequency of the variables. The lower diagonal corresponds to the levels of the variables and the upper diagonal corresponds to the respective AR(1) innovations. Panel B shows the Dickey-Fuller test for unit roots, with no lags. The variables are the BM of the Dow Jones stocks, DJBM, (the negative of) the net equity expansion, (NTIS), and the factor corresponding to their first principal component, PC1.

Panel A: AR(1)

	Mean	SE	AC(1)	(NTIS)	DJBM	PC1
(NTIS)	-0.02	0.01	0.98	.	0.03	0.77
DJBM	0.56	0.14	0.98	0.08	.	0.66
PC1	-0.05	0.34	0.98	0.73	0.74	.

Panel B: Dickey-Fuller tests for unit root

			Interpolated Dickey-Fuller			
		Test statistic	1% critical	5% critical	10% critical	Obs.
PC1	Z(t)	-3.13	-3.43	-2.86	-2.57	1080
	MacKinnon approximate p-value for Z(t) = 0.025					
DJBM	Z(t)	-2.73	-3.43	-2.86	-2.57	1085
	MacKinnon approximate p-value for Z(t) = 0.069					
NTIS	Z(t)	-3.15	-3.43	-2.86	-2.57	1080
	MacKinnon approximate p-value for Z(t) = 0.023					

Table IV Monthly bivariate predictive regressions of the size and value premiums based on the value of the first principal component of the DJBM and NTIS either in the previous month or on the portfolio formation date. The sample is from July 1926 until December 2016. The "Full sample" in the left panel includes the period between July/1932 and June/1933, the 1932 portfolio formation period. The "Ex-1932 formation" subsample, on the right-hand panel, excludes this period. The panels display the estimated $\beta_{i,X}$, the slope of the predictive regressions of the $R_{i,t+1} = \alpha_i + \beta_{i,X} X_t + \varepsilon_{i,t+1},$, where $R_{i,t+1}$ is the time $t+1$ monthly return on the HML or the SMB portfolios. For $X_t = \Lambda_t$, the slope corresponds to the value of the proxy in month t. For $X_t = \Lambda_{f(t)}$, the slope corresponds to the value of the proxy on the portfolio formation date to which month t belongs. The number of observations is the number of months in each subsample.

	Full sample				Ex-1932 formation			
	SMB	SMB	HML	HML	SMB	SMB	HML	HML
Λ_t	0.37***		0.39***		0.31**		0.090	
	(3.86)		(3.77)		(3.25)		(0.93)	
$\Lambda_{f(t)}$		0.42***		0.38***		0.30**		0.096
		(4.69)		(3.87)		(3.02)		(0.94)
Constant	0.22*	0.21*	0.40***	0.38***	0.18*	0.18*	0.34***	0.33***
	(2.24)	(2.13)	(3.78)	(3.61)	(2.03)	(2.03)	(3.63)	(3.55)
Observations	1080	1074	1080	1074	1068	1062	1068	1062
Adjusted R^2	0.013	0.019	0.012	0.013	0.009	0.008	-0.000	-0.000

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table V **The size related premiums and spanning tests in 1926-2016, conditioned on the market price of risk.** In each panel, the first two columns display the mean of the premiums as the constant, α_i, in regressions of the form $R_{i,t} = \alpha_i + \varepsilon_{i,t}$, where $R_{i,t}$ is either the return on the SMB or HML portfolios depending on the panel. The next two columns display the coefficients from the spanning regressions of premium i on premium j, as in $R_{i,t} = \alpha_i + \beta_j R_{j,t} + \varepsilon_{i,t}$, where $R_{j,t}$ is the return on the HML portfolio when $R_{i,t}$ is the return on the SMB portfolio, and vice versa. The final two columns add the market premium as an explanatory variable in the spanning regressions above, as in $R_{i,t} = \alpha_i + \beta_j R_{j,t} + \beta_{MP} R_{mp,t} + \varepsilon_{i,t}$. The estimates are conditioned on the market price of risk on the portfolio formation date being low ($\Lambda_{f(t)} < \Lambda^*$) or high ($\Lambda_{f(t)} \geq \Lambda^*$) according to the first principal component between the DJBM and the NTIS, PC1. In panels A and B, the threshold is the one that minimizes the MSE of the prediction of the size premium, $\Lambda^* = \Lambda^*_{MSE}$. In Panel C, the threshold is relaxed just enough to obtain a significant value premium in the high market price of risk subsample, $\Lambda^* = \Lambda^*_{HML}$. "Observations" is the number of months in each subsample.

