

Long-term discount rates do not vary across firms*

Matti Keloharju

Juhani T. Linnainmaa

Peter Nyberg

July 2019

Abstract

Long-term expected returns appear to vary little, if at all, in the cross section of stocks. We devise a bootstrapping procedure that injects small amounts of variation into expected returns and show that even negligible differences in expected returns, if they existed, would be easy to detect. Markers of such differences, however, are absent from actual stock returns. Our estimates are consistent with behavioral models and production-based asset pricing models such as Berk, Green, and Naik (1999) and Gomes, Kogan, and Zhang (2003) in which firms' risks change over time. We show that long-term reversals in stock returns are the consequence of the rapid convergence in expected returns. Our results imply stock market anomalies have only a limited effect on firm valuations.

JEL classification: G12; G31

*Keloharju is with Aalto University School of Business, CEPR, and IFN. Linnainmaa is with the Dartmouth College and NBER. Nyberg is with the Aalto University School of Business. We thank Jonathan Berk, John Cochrane, Mathias Kronlund, Jon Lewellen, Yan Liu (discussant), Anders Löflund, Owen Lamont, and conference and seminar participants at Florida International University, Lund University, Tinbergen Institute at Amsterdam, University of Illinois at Chicago, University of Illinois at Urbana-Champaign, University of Miami, University of Michigan, LA Finance Day, and Society of Financial Studies 2018 Cavalcade for insightful comments and Jonathan Berk, Timothy Johnson, and Lu Zhang for providing the code for the models simulated in Sections 2 and 5. Linnainmaa and Nyberg are also affiliated with Citadel and Linnainmaa with Research Affiliates. Neither Citadel nor Research Affiliates provided any funding for this research. Financial support from the Academy of Finland and Nasdaq Nordic Foundation is gratefully acknowledged.

1 Introduction

Time-varying risks lie at the heart of production-based asset pricing models. Firms' risks can change either because their asset bases change (Berk, Green, and Naik 1999; Gomes, Kogan, and Zhang 2003) or because firms face systematic productivity shocks and costly reversibility of investment (Zhang 2005; Cooper 2006). If risks change but are stationary, as they are in these models, today's high-risk firms are expected to become less risky and low-risk firms are expected to become riskier. As risks converge, so do expected returns.

The prediction that differences in expected returns die out is not specific to rational models; it is, in fact, inherent to any behavioral explanation for asset pricing anomalies. If a stock's expected return is too high because the stock is underpriced, the return can remain elevated only for as long as the mispricing persist. If, for example, a stock is underpriced by 50% relative to its fair value and its expected return is consequently 10% too high, the mispricing would vanish in 6.9 years.¹ The prediction that expected returns converge is therefore very general. In risk-based models this prediction requires only that the risks that firms face change over time; in behavioral models, it is true by assumption.

In this paper, we measure the extent to which the data align with this prediction of behavioral and production-based asset pricing models. We first show the data strongly support the idea of cross-sectional differences in expected returns converging to zero over time. Even if a stock's expected return today lies far above or below the average, we cannot reject the null that its expected return after five years equals the average. Second, we show the convergence in expected returns is commensurate with

¹Black (1986) writes “[w]e might define an efficient market as one in which price is within a factor of 2 of value. . . it seems reasonable to me, in the light of sources of uncertainty about value and the strength of the forces tending to cause price to return to value. By this definition, I think almost all markets are efficient almost all of the time. ‘Almost all’ means at least 90%.” In the foregoing example, a stock's price should be P and it should earn an expected return of r . If the stock's expected return exceeds r by δ then, with continuous compounding, the mispricing will vanish in t years, with t satisfying $Pe^{rt} = \frac{1}{2}Pe^{(r+\delta)t}$. With $\delta = 0.1$, $t = -\log(\frac{1}{2})/x = 6.93$ years.

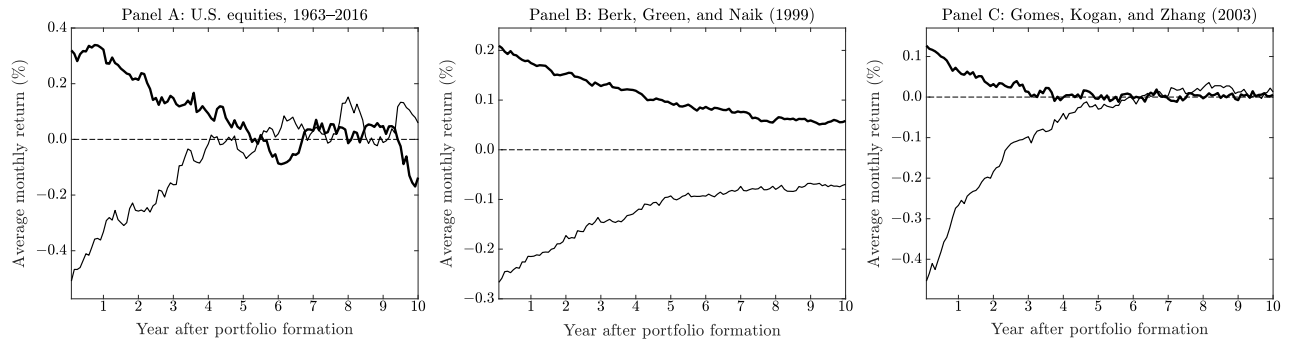


Figure 1: **Average monthly returns on stocks sorted by expected returns.** Panel A assigns U.S. stocks into deciles based on a combination of 34 return predictors that use accounting information (see text for details) and plots the average (non-cumulative) monthly returns for the top and bottom deciles over the next ten years. Panels B and C simulate data from Berk et al.’s (1999) and Gomes et al.’s (2003) models using cross-sectionally demeaned returns. The number of stocks and time periods in the simulations match the average number of stocks and number of months in the U.S. data.

that in risks, regardless of whether we measure them using betas or firm characteristics. Thus, today’s high-risk firms indeed become less risky and low-risk firms become riskier.

In Panel A of Figure 1, we use U.S. stock data and assign stocks into deciles each month from 1963 through 2016 based on estimates of stocks’ expected returns. In this computation, which we detail below, we combine 34 return predictors into a proxy for firms’ expected returns. We report average *noncumulative* returns for the top and bottom deciles over the next ten years; that is, the month- t estimate is the average return in month t after portfolio formation, not the average return from today to month t . Panels B and C illustrate the extent to which the data conform to two models, Berk, Green, and Naik (1999) and Gomes, Kogan, and Zhang (2003). We simulate data from these models using the original papers’ parameters and rank firms into deciles each month based on their expected returns. In both the simulations and actual data, differences in average returns are initially large but collapse to zero as we extend the maturity.² Expected returns behave differently in the two models. In Berk, Green, and Naik (1999), the cross-sectional distribution of expected returns is almost symmetric and

²We compute averages of cross-sectionally demeaned returns to remove market-level variation, which is why the averages in Figure 1 converge towards zero. Averages of non-demeaned returns would converge towards the average return of the equal-weighted market portfolio.

the differences decay slowly; in Gomes, Kogan, and Zhang (2003), the distribution is left-skewed and the differences decay faster.

We illustrate our key finding first with a characteristics-free bootstrap procedure. We take the cross section of monthly stock returns from 1963 through 2016, preserve both the covariance structure and distribution of returns, but set all expected returns to zero. We then inject small cross-sectional differences into expected returns and measure our ability to detect these differences.

If differences in expected returns persist, past stock returns positively predict the cross section of stock returns. To see why, note that a stock’s return, r_{it} , here consists of a constant term, μ_i , plus noise, ε_{it} :

$$r_{it} = \mu_i + \varepsilon_{it}. \tag{1}$$

A stock’s average return therefore measures its expected return:

$$\bar{r}_{itn} = \frac{1}{n} \sum_{j=1}^n r_{i,t-j} = \underbrace{\mu_i}_{\text{expected return}} + \underbrace{\frac{1}{n} \sum_{j=1}^n \varepsilon_{i,t-j}}_{\text{average noise term}}. \tag{2}$$

A cross-sectional regression of stock returns against past average returns is a regression of returns against noisy estimates of expected returns.³ We show returns would be highly predictable even if the persistent differences in expected returns were negligible. Suppose, for example, the cross-sectional standard deviation of the expected monthly stock returns is just 0.4%. This number is small, as it would imply that differences in expected returns would account for just $R^2 = 0.06\%$ of the cross-sectional

³Conrad and Kaul (1998) suggest this mechanism—that the cross section of realized returns measure differences in expected returns—may account for some of the momentum in stock returns: “The repeated purchase of winners from the proceeds of the sale of losers will, on average, be tantamount to the purchase of high-mean securities from the sale of low-mean securities. Consequently, as long as there is some cross-sectional dispersion in the mean returns of the universe of securities, a momentum strategy will be profitable.” See, also, Berk, Green, and Naik (1999, p. 1584).

variance of realized stock returns.⁴ Our bootstrap procedure shows that, among all-but-microcaps, the prior five-year return (ten-year return) would nevertheless predict the cross section of returns with a t -value of 2.03 (3.93). In actual data, however, past returns do not positively predict returns outside the prior one-year period (momentum). In the five-year regression, for example, the t -value is -3.72 .

Why do past returns predict the cross section of returns with a negative sign? We show that production-based asset pricing models predict a positive association between past and future returns *only* if the differences in discount rates are persistent enough. Long-term reversals emerge when (1) there are large differences in expected returns that (2) rapidly converge to zero. Under these conditions, past returns predominantly measure the direction of discount rate shocks. For example, a stock with a positive past return likely is a stock hit with negative discount rate shocks—*not* one whose expected return was high. This interpretation for the long-term reversals is very different from the overreaction explanation proposed by De Bondt and Thaler (1985). The existence of long-term reversals does not imply that high-return stocks become low-return stocks, and vice versa—in the risk-based models, both the high- and low-return firms merely become *average*-return firms. The shocks to expected returns show up in stock returns as discount rate shocks and, when these shocks are frequent and large enough, they induce a significant negative association between past and future returns. The existence of long-term reversals is consistent with our thesis: long-term discount rates do not vary across stocks.

