SDU❧

A critique of momentum anomalies

by

Thiago de Oliveira Souza

Discussion Papers on Business and Economics
No. 5/2019

A critique of momentum anomalies

Thiago de Oliveira Souza

University of Southern Denmark

and Danish Finance Institute

February 15, 2019

Abstract

This paper offers theoretical, empirical, and simulated evidence that momentum regularities in asset prices are not anomalies. Within a general, frictionless, rational expectations, risk-based asset pricing framework, riskier assets tend to be in the loser portfolios after (large) increases in the price of risk. Hence, the risk of momentum portfolios usually decreases with the prevailing price of risk, and their risk premiums are approximately negative quadratic functions of the price of risk (and the market premium) theoretically truncated at zero. The best linear (CAPM) function describing this relation unconditionally has exactly the negative slope and positive intercept documented empirically.

1   Introduction

Price momentum is a tendency of previous winners to outperform previous losers in the future (Jegadeesh and Titman, 1993) and it is documented for several assets classes and markets (Asness et al., 2013). A full explanation of the phenomenon appears so difficult within the traditional risk-based asset pricing paradigm that "asset pricing theorists have mostly seen the task as simply one of deciding which sort of investor irrationality is at work" (Johnson, 2002). Indeed, recent efforts even seem to have shifted from attempting to explain momentum to just offering directions on how to avoid its "crashes", as in Daniel and Moskowitz (2016) or Barroso and Santa-Clara (2015), for example.

In this paper I argue that, rather than anomalies, momentum regularities are consistent with traditional risk-based asset pricing models under very general conditions. I theoretically explain *why* momentum is associated with risk, based on arguments similar to the ones in Berk (1995) and Souza (2018): The size (and momentum) premium exists because market values (and realized returns) are related to risks (Berk, 1995), and the relation changes with the price of risk (Souza, 2018).

Under the assumptions of the CAPM of Sharpe (1964) and Lintner (1965), for example, high beta stocks have higher average returns than low beta stocks. Hence, high risk/beta stocks are both past and future winners on average. This point is made by Conrad and Kaul (1998) to explain why momentum portfolios, which buy winners and sell losers, have positive average returns. However, the negative betas of momentum portfolios seem to contradict this explanation and can be regarded as the true "momentum puzzle". The key to this apparent puzzle is to understand how a stochastic price of risk, as documented in, for example, Cochrane (2011), affects the risk of momentum portfolios.

First, consider a series containing the current and expected price of risk in every future period. The expectation is fully rational, based on perfect information and shared by every investor. Within a standard risk-based asset pricing framework, this series, together with the risk of the asset, determines the price (present value) of any stream of cash flows, all of which are also common knowledge among the investors. As I explain in details in Section

2, the **current** price of risk and two parameters are sufficient statistics for this entire series. One parameter indicates the persistence, and the other is the unconditional mean of the price of risk process. Under the assumption of a stable process, the parameters are constant over time, and the whole series changes only in proportion to the stochastic price of risk realized in each period.

Next, consider the "formation" period over which the realized returns are calculated to split the assets into winners and losers and create the momentum portfolios. For simplicity, each asset only pays a final cash flow in a distant the future. If the new realized price of risk is larger than the previous one, for example, the present values of all cash flows decrease compared to the prices at the beginning of the formation period: All prices fall in proportion to their risks after reflecting also possible news about expected cash flows. These cash flow shocks generate the "noise" in returns mentioned in the simplified analysis that follows.

In the first scenario, the price of risk decreases from the beginning to the end of the formation period. In this situation, the vast majority of winners are riskier than the losers because two different effects bring risky asset returns up: (i) Riskier assets have larger risk premiums ex-ante, and (ii) riskier assets benefit the most from being "repriced" using a lower price of risk series ex-post. The momentum portfolios formed under these circumstances have very large and positive risk exposures because a very large fraction of the riskier assets is in the winner (long) portfolio, even after the cash flow shocks.

In the second scenario, the price of risk realized at the end of the formation period is exactly the same as the one at the beginning of the period. In this case, winners still tend to be riskier, but only due to the first effect above (larger premiums ex-ante). The realized return ranking is less efficient in differentiating risky and safe assets in this case because the noise-to-risk ratio in returns is larger than in the first scenario, implying a relatively lower fraction of risky assets among the winners. As a consequence, the momentum portfolios formed in these states have positive but smaller risk.

Finally, in the third scenario, the price of risk realized at the end of the formation period is (much) larger than the one at the beginning of the period. In this case, winners tend to be safer (than losers), because the riskiest assets realize the most negative returns due to being repriced using a larger price of risk series ex-post. This counterbalances their higher ex-ante risk premiums at least to some extent. Hence, momentum has even lower (and negative) risk after (large) increases in the price of risk.

In summary, the risk of momentum portfolios decreases (eventually becoming negative) with the magnitude of the increases in the price of risk during the portfolio formation period. This implies that the risk of momentum portfolios is also negatively related to the price of risk **level** at the portfolio formation dates, which is essential to understand their CAPM properties.[1]

As an approximation, consider a linear relation between the risk of momentum portfolios and the price of risk level, as in Fig. 1a. The momentum premium – momentum risk multiplied by the price of risk – is a negative quadratic function of the price of risk in this case, as in Fig. 1c. For a constant market portfolio risk exposure – implying the market premium in Fig. 1b (also for simplicity) – momentum is a negative quadratic function of the market premium as well. But the market premium is theoretically positive, and therefore the domain of the function describing momentum in terms of the market premium is the positive real numbers. Visually, this function is a parabola opening downwards with a short left tail and a long right tail, as the one displayed in Fig. 1d. The best linear approximation of this function in this domain is, indeed, one with a negative slope and positive intercept. This explains the unconditional negative CAPM beta and positive intercept of momentum portfolios.

The approximation in Fig. 1d also explains the main results in Daniel and Moskowitz (2016) and Barroso and Santa-Clara (2015). It explains (i) why momentum "crashes" exactly when the market rises abruptly: Negative momentum returns correspond to the largest market premiums in the graph; (ii) why the particular momentum portfolios that

---

[1]This happens because increases in the price of risk are associated with larger prices of risk. For example, within a given stationary series, a current value above the population mean is more likely to have increased to that value from a random past value than to have decreased from an even higher value.

crash have very negative betas: These portfolios correspond to the right portion of the graph with the steepest slope; (iii) why the portfolios formed in the remaining periods do not have positive betas either: These are the portfolios in the top, symmetric section of the graph with zero slope; and (iv) why momentum returns are negatively skewed: Positive returns reach a maximum, but negative returns are unbounded. I illustrate each of these points in the Monte Carlo simulation and confirm them with real data in the empirical section.

The framework also provides theoretical guidance to avoid the downside risk of momentum, as in Daniel and Moskowitz (2016) or Barroso and Santa-Clara (2015), for example. Theoretically, momentum portfolios should only have negative risks (and negative expected returns) when the price of risk increases during the portfolio formation period. Thus, it is possible to avoid these portfolios given a price of risk proxy. Indeed, Fig. 2 shows how this strategy, based on five different price of risk proxies, performs compared to (i) the opposite strategy or to (ii) investing in momentum unconditionally. The five price of risk proxies, which I use later in the empirical section, are: The Dow Jones' book-to-market (DJBM), the net equity expansion (NTIS), the investment-to-capital ratio (IK), the consumption-wealth ratio (CAY), and the lower bound for the equity premium (SVIX).[2] All proxies (except the NTIS) give results highly consistent with the theory, especially the SVIX.

The results in Fig. 12 are particularly revealing: For the period in which the SVIX is available (1997-2012), for example, there are four months with absolute momentum returns above 10%. All of them are negative, in line with a negatively skewed distribution. And all of them happen after observed increases in this price of risk proxy. The other graphs in Fig. 12 display very similar patterns based on the other proxies (and sample periods), while Fig. 11 shows that the same patterns arise in simulations of the theoretical model. This not only confirms that the strategy avoids all momentum "crashes", but it also confirms that it is, indeed, the right tail of the parabola, below the horizontal axis in Fig. 1c, that

---

[2]These are the only four variables found by Welch and Goyal (2008) (updated) to significantly predict the one- and five-year equity premiums in sample within the period 1927-2013, on top of the more recently documented SVIX (Martin, 2017). The NTIS is the only variable that seems inappropriate as a price of risk proxy in the exercises. It could be that the NTIS forecasts cash flows instead of being related to discount rates. In this case, the variable is not appropriate as a price of risk proxy, although it forecasts the equity premium.

generates the negative skewness in the momentum returns distribution, as the theory in the paper predicts.

1.1 Related literature

This paper provides a theoretical explanation for momentum regularities in asset returns, so it is mainly related to the theoretical literature on momentum. However, it also significantly improves the empirical description of momentum and it relates to this literature as well. Jegadeesh and Titman (2011) provide an extensive review of theories of momentum based on irrational agents, including the ones based on frictions, and summarize part of the theoretical literature. Another strand consists of models that empirically "explain" momentum based on its multiple covariances with a number of factors. These factor models are not directly obtained from equilibrium conditions and only address stocks. Thus, they are relatively silent about the economic channel behind the momentum premium and its risks (especially its negative market risk exposure). Examples of empirically driven factors that seem to span momentum in stock returns are given in Carhart (1997) or Hou et al. (2015).[3] In addition, Avramov and Chordia (2006) and Liu and Zhang (2008) provide examples of macroeconomic factors that seem to covary with momentum, although Griffin et al. (2003) find no such evidence in international markets.

The closest literature to the present paper is the one on rational economic theories of momentum. So far, the literature has focused only on stocks and on explaining the positive average return on momentum portfolios. The models are silent about their negative CAPM betas and most other features of the return distribution, for example. These theories are fundamentally grounded on the same idea: Realized returns are a measure of expected returns. Examples of these models are Conrad and Kaul (1998), Berk et al. (1999), and Johnson (2002). Conrad and Kaul (1998) assume constant differences in expected returns among the stocks over time. This is reflected by realized returns (in the portfolio formation period) and continues in the evaluation period. In Berk et al. (1999), asset turnover

---

[3] In fact, Novy-Marx (2015) argues that it is the post earnings announcement drift that prices momentum portfolios and not the ROE factor, which is claimed to be related to q-theory by Hou et al. (2015).

generates variation in expected returns under the assumption that the firm's risk changes over time as the firm takes on different projects. In Johnson (2002), the variation exists because the firms receive shocks to their growth rates, with the extra assumptions that there is a systematic and persistent component to growth rate shocks, and that exposure to this component is associated with positive expected returns.

Regarding the empirical literature, the closest examples are Daniel and Moskowitz (2016), Barroso and Santa-Clara (2015), Barroso (2014), Grundy and Martin (2001), and Kothari and Shanken (1992), who all mention time variation in the risk of momentum portfolios.

The explanation that I provide stands out from the previous literature in several dimensions. First is the lack of restrictive assumptions: There are none regarding irrationality, frictions, asymmetric shocks to cash flows nor the risk of growth options versus assets in place, for example. The explanation is also simple and contains close to no "dark matter" in the words of Chen et al. (2015). Second is completeness: The model clearly identifies the economic channel generating several crucial properties of momentum at once. It explains the negative CAPM beta for stock returns (the key ingredient of the puzzle, yet unexplained by previous theories), negative skewness, the relation between crashes and the market return, conditional betas, etc. Third, and related to the second, this single theory explains momentum in the returns on any financial asset, not only stocks. Finally, the detailed empirical description of conditional momentum and its forecastability is supported by a theoretical background that is absent in previous empirical work.

2 Theory

Let $\zeta = (\zeta_t)$ be the unique stochastic discount factor (SDF) that follows the continuous-time stochastic process

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

where dz_t is a one-dimensional standard Brownian motion, r_f is the constant risk-free rate, and the stochastic process λ_t is the price of risk. Let asset i, with price $P_i = (P_{i,t})$, pay a unique and final cash flow, $D_{i,T}$, at time T, and let its expectation, given by $x_t = E_t[D_{i,T}]$, follow the process

$$dx_t = x_t[0 \, dt + \sigma_{i,t} \, dz_t + \tilde{\boldsymbol{\sigma}}_{i,t}^{\top} \, d\tilde{\mathbf{z}}_t], \tag{2}$$

where $\sigma_{i,t}$ is a one-dimensional stochastic process representing the effective risk of the asset, $d\tilde{\mathbf{z}}_t$ is a multi-dimensional standard Brownian motion independent of dz_t, $\tilde{\boldsymbol{\sigma}}_{i,t}$ is a multi-dimensional stochastic process representing unpriced volatility, and $^{\top}$ is the transposition sign. Without intermediate dividends, the expected excess rate of return on this asset is

$$\mu_{i,t} - r_f = \sigma_{i,t}\lambda_t, \tag{3}$$

which highlights that the premium depends on the risk of the asset, $\sigma_{i,t}$, and also on the overall market price of risk, λ_t.

With constant volatility, σ_i, and in discrete time, the price of this asset at time t is

$$P_{i,t} = E_t[D_{i,T}]e^{-r_f(T-t)}e^{-\sigma_i\lambda_t}e^{-\sigma_i E_t\left[\sum_{i=t+1}^{T}\lambda_i\right]}, \tag{4}$$

where the first term multiplying the expected cash flow corresponds to discounting at the risk-free rate, the second adds the risk premium required in the current period (with the price of risk at time t, λ_t), and the last one adds the expected risk premium required in all future periods.