Panel A: The size premium and span tests conditioned on Λ^*_{MSE}

	SMB $< \Lambda^*_{MSE}$	SMB $\geq \Lambda^*_{MSE}$	SMB $< \Lambda^*_{MSE}$	SMB $\geq \Lambda^*_{MSE}$	SMB $< \Lambda^*_{MSE}$	SMB $\geq \Lambda^*_{MSE}$
HML			0.03	0.3***	0.004	0.2*
			(0.81)	(4.14)	(0.13)	(2.35)
MP					0.2***	0.1*
					(10.08)	(2.01)
Constant	0.1	1.3**	0.09	1.2**	-0.02	1.1*
	(1.07)	(2.73)	(0.95)	(2.69)	(-0.21)	(2.55)
Observations	966	108	966	108	966	108
Adjusted R^2	0.000	0.000	-0.000	0.131	0.094	0.155

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Panel B: The value premium and span tests conditioned on Λ^*_{MSE}

	HML $< \Lambda^*_{MSE}$	HML $\geq \Lambda^*_{MSE}$	HML $< \Lambda^*_{MSE}$	HML $\geq \Lambda^*_{MSE}$	HML $< \Lambda^*_{MSE}$	HML $\geq \Lambda^*_{MSE}$
SMB			0.03	0.5***	0.004	0.3*
			(0.81)	(4.14)	(0.13)	(2.35)
MP					0.04*	0.3***
					(2.07)	(5.79)
Constant	0.4***	0.4	0.4***	-0.3	0.4***	-0.3
	(4.16)	(0.59)	(4.13)	(-0.45)	(3.85)	(-0.55)
Observations	966	108	966	108	966	108
Adjusted R^2	0.000	0.000	-0.000	0.131	0.003	0.335

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Panel C: The value premium and span tests conditioned on Λ^*_{HML}

	HML $< \Lambda^*_{HML}$	HML $\geq \Lambda^*_{HML}$	HML $< \Lambda^*_{HML}$	HML $\geq \Lambda^*_{HML}$	HML $< \Lambda^*_{HML}$	HML $\geq \Lambda^*_{HML}$
SMB			-0.10**	0.5***	-0.1**	0.4***
			(-2.99)	(7.13)	(-3.11)	(4.71)
MP					0.02	0.3***
					(0.86)	(7.13)
Constant	0.3**	0.7*	0.3**	0.3	0.3**	0.1
	(3.11)	(2.11)	(3.18)	(0.97)	(3.07)	(0.52)
Observations	828	246	828	246	828	246
Adjusted R^2	0.000	0.000	0.010	0.169	0.009	0.310

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table VI Spreads between the value and the size premiums in 1926–2016 conditioned on the market price of risk. The first two rows display the mean spreads and t statistics between the value and the size premiums, based on $SP_t = R_{hml,t} - R_{smb,t}$, calculated exclusively for portfolios formed in low, $\Lambda_{f(t)} < \Lambda^*_{MSE}$, or high market price of risk states, $\Lambda_{f(t)} \geq \Lambda^*_{MSE}$. The third row displays the (combined) unconditional estimates, and the last row shows the difference in the mean spreads for portfolios formed in low and high market price of risk states, $\Delta_{SP} = Mean(SP_{<\Lambda^*_{MSE}}) - Mean(SP_{\geq\Lambda^*_{MSE}})$. The t test with unequal variances – in brackets – based on $H_0\colon \Delta_{SP} \leq 0$ and $H_a\colon \Delta_{SP} > 0$ gives the significance of the estimation (equal variances in parentheses). "Obs." is the number of months in each subsample and $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

	Mean	Obs.
$SP_{<\Lambda^*_{MSE}}$	0.29*	966
	(2.20)	
$SP_{\geq\Lambda^*_{MSE}}$	-0.88	108
	(-1.39)	
Combined	0.18	1074
	(1.29)	
Δ_{SP}	1.18*	
	(2.60)	
	[1.81]	