As in Figure 1, we complement our bootstrapping results by sorting stocks into portfolios by different combinations of return predictors. Many variables predict the cross section of stock returns,⁵ but most are short-lived. Although some predict returns a few years out, none of them identify long-term differences in expected returns. We show this property extends to combinations of predictors as well.

⁴The average cross-sectional variance of U.S. stock returns between 1963 and 2016 is 0.0286. If the cross-sectional standard deviation of expected returns is 0.4%, this variation in expected returns would account for $\frac{(0.004)^2}{0.0286} = 0.06\%$ of the cross-sectional variation in stock returns; or, conversely, 99.94% of the cross-sectional variation in monthly stock returns would be unrelated to persistent differences in expected returns.

⁵See, for example, Hou, Xue, and Zhang (2015), McLean and Pontiff (2016), and Harvey, Liu, and Zhu (2016).

When we sort stocks into portfolios by a combination of 46 predictors, the average return difference between the top and bottom deciles is 0.74% in the year following portfolio formation (t -value = 4.48), but in the second year, the return difference is just 0.25% with a t -value of 1.55. Even when we select ex-post the most persistent predictors—which turn out to be those that use accounting rather than price, volume, and return information—the combination of these predictors loses its predictive power after approximately three years.

The long-run return estimates are precise enough to bound the amount of cross-sectional variation in expected returns. Consider, for example, the CAPM alphas. In the year following portfolio formation, the 95% confidence interval for the CAPM alpha for the difference between the top and bottom deciles runs from 77 basis points to 129 basis points. In year eight, this confidence interval runs from -14 basis points to 31 basis points. We can thus reliably identify significant differences in short-term expected returns but, as Figure 1 shows, these differences evaporate quickly.

In production-based asset pricing models, expected returns converge because risks mean revert. If there are persistent differences in characteristics, they cannot associate with differences in expected returns if there are no persistent differences in expected returns to begin with. Whether this prediction holds in the data is not a foregone conclusion in the light of existing literature, which generally does not differentiate between persistent and transitory sources of risk. For example, in his Presidential Address to the American Finance Association, Cochrane (2011) points out that “[A]n implicit assumption underlies everything we do: expected returns, variances, and covariances are *stable* functions of characteristics such as size and book-to-market ratio” (p. 1062, emphasis ours). When further discussing the valuation effects of various return predictors, he suggests that “long-lasting characteristics are likely to be more important” than short-run predictors such as momentum which lasts less than a year (p. 1064).

We first study whether risks mean revert and whether there are persistent differences in risk. Without taking a stance on whether betas or characteristics better measure risk, we show both converge in expectation toward the mean. However, whereas differences in average returns collapse to zero in five years, those in betas and characteristics do not.

We then test whether persistent differences in characteristics associate with differences in expected returns. To do that, we decompose firm characteristics into permanent and transitory components. We find that the transitory differences in characteristics command premiums and discounts, and these differences vanish in approximately five years. However, persistent cross-sectional differences in characteristics are not associated with any differences in average returns.

These results tie back to those in Cohen and Polk (1996), Asness, Porter, and Stevens (2000), and Novy-Marx (2013), who find that industry-level (and therefore persistent) differences in characteristics such as value are not associated with differences in average returns. We generalize this result by showing that persistent differences in characteristics, regardless of the source, do not correlate with differences in expected returns.

Existing literature documents sizable short-term differences in expected returns, sometimes assuming any differences to persist for a long time.⁶ We contribute to the literature by showing a wide gap between short-term and long-term returns—and little evidence of persistent differences in expected returns between firms. Because any differences in expected returns are short-lived, they carry little weight in the average discount rate and in firm valuation.⁷ These results are consistent with financial markets valuing the equity of different firms using approximately the same discount rate.

⁶For example, Martin and Wagner (2019) estimate that “there is considerably more variation in expected returns... than has previously been acknowledged.” MacKinlay and Pastor (2000) and DeMiguel, Garlappi, and Uppal (2009) model portfolio selection strategies assuming the moments of returns are constant over time. Kan and Zhou (2007) estimate expected returns using data for up to 480 months.

⁷If a firm’s discount rate is r over the first year and r' thereafter, its k -year discount rate is $[(1+r)(1+r')^{k-1}]^{1/k} - 1$. Differences in short-run discount rates carry over to long-run discount rates, but increasing the horizon k dilutes their role.

Although the data are consistent with the predictions of production-based asset pricing models, we cannot rule out other theories. Consumption-based asset pricing models are the flip side of production-based asset pricing models (Cochrane 1991). If firms’ covariances with consumption growth change over time and their risks are stationary (i.e., they mean revert in the cross section), firms’ expected returns will, on average, converge toward the mean. The data also do not shut out behavioral theories. If differences in expected returns reflect mispricing, and arbitrageurs profit by trading against mispricing, cross-sectional differences in expected returns can be expected to vanish over time. What the data reject are models in which differences in expected returns persist to a significant degree.

Assuming all anomalies reflect mispricing gives insight into mispricing’s potential economic significance. As van Binsbergen and Opp (2019) point out, more persistent anomalies are likely to generate greater distortions in the economy. An anomaly that generates an expected return of 10% per annum and lasts for a year causes prices to be wrong by 10%; an otherwise similar anomaly lasting only for one month causes prices to be wrong by less than one percent. If most anomalies are short-lived—as our estimates suggest—the stock market can be inefficient in returns but close to efficient in prices. Its perceived efficiency depends on the investment horizon. An arbitrageur could reap great rewards by trading anomalies but, at the same time, the market would be close to efficient to a buy-and-hold investor. This conclusion parallels that of Cohen, Polk, and Vuolteenaho (2009): “price is (almost) right.”

2 Expected returns in production-based asset pricing models

Many production-based asset pricing theories seek to explain cross-sectional return patterns associated with, for example, size and book-to-market. These models feature time-varying risks. In this section, we describe and simulate data from four models. We show that—and explain why—cross-

sectional firm-level differences in expected returns in these models tend to vanish.

2.1 Models and mechanisms

Model 1: Berk, Green, and Naik (1999): assets in place versus growth options. Firms encounter new projects each period and they accept those with positive NPVs. Projects differ in their amount of systematic risk, and old projects turn obsolete at random. Because new projects are drawn from the same distribution, firms are asymptotically identical. In other words, high-risk firms will, on average, encounter projects that lower their risk, and low-risk firms will tend to encounter projects that increase their risk. Firm valuation in the model buttresses this point. A firm's value is the sum of the value of the assets in place and the value of future growth options. Of these two terms, the value of the future growth options is the same across firms, because all firms expect to encounter the same projects.

Model 2: Gomes, Kogan, and Zhang (2003): general equilibrium with growth options. Gomes, Kogan, and Zhang (2003) take the investment mechanism of Berk, Green, and Naik (1999) to general equilibrium. Whereas Berk et al. assume the process describing the pricing kernel, Gomes et al. model the household sector and let the markets clear. Returns in this model are completely described by a conditional CAPM; size and book-to-market predict returns because they correlate with the true conditional market betas. Also similar to Berk et al., new projects are distributed randomly across all firms with equal probabilities. Therefore, all firms derive the same value from future growth options, and expected returns perfectly converge over time.

Model 3: Zhang (2005) and Lin and Zhang (2013): production shocks and costly reversibility of investment. Production in this model requires capital, and firm-level productivity is subject to aggregate and idiosyncratic shocks. Firms have to pay to install new capital and to adjust it. The pricing kernel is parametrized directly to have a countercyclical price of risk. Productivity shocks

alter firms' riskiness; in bad times, when the price of risk is high, low-productivity firms find it costly to shed unproductive capital. Because productivity shocks mean revert, firms' expected returns converge over time.

Model 4: Hackbarth and Johnson (2015) and Gu, Hackbarth, and Johnson (2017): operating leverage and real options. Similar to Zhang (2005), production in this model requires capital, and productivity is subject to aggregate and idiosyncratic shocks. Firms face both quasi-fixed and variable costs for both upward and downward adjustments to capital. The random shocks cause risks and, by extension, expected returns, to mean revert.

2.2 Changes in expected returns

We simulate 1,000 months of return data from the models described above using the parameters used in the original studies; when a study considers multiple sets of parameters, we use those of the baseline specification. Studies use different methods to choose the parameters. They are typically "fixed" based on prior literature, directly estimated, or calibrated to match some features of the data, such as the levels and volatilities of equity premium and interest rate. We discard the first 400 months of data to ensure the simulations stabilize.

We assign stocks into deciles based on expected returns at the end of month t , and compute average returns for stocks in these deciles over the next 15 years. We then average the estimates over all starting months t . We repeat each simulation 1,000 times to reduce simulation-specific noise; this is different from the graphs in the introduction's Figure 1 in which we plot the data from single runs.