The log return on this asset from time t to $t+1$ is, therefore,

$$r_{i,t+1} = \ln\left(\frac{P_{i,t+1}}{P_{i,t}}\right) = r_f + \sigma_i\left(\lambda_t + E_t\left[\sum_{i=t+1}^{T}\lambda_i\right] - E_{t+1}\left[\sum_{i=t+1}^{T}\lambda_i\right]\right) + \varepsilon_{i,t+1}, \qquad (5)$$

where $\varepsilon_{i,t+1} \equiv \ln\left(\frac{E_{t+1}[D_{i,T}]}{E_t[D_{i,T}]}\right)$ is an "error" term, given by the parameters in Eq. (2), and corresponds to returns from "cash flow news", unrelated to the risk of the asset. The term in parentheses in Eq. (5) includes the **surprises** about the price of risk: A larger than expected price of risk reduces the return because the asset is repriced using a larger risk premium.

In each period, the price of risk is expected to remain the same with probability p or change with probability $1-p$, in which case the new price of risk corresponds to a random draw from its unconditional distribution. Therefore, its expectation one period ahead is

$$E_t[\lambda_{t+1}] = p\,\lambda_t + (1-p)\mu_\lambda, \qquad (6)$$

where $\mu_\lambda < \infty$ is the unconditional (finite) mean of the stochastic process for the price of risk in Eq. (1). In general,

$$E_0[\lambda_t] = p^t\,\lambda_0 + (1-p^t)\mu_\lambda, \qquad (7)$$

implying that, from one period to the next (for $T \to \infty$),

$$E_t\left[\sum_{i=t+1}^{T}\lambda_i\right] - E_{t+1}\left[\sum_{i=t+1}^{T}\lambda_i\right] = \frac{E_t[\lambda_{t+1}] - \lambda_{t+1}}{1-p}, \qquad (8)$$

and that the return on the asset in Eq. (5) is

$$r_{i,t+1} = r_f + \sigma_i\left(\mu_\lambda + \frac{\lambda_t - \lambda_{t+1}}{1-p}\right) + \varepsilon_{i,t+1}. \qquad (9)$$

Eq. (9) shows two components of the premium: The first, $\sigma_i\mu_\lambda$, is the average premium on this asset (corresponding to an average price of risk). The second, $\sigma_i\frac{\lambda_t - \lambda_{t+1}}{1-p}$, is the return related to changes in the price of risk. The change is scaled by $1-p$ because the impact in

8

the price of the asset is stronger if the price of risk is more persistent (as p increases) than if it is expected to quickly revert back to the mean.

2.1   Rationality and arbitrage opportunities

In the sense of Muth (1961) and Lucas (1972), the model features fully rational expectations: Eq. (9) is obtained under the assumption that the investors know the parameters for the price of risk process, p and μ_λ, the risk of the asset, σ_i, and observe the price of risk realizations, λ_t.

In fact, the asset prices are perfectly efficient because they reflect all available information at every point in time. In particular, there is no asymmetric information, slow information diffusion, "noise traders", miscalculated probabilities (behavioral biases), nor any type of frictions. The mean reversion in the price of risk process is known and fully accounted for in the agents' expectations, which are always correct ex-ante. Hence, the model is internally consistent because the prices and expectations contain no systematic biases over time. As a consequence, there are no risk-adjusted profits to be made and hence no incentives for any investor to take advantage of the higher returns on certain assets: Higher returns are exactly offset by higher risks.

Regarding the discussion about premiums of this magnitude being generated by risk, the average return on momentum portfolios of stocks, for example, is around 8%. This is close to the historical market premium and illustrates that, in terms of the CAPM, an average spread in CAPM betas between the typical winner and loser as small as one already generates returns of similar magnitudes.

2.2 Time-varying momentum risk and the price of risk

Let us consider two types of assets with constant risks over time (or at least constant during portfolio formation and evaluation): Assets r are riskier and have volatility term $\sigma_{i,t} = \sigma_r$, while assets s are safer, $\sigma_{i,t} = \sigma_s$, such that

$$\sigma_r > \sigma_s. \tag{10}$$

According to Eq. (9), the realized one period return spread between these two assets is

$$r_{r,t} - r_{s,t} = (\sigma_r - \sigma_s)\left(\mu_\lambda + \frac{\lambda_{t-1} - \lambda_t}{1 - p}\right) + \varepsilon_{r-s,t}, \tag{11}$$

where $\varepsilon_{r-s,t} \equiv \varepsilon_{r,t} - \varepsilon_{s,t}$ is the difference in returns realized from cash flow news, unrelated to risk. The equation shows that both the initial and final prices of risk are crucial to obtain the correct signal about the risk of the assets based on the spread in their realized returns. The equation also shows that cash flow news act as noise in this process.

If the price of risk remains constant, $\lambda_{t-1} = \lambda_t$, the spread should be positive (subject to random cash flow shocks). In this case, risky assets, r, tend to be in the winner portfolio and so the momentum portfolios should have positive risk.

If the price of risk decreases, $\lambda_{t-1} > \lambda_t$, the term in brackets in Eq. (11) becomes even larger, so the spread should be even more positive. The repricing effect magnifies the difference in returns between the risky and safe assets because $\sigma_r > \sigma_s$. In this case, the risky assets, r, are even more likely to be in the winner portfolio and the positive risk of the momentum portfolios should be even larger as a consequence.

Finally, if the price of risk increases, $\lambda_{t-1} < \lambda_t$, the term in brackets in Eq. (11) becomes smaller (or negative for large changes in λ). In this case, the risky assets, r, are less likely to be in the winner portfolio compared to the other situations (in fact, they are more likely to be in the loser portfolio for a substantial increase in the price of risk). The risk of the momentum portfolios formed in these states should be smaller (and possibly negative).

In general, we can represent the risk of momentum portfolios, σ_{mom}, at time t as

$$\sigma_{mom,t} = f(\Delta\lambda_{t-1}), \quad f(0) > 0, \quad f'(\Delta\lambda_{t-1}) < 0 \quad \forall \Delta\lambda_{t-1}, \tag{12}$$

where $f(.)$ is a certain decreasing function that becomes negative for large values of $\Delta\lambda_{t-1}$, and

$$\Delta\lambda_{t-1} \equiv \lambda_{t-1} - \lambda_{t-h} \tag{13}$$

is the increase in the price of risk during the portfolio formation period over which the realized returns are calculated to classify the assets into winners or losers. For example, $h = 12$ months in Jegadeesh and Titman (1993).

2.2.1 Momentum risk and the price of risk levels

Positive changes in the price of risk cause negative changes in the risk of momentum portfolios, as explained above. But positive changes also correlate with higher price of risk **levels**, as long as the price of risk process is stationary.[4] Hence, there is negative covariance between the risk of momentum portfolios and the price of risk levels on the portfolio formation date,

$$cov\left(\sigma_{mom,t}, \lambda_t\right) < 0. \tag{14}$$

For example, a price of risk below its median is more likely to have decreased to this level than to have increased from an even lower random starting point. Therefore, low price of risk states correlate with decreases in the price of risk and, thus, with momentum portfolios with large and positive risks. The opposite happens in "high" price of risk states, which correlate with increases in the price of risk and, thus, with momentum portfolios with low and possibly negative risks. Finally, a price of risk close to its median, $\overline{\lambda}$, is correlated

---

[4]This follows from the fact that the variance of the SDF is finite, which implies that λ_t in Eq. (1) cannot grow indefinitely.

with no changes in the price of risk, and, thus, with momentum portfolios that have positive but not particularly large risk exposures.

Fig. 1a summarizes the arguments above representing the risk of momentum portfolios approximately as a linear function of the price of risk,

$$\sigma_{mom,t+1} = a - b\lambda_t, \qquad (15)$$

where a and b are positive constants. This implies that the momentum risk premium between times t and $t+1$, given by Eq. (3), is

$$\mu_{mom,t+1} = (a - b\lambda_t)\lambda_t, \qquad (16)$$

as in Fig. 1c. This approximation already explains several momentum features described in Daniel and Moskowitz (2016) and Barroso and Santa-Clara (2015), apart from explaining why past winners are also future winners. For example, the return on momentum portfolios given by this function is bounded from above and unbounded from below. Therefore, the expected returns of the largest magnitudes are negative and the distribution is negatively skewed. These big expected losses tend to happen exactly when the price of risk and, therefore, the market premium are the highest. For lower prices of risk, the portfolio tends to deliver (smaller) positive returns. This is why momentum has an "option-like payoff" and why its "crashes" are associated with market rebounds.

2.2.2 Positive momentum and the case of Japan

In large samples, momentum is usually positive because the risky assets tend to be in the winner portfolios more often than in the loser portfolios and because there is relatively large cross-sectional dispersion in risk among the assets. But the theory is also consistent with negative or zero momentum premiums, especially in small samples.

Let us start with the unconditional expectation of momentum premium from Eq. (3),

$$\mathrm{E}\left[\mu_{mom,t+1}\right] = \mathrm{E}\left[\sigma_{mom,t}\lambda_t\right] \tag{17}$$

$$= \mathrm{E}\left[\sigma_{mom,t}\right]\mu_\lambda + cov\left(\sigma_{mom,t},\lambda_t\right). \tag{18}$$

As explained in the previous section and Eq. (14), the covariance term in Eq. (18) is negative. Therefore, Eq. (18) is positive as long as the magnitude of this (negative) covariance is not very large relative to the premium of the average momentum portfolio, $\mathrm{E}\left[\sigma_{mom,t}\right]\mu_\lambda$. This first term in Eq. (18) is positive because the unconditional expectation of momentum risk is positive,

$$\mathrm{E}\left[\sigma_{mom,t}\right] > 0, \tag{19}$$

given that the unconditional expectation of the realized return spread in Eq. (11) is also positive,

$$\mathrm{E}\left[r_{r,t} - r_{s,t}\right] = (\sigma_r - \sigma_s)\mu_\lambda. \tag{20}$$

However, it is theoretically possible for momentum to be zero or even negative depending on the combination between the cross-sectional dispersion in risk among the assets, affecting $\mathrm{E}\left[\sigma_{mom,t}\right]$, the average price of risk, μ_λ, or the covariance between the risk of momentum portfolios and the price of risk, $cov\left(\sigma_{mom,t},\lambda_t\right)$ in Eq. (18). For example, a relatively low cross-sectional dispersion in risk among the assets or a volatile price of risk process that makes the covariance term in Eq. (18) more negative could generate the positive but insignificant momentum premium in Japan reported by Asness et al. (2013). Alternatively, highly volatile cash flow news in Eq. (5), $\varepsilon_{i,t+1}$, relative to the cross-sectional dispersion in risks should also lower the significance of momentum in finite samples. This could also explain the positive but insignificant momentum premium in Japan.

2.3 The typical negative CAPM betas of momentum portfolios

Momentum portfolios have negative unconditional CAPM betas as long as the price of risk process is such that (large) values relatively far from the mean are common. The market premium, given by Eq. (3), is

$$\mu_{mp,t+1} - r_f = \sigma_{m,t}\lambda_t, \tag{21}$$

where $\sigma_{m,t} > 0$ is the risk of the market portfolio at time t. Continuing with the previous simplifying assumptions, we can consider a constant risk for the market portfolio, $\sigma_{m,t} = \sigma_m$, substitute out the price of risk, λ_t, in Eq. (16), and obtain

$$\mu_{mom,t+1} = \frac{a}{\sigma_m}\left(\mu_{mp,t+1} - r_f\right) - \frac{b}{\sigma_m^2}\left(\mu_{mp,t+1} - r_f\right)^2. \tag{22}$$

Fig. 1d plots this function and illustrates the main CAPM properties of momentum portfolios. For example, the unconditional CAPM beta estimates for the momentum portfolios tend to be negative, and the intercept tends to be positive because the domain of the function starts at zero market premium, $\mu_{mp,t} - r_f > 0$. This implies that the decreasing portion of the function is usually larger than the increasing one, dominating the unconditional estimations.

Hence, the best linear fit between the two premiums usually corresponds to a negative slope and positive intercept. This explains why Daniel and Moskowitz (2016) report highly negative CAPM betas for momentum in "distressed periods" (when the price of risk is high and the expected market premium is large), which correspond to the right portion of the graph, while reporting close to zero betas in the remaining periods that correspond to the top section of the graph.