Appendix I. Extended theoretical framework

I start by constructing a theoretical framework in which the most restrictive assumption is the one of a time-varying market price of risk.[11] Let $\zeta = (\zeta_t)$ be the unique stochastic discount factor (SDF) that follows the continuous-time stochastic process

$$d\zeta_t = -\zeta_t[r_t^f\, dt + \lambda_t\, dz_{1t}], \tag{45}$$

where dz_{1t} is a one-dimensional standard Brownian motion and the stochastic processes, r_t^f and λ_t, are the risk free rate and the market price of risk, respectively.

Let $P_i = (P_{it})$ be the price process of portfolio i, such that

$$dP_{it} = P_{it}[\mu_{it}\, dt + \sigma_{it}\, dz_{1t} + \tilde{\boldsymbol{\sigma}}_{it}^\top\, d\boldsymbol{z}_t], \tag{46}$$

where μ_{it} and σ_{it} are one-dimensional stochastic processes, $d\boldsymbol{z}_t$ is a multi-dimensional standard Brownian motion independent of dz_{1t}, $\tilde{\boldsymbol{\sigma}}_{it}$ is a multi-dimensional stochastic process, and $^\top$ is the transposition sign. μ_{it} and σ_{it} represent, respectively, the expected returns and the sensitivity of the returns on the portfolio to the exogenous (priced) shocks to the economy. So σ_{it} gives the effective risk of the portfolio. $\tilde{\boldsymbol{\sigma}}_{it}$ represents the unpriced return volatility of the portfolio (the sensitivity of the returns on the portfolio to the exogenous unpriced shocks). Without intermediate dividends, the expected excess rate of return on the portfolio is

$$\mu_{it} - r_t^f = \sigma_{it}\lambda_t. \tag{47}$$

So the overall market price of risk, λ_t, and the risk of the portfolio, σ_{it}, **combined**

[11] There is a myriad of models based on different assumptions that generate time variation in the market price of risk. Examples of these models include Rietz (1988), Epstein and Zin (1989), Constantinides and Duffie (1996), Campbell and Cochrane (1999), Hansen and Sargent (2001), Bansal and Yaron (2004), Barro (2006), Piazzesi et al. (2007), Brunnermeier (2009), Bansal et al. (2012), Shiller (2014), and Garleanu and Panageas (2015).

determine the time t expected return on the portfolio.

Pricing a future uncertain cash flow: For tractability, assume that the risk-free rate and the market price of risk are constant between time t and time T. So the SDF follows a one-dimensional geometric Brownian motion process,

$$d\zeta_t = -\zeta_t[r^f\, dt + \overline{\lambda}_t\, dz_t], \tag{48}$$

where (the now constants) $\overline{\lambda}_t$ and r^f are, respectively, the market price of risk and the risk free rate prevailing from time t to time T.

As in Eq. (46), let $P_i = (P_{it})$ be the price process of the equity in the firm i. The firm makes only a final uncertain lump sum dividend payment, $D_{i,T}$, at time T, which is modeled through $x_{i,t} = \mathrm{E}_t[D_{i,T}]$ following the process

$$dx_{i,t} = x_{i,t}\left[0\, dt + \overline{\sigma}_{it}\, dz_{1t} + \tilde{\boldsymbol{\sigma}}_{\overline{it}}^{\top}\, d\boldsymbol{z}_t\right], \tag{49}$$

where $\overline{\sigma}_{it}$ is a (one-dimensional) constant and $\tilde{\boldsymbol{\sigma}}_{\overline{it}}$ is a vector of constants. Under these conditions, the time t value of the payoff, $D_{i,T}$, is

$$P_{it} = \mathrm{E}_t\left[D_{i,T}\frac{\zeta_T}{\zeta_t}\right] = \mathrm{E}_t[D_{i,T}]e^{-(r^f + \overline{\sigma}_{it}\overline{\lambda}_t)(T-t)}, \tag{50}$$

which depends positively on the expected payoff, $\mathrm{E}_t[D_{i,T}]$, and negatively on the risk of the payoff, $\overline{\sigma}_{it}$, and on the market price of risk, $\overline{\lambda}_t$, (apart from the risk free rate, r^f, and the time interval, $T - t$).