Figure 2 shows average, cross-sectionally demeaned monthly returns from the four models discussed above. The models differ in the amount of dispersion in expected returns. In Berk, Green, and Naik (1999), for example, the difference in expected monthly returns between the top and bottom deciles

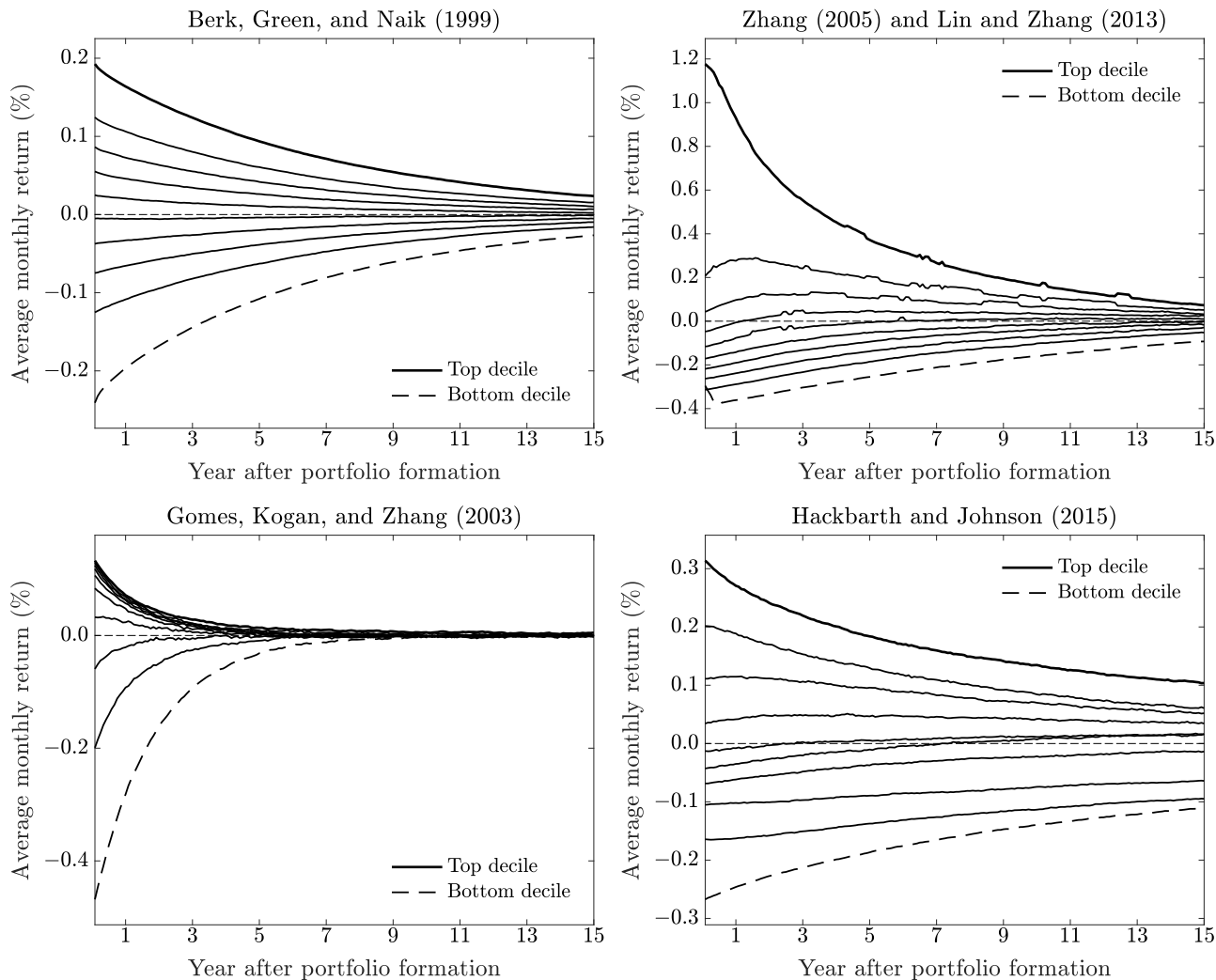


Figure 2: **Average monthly returns on stocks sorted by expected returns.** We simulate 1,000 months of return data from the four models described in Section 2. We run these simulations using the same parameters as those used in the original studies. We discard the first 400 months and then begin ranking stocks into deciles based on expected returns. We report the average cross-sectionally demeaned monthly returns for these deciles over the next 15 years after portfolio formation.

is just over 40 basis points; in Zhang (2005), this difference is over 140 basis points. In some models, such as Berk, Green, and Naik (1999), the cross-sectional distribution of expected returns is nearly symmetric; in others, such as Gomes, Kogan, and Zhang (2003), it is considerably left-skewed. The common element of these models, however, is the convergence in expected returns. Although the models differ in the speed of convergence—they model different economic mechanisms, and they are

parametrized differently—expected returns converge in all of them.⁸ Figure A1 in the appendix shows that as expected returns converge toward the mean, so do market betas.

The pattern in Figure 2 is not specific to the models we consider. If risks change but are stationary, they must mean-revert. Moreover, unless a model builds in *permanent* cross-sectional differences in, for example, production technology, each firm’s risk (and its expected return) must be expected to converge toward the common mean. We are not aware of any study on production-based asset pricing models that has found the need to build in persistent differences across firms.⁹

3 Data

We use the daily and monthly CRSP return data from January 1963 through December 2016 on stocks listed on the NYSE, Amex, and NASDAQ. We exclude securities other than ordinary common shares. We also exclude financials, which are identified as firms with SIC codes between 6000 and 6999. We use CRSP delisting returns; if a delisting return is missing and the delisting is performance-related, we impute a return of -30% for NYSE and Amex stocks (Shumway 1997) and -55% for Nasdaq stocks (Shumway and Warther 1999).

We use balance sheet and income statement information from the annual and quarterly Compustat files to construct various return predictors that have been proposed in the literature. We describe these predictors in Section 6.1.

⁸For example, Hackbarth and Johnson (2015, Table 2) impose significantly more firm-level persistence than the other models or what is seen in the data. In their baseline model, the autocorrelation in profitability is 0.97, whereas it is 0.51 in the data.

⁹In the risk-based models, expected returns converge because of discount rate shocks. Although the market rationally expects expected returns to converge, every discount rate shock comes as a surprise. In Berk, Green, and Naik (1999), for example, a firm’s expected return changes when the firm’s old project (randomly) dies off or it (randomly) encounters a new project it chooses to undertake. Expected returns could also converge deterministically—an asset’s expected return could, for example, be 10%, 9%, 8%, . . . over its life—and the market would discount the cash flows given these expected changes. We use the term “convergence in expected returns” to refer to the convergence induced by discount rate shocks.

4 Detecting differences in long-term discount rates: A bootstrap approach

4.1 Cross-sectional regressions

In this section, we estimate regressions that predict the cross section of stock returns with each stock's average past return. For each month t , we estimate a cross-sectional regression

$$r_{it} = a_t + b_t \times \bar{r}_{i,t-k_1,t-k_2} + e_{it}, \quad (3)$$

where $\bar{r}_{i,t-k_1,t-k_2}$ is stock i 's average return from month $t - k_2$ to $t - k_1$. If stock returns contain persistent differences in expected returns, as in equation (1), the slope estimate from these regressions is proportional to cross-sectional variance of these differences, $\hat{b} \sim \hat{\sigma}_\mu^2$. For example, if expected returns are constant and return innovations IID, then

$$\hat{b}_t = \frac{\text{cov}^{\text{cs}}(r_{it}, \bar{r}_{i,t-k_1,t-k_2})}{\text{var}^{\text{cs}}(\bar{r}_{i,t-k_1,t-k_2})} = \frac{\text{cov}^{\text{cs}}(\mu_i + \varepsilon_{i,t}, \mu_i + \frac{1}{k_2-k_1+1} \sum_{t'=t-k_2}^{t-k_1} \varepsilon_{i,t'})}{\text{var}^{\text{cs}}(\mu_i + \frac{1}{k_2-k_1+1} \sum_{t'=t-k_2}^{t-k_1} \varepsilon_{i,t'})} = \frac{\hat{\sigma}_\mu^2}{\hat{\sigma}_\mu^2 + \frac{1}{k_2-k_1+1} \hat{\sigma}_\varepsilon^2} \geq 0. \quad (4)$$

We estimate these regressions using actual and bootstrapped stock return data. We construct the bootstrapped data to preserve both the distributions of returns and their covariance structure.

4.2 Methodology

We generate each draw of simulated data in five steps:

1. We draw each stock's expected return from a normal distribution with a mean of zero and a standard deviation σ_μ . We denote stock i 's expected return by μ_i .

2. We cross-sectionally demean month- t returns and divide them by their standard deviation to generate a vector of studentized residuals with unit variance. We denote this $N_t \times 1$ vector of residuals by ϵ_t , where N_t is the number of stocks in month t .
3. We estimate the $N_t \times N_t$ covariance matrix of stock returns, Σ_t , using monthly data from month $t - 30$ to month $t + 30$, that is, five years plus one month. We estimate pairwise covariances; that is, we do not require all stocks to have non-missing returns for the entire estimation period. Because the number of stocks exceeds the length of the time series, Σ_t is singular.
4. We replace the covariance matrix Σ_t with its nearest positive definite matrix using the algorithm of Higham (2002). We denote this positive definite matrix by S_t .
5. We generate month- t stock-specific shocks ε_t by randomizing the elements of ϵ_t and by post-multiplying them by the Cholesky factor of S_t . Stock i 's return in month t is then $\mu_i + \varepsilon_{it}$.

These simulated returns have appealing properties. First, the resulting data matrix has the same dimensions as the actual return data; that is, it has the same number of months and the same number of stocks each month. Second, the factor structure of returns is the same as that in the actual data and, given the rolling estimates of covariances, this factor structure changes over time as it does in the data. Third, because we extract the shocks from actual stock returns, the distribution of stock returns each month resembles the actual distribution of returns.

We draw 100 random samples of returns using this procedure. Using each simulated dataset, we estimate the cross-sectional regressions and record the average regression slopes and their t -values. We then compare these estimates with those from the actual data, and examine how the estimates change as we increase the amount of variation in expected returns, σ_μ .

Table 1: Fama-MacBeth regressions with actual and bootstrapped data

This table reports average coefficients and t -values from regressions to predict the cross section of monthly stock returns. The explanatory variable is the stock’s average return over the window specified in the first column. Column “Actual data” uses returns on common stocks listed on NYSE, Amex, and NASDAQ from 1963 through 2016. The sample excludes financials, which are identified as firms with SIC codes between 6000 and 6999. Columns “Bootstrapped data” use data that are first randomized to set the variation in expected returns to zero. These data are generated in five steps. First, month- t returns are demeaned and divided by their standard deviation to generate a vector of residuals, ϵ_t , with unit variance. Second, the covariance matrix of individual stock returns, Σ_t , is estimated using data from 30 months prior to 30 months after month t . Third, this covariance matrix is transformed to the nearest positive definite matrix S_t using the algorithm of Higham (2002). Fourth, month- t returns are generated by randomizing the elements of ϵ_t and post-multiplying this vector by the Cholesky factor of S_t . Fifth, μ_i is added to stock i ’s return each month, where μ_i s are drawn from a distribution with a mean of zero and a standard deviation of σ_μ . The simulations vary the σ_μ from 0 to 1.4%. We repeat the bootstrapping procedure 100 times for each value of σ_μ ; this table reports the average coefficients and t -values across these simulations. Panel A uses data on all stocks; Panel B uses all-but-microcaps. All-but-microcaps are stocks with market capitalizations above the 20th percentile in the NYSE distribution.