2.3.1 Similar conclusions with fewer assumptions

The explanations until this point do not rely on the simplifying assumptions made to illustrate them. Let us start by using Eq. (21) to inversely define the price of risk as a function of the market premium, $MP_t \equiv \mu_{mp,t} - r_f$,

$$\lambda_t = h(MP_t), \quad h(0) = 0, \quad h'(MP) = \frac{1}{\sigma_{m,t}} > 0 \quad \forall MP_t. \tag{23}$$

Second, let us generalize Eq. (15) to

$$\sigma_{mom,t} = g(\lambda_t), \quad g(0) > g(\overline{\lambda}) > 0, \quad g'(\lambda_t) < 0 \quad \forall \lambda_t > 0, \tag{24}$$

which implies that the momentum premium (as a function of the price of risk), given by Eq. (3), is

$$\mu(\lambda_t)_{mom,t} = g(\lambda_t)\lambda_t, \quad \mu(\overline{\lambda})_{mom,t} > 0, \quad \mu(0)_{mom,t} = 0. \tag{25}$$

Now, the changing CAPM betas of the momentum portfolios are given by

$$\frac{\partial \mu_{mom,t}}{\partial MP_t} = \frac{\partial \mu_{mom,t}}{\partial \lambda_t} \frac{\partial \lambda_t}{\partial MP_t} = \left(\sigma_{mom,t} + \lambda_t g'(\lambda_t) \right) h'(MP_t), \tag{26}$$

where the first partial derivative comes from Eq. (25) and the second comes from Eq. (23). The term in Eq. (23), $h'(MP_t)$, is always positive, so the sign of the derivative is given by the term in brackets. The second term inside the brackets is always negative, $\lambda_t g'(\lambda_t) < 0$, because $\lambda_t > 0$ (by definition) and $g'(\lambda_t) < 0$ from Eq. (24). The other term inside the brackets is the risk of the momentum portfolio, $\sigma_{mom,t}$, which is positive for a zero price of risk, $\lambda_t = 0$, and also for an average price of risk, $\lambda_t = \overline{\lambda}$, but eventually becomes negative as the price of risk increases further, as given by Eq. (24). According to Eq. (25), the momentum premium must increase from zero (when $\lambda_t = 0$) to positive (when $\lambda_t = \overline{\lambda} > 0$). Therefore, as long as the magnitude of $g'(\lambda_t)\lambda_t$ is non-decreasing, the term inside the brackets must be maximum and positive when the price of risk and the market premium are zero, $\sigma_{mom,t} - \lambda_t g'(\lambda_t) > 0$ because $\sigma_{mom,t}$ is decreasing in λ_t.

As a consequence, the momentum premium in terms of the market premium is a function that starts at zero, increases, reaches a maximum, and decreases (relatively fast after it reaches a point at which $\sigma_{mom,t} < 0$). The function has a longer downward portion because its domain is the positive values for the market premium. This is equivalent to the function plotted in Fig. 1d and leads to the same conclusions. The part above the axis in Fig. 1d is also, at least, a good approximation describing the increasing and decreasing sections of the function. So, in general, it also leads to the same conclusions.

3 Simulation, real data, and variables

3.1 Real return data

The monthly returns data come from Kenneth French's website between January 1927 and February 2017: I obtain the monthly market premium (MP) and the return on 10 portfolios formed on past returns (using NYSE prior (2-12) return decile breakpoints) and calculate the return on momentum (mom) portfolios as the difference between equal weighted averages of the five winning portfolios and the five losing ones.

3.2 Real price of risk proxies

Most price of risk proxies come from Amit Goyal's website. These proxies are the only four individual variables that significantly predict both the one- and five-year equity premiums in sample during the period 1927-2013, according to the updated version of Welch and Goyal (2008). These series end in December 2016, and I mention their starting dates in the respective brackets.

The proxies are the Dow Jones' book-to-market of Pontiff and Schall (1998) (DJBM, starting in March 1921), the net equity expansion of Boudoukh et al. (2007) (NTIS, December 1926), the investment to capital ratio of Cochrane (1991) (IK, January 1947), and the consumption-wealth-income ratio of Lettau and Ludvigson (2001) (CAY, January

1952). I also obtain the lower bound for the equity premium of Martin (2017) (SVIX, available from January 1996 to January 2012) from Ian Martin's website.

All series are monthly, except for the quarterly IK series. I change the sign of the IK and the NTIS, so that these variables forecast positive market premiums and can also be interpreted as prices of risk. For the same reason, I change the values of all series so that they all have a minimum at zero: For each of the five series, I subtract their minimum value from each observation.

As mentioned earlier, the only variable that seems inappropriate as a price of risk proxy is the NTIS because it seems to forecast cash flow news instead of being associated with discount rates. I report these results for completeness.

3.3 Simulated data

From Eq. (9), I generate 1000 periods of 100 years of monthly returns each $(1,200,000$ observations) for 200 assets. The risk-free rate is zero, $r_f = 0$, $p = 0.5$, and the cash flow shocks have two components,

$$\varepsilon_{i,t} = 0.5\varepsilon_{id_i,t} + 0.5\varepsilon_{sys,t}, \quad \varepsilon_{id_i,t} \sim \varepsilon_{sys,t} \sim \mathcal{N}(0, \mu_\lambda^2) \tag{27}$$

where the systematic shock is the same for every asset, $\varepsilon_{sys,t}$, and the idiosyncratic component, $\varepsilon_{id_i,t}$, is particular to each asset. The standard deviation of both normally distributed shocks is equal to the mean of the price of risk process, μ_λ.

I assume that dz_t in Eq. (1) has dimension 200 (all independent) and each asset is only exposed to one of these dimensions.[5] Half of the 200 assets are "safe", $\sigma_{i,s} = 0.5$, and the others are risky, $\sigma_{i,r} = 2.5$.[6]

The multi-dimensional log-normal price of risk process is given by

$$\lambda_{i,0} = e^{z_0}, \quad z_0 \sim \mathcal{N}(0, 1), \tag{28}$$

$$\lambda_{i,t+1} = \begin{cases} \lambda_{i,t}, & \text{for } 0 \leq \mathcal{U}(0,1) \leq p \\ e^{z_t}, & \text{for } p < \mathcal{U}(0,1) \leq 1, \quad z_t \sim \mathcal{N}(0, 1), \end{cases} \tag{29}$$

where $\lambda_{i,t}$ is the ith component of the price of risk vector with dimension $I = 200$ at time t and $\mathcal{U}(0,1)$ is a random variable uniformly distributed between 0 and 1. Later in the analysis, I refer to the average of these 200 components as the (single) measure of the price of risk in a given period.

The return on the market is an equal-weighted average of the return on these 200 assets. The return on the momentum portfolio is the difference in equal-weighted returns between

---

[5]It is important to assume multi-dimensional shocks because otherwise the returns on all assets are perfectly correlated.

[6]More precisely, the risk of each asset is a vector with zeros in every position except the ith position, such as $\sigma_{i,s} = (0, 0, ..., 0.5, ..., 0, 0)^\top$ and $\sigma_{i,r} = (0, 0, ..., 2.5, ..., 0, 0)^\top$.

the assets that were the 100 winners and the ones that were the 100 losers in the previous period.

3.4   The timing of the changes in the price of risk (proxies)

The momentum portfolios in Kenneth French's data library are formed on past returns using NYSE prior (2-12) return decile breakpoints. For this data set, the relevant change in the price of risk is calculated with respect to the value of the proxy two months before the portfolio is formed,

$$\Delta\Lambda_{t-2} = \Lambda_{t-2} - \Lambda_{t-12}. \tag{30}$$

On the other hand, for the simulated data, the relevant data are simply the value of the price of risk immediately before the portfolio is formed,

$$\Delta\lambda_{t-1} = \lambda_{t-1} - \lambda_{t-12}. \tag{31}$$

19

4 Empirical and simulated results

4.1 CAPM estimates for momentum and changes in the price of risk

The parabola that approximately describes the theoretical relation between momentum and the market premium, shown in Fig. 1d, follows from the prediction that the risk of momentum portfolios decreases with the price of risk level (a point that I also show later based on simulated data).

For example, if the sample includes only the momentum portfolios with positive risk the CAPM beta should be close to zero and the intercept should be positive. These are the portfolios in the symmetric part of the function in Fig. 1d, above the market premium axis. On the other hand, if the sample only includes the portfolios with negative risk, the estimated beta should be highly negative, and the intercept should still be positive because these are the portfolios on the right tail of the graph.

The problem is to separate these portfolios according to positive or negative risks. In theory, the momentum portfolios have positive risks if the price of risk does not increase during the portfolio formation period. However, the opposite is not true,

$$\Delta \lambda_{t-1} \leq 0 \implies \sigma_{mom,t} > 0, \tag{32}$$

$$\Delta \lambda_{t-1} > 0 \notimplies \sigma_{mom,t} < 0. \tag{33}$$

which means that the method is not perfect.

In terms of Fig. 1d, all the momentum portfolios for which Eq. (32) holds should be on the upper part of the graph, above the market premium axis. The remaining portfolios should mainly be located on the lower part, but they could also be located (particularly) on the (right) upper part, too. Hence, the conservative selection of portfolios implied by Eq. (32) can generate positive CAPM betas for the portfolios with positive risks by excluding the portfolios closer to the axis on the right-hand side. This makes the graph less symmetric, while lowering the estimated betas in the rest of the sample, too.

By showing only the subsample in which the price of risk proxies increased or decreased, respectively in Fig. 3 and Fig. 4, the graphs suggest that the predictions in the paper are correct based on real data. The patterns in these graphs are very similar to the ones based on simulated data, in Fig. 5 and Fig. 6, respectively, which also suggests consistency with the theory.[7] The data in these graphs come from the first four (out of 1000) simulated periods of 1200 observations that do not contain outliers.[8]

I formally confirm these predictions by running the CAPM regressions,

$$R_{mom,t} = \alpha + \beta \left(R_{m,t} - R_{f,t} \right) + \varepsilon_t, \tag{34}$$

where $R_{m,t}$ is the return on the market, $R_{f,t}$ is the risk-free return, and ε_t is the error term at time t. I run these regressions in two separate subsamples: (i) With the momentum portfolios formed in periods in which the price of risk proxy did not increase, $\Delta \Lambda_{t-2} \leq 0$, as shown in Table 1a, and (ii) with the remaining portfolios, in which $\Delta \Lambda_{t-2} > 0$, as shown in Table 1b.

For the momentum portfolios formed after no increases in the price of risk proxies, $\Delta \Lambda_{t-2} \leq 0$, Table 1a shows that the betas based on every proxy (except the NTIS) have positive point estimates. In fact, the beta estimate based on the DJBM is significantly positive. Every intercept is also significantly positive.[9] Therefore, the evidence is consistent with the theory.

For the momentum portfolios formed after increases in the price of risk proxies, $\Delta \Lambda_{t-2} > 0$, Table 1b shows, in line with the theory, that the betas based on every proxy are highly significant and negative. The intercepts still tend to be significant (but to a lower degree) which, again, is consistent with the theory.

Fig. 8 shows that the 1000 simulated samples of 1200 observations produce very similar results to the ones predicted by the theory and observed in the real data: The graphs in

---

[7]In fact, for completeness, Fig. 7 also suggests that the simulated data deliver the unconditional negative CAPM betas that are empirically observed for momentum portfolios in general.

[8]The outliers are also typically in line with the theory but they make the distribution of data points more difficult to visualize.

[9]The intercept based on the SVIX is only marginally significant, but the sample is a lot smaller.

the first column shows that, unconditionally, the CAPM betas from the simulations are negative and the intercepts are positive. The betas become somewhat more negative after increases in the price of risk, with the intercepts also increasing as the graphs in the second column show. After decreases in the price of risk, corresponding to the graphs in the third column, the betas are close to zero but the intercepts remain significantly positive (although smaller). Each of these values is consistent with the simplified description of the momentum returns and its relation with the market premium in Fig. 1d and with the overall framework presented in the paper.

4.2   Risk of momentum portfolios and the price of risk

One of the most important theoretical predictions of the framework in the paper is that the risk of momentum portfolios correlates with decreases in the price of risk level. This cannot be easily confirmed with real data because the risk of the assets in non-observable. However, this can be checked in the simulated samples and, indeed, the graph in Fig. 9 suggests that the prediction is confirmed.

The graphs show that the proportion of risky assets (with $\sigma_r = 2.5$ compared to the $\sigma_s = 0.5$ of the safe assets) among the assets that comprise the past winners (and therefore the long position in the momentum portfolios formed each period) indeed decreases with the average price of risk level. The implication is that the momentum portfolio is riskier (has a larger position in the risky assets) when the price of risk is lower. When the price of risk increases, there are fewer risky assets in the long part of the portfolio (and more in the short part), so the portfolio has lower risk.

On the other hand, the proportion of risky assets among the winners in the graphs is always high (above 50%), which seems to indicate that the risk of momentum portfolios never becomes negative in these samples. However, this is only a partial description of risk (which also depends on the volatility of the price of risk). For example, if the ith component of the price of risk vector in Eq. (29) jumps to a very positive value, the asset exposed to this component realizes a very negative return. But because this price of risk component

22

reverts to the mean in relatively large steps, the subsequent asset returns also tend to be large. In summary, this asset is riskier than one that did not experience a large price of risk change in the past even if both have the same volatility coefficient, $\sigma_r = 2.5$, for example. This volatility clustering effect is not reflected in the volatility coefficient that appears in the graphs in Fig. 9.

The data in this graph (which are the same used in the other similar simulated graphs) correspond to the first four simulated samples of 1200 observations each that did not contain outliers. To confirm these conclusions, the third graph in Fig. 10 shows the distribution of the estimated slopes from regressing the proportion of high risk stocks in the winner portfolio, W_{risky}, on the previous price of risk,

$$W_{risky,t} = \alpha + \beta \lambda_{t-1} + \varepsilon_t, \tag{35}$$

based on each one of the 1000 simulated samples. The negative coefficients in the graph confirm the predictions of the model.