The SMB portfolio's risk: Eq. (47) and Eq. (50) taken together summarize the static framework in Berk (1995): Given two firms with the same expected cash flows, the riskiest one (with the largest $\overline{\sigma}_{it}$) has the lowest market value, P_{it}, given by Eq. (50) and the highest market premium, given by Eq. (47). This connects the price ranking, based on P_{it}, to the risk ranking, based on $\overline{\sigma}_{it}$. However, the price ranking does not reflect the risk ranking perfectly because the prices also depend on the expected cash flows, $\mathrm{E}_t[D_{i,T}]$ in Eq. (50).

The role of the market price of risk: The SMB portfolio is long on the stocks with ME below its median and short on the others. In terms of Eq. (47), we have that

$$\sigma_{smb,t} = \sigma_{small,t} - \sigma_{big,t}, \tag{51}$$

$$\mu_{smb,t} = \sigma_{smb,t}\lambda_t, \tag{52}$$

where $\sigma_{small,t}$ and $\sigma_{big,t}$ are, respectively, the risks of the portfolios of small stocks and big stocks, corresponding to σ_{it} in Eq. (46), and $\mu_{smb,t}$ is the time t expected return on the SMB portfolio. Given a positive market price of risk, the expected size premium can only be positive when the small stocks are riskier than the big stocks, so that $\sigma_{small,t} - \sigma_{big,t} > 0$.

For example, if the market price of risk is zero then the expected cash flows of all firms are discounted at the risk free rate in Eq. (50). The ME ranking is exactly the same as the expected cash flows' ranking. Under the assumption that there is no relation between the expected cash flow amount and its risk (as suggested by Berk, 1996),

$$\overline{\sigma}_{it} \perp\!\!\!\perp E[CF_{iT}], \tag{53}$$

the distribution of risks among the two portfolios formed by size in this market price of risk state is equal. So the size premium is zero.

As the market price of risk increases, the expected cash flows of all firms are discounted at a higher rate, and all prices fall. However, the prices of the riskier stocks fall relatively more, and they become over-represented in the small stocks group. This generates the size premium **in these states only**. Hence, the first rejectable prediction of the model is:

Hypothesis 1 (H1): *The size premium is positive if and only if the market price of risk is above a certain threshold, $\lambda_t \geq \lambda^*$, being zero otherwise.*

In terms of Eq. (51) and Eq. (52), we can express this hypothesis as

$$\sigma_{smb,t} = \begin{cases} 0 & \lambda_t < \lambda^* \\ f(\lambda_t) > 0 & \lambda_t \geq \lambda^* \end{cases} \implies \mu_{smb,t} = \begin{cases} 0 & \lambda_t < \lambda^* \\ f(\lambda_t)\lambda_t > 0 & \lambda_t \geq \lambda^* \end{cases} \tag{54}$$

where $f(\lambda_t)$ is a non decreasing function of the market price of risk, λ_t.

Appendix I.A. The determinants of the market price of risk threshold, λ^*

Consider a pair of stocks in which the cash flows of firm r are riskier, with volatility term $\overline{\sigma}_{rt}$, and the cash flows of firm s are less risky, with volatility term $\overline{\sigma}_{st}$, such that

$$\overline{\sigma}_{rt} > \overline{\sigma}_{st}. \tag{55}$$

The price and the risk rankings align for these stocks if and only if the price of the riskier stock is smaller than the price of the safer stock, $P_{r,t} < P_{s,t}$. In terms of the parameters in Eq. (50), we obtain

$$P_{r,t} < P_{s,t} \iff \overline{\lambda}_t > \ln\left(\frac{\mathrm{E}_t[D_{r,T}]}{\mathrm{E}_t[D_{s,T}]}\right) \frac{1}{(\overline{\sigma}_{rt} - \overline{\sigma}_{st})\,(T-t)} \equiv \lambda^*_{rs}, \tag{56}$$

where λ^*_{rs} is the market price of risk threshold specific to this pair of stocks.