Panel A: All stocks

Historical return horizon	Actual data	Bootstrapped data							
		Cross-sectional variation in expected returns, σ_μ							
		0.0%	0.2%	0.4%	0.6%	0.8%	1.0%	1.2%	1.4%

Fama-MacBeth coefficient estimates

$[-1, -1]$	-0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01
$[-12, -2]$	0.04	0.00	0.00	0.01	0.02	0.03	0.04	0.06	0.08
$[-60, -13]$	-0.17	-0.00	0.00	0.03	0.07	0.11	0.16	0.22	0.27
$[-120, -61]$	-0.22	-0.00	0.02	0.08	0.16	0.24	0.32	0.40	0.47
$[-120, -2]$	-0.28	-0.01	0.02	0.07	0.15	0.23	0.31	0.38	0.45
$[-240, -121]$	-0.24	-0.00	0.05	0.17	0.30	0.42	0.52	0.60	0.66
$[-240, -2]$	-0.29	-0.01	0.05	0.17	0.30	0.41	0.51	0.59	0.66

t -values

$[-1, -1]$	-12.12	0.05	0.07	0.15	0.29	0.48	0.72	1.02	1.36
$[-12, -2]$	2.54	0.12	0.21	0.48	0.94	1.58	2.37	3.34	4.46
$[-60, -13]$	-6.43	-0.15	0.15	1.02	2.46	4.36	6.68	9.38	12.42
$[-120, -61]$	-4.85	-0.09	0.52	2.18	4.77	8.07	11.95	16.45	21.45
$[-120, -2]$	-7.17	-0.19	0.48	2.33	5.18	8.82	13.06	17.97	23.39
$[-240, -121]$	-3.59	-0.11	1.11	4.24	8.65	14.10	20.38	27.78	35.82
$[-240, -2]$	-4.85	-0.18	1.17	4.62	9.49	15.45	22.21	30.20	39.00

Panel B: All-but-microcaps

Historical return horizon	Actual data	Bootstrapped data							
		Cross-sectional variation in expected returns, σ_μ							
		0.0%	0.2%	0.4%	0.6%	0.8%	1.0%	1.2%	1.4%
Fama-MacBeth coefficient estimates									
[-1, -1]	-0.02	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.02
[-12, -2]	0.10	0.00	0.01	0.02	0.04	0.06	0.09	0.12	0.15
[-60, -13]	-0.11	-0.00	0.01	0.06	0.12	0.20	0.28	0.35	0.42
[-120, -61]	-0.10	-0.00	0.04	0.13	0.25	0.37	0.47	0.55	0.62
[-120, -2]	-0.17	-0.00	0.03	0.12	0.24	0.35	0.45	0.53	0.60
[-240, -121]	-0.12	-0.01	0.07	0.24	0.41	0.54	0.64	0.71	0.77
[-240, -2]	-0.18	-0.01	0.07	0.23	0.40	0.53	0.63	0.71	0.77
t-values									
[-1, -1]	-3.51	0.04	0.10	0.27	0.56	0.95	1.46	2.08	2.81
[-12, -2]	4.88	0.10	0.29	0.86	1.81	3.11	4.76	6.71	8.95
[-60, -13]	-3.72	-0.14	0.41	2.03	4.62	8.09	12.23	16.98	22.30
[-120, -61]	-1.76	-0.04	0.95	3.76	8.09	13.60	20.09	27.42	35.58
[-120, -2]	-3.58	-0.07	0.96	3.93	8.48	14.28	21.06	28.70	37.19
[-240, -121]	-1.59	-0.18	1.61	6.44	13.44	22.00	32.12	43.51	55.51
[-240, -2]	-2.67	-0.25	1.66	6.77	14.13	23.08	33.53	45.23	57.78

4.3 Estimates

Table 1 reports results from cross-sectional regressions that we estimate using both actual and simulated data. Panel A uses data on all stocks; Panel B restricts the sample to all-but-microcaps. All-but-microcaps are stocks with market capitalizations above the 20th percentile of the NYSE distribution.

The estimates in the first column of Table 1 use actual data; the remaining columns simulate data with values of σ_μ ranging from 0% to 1.4%. The average cross-sectional standard deviation of realized monthly stock returns is 16.9%. Therefore, when $\sigma_\mu = 1\%$, cross-sectional variation in expected returns explains $(0.01)^2 / (0.169)^2 = 0.35\%$ of the cross-sectional variation in realized returns.

We estimate the cross-sectional regressions in equation (3) using various past return windows. The first row predicts the cross section of month t returns using month $t - 1$ returns; the last row uses prior

20-year returns skipping a month. Each regression contains stocks that have non-missing returns in month t and for at least half of the past return period. The regressions on the last row, for example, include stocks that have at least ten years of non-missing returns.

The estimates that use actual return data show the effects of short-term reversals (Jegadeesh 1990), momentum (Jegadeesh and Titman 1993), and long-term reversals (De Bondt and Thaler 1985). Short-term reversals show up as negative slope coefficients in cross-sectional regressions of month t returns against prior month returns. The estimates in the full and all-but-microcaps samples are -0.05 (t -value = -12.12) and -0.02 (t -value = -3.51). Momentum registers as positive coefficients in regressions against prior one-year returns skipping a month. The estimates in the full and all-but-microcaps samples are 0.04 and 0.10 , and these estimates are associated with t -values of 2.54 and 4.88 . Long-term reversals appear as negative slope coefficients when we lag average returns by more than a year. For example, in regressions that predict returns using prior five-year returns skipping a year, the slope coefficients are -0.17 and -0.11 , and the t -values are -6.43 and -3.72 .

Data in which expected returns are constant over time, vary across firms, and stock-specific innovations are serially and cross-serially uncorrelated cannot produce *negative* regression slopes. The inevitability of positive regression coefficients under these assumptions is best illustrated by considering the Lo and MacKinlay (1990) decomposition of trading profits.¹⁰ They consider a strategy that weighs stock i by

$$w_{i,t} = \frac{1}{N} (\bar{r}_{i,t-k_1,t-k_2} - \bar{r}_{m,t-k_1,t-k_2}), \quad (5)$$

where $\bar{r}_{i,t-k_1,t-k_2}$ is stock i 's average return from month $t - k_2$ to $t - k_1$, and $\bar{r}_{m,t-k_1,t-k_2}$ is the return on the equal-weighted market index. The expected trading profit in month t , $E(\pi_t) = E\left(\sum_{i=1}^N w_{i,t} r_{i,t}\right)$,

¹⁰Lo and MacKinlay (1990) decompose profits to strategies that trade long-term reversals. Lewellen (2002) considers profits to strategies that trade the momentum in individual stocks or portfolios of stocks.

then decomposes into

$$E(\pi_t) = \underbrace{\frac{N-1}{N^2} \text{tr}(\Sigma)}_{\text{Autocovariances}} - \underbrace{\frac{1}{N^2} [1'\Sigma 1 - \text{tr}(\Sigma)]}_{\text{Cross-serial covariances}} + \underbrace{\sigma_\mu^2}_{\text{Persistent variation in means}}, \quad (6)$$

where $\text{tr}(\Sigma)$ is the trace of the covariance matrix, $1'\Sigma 1 - \text{tr}(\Sigma)$ is the sum of the cross-serial covariances, and σ_μ^2 is the cross-sectional variance of unconditional expected returns. The last term is always positive; persistent cross-sectional variation in mean returns increases the profitability of trading strategies in which weights increase in realized returns. Profits can therefore be negative only if stock returns are negatively autocorrelated (the first term of equation (6) is negative) or positively cross-serially correlated (the second term of equation (6) after the minus sign is positive). That is, a negative slope coefficient emerges only if a high return on an asset predicts a low return on that asset or high returns on other assets.

The estimated regression slope in equation (4), which assumes IID innovations, is nonnegative; the IID assumption sets the first two terms of the decomposition in equation (6) to zero. Therefore, in Table 1's bootstrapped data, the slope coefficients start at zero when $\sigma_\mu = 0$ and become increasingly positive as we increase the amount of cross-sectional variation in expected returns. When $\sigma_\mu = 0.4\%$, average past returns over various horizons begin to be identified as statistically significant predictors of returns in the full sample. For the $[-240, -2]$ window in Panel A, for example, the average slope coefficient is 0.17 and the t -value associated with this estimate is 4.62. In the all-but-microcaps sample, the estimates are even more statistically significant when $\sigma_\mu = 0.4\%$ or greater.

The simulations in Table 1 show that if even a small amount of persistent variation was present in expected returns, our regressions would have the power to detect it. A volatility parameter of $\sigma_\mu = 0.4\%$, for example, corresponds to a world in which a cross-sectional regression of *realized* month- t returns r_{it} against expected returns μ_i has an R^2 of 0.06%. Moreover, if expected returns remained constant in

the cross section, past returns should become more informative about returns as we increase the length of the past-return window; a wider window yields more precise estimates of expected returns. In the simulations, the t -values indeed increase in the length of the estimation window. In the data, however, the slope coefficients are and remain negative after the one-year momentum period.

These comparisons suggest the amount of persistent cross-sectional variation in expected returns must be negligible; whatever the variation might be, it is completely overshadowed by long-term reversals in individual stock returns. The term *persistent* can also be interpreted loosely here. If stocks' expected returns change over time, but slowly, a regression against five-year average returns would still typically return a positive coefficient. In the data, the cutoff is one year.