4.3 Negative skewness

Another theoretical prediction of the framework in the paper is that momentum returns are negatively skewed. This is also a stylized fact about momentum mentioned in, for example, Barroso and Santa-Clara (2015) or Daniel and Moskowitz (2016). The first graph in Fig. 10 shows that the simulated data also deliver negative skewness in momentum returns: The graph displays the distribution of the (always negative) skewness estimates from the 1000 simulated samples with 1200 observations each.

4.4 Average return on momentum portfolios and changes in the price of risk

In theory, Eq. (3) and Eq. (12) tell that the expected momentum premium from time t to $t+1$ is

$$\mu_{mom,t+1} = f(\Delta\lambda_t)\lambda_t, \tag{36}$$

and, in particular,

$$f(\Delta\lambda_t \leq 0) > 0 \implies \mu_{mom,t+1} > 0. \tag{37}$$

The expected premium is decreasing and eventually becomes negative for large enough values of $\Delta\lambda_t$. Therefore, the theory predicts that it is possible to avoid negative expected premiums by not investing in momentum portfolios when the price of risk increases. This generates managed portfolios similar to the ones in Barroso and Santa-Clara (2015) and Daniel and Moskowitz (2016), for example. The portfolios otherwise formed after increases in the price of risk should concentrate more (and sometimes large) negative return realizations.

The graphs in Fig. 11 show the distribution of simulated momentum returns and changes in the price of risk, λ_t, for the first four simulated periods of 1200 observations that do not contain outliers. Fig. 12 displays the distribution of momentum return realizations and changes in the five price of risk proxies, Λ_t. The similarity between the graphs based on simulated and real data suggest that increases in the price of risk are, indeed, associated with more negative (and sometimes large) returns both theoretically and empirically.

I formally test the hypothesis of differences in average returns in the (real) data by creating, for each price of risk proxy, an indicator for the months in which there are no increases in the price of risk proxy,

$$I_{\Delta\Lambda_{t-2}\leq 0} = \begin{cases} 1 & \Delta\Lambda_{t-2} \leq 0 \\ 0 & \Delta\Lambda_{t-2} > 0 \end{cases}, \tag{38}$$

24

and run the regression

$$R_{mom,t} = \alpha_\Lambda + \beta_\Lambda I_{\Delta\Lambda_{t-2}\leq 0} + \varepsilon_t, \tag{39}$$

where $R_{mom,t}$ is the return on the momentum portfolio in time t, α_Λ and β_Λ are the coefficients to be estimated based on each proxy, and ε_t is an error term.

The evidence in Table 2a strongly supports the theoretical prediction of larger momentum when there are no increases in the price of risk, given the significant β_Λ coefficients based on both the DJBM and the IK proxies. In fact, the hypothesis is not rejected based on any of the proxies. Table 2b complements the evidence by displaying the estimated average return on the portfolios formed in periods in which the price of risk does not increase,

$$E\left[R_{mom,t} \mid \Delta\Lambda_{t-2} \leq 0\right] = \frac{\sum_{m=1}^{M}\left(R_{mom,t}\,I_{\Delta\Lambda_{t-2}\leq 0}\right)}{\sum_{m=1}^{M} I_{\Delta\Lambda_{t-2}\leq 0}}, \tag{40}$$

where M is the total number of months in the sample. The table shows that momentum returns are more significant and have larger point estimates for the portfolios formed when there are no increases in the price of risk based on every proxy (except the NTIS).

Equivalently, the second graph in Fig. 10 shows the distribution of the estimated coefficients of the dummy for decreases in the price of risk, $I_{\Delta\lambda_{t-1}\leq 0}$, in equation

$$R_{mom,t} = \alpha + \beta I_{\Delta\lambda_{t-1}\leq 0} + \varepsilon_t, \tag{41}$$

from the 1000 simulated samples of 1200 observations each. The results from the simulation are also in line with the theory and with the conclusions above, especially considering the DJBM and the IK proxies. However, the simulations also show that sometimes the coefficient can be negative, which is in line with the lack of significance based on the remaining proxies.

25

4.5 Forecasting momentum returns

Eq. (36) theoretically motivates a predictive relation for momentum returns even if the form of $f(.)$ is not strictly determined: The quantities $\Delta\lambda_t$, λ_t, and especially their product should forecast the returns on momentum portfolios (which is, indeed, suggested by the plots in Fig. 13).

I start to investigate this hypothesis by running the following predictive regressions individually, based on the five price of risk proxies:

$$R_{mom,t} = \alpha_\Lambda + \beta_\Lambda \Delta\Lambda_{t-2} + \varepsilon_t, \tag{42}$$

$$R_{mom,t} = \alpha_\Lambda + \beta_\Lambda \Lambda_{t-1} + \varepsilon_t, \tag{43}$$

$$R_{mom,t} = \alpha_\Lambda + \beta_\Lambda (\Delta\Lambda_{t-2})\Lambda_{t-1} + \varepsilon_t. \tag{44}$$

We can analyze the results in terms of Eq. (3) and Eq. (36), which indicate that the risk and the price of risk together determine the return on momentum portfolios. This formulation is equivalent to the regression in Eq. (44).

Indeed, Table 3c shows that Eq. (44) provides the best forecast for momentum based on every proxy (except for the NTIS). The significantly negative coefficients also conform with the theory: The momentum premium at time t is given by momentum risk, which is negatively related to the changes in the price of risk, $\Delta\Lambda_{t-2}$, multiplied by a positive price of risk, Λ_{t-1}, so that the coefficient in this case is negative.

Table 3a (corresponding to Eq. (42)) confirms that momentum risk forecasts the premium, even without the price of risk in the estimation, with the theoretically expected negative sign for the changes in the price of risk, $\Delta\Lambda_{t-2}$, for every proxy (except NTIS).

On the other hand, the price of risk has a negative coefficient when included alone in the regression (Table 3b corresponding to Eq. (43)). This happens because Eq. (43) does not control for increases in the price of risk (which determine the risk of the momentum portfolios). As explained earlier, large prevailing prices of risk tend to be associated with increases in the price of risk. The negative sign for the price of risk alone indicates that the

risk of momentum portfolios is more important than the price of risk for determining the premium of momentum portfolios. This is natural because momentum can both increase and decrease with the price of risk depending on the positive or negative risk of the momentum portfolios. This omitted variable problem is solved in Eq. (44).[10]

The results from the simulation confirm these conclusions. Fig. 14 shows the distribution of the 1000 estimated slopes, β, in the equivalent predictive regressions

$$R_{mom,t} = \alpha + \beta \lambda_{t-1} + \varepsilon_t, \tag{45}$$

$$R_{mom,t} = \alpha + \beta \Delta \lambda_{t-1} + \varepsilon_t, \tag{46}$$

$$R_{mom,t} = \alpha + \beta (\Delta \lambda_{t-1}) \lambda_{t-1} + \varepsilon_t. \tag{47}$$

The graphs show the same negative values that we find in the real data, in line with the theory.

5   Summary

In the words of Berk (1995), an "empirical anomaly" is, by definition, an empirical fact that cannot be supported by the prevailing theory. In this paper we learn that the presence of momentum in asset returns cannot be regarded as such: A standard risk-based asset pricing framework theoretically explains why the average return on momentum portfolios is positive, why the return distribution is negatively skewed, why the portfolios have negative CAPM betas and positive intercepts, why the strategy "crashes", and why this happens exactly when the market rebounds.

We also learn about a theoretically motivated strategy to manage these crashes based on changes in the price of risk. In addition, we learn about further evidence supporting the framework of the paper in terms of forecasting momentum returns, confirming the importance of the changes in the price of risk for determining the risk of momentum portfolios, and for determining the CAPM properties of these portfolios.
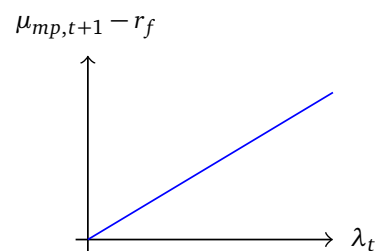
---

[10]There are no negative values among the price of risk proxies, as explained earlier.
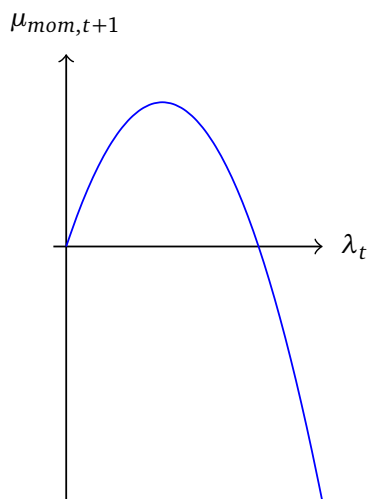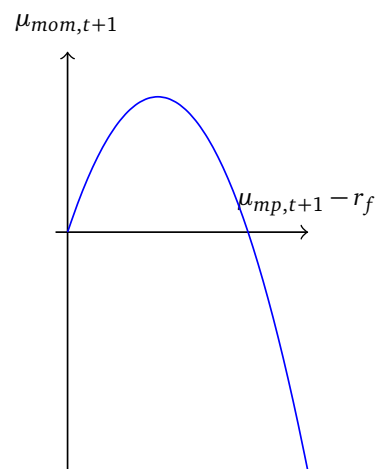
6 Figures
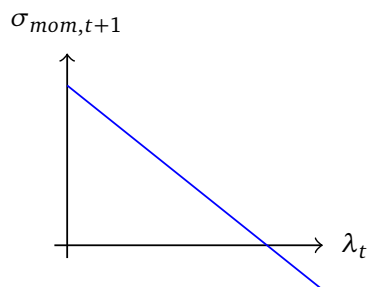
(a) The risk of momentum portfolios, $\sigma_{mom,t+1} = a - b\lambda_t$, as a function of the price of risk at time t, λ_t.

(b) The market premium, $\mu_{mp,t+1} - r_f = \sigma_m \lambda_t$, as a function of the price of risk at time t, λ_t.
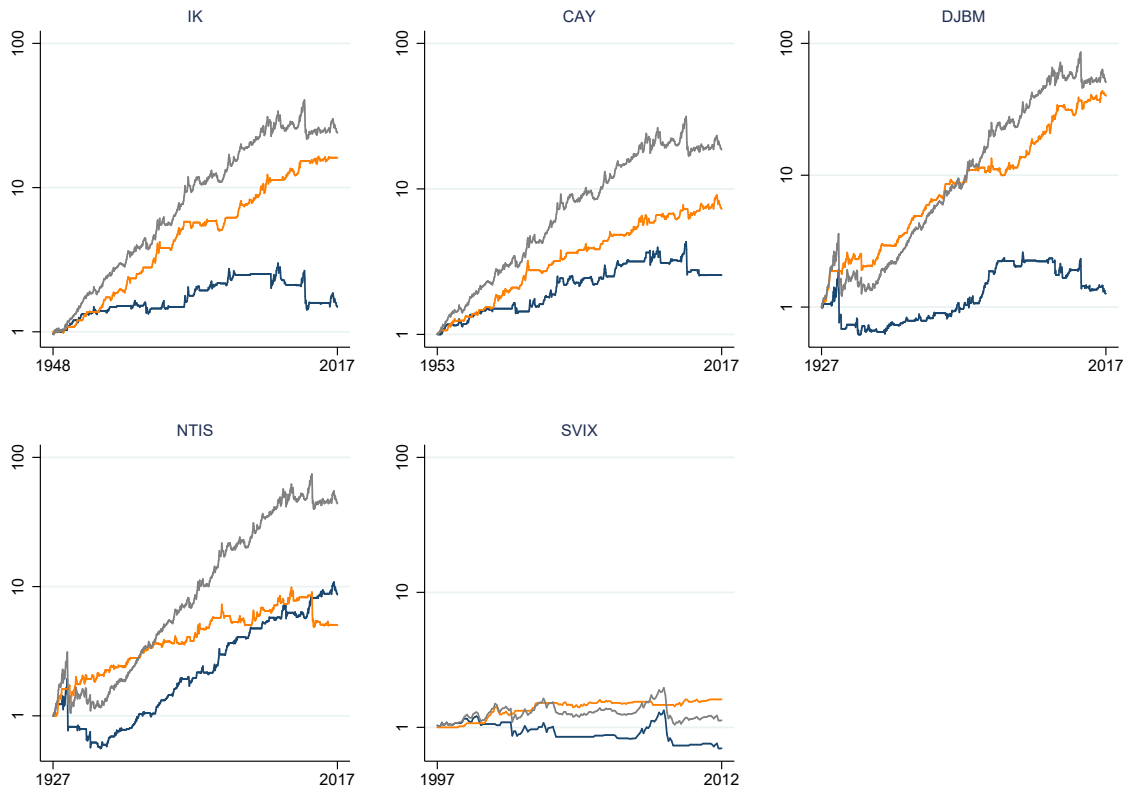

(c) The risk premium of momentum portfolios, $\mu_{mom,t+1} = (a - b\lambda_t)\lambda_t$, as a function of the price of risk at time t, λ_t.