I will loosely use the term "statistically aligned" for a market price of risk threshold, λ^*, that is large enough so that Eq. (56) holds for several pairs of stocks, (r, s), and the size premium is statistically significant. This is not a strict definition because λ^* changes with the sample size and other characteristics of the stocks.[12] But in general, λ^* is smaller and the size ranking is more likely to be statistically aligned with the risk ranking when:

1. The cash flows are received at a more distant point into the future, so they have higher "duration", corresponding to larger $(T-t)$.
2. There is more dispersion in the distribution of risks among the firms, which corresponds to larger risk differences, $\overline{\sigma}_{rt} - \overline{\sigma}_{st}$.
3. There is less dispersion in the expected cash flows among the firms, which corresponds to lower cross-sectional variance of the (log of the) expected cash flows among the firms, $\mathrm{Var}(\ln \mathrm{E}_t[D_{i,T}])$, so the ratio of expected cash flows among two random firms is close to one.

Appendix I.B. The HML portolio's risk (and other scaled-price rankings)

Consider the same pair of stocks, r and s, but now ranked by a scaled-price ratio. Without loss of generality, condider the BE as the scaling variable. This creates the price-to-book (PB) ratio that generates the HML portfolio.[13] The value and the risk rankings are aligned if and only if the PB of the risky stock is smaller than the PB of the safe stock,

$$\frac{P_{r,t}}{B_{r,t}} < \frac{P_{s,t}}{B_{s,t}}, \tag{57}$$

where $B_{i,t}$ (the time t BE of firm i in this case) divides the market value calculated in Eq. (50) for each stock. In terms of the parameters in Eq. (56),

$$\frac{P_{r,t}}{B_{r,t}} < \frac{P_{s,t}}{B_{s,t}} \Longleftrightarrow \overline{\lambda}_t > \ln\left(\frac{\mathrm{E}_t[D_{r,T}]}{\mathrm{E}_t[D_{s,T}]}\frac{B_{s,t}}{B_{r,t}}\right)\frac{1}{(\overline{\sigma}_{rt} - \overline{\sigma}_{st})(T-t)} \equiv \lambda^*_{PB,rs}, \tag{58}$$

[12]Indeed, this only makes sense in finite samples. Asymptotically, if the time series is infinite, the size premium is significant if there is any (persistent) imbalance between the risks of small and big stocks

[13]The conclusions are similar for other scaling variables, such as dividends, earnings, or cash flows. This is expected because the value premium spans several other premiums associated with scaled-price ratios, as shown in Fama and French (1996).

which is smaller than the threshold necessary to align the price and the risk rankings, $\lambda^*_{PB,rs} < \lambda^*_{rs}$, in two cases:

1. If the BE of the risky stocks is larger than the BE of the safer stocks, $B_r > B_s$, the term inside the log term is closer to 1 and the threshold, $\lambda^*_{PB,rs}$, is closer to zero.
2. If the BE is a proxy for expected cash flows, for example, in the limit with $B_{i,t} = \mathrm{E}_t[D_{i,T}]$, the term inside the log term is 1 so that $\overline{\lambda}_t > \lambda^*_{PB,rs} = 0$.

I model the BE as a proxy for expected cash flows, which is both in line with the evidence in Ball et al. (2017) and with the theoretical framework in Berk (1995). On the other hand, alternative 1 implies that the BE characteristic is positively related to the risk of the stocks. It also implies that the BE ranking aligns with the risk ranking (unconditionally): There should be a “BE premium” (similar to the size premium), which we do not observe in the data.

The BE as a proxy for expected cash flows: Let us assume that the BE is a good proxy for the expected cash flows of at least one pair of firms, $i = \{r, s\}$, in the universe of all firms, I:

$$\exists\, i \in I \mid \mathrm{E}_t[D_{i,T}] \approx BE_{i,t} \equiv B_{i,t}. \tag{59}$$

If the risker firm in this pair has larger expected cash flows, there is a range for the market price of risk in which the value ranking aligns with the risk ranking while the price ranking does not,

$$\exists\, (r, s) \in I \mid \lambda^*_{rs} > \overline{\lambda}_t > \lambda^*_{PB,rs}. \tag{60}$$

Depending on how well the BE proxies for the expected cash flows, the threshold in Eq. (58) can even be zero, $\lambda^*_{PB,rs} = 0$.