It is instructive to consider momentum to understand the amount of predictive power we would expect to find at longer lags if cross-sectional differences in expected returns persisted. In Panel B's all-but-microcaps sample, the slope coefficient on momentum is 0.10 with a t -value of 4.88. In the bootstrapped data, we approximately match this coefficient when $\sigma_\mu = 1.0\%$. Here, the slope coefficient is 0.09 with a t -value of 4.76. If we were to attribute momentum to persistent cross-sectional variation in mean returns, long-horizon past returns would be tremendously powerful predictors of the cross section of stock returns. For example, in regressions against the prior ten-year returns skipping a month, the t -value would be 21.06.

5 Production-based models, discount rate shocks, and long-term reversals

5.1 Intuition

The Fama-MacBeth regressions in Section 4 show that past returns predict the cross section of average returns with negative signs. These negative associations correspond to the long-term reversals identified by De Bondt and Thaler (1985). The negative sign, based on the Lo and MacKinlay (1990) logic, must be due either to negative autocorrelations or positive cross-serial correlations in stock returns. The lack of persistent differences in expected returns is a potential source of negative autocorrelations. If differences in expected returns are short-lived, some of the stock return variation is, by definition, due to discount rate shocks. Long-term reversals could emerge if there are large but short-lived differences in expected returns; in that case, past returns predict future returns with a negative sign because the return variation mostly reflects discount rate shocks.

Figure 3 gives a textbook illustration of the association between past and expected returns in the presence of discount rate shocks.¹¹ We assume all firms pay just one cash flow to shareholders at time $t = 5$ and the expected value of this terminal cash flow is $E(\widetilde{CF}_1) = 100$ at time $t = 0$. A firm's market value at any date prior to the terminal date is the present value of the expected cash flow. In Panel A, all return variation is due to cash flow shocks. Because the discount rate is constant, differences in past returns are uninformative about differences in expected returns. If a firm's market value decreases significantly—as it does at time $t = 0$ —it must do so because the expected cash flow decreased. In Panel B all return variation is due to discount rate shocks. If a firm's market value decreases significantly, it must be because the firm's expected return increased. In this case, past returns correlate negatively

¹¹Cochrane (2011) provides a detailed discussion and analysis of the variation in discount rates at the market level.

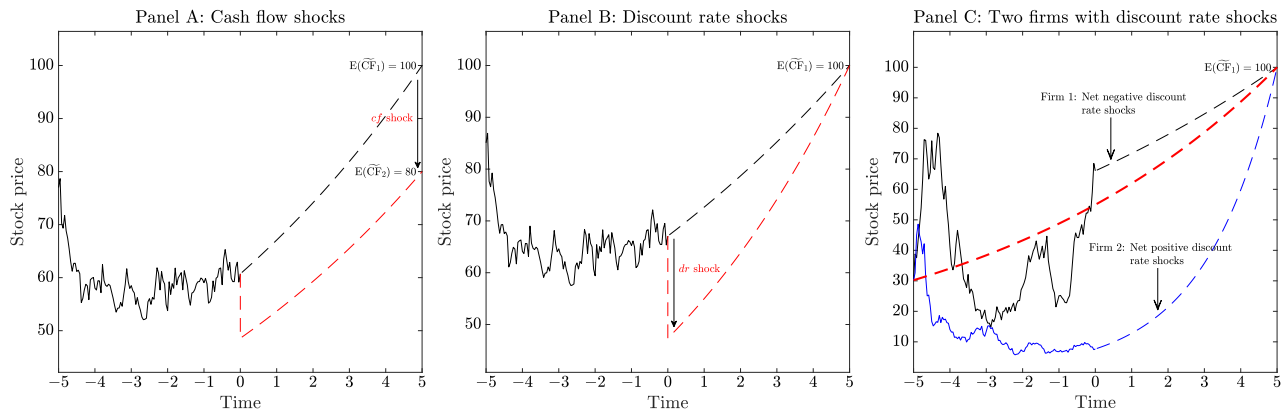


Figure 3: **Cash flow shocks, discount rate shocks, and the information content of past returns.** This figure illustrates the association between past and expected returns when return variation is due to cash flow shocks (Panel A) or discount rate shocks (Panel B). When return variation is due to cash flows shocks, expected returns are unrelated to past returns. When return variation is due to discount rate shocks, the relation between past and expected returns is negative. In Panel C two firms have identical discount rates at time $t = -5$. The differences in these firms' discount rates at time $t = 0$ correlate perfectly with the differences in past returns because their return variation is only due to discount rate shocks.

with expected returns. A negative realized return signifies an increase in the expected return.

In Panel C, we endow two firms the same expected return at time $t = -5$ and randomly shock these firms' discount rates until time $t = 0$. Because one firm is hit with net-positive discount rate shocks and the other with net-negative discount rate shocks, one firm's expected return is higher than that of the other at time $t = 0$. Moreover, because the firms started with identical expected rates at time $t = -5$, the differences in expected returns at time 0 correlate perfectly with differences in past returns.¹² Past returns' informativeness about expected returns depends on the intensity of discount rate shocks relative to cash flow shocks. If firms' expected returns rapidly converge toward a common mean, past returns may mostly measure discount rate shocks. Here, past returns no longer measure differences in the *levels* of discount rates.

¹²The log return on firm i from $t - 5$ to t is $\ln(P_{i,t}/P_{t-5}) = p_{i,t} - p_{t-5}$. The expected log return from t to $t + 5$ is $\ln(E(\widetilde{CF}_1)/P_{i,t}) = \ln(E(\widetilde{CF}_1)) - p_{i,t}$. Because p_{t-5} and $\ln(E(\widetilde{CF}_1))$ are common across firms, the past and expected returns encode the same information and correlate perfectly in the time-series (within a firm) and the cross section (across multiple firms).

5.2 Evidence from production-based models

In Table 2 we report estimates from cross-sectional regressions similar to those in Table 1. However, instead of using the actual or bootstrapped data, we estimate these regressions using data simulated from the four models described in Section 2. We simulate data using each model’s baseline parameters and generate hypothetical datasets that match the length and size of the cross section of the actual return data. With the Berk, Green, and Naik (1999) model, we also alter the model’s depreciation rate and run additional simulations. We estimate cross-sectional regressions using each simulated sample and record the average slope estimates and their t -values. Table 2 reports averages of the estimates and t -values over 10,000 simulations for each specification.

The first four columns in Table 2 use data simulated from Berk et al. (1999). The original parametrization of the model uses a depreciation rate of $1 - \pi = 1\%$. Berk et al. (1999, p. 1573) determine this rate based on the calibrations of real business cycle models (Kydland and Prescott 1982; Christiano and Eichenbaum 1992). In the model, the depreciation rate accounts for the probability that a firm’s project becomes obsolete and ceases to generate cash flows. We vary this rate because it affects the rate of convergence of expected returns and, through the discount-rate mechanism, the relation between past and future returns.

Figure 4 shows how the convergence of expected returns in the Berk et al. (1999) model depends on the depreciation rate. When the depreciation rate increases, firms’ risks converge faster toward the common mean; when we decrease the rate below that in the baseline model, expected returns become more persistent. An increase in the depreciation rate therefore increases the amount of return variation that is due to discount rate shocks.

Table 2 shows that, in the baseline specification of Berk et al. (1999), past returns do not significantly

Table 2: Long-term reversals in production-based asset pricing models

This table reports average coefficients and t -values from regressions to predict the cross section of monthly stock returns. The explanatory variable is the stock's average return over the window specified in the first column. We estimate these regressions using data simulated from the models of Berk, Green, and Naik (1999), Gomes, Kogan, and Zhang (2003), Lin and Zhang (2013), and Hackbarth and Johnson (2015). Each simulated sample matches the dimensions of the actual stock data in Table 1. We use each model's baseline parameters. With the Berk et al. (1999) model, we also vary the depreciation rate, $1 - \pi$, from the 1% rate of the baseline model. We repeat the simulations 10,000 times for each specification and report the average estimates and t -values.

Historical return horizon	Model						
	Berk, Green, and Naik				Gomes, Kogan, and Zhang	Lin and Zhang	Hackbarth and Johnson
	Depreciation rate, $1 - \pi =$						
0.5%	1.0%	2.0%	3.0%				
Coefficient estimates							
$[-1, -1]$	-0.00	-0.00	-0.00	-0.01	-0.01	-0.01	0.00
$[-12, -2]$	-0.00	0.00	-0.03	-0.06	-0.07	-0.05	0.02
$[-60, -13]$	0.01	0.01	-0.08	-0.10	-0.10	-0.13	0.05
$[-120, -61]$	0.01	0.01	-0.03	-0.02	-0.01	-0.09	0.04
$[-120, -2]$	0.02	0.02	-0.16	-0.19	-0.21	-0.31	0.10
$[-240, -121]$	0.01	0.01	-0.01	-0.00	-0.00	-0.09	0.05
$[-240, -2]$	0.03	0.03	-0.18	-0.19	-0.22	-0.49	0.14
t-values							
$[-1, -1]$	-0.64	-0.39	-3.25	-4.64	-9.11	-2.96	2.71
$[-12, -2]$	0.10	0.53	-4.89	-6.79	-17.27	-3.20	5.77
$[-60, -13]$	1.29	0.97	-4.77	-6.50	-14.25	-3.29	6.82
$[-120, -61]$	1.21	1.05	-3.55	-3.47	-2.45	-3.22	6.14
$[-120, -2]$	1.94	1.07	-4.85	-6.48	-16.28	-3.39	7.32
$[-240, -121]$	1.13	0.85	-1.29	-0.51	-0.22	-2.95	5.70
$[-240, -2]$	2.08	1.13	-4.73	-6.21	-14.16	-3.36	7.36

predict the cross section of average returns. However, outside of this Goldilocks point, the association is statistically significant. If we lower the depreciation rate, past returns positively predict future returns. In this case, discount rate shocks are rare, and past returns mostly measure differences in (relatively) persistent expected returns. If we increase the depreciation rate, past returns negatively predict future returns. This parametrization is close to the world of Panel C of Figure 3 in which all return variation