(d) The risk premium of momentum portfolios, $\mu_{mom,t+1} = \frac{a}{\sigma_m}\left(\mu_{mp,t+1} - r_f\right) - \frac{b}{\sigma_m^2}\left(\mu_{mp,t+1} - r_f\right)^2$, as a function of the market premium, $\mu_{mp,t+1} - r_f$.

Figure 1: Illustration of the theoretical relations between the risks and premiums of momentum and market portfolios, and the price of risk.

Figure 2: Evolution of 1$ invested in managed or unmanaged momentum strategies for different periods (depending on the price of risk proxy availability). The gray line corresponds to the unmanaged, buy-and-hold, strategy for the momentum portfolio. The orange line corresponds to investing in momentum exclusively when the price of risk proxy does not increase during the portfolio formation period. The blue line corresponds to investing in momentum in the remaining months. The price of risk proxies are the investment-to-capital ratio (IK), the consumption-wealth ratio (CAY), the Dow Jones book-to-market (DJBM), the net equity expansion (NTIS), and Ian Martin's lower bound on the equity premium (SVIX).
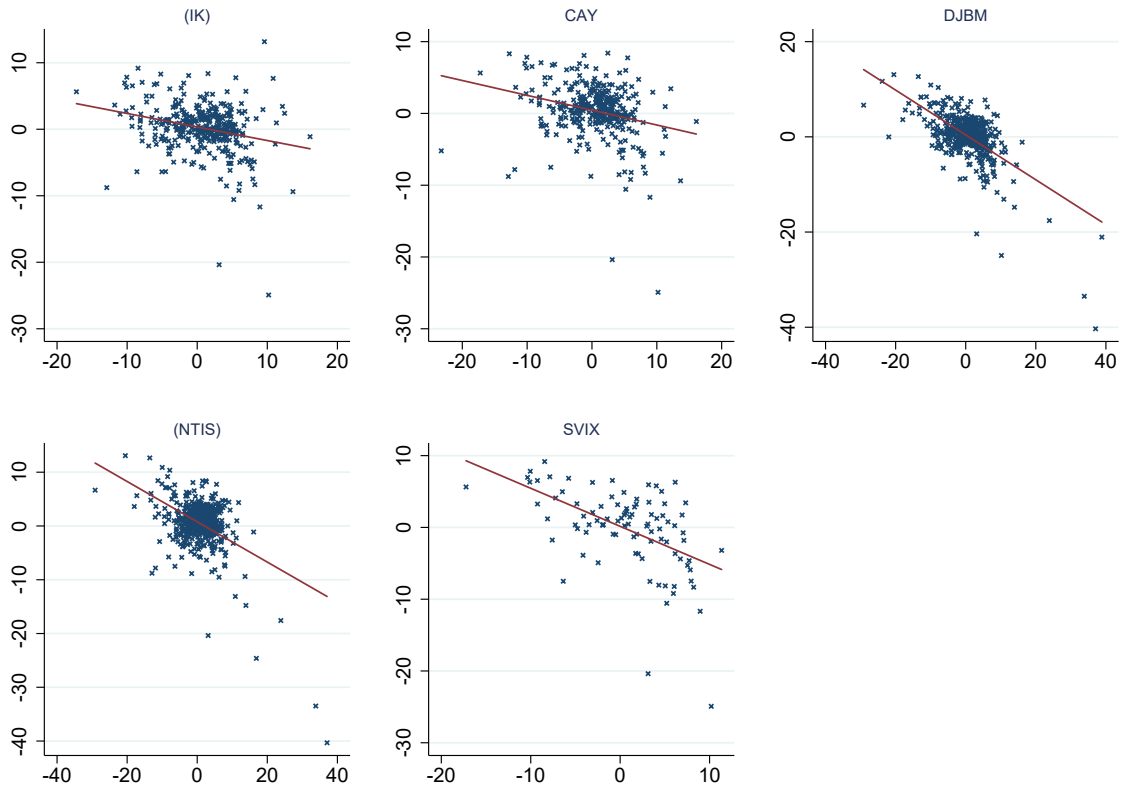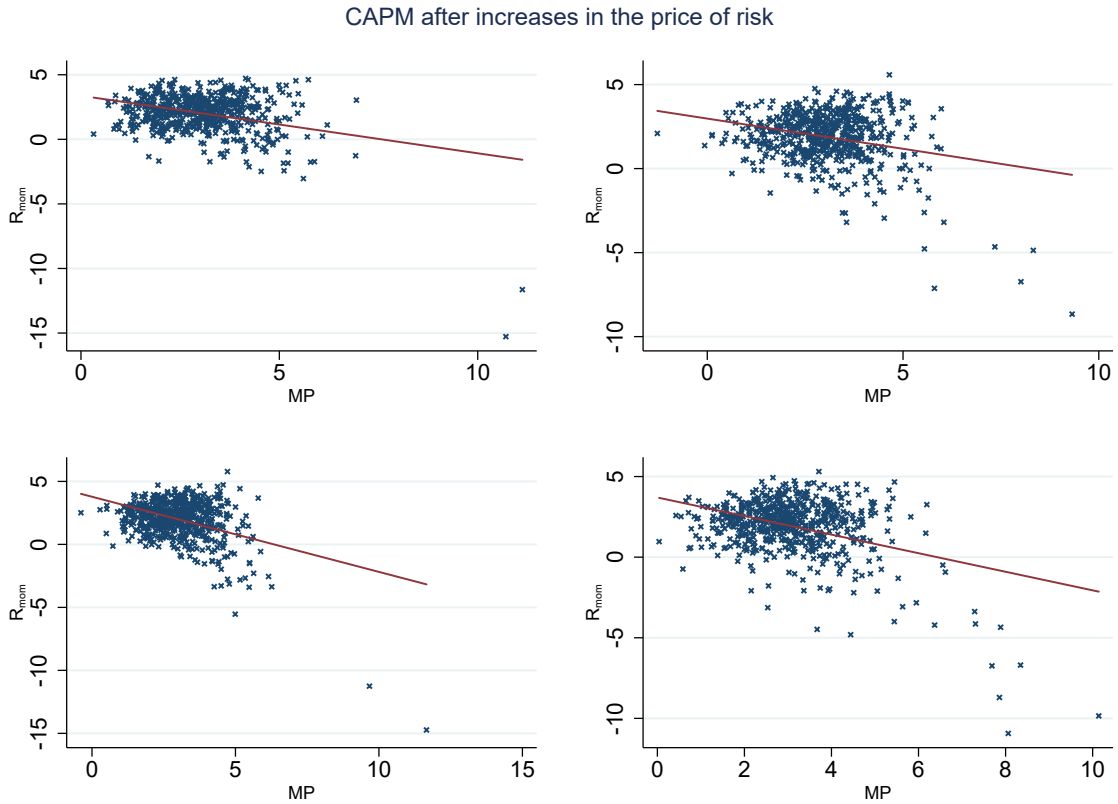
Figure 3: Return on momentum portfolios (vertical axis) vs. market premium, conditional on increases in the price of risk. Each marker corresponds to a month in which the price of risk increased during the portfolio formation period. The red line shows the best linear "CAPM" fit between the variables. The price of risk proxies (with their time spans in brackets) are the (negative of the) investment-to-capital ratio (IK, 1948–2017), the consumption-wealth ratio (CAY, 1953–2017), the Dow Jones book-to-market (DJBM, 1927–2017), the (negative of the) net equity expansion (NTIS, 1927–2017), and Ian Martin's lower bound on the equity premium (SVIX, 1997–2012).
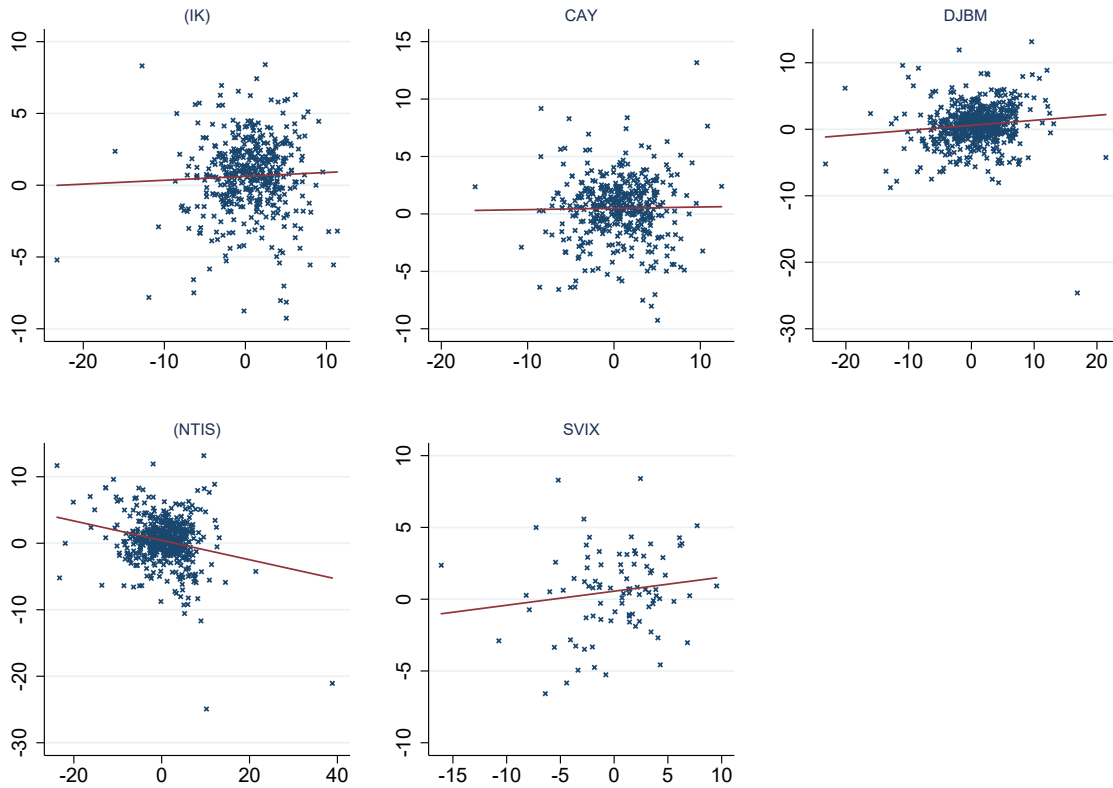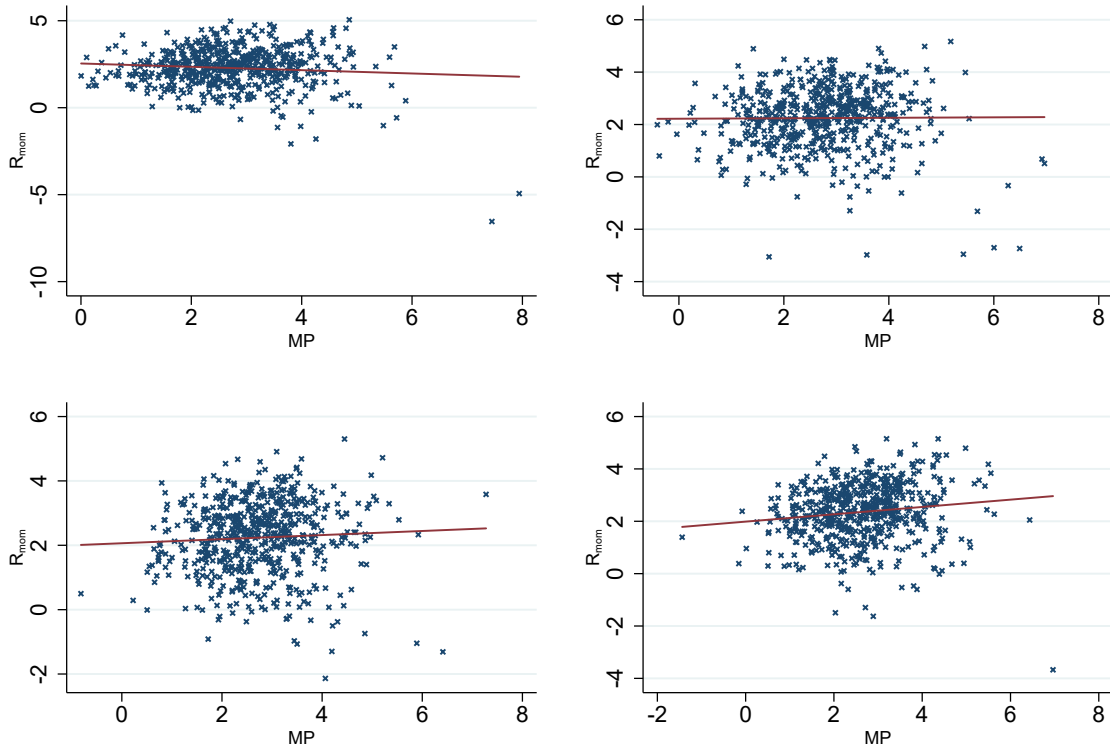
Figure 4: Return on momentum portfolios (vertical axis) vs. market premium, conditional on no increases in the price of risk. Each marker corresponds to a month in which the price of risk did not increase during the portfolio formation period. The red line shows the best linear "CAPM" fit between the variables. The price of risk proxies (with their time spans in brackets) are the (negative of the) investment-to-capital ratio (IK, 1948–2017), the consumption-wealth ratio (CAY, 1953–2017), the Dow Jones book-to-market (DJBM, 1927–2017), the (negative of the) net equity expansion (NTIS, 1927–2017), and Ian Martin's lower bound on the equity premium (SVIX, 1997–2012).