If Eq. (59) holds for several pairs of stocks, then a large portfolio of value stocks (with low PB) should contain a disproportional number of risky stocks, so $\sigma_{value,t} > \sigma_{growth,t}$

equivalent to Eq. (51), even if the prevailing market price of risk is lower than the one necessary to generate the size premium. The equivalent to Eq. (54) (with $\lambda^*_{PB} < \lambda^*$) is

$$\sigma_{hml,t} = \begin{cases} f(\lambda_t) > 0 & \lambda_t < \lambda^* \\ f(\lambda_t) > 0 & \lambda_t \geq \lambda^* \end{cases} \implies \mu_{hml,t} = \begin{cases} f(\lambda_t)\lambda_t > 0 & \lambda_t < \lambda^* \\ f(\lambda_t)\lambda_t > 0 & \lambda_t \geq \lambda^* \end{cases}, \tag{61}$$

which is the mathematical representation of the second testable prediction of the model:

Hypothesis 2 (H2): *Given that the BE is a proxy for expected cash flows: The value premium increases with the market price of risk, being small, but still positive, even if the market price of risk is lower than the threshold below which the size premium is zero,* $\lambda_t \leq \lambda^*$.

The relative magnitudes of the size and the value premiums: On the other hand, scaling by the BE can also work in the direction of randomizing the ranking and breaking the alignment with the risk ranking. Intuitively, any adjustment remotely related to the expected cash flows helps if the market price of risk is low because the price ranking is mostly unrelated to risk in this case. But a less than perfect adjustment to the price ranking should disturb the alignment with the risk ranking in case this alignment was already good. The proportion of risky firms in the portfolio of small stocks is high if the market price of risk is large. Hence, reshuffling the ranking by some variable that is only an imperfect proxy for expected cash flows is likely to make the distribution of risk more uniform among the (value) portfolios and reduce the (value) premium.

Appendix I.C. The PB characteristic could be related to a risk orthogonal to size

In a multidimensional setting, the SDF in Eq. (45) becomes

$$d\zeta_t = -\zeta_t[r_t^f \, dt + \boldsymbol{\lambda}_t^\top \, d\boldsymbol{z}_{1t}], \tag{62}$$

where $d\boldsymbol{z}_{1t}$ is a multi-dimensional standard Brownian motion and $\boldsymbol{\lambda}_t$ is the multi-dimensional stochastic process representing the time t market price of risk associated with each risk source. We also adjust the price process in Eq. (46) to

$$dP_{it} = P_{it}[\mu_{it} \, dt + \boldsymbol{\sigma}_{it}^\top, d\boldsymbol{z}_{1t} + \tilde{\boldsymbol{\sigma}}_{it}^\top \, d\boldsymbol{z}_t], \tag{63}$$

where $\boldsymbol{\sigma}_{it}$ is now a multi-dimensional stochastic process, and $d\boldsymbol{z}_t$ is independent of $d\boldsymbol{z}_{1t}$. The multi-dimensional equivalent to the expected excess rate of return in Eq. (47) is now

$$\mu_{it} - r_t^f = \boldsymbol{\sigma}_{it}^\top \boldsymbol{\lambda}_t, \tag{64}$$

and the equivalents of Eq. (52) are

$$\mu_{smb,t} = \boldsymbol{\sigma}_{smb,t}^\top \boldsymbol{\lambda}_t, \tag{65}$$

$$\mu_{hml,t} = \boldsymbol{\sigma}_{hml,t}^\top \boldsymbol{\lambda}_t, \tag{66}$$

where $\boldsymbol{\lambda}_t$ is the multi-dimensional market price of risk process, and $\boldsymbol{\sigma}_{smb,t}$ and $\boldsymbol{\sigma}_{hml,t}$ are the multi-dimensional stochastic sensitivities of the returns on the SMB and on the HML portfolios to the (independent) exogenous priced shocks to the economy, respectively.

The main difference from the one-dimensional formulation is that with a single risk source the premiums on small and value stocks (and indeed all risk premiums) must be perfectly correlated given that there is only one priced risk in the economy.