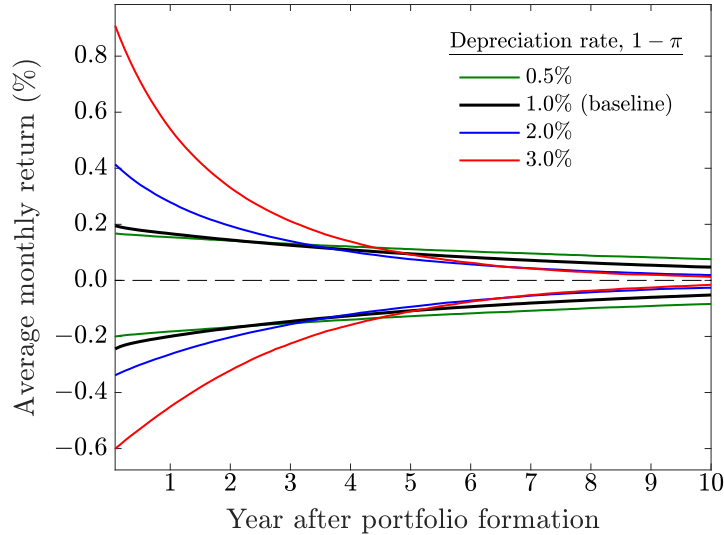


Figure 4: **Average monthly returns in the Berk, Green, and Naik (1999) model with different depreciation rates.** We simulate 1,000 months of return data from the model of Berk et al. (1999) using four alternative values for the depreciation rate, $1 - \pi$. All other parameters of the model are set to the values used in the original study. We discard the first 400 months from each simulation and then begin ranking stocks into deciles based on expected returns. We compute the average cross-sectionally demeaned monthly returns for these deciles over the next 10 years after portfolio formation. This figure averages over 10,000 simulations and reports the average monthly returns for the top and bottom deciles.

is due to discount rate shocks. Even when the depreciation rate is very high, the association between past returns and future returns is weak only when the historical returns are over ten years old. What happened between 10 and 20 years ago is uninformative about a firm's expected return today because these returns have already been followed by another 10 years of discount rate shocks.

The estimates from the other three models in Table 2 suggest the strength (or lack thereof) of long-term reversals depends on the rate at which the discount rates converge. In Figure 2 the rate of convergence varies across the three models. The rate of convergence is the fastest in Gomes et al. (2003); it is the second fastest in Lin and Zhang (2013); and, in Hackbarth and Johnson (2015), discount rates are very persistent. Consistent with this ordering, the estimates in Table 2 show the strongest long-term reversals for Gomes et al. (2003), followed by Lin and Zhang (2013). In Hackbarth and Johnson (2015),

expected returns are so persistent that past returns positively predict future returns.

Figure 3 and Table 2 provide a compelling explanation for the negative slope estimates found in the cross-sectional returns-on-past-returns regressions. These estimates would be positive if cross-sectional differences in expected returns were very persistent. The apparent reason for the negative estimates is that expected returns are not very persistent. Expected returns vary dramatically across firms (Martin and Wagner 2019), but these differences, on average, converge rapidly to zero. It is this convergence—the discount rate shocks—that imprints the negative serial correlation into the data. Long-term reversals are therefore the consequence of the lack of differences in long-term discount rates.

6 Portfolio sorts

6.1 Return predictors

We complement the bootstrap analysis by using 46 predictors of stock returns to sort stocks into portfolios. We use different combinations of these predictors to form portfolios, and then examine the persistence in these portfolios' average returns. The predictors we include are among those analyzed in McLean and Pontiff (2016). We also take those few additional predictors from Linnainmaa and Roberts (2018) that were published after McLean and Pontiff (2016) created their list.

Table 3 reports monthly average returns and CAPM and three-factor model alphas for HML-style factors based on each of the 46 return predictors. We sort stocks into six portfolios by market capitalization and the predictor, and then compute value-weighted returns on these portfolios. The breakpoint for size is the 50th NYSE percentile, and the breakpoints for the predictor are the 30th and 70th NYSE percentiles. A factor's return is the average return on the two high portfolios minus the average return on the two low portfolios.

Table 3: Average returns, CAPM and three-factor model alphas, and persistence of 46 return predictors

This table reports average returns and CAPM and three-factor model alphas for 46 return predictors. Each return predictor is used to construct an HML-style factor by sorting stocks into six portfolios by size and predictor. These sorts are independent and use NYSE breakpoints; the size breakpoint is the median, and the predictor breakpoints are the 30th and 70th percentiles. A predictor’s return is the average return on the two value-weighted high portfolios minus the average return on the two value-weighted low portfolios. The high and low portfolios are determined so that “high” corresponds to those stocks that the initial study identified as earning higher returns. “Accounting-based predictors” are rebalanced annually at the end of each June; “Return-based predictors” are rebalanced monthly. “Persistence” is the last holding period after which the average return or the CAPM or the three-factor model alpha has a p -value above 0.05 for three consecutive months. A value of 50, for example, indicates the factor’s average return or alpha remains statistically significantly different from zero for 50 months after portfolio formation.

Predictor	Start year	Average return		CAPM alpha		FF3 alpha		Persist- ence, months
		\bar{r}	t	$\hat{\alpha}$	t	$\hat{\alpha}$	t	
Accounting-based predictors								
Earnings to price	1963	0.35	2.78	0.52	3.56	0.20	2.15	50
Enterprise multiple	1963	0.35	3.54	0.46	3.86	0.17	2.40	44
Gross profitability	1963	0.30	3.35	0.28	2.53	0.41	4.19	54
Inventory growth	1963	0.20	3.14	0.26	3.85	0.18	2.94	7
Piotroski’s F score	1963	0.28	3.73	0.37	4.95	0.34	5.24	12
Abnormal investment	1966	0.15	2.75	0.14	2.53	0.17	3.33	12
Accruals	1963	0.19	3.08	0.22	3.05	0.17	2.63	7
Accruals and book-to-market	1963	0.57	3.00	0.62	2.85	0.06	0.36	33
Net operating assets	1963	0.31	4.44	0.27	3.55	0.34	4.56	19
Net working capital changes	1963	0.18	3.07	0.22	3.26	0.20	3.17	7
O-score	1963	0.06	0.87	0.12	1.59	0.21	3.39	81
Profit margin	1963	-0.01	-0.13	0.10	1.00	0.18	2.19	1
Asset growth	1963	0.26	3.23	0.34	3.94	0.13	1.88	19
Sales to price	1963	0.42	3.70	0.45	3.08	0.06	0.58	26
One-year share issuance	1963	0.21	2.77	0.27	3.33	0.18	2.50	53
Five-year share issuance	1963	0.25	3.40	0.32	4.55	0.23	3.55	64
Sustainable growth	1963	0.16	2.01	0.24	2.82	0.04	0.50	7
Total external financing	1972	0.34	3.48	0.48	4.96	0.39	5.11	12
Z-score	1965	-0.01	-0.11	-0.08	-0.69	0.17	2.25	3
Industry-adjusted CAPX growth	1965	0.22	4.42	0.25	5.30	0.20	4.45	28
Sales-minus-inventory growth	1965	0.20	3.80	0.19	3.60	0.22	3.89	12
Investment to capital	1964	0.17	1.57	0.32	3.28	0.11	1.54	5
Investment growth rate	1964	0.20	3.48	0.24	4.03	0.15	2.68	18
Investment to assets	1963	0.25	3.94	0.29	4.38	0.19	2.94	28

QMJ: Profitability	1964	0.25	3.14	0.31	3.43	0.45	5.56	80
Distress	1975	0.33	2.10	0.55	3.23	0.69	5.09	27
Book to market	1963	0.35	3.02	0.46	3.35	0.03	0.77	25
Operating profitability	1963	0.24	2.61	0.29	2.54	0.32	3.03	11
Organization capital	1963	0.25	4.02	0.32	5.66	0.36	6.76	29
Cashflow to equity	1963	0.40	3.19	0.53	3.71	0.14	2.13	47
Return on assets	1973	0.27	2.25	0.37	2.77	0.53	4.29	15
Return on equity	1970	0.29	2.38	0.40	3.03	0.50	3.92	4
Asset turnover	1963	0.21	3.87	0.23	3.91	0.20	3.57	22
Debt issuance	1972	0.10	1.01	0.01	0.10	0.16	2.01	1

Return-based predictors

52-week high	1963	0.52	3.10	0.76	5.11	0.82	6.39	11
Amihud's illiquidity	1963	0.29	1.44	0.41	2.23	0.21	1.21	1
Beta	1963	0.03	0.17	0.39	2.57	0.30	2.62	3
Idiosyncratic volatility	1963	0.23	1.21	0.59	3.65	0.50	4.37	10
Industry momentum	1963	0.27	1.66	0.36	2.24	0.37	2.21	7
Long-term reversals	1963	0.31	3.15	0.32	2.76	0.07	0.70	14
Maximum daily return	1963	0.29	1.77	0.58	4.11	0.48	4.62	5
Momentum	1963	0.64	3.85	0.69	4.39	0.84	5.49	10
Intermediate momentum	1963	0.60	4.87	0.59	4.16	0.72	5.44	5
Seasonality	1963	0.60	7.83	0.57	7.00	0.63	7.67	1
Short-term reversals	1963	0.48	3.90	0.38	3.36	0.34	2.75	1
High-volume return premium	1963	0.52	8.16	0.56	9.14	0.51	8.00	3

We divide the predictors into accounting- and return-based predictors. A predictor is accounting-based if it uses any information from either the balance sheet or income statement; the return-based predictors use only return, price, or volume information. We recompute and rebalance the accounting-based factors annually at the end of June and the return-based factors monthly.

The 46 return predictors that we examine explain differences in average returns over our 1963 through 2016 sample period. Each factor earns a statistically significant average return, CAPM alpha, or three-factor model alpha. The last column reports estimates of how long a factor retains its statistically significant average return or alpha. We define persistence as the number of months that can be skipped after portfolio formation without losing statistical significance for three consecutive months. Inventory growth, for example, has an estimated persistence of seven months. That is, this factor's returns

are statistically significant in months one through seven after portfolio formation, but statistically insignificant afterwards.¹³ We later use these persistence estimates to group together the ex-post most persistent predictors.