CAPM after increases in the price of risk

Figure 5: Simulated returns on momentum portfolios (vertical axis) vs. market premiums, conditional on increases in the price of risk. Each marker represents one observation (month) in which the price of risk increased during the portfolio formation period. The red line shows the best linear "CAPM" fit between the variables. The data (the same used in the other graphs) correspond to the first four simulated samples of 1200 observations each that did not contain outliers (which simply magnify the effects displayed).

Figure 6: Simulated returns on momentum portfolios (vertical axis) vs. market premiums, conditional on no increases in the price of risk. Each marker represents one observation (month) in which the price of risk did not increase during the portfolio formation period. The red line shows the best linear "CAPM" fit between the variables. The data (the same used in the other graphs) correspond to the first four simulated samples of 1200 observations each that did not contain outliers (which simply magnify the effects displayed).
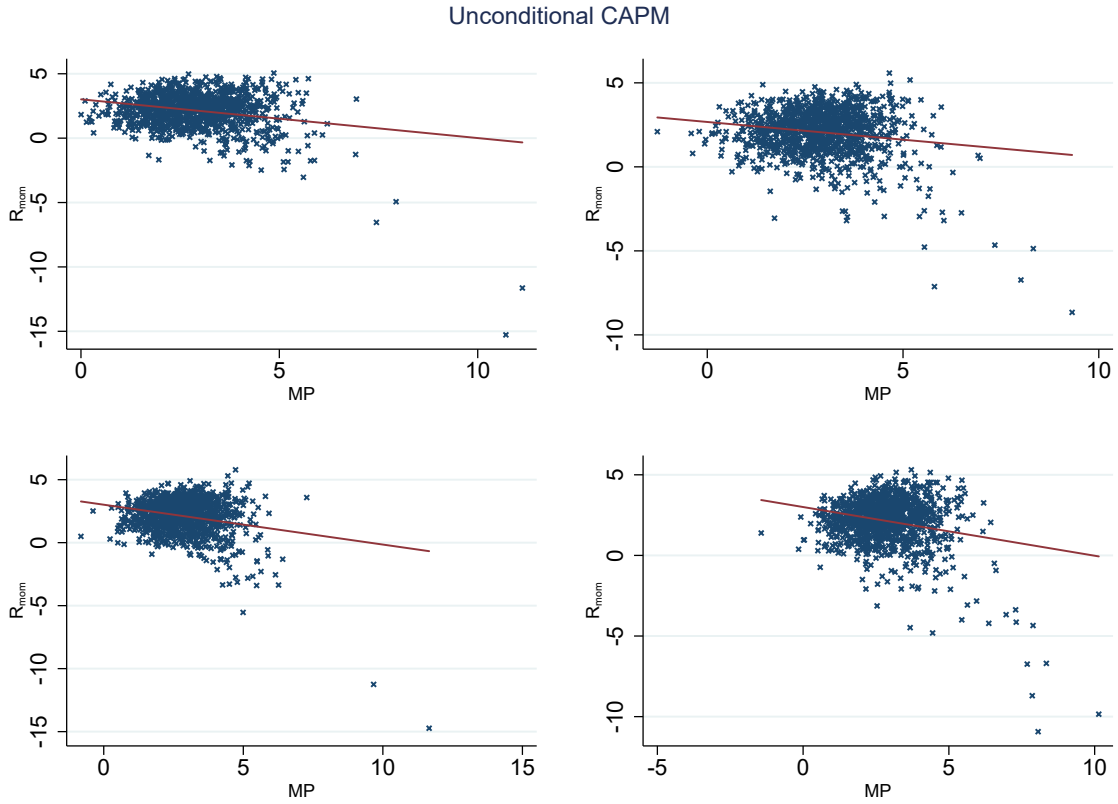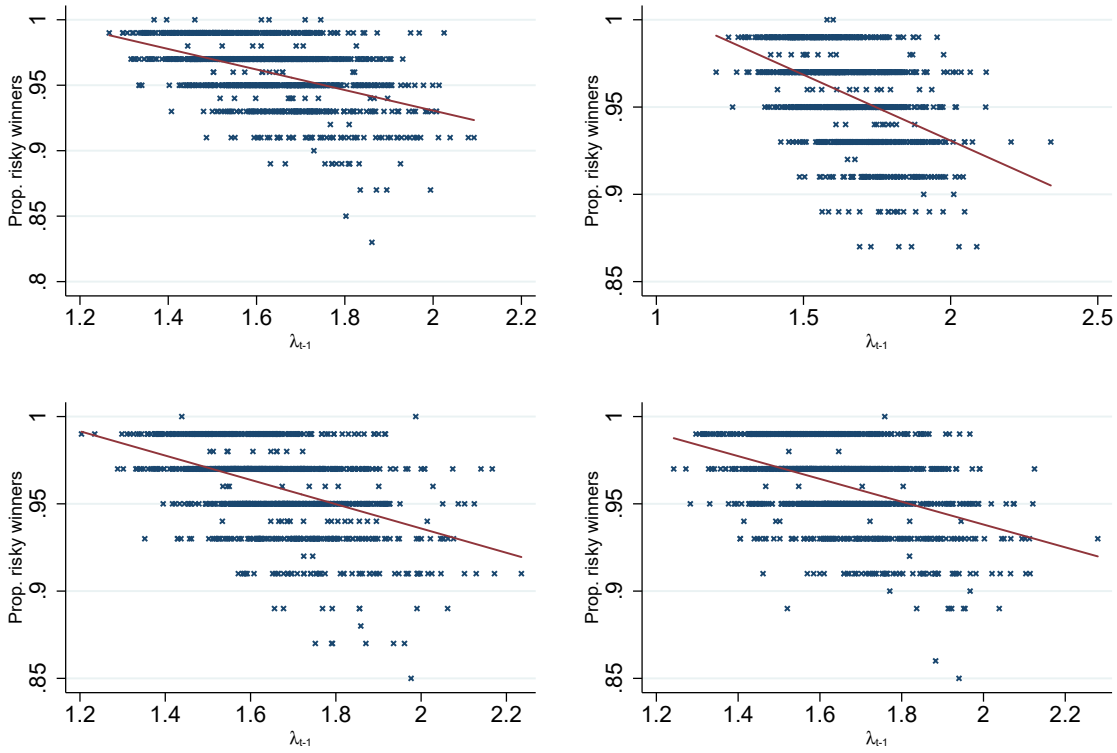
Figure 7: All (unconditional) simulated returns on momentum portfolios (vertical axis) vs. market premiums. Each marker represents one observation (month) and the red line shows the best linear "CAPM" fit between the variables. The data (the same used in the other graphs) correspond to the first four simulated samples of 1200 observations each that did not contain outliers (which simply magnify the effects displayed).
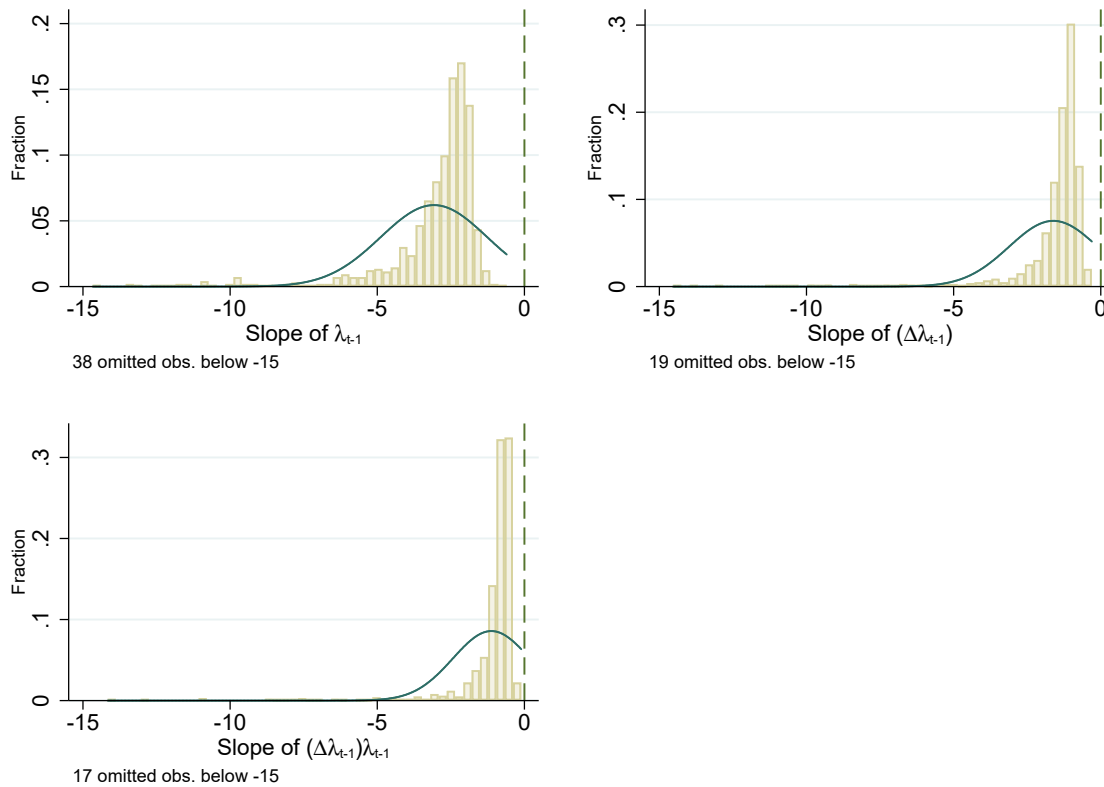
**Figure 8: Distribution of 1000 conditional and unconditional CAPM estimates for momentum portfolios based on simulated data.** The two graphs on the left column display the unconditional estimates for the CAPM intercept (α) and slope (β) for momentum portfolios. The middle (right) column displays equivalent estimates but only for the months following increases (decreases) in the price of risk, λ.

Figure 9: Momentum risk (on the vertical axis) vs. the price of risk at the end of the previous period, λ_{t-1}. Momentum risk is given by the proportion of risky stocks among the past winners (which comprise the long position in the momentum portfolios formed each period). Each marker represents one observation (month) and the red line shows the best linear fit between the variables. The data (the same used in the other graphs) correspond to the first four simulated samples of 1200 observations each that did not contain outliers.
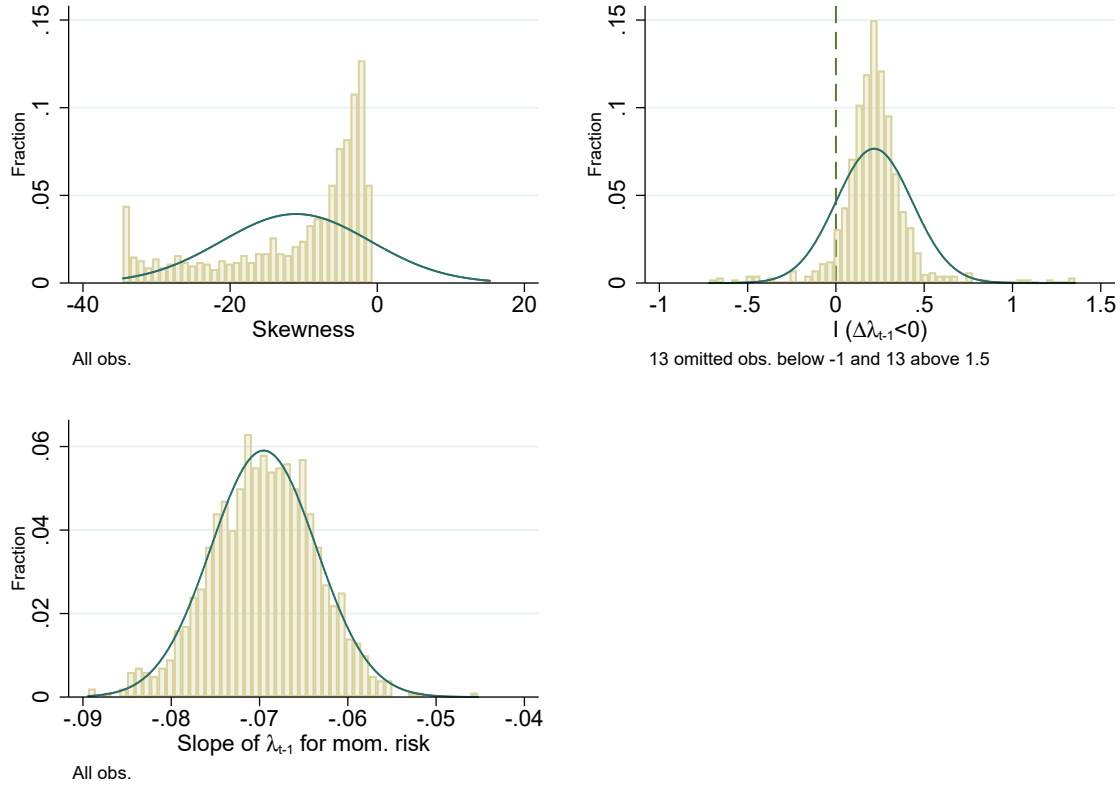
Figure 10: Other results for the return on momentum portfolios. Starting from the top left, the graphs display the distribution of 1000 estimated values for (i) the skewness of the returns on momentum portfolios, (ii) the coefficient of the dummy for decreases in the price of risk, $I_{\Delta\lambda_{t-1}\leq 0}$, in equation