Let the k^{th} element in $\boldsymbol{\sigma}_{smb,t}$ be equal to zero while being different from zero in $\boldsymbol{\sigma}_{hml,t}$. This means that the PB ranking captures this risk, but the price ranking does not. Consequently, it is impossible to find a constant, a, that multiplied by the risk of the SMB portfolio, $\boldsymbol{\sigma}_{smb,t}$, gives the risk of the HML portfolio, $\boldsymbol{\sigma}_{hml,t}$:

$$\nexists\, a \in \mathbb{R} \mid \boldsymbol{\sigma}_{hml,t} = a\boldsymbol{\sigma}_{smb,t}. \tag{67}$$

And therefore, it is also (usually) impossible to represent the expected return on the HML portfolio as a multiple of the expected return on the SMB portfolio:[14]

$$\nexists\, a \in \mathbb{R} \mid \mu_{hml,t} = a\boldsymbol{\sigma}_{smb,t}^{\top}\boldsymbol{\lambda}_t = a\mu_{smb,t}. \tag{68}$$

If the PB is really related to this particular risk, then Eq. (68) should always hold. In particular, it should hold regardless of the prevailing market price of risk when the (PB) portfolios are formed, $\lambda_{f,t}$.

The alternative is that the BE is a proxy for expected cash flows. In this case, Eq. (68) still holds when the market price of risk is low because the size and the risk rankings are not aligned in these states anyway, as given by Eq. (54). However, the two portfolios reflect similar risks when the market price of risk is high, $\lambda_{f,t} > \lambda^*$, and therefore the constant, a, in Eq. (54) could exist.[15] Testing this hypothesis is equivalent to testing whether such a constant, a, exists depending on the market price of risk on the portfolio formation date, $\lambda_{f,t}$:

Hypothesis 3 (H3): *Given that the BE is a proxy for expected cash flows, and the PB characteristic is* ***not*** *related to risks that are orthogonal to size: The value premium can*

[14] Except for very particular combinations of $\boldsymbol{\sigma}_{smb,t}$ and $\boldsymbol{\lambda}_t$.

[15] In fact, the constant a only exists in general if the size and the value rankings have the same relative exposures to the different priced shocks to the economy. So this is a restrictive test.

be a multiple of (be spanned by) the size premium, $\mu_{hml,t} = a\mu_{smb,t}$, if and only if the market price of risk is high, $\lambda_t > \lambda^*$.

REFERENCES

Ball, Ray, Joseph J. Gerakos, Juhani T. Linnainmaa, and Valeri V. Nikolaev, 2017, Earnings, retained earnings, and book-to-market in the cross section of expected returns, *Unpublished manuscript, SSRN* .

Bansal, Ravi, Dana Kiku, and Amir Yaron, 2012, An empirical evaluation of the long-run risks model for asset prices, *Critical Finance Review* 1, 183–221.

Bansal, Ravi, and Amir Yaron, 2004, Risks for the long run: A potential resolution of asset pricing puzzles, *Journal of Finance* 59, 1481–1509.

Barro, Robert J., 2006, Rare disasters and asset markets in the twentieth century, *The Quarterly Journal of Economics* 121, 823–866.

Berk, Jonathan, 1996, An empirical re-examination of the relation between firm size and return, *Unpublished manuscript, University of British Columbia* .

Berk, Jonathan B., 1995, A critique of size-related anomalies, *Review of Financial Studies* 8, 275–286.

Boudoukh, Jacob, Roni Michaely, Matthew Richardson, and Michael R. Roberts, 2007, On the importance of measuring payout yield: Implications for empirical asset pricing, *Journal of Finance* 62, 877–915.

Brunnermeier, Markus K., 2009, Deciphering the liquidity and credit crunch 2007-2008, *Journal of Economic Perspectives* 23, 77–100.

Campbell, John Y., and John H. Cochrane, 1999, By force of habit: A consumption-based explanation of aggregate stock market behavior, *Journal of Political Economy* 107, 205–251.

Campbell, John Y., Jens Hilscher, and Jan Szilagyi, 2008, In search of distress risk, *Journal of Finance* 63, 2899–2939.

Chava, Sudheer, and Amiyatosh Purnanandam, 2010, Is default risk negatively related to stock returns?, *Review of Financial Studies* 23, 2523–2559.