6.2 Average long-term returns

In Figure 5 and Table 4, we form value-weighted portfolios using different combinations of return predictors and then examine differences in average returns at various horizons. In these computations, the portfolios are formed at date t and then held unchanged for up to ten years. In addition to using all 46 returns predictors, we form portfolios based on three subsets of predictors: (a) the 34 accounting-based predictors, (b) the 12 return-based predictors, and (c) the 21 ex-post most persistent predictors. These subsets comprise the predictors that retain their statistical significance for at least over a year according to Table 3’s last column.

We form the portfolios by computing the average studentized predictor for each stock. We first convert each predictor into a z -score by subtracting the cross-sectional average and by then dividing by the cross-sectional standard deviation (see, e.g., Asness, Frazzini, and Pedersen (2019)). Similar to the factors in Table 3, we sign each predictor so that high values correspond to high average returns based on the original study. A stock’s signal is the average of its non-missing z -scores.

Figure 5 suggests little, if any, persistent differences exist in average returns. In this figure, we take the difference between the top and bottom deciles and report the average value-weighted monthly returns (Panel A) and CAPM alphas (Panel B) for months 1 to 120 following portfolio formation. These returns are non-cumulative. That is, the return at horizon k is the average return that an investor would have earned in month k after portfolio formation (“forward rate”). Table 4 reports average returns and

¹³Predictors can regain statistical significance later by chance or, in the case of seasonality, by nature (Heston and Sadka 2008). If a predictor is statistically insignificant for three consecutive months, Table 3 classifies it as having lost its predictive power.

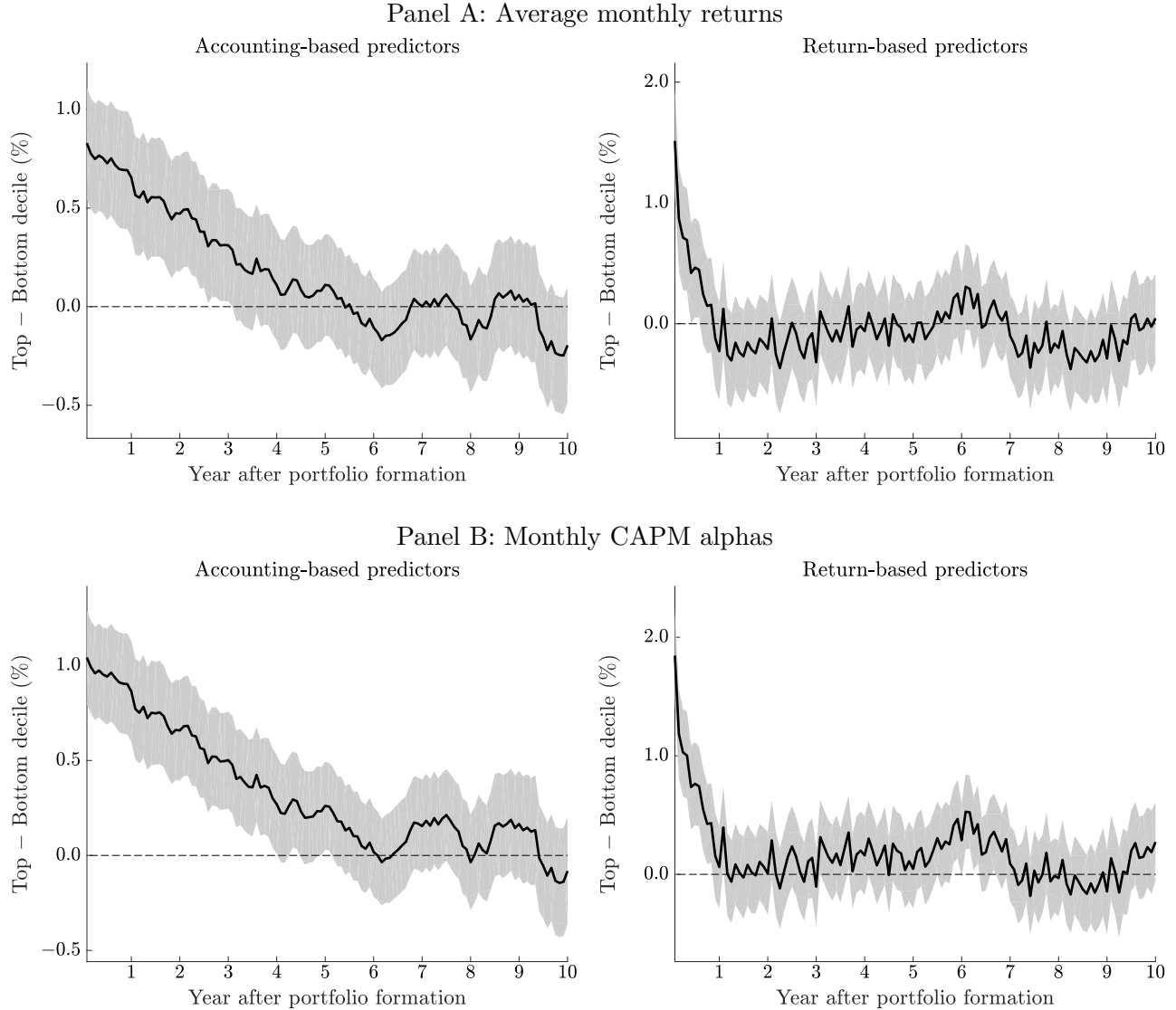


Figure 5: **Average monthly returns and CAPM alphas on long-short strategies in months 1–120 after portfolio formation.** We form decile portfolios at the end of each month by sorting stocks into portfolios by different combinations of the 46 return predictors listed in Table 3. The left-hand side of each panel uses the 34 predictors that use income statement or balance sheet information; the right-hand side of each panel uses the 12 predictors that use price, return, or volume information. We convert each predictor into a z -score by subtracting the cross-sectional average from the predictor and dividing the difference by the cross-sectional standard deviation. Each stock’s predictor is the average of its non-missing z -scores. We construct value-weighted portfolios each month and hold these portfolios for up to ten years. Panel A plots the average return difference between the top and bottom deciles, whereas Panel B plots the monthly CAPM alphas for the return difference between the top and bottom deciles. The average return in month t is its month t return (“forward rate”), not its average return from today to month t . The shaded areas indicate 95% confidence intervals.

Table 4: Annual post-formation spread portfolio returns

We form decile portfolios at the end of each month by sorting stocks into portfolios by different combinations of the 46 return predictors listed in Table 3. Row “All” uses all predictors; “Accounting-based” uses the 34 predictors that use some income statement or balance sheet information; “Return-based” uses the 12 predictors that use price, return, or volume information; and “Ex-post long-lived” uses the 21 predictors in Table 3 that, when used in isolation, persist for at least a year and a month. We convert each predictor into a z -score by subtracting the cross-sectional average from the predictor and dividing the difference by the cross-sectional standard deviation. Each stock’s predictor is the average of its non-missing z -scores. We construct value-weighted portfolios each month and hold these portfolios for up to eight years. Panel A reports average monthly returns and t -values for value-weighted high-minus-low portfolios for different horizons following portfolio formation; Panel B reports monthly CAPM alphas and t -values. The holding periods range from one month after portfolio formation to year ten after portfolio formation. We adjust standard errors for overlapping returns using the Jegadeesh and Titman (1993) method.

Panel A: Average returns

Return predictor	Horizon								
	1 month	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	Year 8
Average monthly long-short returns									
All	1.23	0.74	0.25	0.20	0.05	0.02	0.16	0.06	-0.11
Accounting-based	0.83	0.73	0.53	0.40	0.18	0.08	0.00	-0.10	0.00
Price-based	1.51	0.46	-0.18	-0.20	-0.05	-0.06	0.05	0.11	-0.19
Long-lived	0.72	0.61	0.51	0.38	0.16	0.13	0.11	0.01	0.16
t-values									
All	6.62	4.48	1.55	1.24	0.35	0.11	1.15	0.47	-0.83
Accounting-based	5.75	5.20	3.66	2.88	1.40	0.65	-0.04	-0.86	-0.02
Price-based	6.71	2.46	-1.06	-1.21	-0.30	-0.39	0.34	0.77	-1.32
Long-lived	3.99	3.55	3.02	2.31	1.02	0.86	0.78	0.07	1.19

t -values for the month following portfolio formation, as well as the average monthly returns and CAPM alphas in years 1, 2, . . . , 10 following portfolio formation. We adjust standard errors for overlapping observations using the Jegadeesh and Titman (1993) procedure.

Panel A of Table 4 shows the average return differences between the top and bottom quintiles rapidly vanish when we estimate expected returns using all 46 predictors.¹⁴ One month after portfolio

¹⁴Figure 5 and Table 4 use value-weighted portfolios. Table A1 in the Appendix reports the estimates for equal-weighted portfolios.

Panel B: CAPM alphas

Return predictor	Horizon								
	1 month	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	Year 8
CAPM alphas									
All	1.55	1.03	0.49	0.45	0.29	0.23	0.35	0.24	0.08
Accounting-based	1.04	0.94	0.72	0.59	0.36	0.24	0.13	0.03	0.15
Price-based	1.84	0.76	0.06	0.04	0.17	0.14	0.25	0.30	0.00
Long-lived	0.99	0.88	0.75	0.60	0.37	0.32	0.28	0.17	0.33
<i>t</i>-values									
All	10.11	7.79	3.73	3.41	2.36	1.78	3.00	2.10	0.71
Accounting-based	8.29	7.74	5.67	4.71	3.08	2.03	1.16	0.22	1.29
Price-based	9.55	5.06	0.42	0.30	1.30	1.07	1.97	2.38	-0.03
Long-lived	6.34	5.95	5.04	4.12	2.64	2.43	2.11	1.31	2.68

formation, the average return for the long-short portfolio is 123 basis points (t -value = 6.62); in the year following portfolio formation, the return difference is 74 basis points (t -value = 4.48); and, in year two, it is just 25 basis points (t -value = 1.55). After year two, the differences in average returns are economically small and statistically insignificant. Taken together, this full set of predictors is not informative about long-term differences in average returns.