$$R_{mom,t} = \alpha + \beta I_{\Delta\lambda_{t-1}\leq 0} + \varepsilon_t,$$

and (iii) the slope in the regression of the proportion of high risk stocks in the winner portfolio, W_{risky}, on the previous price of risk:

$$W_{risky,t} = \alpha + \beta \lambda_{t-1} + \varepsilon_t$$

Figure 11: Simulated returns on momentum portfolios (vertical axis) as a function of the increases in the price of risk. Each marker represents one observation (month), the dotted vertical green line (on the horizontal axis) represents a constant price of risk during the portfolio formation, and the red line shows the best linear fit between the variables. The increase in the price of risk is given by $\Delta\lambda_{t-1} \equiv \lambda_{t-1} - \lambda_{t-12}$, where λ_t is the price of risk at time t. The data (the same used in the other graphs) correspond to the first four simulated samples of 1200 observations each that did not contain outliers (which simply magnify the effects displayed).
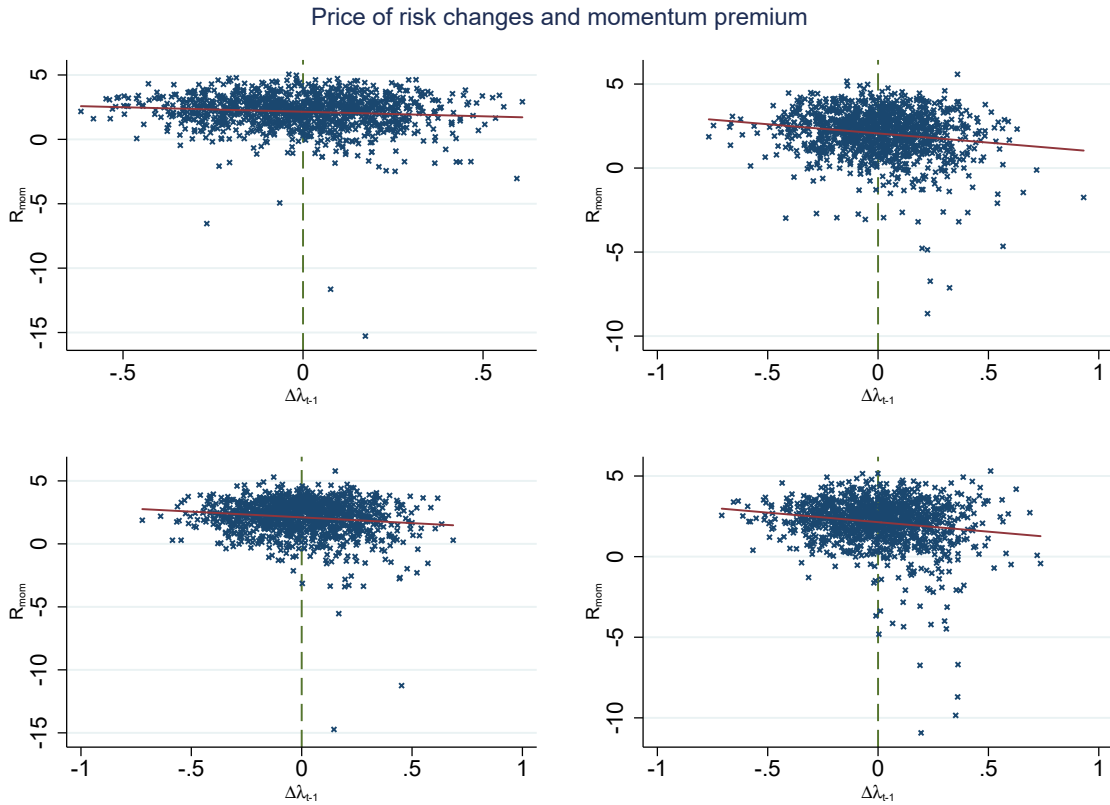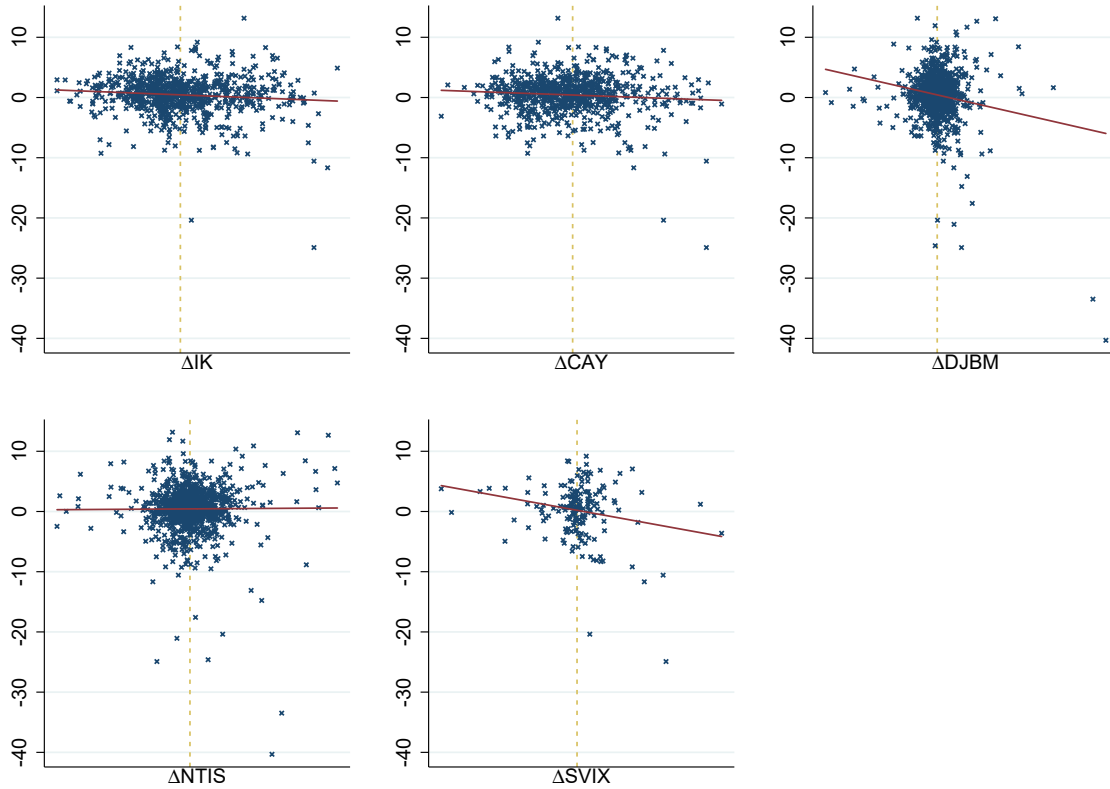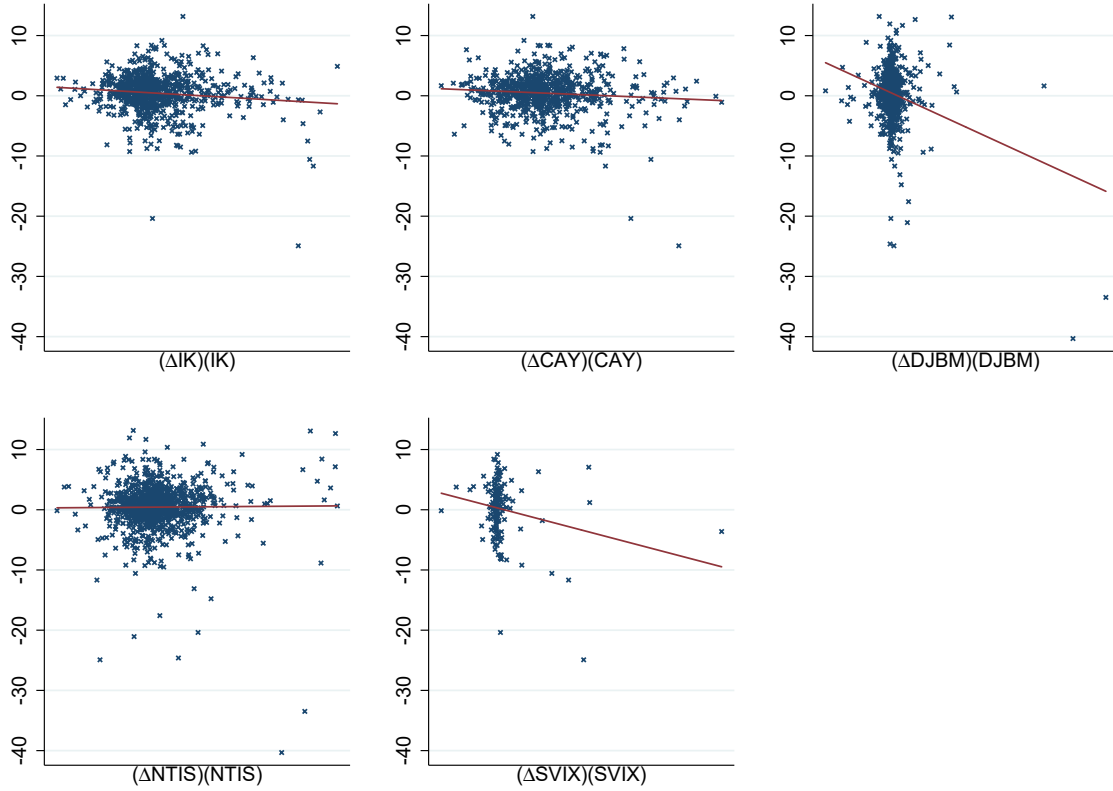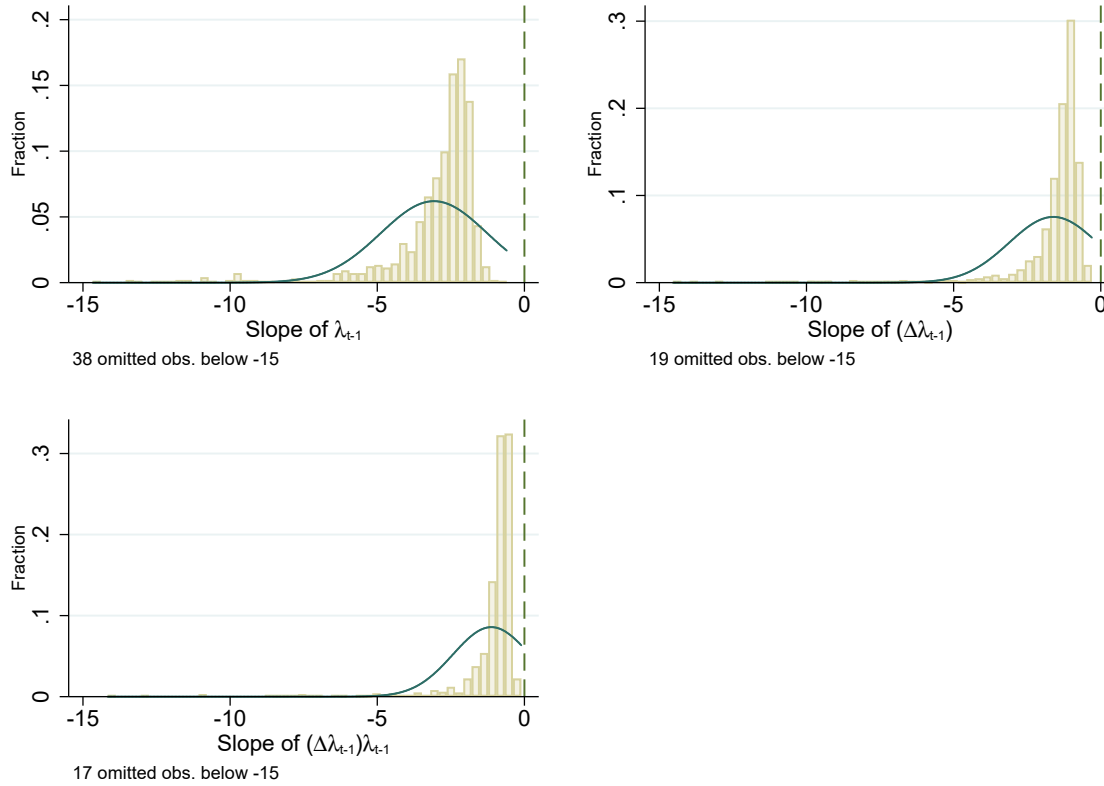
Figure 12: The return on momentum portfolios (vertical axis) as a function of the increases in the price of risk proxies. Each marker corresponds to one month, the dotted vertical yellow line corresponds to a constant price of risk during the portfolio formation (on the horizontal axis) and the red line shows the best linear fit between the variables. The increase in the price of risk is given by $\Delta\Lambda_{t-2} \equiv \Lambda_{t-2} - \Lambda_{t-12}$, where Λ_t is the price of risk proxy at time t. The price of risk proxies (with their time spans in brackets) are the (negative of the) investment-to-capital ratio (IK, 1948–2017), the consumption-wealth ratio (CAY, 1953–2017), the Dow Jones book-to-market (DJBM, 1927–2017), the (negative of the) net equity expansion (NTIS, 1927–2017), and Ian Martin's lower bound on the equity premium (SVIX, 1997–2012).