Chordia, Tarun, Amit Goyal, and Jay A. Shanken, 2015, Cross-sectional asset pricing with individual stocks: betas versus characteristics, *Unpublished working paper, available at SSRN 2549578* .

Constantinides, George M., and Darrell Duffie, 1996, Asset pricing with heterogeneous consumers, *Journal of Political Economy* 104, 219–240.

Daniel, Kent, and Sheridan Titman, 1997, Evidence on the characteristics of cross sectional variation in stock returns, *Journal of Finance* 52, 1–33.

Davis, James L., Eugene F. Fama, and Kenneth R. French, 2000, Characteristics, covariances, and average returns: 1929 to 1997, *Journal of Finance* 55, 389–406.

Dittmar, Robert F., and Christian T. Lundblad, 2017, Firm characteristics, consumption risk, and firm-level risk exposures, *Journal of Financial Economics* 125, 326–343.

Epstein, Larry G., and Stanley E. Zin, 1989, Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework, *Econometrica* 57, 937–969.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55–84.

Fama, Eugene F., and Kenneth R. French, 1997, Industry costs of equity, *Journal of Financial Economics* 43, 153–193.

Favilukis, Jack, and Xiaoji Lin, 2016, Wage rigidity: A quantitative solution to several asset pricing puzzles, *Review of Financial Studies* 29, 148–192.

Garleanu, Nicolae, and Stavros Panageas, 2015, Young, old, conservative, and bold: The implications of heterogeneity and finite lives for asset pricing, *Journal of Political Economy* 123, 670–685.

Gompers, Paul A., and Andrew Metrick, 2001, Institutional investors and equity prices, *The Quarterly Journal of Economics* 116, 229–259.

Green, Jeremiah, John R. M. Hand, and Frank Zhang, 2017, The characteristics that provide independent information about average U.S. monthly stock returns, *Review of Financial Studies,* forthcoming.

Griffin, John M., and Michael L. Lemmon, 2002, Book-to-market equity, distress risk, and stock returns, *Journal of Finance* 57, 2317–2336.

Hansen, Lars Peter, and Thomas J. Sargent, 2001, Robust control and model uncertainty, *American Economic Review* 91, 60–66.

Kapadia, Nishad, 2011, Tracking down distress risk, *Journal of Financial Economics* 102, 167–182.

Kogan, Leonid, and Dimitris Papanikolaou, 2013, Firm characteristics and stock returns: The role of investment-specific shocks, *The Review of Financial Studies* 26, 2718–2759.

Lee, Charles M. C., Andrei Shleifer, and Richard H. Thaler, 1991, Investor sentiment and the closed-end fund puzzle, *Journal of Finance* 46, 75–109.

Piazzesi, Monika, Martin Schneider, and Selale Tuzel, 2007, Housing, consumption and asset pricing, *Journal of Financial Economics* 83, 531–569.

Piotroski, Joseph D., and Eric C. So, 2012, Identifying expectation errors in value/glamour strategies: A fundamental analysis approach, *The Review of Financial Studies* 25, 2841–2875.

Pontiff, Jeffrey, and Lawrence D. Schall, 1998, Book-to-market ratios as predictors of market returns, *Journal of Financial Economics* 49, 141–160.

Rietz, Thomas A., 1988, The equity risk premium a solution, *Journal of Monetary Economics* 22, 117–131.

Roll, Richard, 1977, A critique of the asset pricing theory's tests Part I: On past and potential testability of the theory, *Journal of Financial Economics* 4, 129–176.

Shiller, Robert J., 2014, Speculative asset prices, *American Economic Review* 104, 1486–1517.

Souza, Thiago de O., 2017, The size premium is not a puzzle, but its predictability is, *Unpublished working paper. University of Southern Denmark* .

Welch, Ivo, and Amit Goyal, 2008, A comprehensive look at the empirical performance of equity premium prediction, *Review of Financial Studies* 21, 1455–1508.

Yogo, Motohiro, 2006, A consumption-based explanation of expected stock returns, *Journal of Finance* 61, 539–580.

Zhang, Lu, 2005, The value premium, *Journal of Finance* 60, 67–103.