Figure 13: The return on momentum portfolios (vertical axis) as a function of the increase in the price of risk multiplied by the (lagged) price of risk proxies. Each marker corresponds to one month and the red line shows the best linear fit between the variables. The horizontal axis corresponds to the increase in the price of risk, $\Delta\Lambda_{t-2} \equiv \Lambda_{t-2} - \Lambda_{t-12}$, multiplied by the price of risk proxy at time $t-1$, Λ_{t-1}. The price of risk proxies (with their time spans in brackets) are the (negative of the) investment-to-capital ratio (IK, 1948–2017), the consumption-wealth ratio (CAY, 1953–2017), the Dow Jones book-to-market (DJBM, 1927–2017), the (negative of the) net equity expansion (NTIS, 1927–2017), and Ian Martin's lower bound on the equity premium (SVIX, 1997–2012).

Figure 14: Distribution of 1000 slopes from predictive regressions for the return on momentum portfolios, R_{mom}, given the price of risk, λ, (and its changes) based on simulated data. Starting from the top left to the right, the graphs display the distributions of the estimated slopes (β) for the following predictive regressions, respectively:

$$R_{mom,t} = \alpha + \beta \lambda_{t-1} + \varepsilon_t,$$

$$R_{mom,t} = \alpha + \beta \Delta \lambda_{t-1} + \varepsilon_t,$$

$$R_{mom,t} = \alpha + \beta (\Delta \lambda_{t-1}) \lambda_{t-1} + \varepsilon_t.$$

7 Tables

Table 1: Estimated CAPM coefficients for the momentum portfolios conditioned on realized increases or decreases in the price of risk. Both panels report the estimated coefficients, α and β, from

$$R_{mom,t} = \alpha + \beta \left(R_{m,t} - R_{f,t} \right) + \varepsilon_t,$$

where $R_{mom,t}$ is the return on the momentum portfolio, $R_{m,t}$ is the return on the market, $R_{f,t}$ is the risk-free return, and ε_t is the error term at time t. For each price of risk proxy, panel (a) reports the estimates for the months in which there was no increase in the price of risk proxy observed in the portfolio formation period, $\Delta\Lambda_{t-2} \equiv \Lambda_{t-2} - \Lambda_{t-12} \leq 0$, while panel (b) reports the coefficients for the months in which the proxy increased. The proxies, with their time spans in brackets, are the (negative of the) investment-to-capital ratio (IK, 1948–2017), the consumption-wealth ratio (CAY, 1953–2017), the Dow Jones book-to-market (DJBM, 1927–2017), the (negative of the) net equity expansion (NTIS, 1927–2017), and Ian Martin's lower bound on the equity premium (SVIX, 1997–2012). The table shows t statistics in parentheses, $^*\,p < 0.05$, $^{**}\,p < 0.01$, $^{***}\,p < 0.001$.

(a) Estimates for momentum portfolios formed in periods in which the price of risk did not increase, $\Delta\Lambda_{t-2} \leq 0$.

	(IK)	CAY	DJBM	(NTIS)	SVIX
$\beta_{\Delta\Lambda_{t-2}\leq 0}$	0.0270	0.0113	0.0755**	-0.146***	0.0977
	(0.88)	(0.33)	(2.85)	(-5.43)	(1.37)
$\alpha_{\Delta\Lambda_{t-2}\leq 0}$	0.618***	0.490***	0.592***	0.442**	0.559
	(5.10)	(3.72)	(4.90)	(3.01)	(1.86)
Months	467	431	610	541	90
R-squared	0.002	0.000	0.013	0.052	0.021

(b) Estimates for momentum portfolios formed in periods in which the price of risk increased, $\Delta\Lambda_{t-2} > 0$.

	(IK)	CAY	DJBM	(NTIS)	SVIX
$\beta_{\Delta\Lambda_{t-2}>0}$	-0.204***	-0.207***	-0.471***	-0.375***	-0.530***
	(-5.12)	(-5.15)	(-17.46)	(-12.35)	(-5.81)
$\alpha_{\Delta\Lambda_{t-2}>0}$	0.345	0.456*	0.408*	0.783***	0.155
	(1.83)	(2.33)	(2.38)	(4.80)	(0.31)
Months	363	339	472	530	93
R-squared	0.068	0.073	0.393	0.224	0.271

Table 2: Differences in average returns on momentum portfolios conditioned on realized increases or decreases in the price of risk proxies. In both panels, $R_{mom,t}$ is the return on the momentum portfolio at time t. For the regressions in panel (a), Λ_t is the price of risk proxy at time t, $I_{\Delta\Lambda_{t-2}\leq 0}$ is the indicator function, which is equal to one if $\Delta\Lambda_{t-2} \equiv \Lambda_{t-2} - \Lambda_{t-12} \leq 0$ (i.e., the price of risk proxy did not increase), α_Λ and β_Λ are the reported least squares estimates of the coefficients based on each proxy, and ε_t is the error term. Panel (b) shows the estimated momentum risk premium when the price of risk proxy does not increase. The proxies, with their time spans in brackets, are the (negative of the) investment-to-capital ratio (IK, 1948–2017), the consumption-wealth ratio (CAY, 1953–2017), the Dow Jones book-to-market (DJBM, 1927–2017), the (negative of the) net equity expansion (NTIS, 1927–2017), and Ian Martin's lower bound on the equity premium (SVIX, 1997–2012). The table shows t statistics in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

(a) $R_{mom,t} = \alpha_\Lambda + \beta_\Lambda I_{\Delta\Lambda_{t-2}\leq 0} + \varepsilon_t$

	(IK)	CAY	DJBM	(NTIS)	SVIX
$I_{\Delta\Lambda_{t-2}\leq 0}$	0.453*	0.148	0.474*	-0.146	0.787
	(2.09)	(0.64)	(2.01)	(-0.62)	(1.19)
α_Λ	0.178	0.350*	0.177	0.509**	-0.215
	(1.09)	(2.02)	(1.00)	(3.03)	(-0.46)
Months	830	770	1082	1071	183

(b) $E\left[R_{mom,t} \mid \Delta\Lambda_{t-2} \leq 0\right]$

	(IK)	CAY	DJBM	(NTIS)	SVIX
$\mu_{\Delta\Lambda_{t-2}\leq 0}$	0.631***	0.497***	0.651***	0.363*	0.571
	(5.25)	(3.84)	(5.44)	(2.42)	(1.89)
Months	467	431	610	541	90

Table 3: Predictive regressions of the returns on momentum portfolios based on changes in and/or lagged values of the price of risk proxies. The predictive regressions are displayed in each panel, where $R_{mom,t}$ is the return on momentum portfolios at time t, Λ_t is the price of risk proxy at time t, $\Delta\Lambda_{t-2} \equiv \Lambda_{t-2} - \Lambda_{t-12}$ is the change in the price of risk proxy during the portfolio formation period, α_Λ and β_Λ are the reported least squares estimates of the coefficients based on each proxy, and ε_t is the error term. The proxies, with their time spans in brackets, are the (negative of the) investment-to-capital ratio (IK, 1948–2017), the consumption-wealth ratio (CAY, 1953–2017), the Dow Jones book-to-market (DJBM, 1927–2017), the (negative of the) net equity expansion (NTIS, 1927–2017), and Ian Martin's lower bound on the equity premium (SVIX, 1997–2012). The table shows t statistics in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

(a) $R_{mom,t} = \alpha_\Lambda + \beta_\Lambda \Delta\Lambda_{t-2} + \varepsilon_t$

	(IK)	CAY	DJBM	(NTIS)	SVIX
$\Delta\Lambda_{t-2}$	-139.0**	-25.18*	-4.664***	1.374	-18.47**
	(-2.87)	(-2.31)	(-5.81)	(0.23)	(-2.93)
α_Λ	0.437***	0.414***	0.439***	0.434***	0.205
	(4.07)	(3.60)	(3.80)	(3.67)	(0.63)
Months	830	770	1082	1071	183
R-squared	0.010	0.007	0.030	0.000	0.045

(b) $R_{mom,t} = \alpha_\Lambda + \beta_\Lambda \Lambda_{t-1} + \varepsilon_t$

	(IK)	CAY	DJBM	(NTIS)	SVIX
Λ_{t-1}	-61.36*	-1.213	-1.006*	-15.82***	-19.69*
	(-2.05)	(-0.26)	(-2.28)	(-3.47)	(-2.51)
α_Λ	0.991***	0.514	0.900***	2.969***	0.933*
	(3.46)	(1.63)	(3.89)	(4.03)	(2.13)
Months	840	780	1081	1081	193
R-squared	0.005	0.000	0.005	0.011	0.032

(c) $R_{mom,t} = \alpha_\Lambda + \beta_\Lambda (\Delta\Lambda_{t-2})\Lambda_{t-1} + \varepsilon_t$

	(IK)	CAY	DJBM	(NTIS)	SVIX
$(\Delta\Lambda_{t-2})\Lambda_{t-1}$	-15867.4***	-400.3*	-6.766***	11.76	-164.5**
	(-3.38)	(-2.48)	(-8.72)	(0.31)	(-3.07)
α_Λ	0.483***	0.447***	0.516***	0.433***	0.349
	(4.46)	(3.90)	(4.54)	(3.64)	(1.05)
Months	829	769	1081	1070	182
R-squared	0.014	0.008	0.066	0.000	0.050

References

Asness, Clifford S., Tobias J. Moskowitz, and Lasse H. Pedersen, 2013, Value and momentum everywhere, *Journal of Finance* 68, 929–985.

Avramov, Doron, and Tarun Chordia, 2006, Asset pricing models and financial market anomalies, *Review of Financial Studies* 19, 1001–1040.

Barroso, Pedro, 2014, The bottom-up beta of momentum, *Unpublished working paper. 29th Australasian Finance and Banking Conference 2016* .

Barroso, Pedro, and Pedro Santa-Clara, 2015, Momentum has its moments, *Journal of Financial Economics* 116, 111–120.

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

Berk, Jonathan B., Richard C. Green, and Vasant Naik, 1999, Optimal investment, growth options, and security returns, *Journal of Finance* 54, 1553–1607.

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.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.

Chen, Hui, Winston Wei Dou, and Leonid Kogan, 2015, Measuring the "dark matter" in asset pricing models, *Unpublished working paper* .

Cochrane, John H., 1991, Production-based asset pricing and the link between stock returns and economic fluctuations, *Journal of Finance* 46, 209–237.

Cochrane, John H., 2011, Presidential Address: Discount rates, *Journal of Finance* 66, 1047–1108.

Conrad, Jennifer, and Gautam Kaul, 1998, An anatomy of trading strategies, *Review of Financial Studies* 11, 489–519.

Daniel, Kent, and Tobias J. Moskowitz, 2016, Momentum crashes, *Journal of Financial Economics* 122, 221–247.

Griffin, John M., Xiuqing Ji, and Spencer J. Martin, 2003, Momentum investing and business cycle risk: Evidence from pole to pole, *Journal of Finance* 58, 2515–2547.

Grundy, Bruce D., and Spencer J. Martin, 2001, Understanding the nature of the risks and the source of the rewards to momentum investing, *Review of Financial Studies* 14, 29–78.

Hou, Kewei, Chen Xue, and Lu Zhang, 2015, Digesting anomalies: An investment approach, *Review of Financial Studies* 28, 650–705.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65–91.

Jegadeesh, Narasimhan, and Sheridan Titman, 2011, Momentum, *Annual Review of Financial Economics* 3, 493–509.

Johnson, Timothy C., 2002, Rational momentum effects, *Journal of Finance* 57, 585–608.

Kothari, S.P., and Jay Shanken, 1992, Stock return variation and expected dividends: A time-series and cross-sectional analysis, *Journal of Financial Economics* 31, 177–210.

Lettau, Martin, and Sydney Ludvigson, 2001, Consumption, aggregate wealth, and expected stock returns, *Journal of Finance* 56, 815–849.

Lintner, John, 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *Review of Economics and Statistics* 47, 13–37.

Liu, Laura Xiaolei, and Lu Zhang, 2008, Momentum profits, factor pricing, and macroeconomic risk, *Review of Financial Studies* 21, 2417–2448.

Lucas, Robert E., 1972, Expectations and the neutrality of money, *Journal of Economic Theory* 4, 103–124.

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

Muth, John F., 1961, Rational expectations and the theory of price movements, *Econometrica* 29, 315–335.

Novy-Marx, Robert, 2015, How can a q-theoretic model price momentum?, Working Paper 20985, National Bureau of Economic Research.

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

Sharpe, William F., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425–442.

Souza, Thiago de O., 2018, Size-related premiums, *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.