Figure 5 shows that accounting-based predictors are significantly more informative about differences in long-term returns than return-based predictors. The accounting-based predictors generate a return difference of 40 basis points per month in year three (t -value = 2.88), whereas the return-based predictors are statistically significant only in the first year. The difference is consistent with the persistence estimates in Table 3. Moreover, this difference between the two groups of predictors is also consistent with prior research. Some return-based predictors, such as short-term reversals, are known to last only for about one month (Jegadeesh 1990; Goyal and Wahal 2015). Others, such as momentum, turn into reversals at longer lags (Jegadeesh and Titman 1993). Although other predictors, such as idiosyncratic volatility, could contain information about long-term returns, they do not.

The last row of Table 4 shows that even when we select the ex-post most persistent predictors—those that, on their own, are statistically significantly different from zero for more than a year following portfolio formation—we cannot reliably detect differences in average returns between the top and bottom quintiles after year three. This test purposefully stacks the playing field in favor of finding differences in average returns; after all, we use the same data to both identify the most promising predictors and to test their performance. Nevertheless, this set of predictors displays only the same amount of persistence as the full set of accounting-based predictors.

Panel B of Figure 5 and Panel B of Table 4 report monthly CAPM alphas for the long-short strategies from Panel A. Statistically significant differences in CAPM alphas persist longer than those in average returns—they converge to zero seven years after portfolio formation. Market adjustment helps both by increasing point estimates and, by the virtue of removing market-wide variation, lowering standard errors. When we use all 46 return predictors, the CAPM alpha in the first year after portfolio formation is 103 basis points with a standard error of 13 basis points; in Panel A, by contrast, the average return in year one is 74 basis points with a standard error of 17 basis points. The fact that CAPM alphas are higher than average returns is consistent with Table 3’s result that most anomalies are stronger on a market risk-adjusted basis.

Average returns (Panel A) and market-adjusted returns (Panel B) are both important, but for different purposes. Market-adjusted returns are relevant from an investing viewpoint, whereas average returns are the ones that matter for capital budgeting. An investor could manage his portfolio to keep it market neutral, thereby aiming to earn the market-adjusted returns reported in Panel B.¹⁵ Firms, however, do not discount cashflows at market-adjusted rates. Instead, Panel A’s average returns represent the appropriate estimates of the differences in firms’ discount rates.

¹⁵A statistically significant CAPM alpha indicates an investor who currently holds the market could earn a higher Sharpe ratio by additionally taking a position in the left-hand-side strategy (Huberman and Kandel 1987).

Figure 4 shows that our inability to detect differences in future average returns is unlikely due to lack of power. If a lack of power were the issue, we would expect the standard errors to increase as a function of time from portfolio formation. Yet they do not: in Figure 5, the width of the confidence interval remains approximately unchanged while the point estimates converge toward zero in all specifications. In fact, we can estimate differences in average returns and CAPM alphas with enough precision to bound the amount of variation in long-term average returns. Consider, for example, the estimates in Panel A of Table 4 where we use accounting-based predictors to form the portfolios. The 95% interval for the average return in year one runs from 46 basis points to 101 basis points; in year six, by contrast, it runs from -24 basis points to 23 basis points. Therefore, strong statistical evidence suggests differences in expected returns change as a function of time from portfolio formation.

Our results on the short-livedness of cross-sectional differences in expected returns are not specific to any one part of our sample. Figure A2 in the appendix divides the sample period into two halves and shows our results are similar before and after 1990. Concerns related to firm survival also cannot plausibly explain our results. Much of the convergence in expected returns happens already within the first year, when any expected return effects due to a change in the set of firms must be small. Our results are also highly similar in a sample restricted to firms that will survive for ten years (Figure A3) and in the unrestricted sample (Panel A of Figure 5).

In principle, this section’s evidence of the lack of persistence in predicted returns could be specific to the chosen set of predictors, but such a scenario is unlikely. First, Table 3 shows that, individually, none of the predictors contain information about differences in average returns beyond year five, and we are not aware of any other predictors outside this list that do so.¹⁶ Second, Section 4’s bootstrap

¹⁶Some of the *alphas* in Table 3 are more persistent—such as the three-factor model alpha of the O-score anomaly, which persists for 81 months—but the persistence in alphas is not the same as the persistence in average returns; an anomaly’s alpha can remain significant because the anomaly’s loadings against the factors change over time. This distinction is important: an investor at date 0 cannot construct a buy-and-hold portfolio based on O-scores that earns positive alphas for 81 consecutive months.

procedure suggests the search for such predictors should be futile. If persistent differences existed in expected returns, the Fama-MacBeth regression slopes would be positive. In the actual data, they are not.

7 Changes in betas and characteristics over time

7.1 Estimates

In production-based asset pricing models, expected returns change as firms' risks change. In a model such as Gomes, Kogan, and Zhang (2003), returns are completely described by a single-factor conditional CAPM. However, precisely because risks change—and, therefore, true betas as well—firm characteristics may better predict returns than estimated betas (Lin and Zhang 2013). In Figure 6, we measure changes in both market betas (Panel A) and firm characteristics (Panel B).

We estimate betas using one quarter of daily data. We assign stocks into deciles using quarter- q estimates and then report average estimates for the next 40 quarters. By using non-overlapping data, the estimation errors at the time of the portfolio sort are uncorrelated with those in the measurement period.

Panel A shows that market betas converge toward the cross-sectional mean. In the first quarter after portfolio formation, the average beta of the firm in the top decile is 1.63; that of the firm in the bottom decile is 0.62. This initial beta gap of 1.01 narrows rapidly at first and then at a slower pace. The gap is half of its initial value after five years, and after ten years, it is 0.36. That is, 65% of the initial differences in market betas disappear in ten years.

In Panel B, we assign firms into deciles at the end of each June by a combination of the 34 accounting-based return predictors listed in Table 3. As before, we convert each predictor into a z -score by

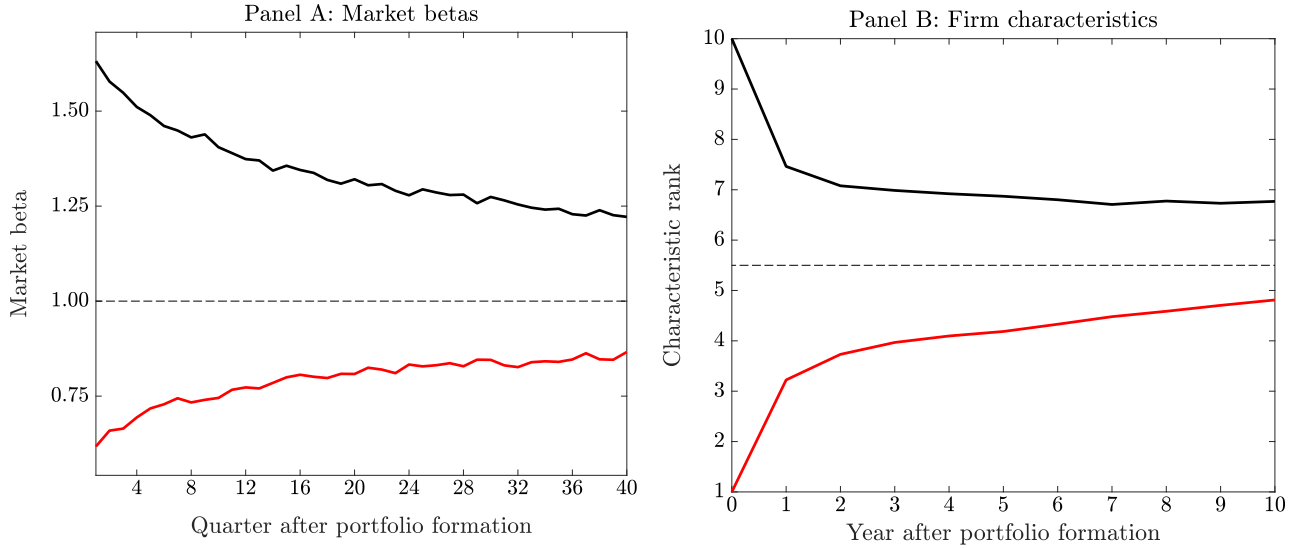


Figure 6: **Changes in betas and firm characteristics.** This figure reports average market betas and firm-characteristic ranks for 10 years after portfolio formation. In Panel A we assign firms into deciles by quarter- q market betas and report average betas starting in quarter $q + 1$. We estimate each market beta using one quarter of daily data so that the quarterly periods do not overlap. In Panel B, we assign firms into deciles at the end of each June by a combination of the 34 accounting-based return predictors listed in Table 3. We convert each predictor into a z -score by subtracting the cross-sectional average from the predictor, and then scale by the cross-sectional standard deviation. A firm’s decile is determined by the average of its non-missing z -scores.

subtracting the cross-sectional average from the predictor and scaling the difference by the cross-sectional standard deviation. A firm’s decile is determined by the average of its non-missing z -scores. By construction, the average rank difference between the top and bottom deciles is 9 at the time of portfolio formation. One year later, this difference is 4.2; after five years, it is 2.7; and after ten years, it is 2.0. Thus, similar to market betas, differences in firm characteristics evaporate rapidly at first. However, although betas and characteristics of stocks in the top and bottom deciles converge toward each other, large differences remain even after ten years.

7.2 Returns on permanent and transitory characteristics

In Figure 7, we show that firm characteristics are so persistent that we would expect them to be associated with meaningful differences in average long-term returns. The black line in this figure is the same as that in Panel A of Figure 5; it shows the difference in the average monthly returns between the top and bottom deciles generated using accounting-based predictors. The shaded red line plots the same top-minus-bottom return difference after replacing each stock with its characteristics-matched portfolio of stocks. We construct this return series by tracking, as in Panel B of Figure 6, a stock’s decile assignment over time. If a stock’s decile in month t after portfolio formation is d , we replace the stock’s actual return with the value-weighted return of stocks that today belong to decile d .

The difference between the black and red lines in Figure 7 measures the difference in the premiums and discounts commanded by persistent and transitory characteristics. Figure 7 shows that cross-sectional differences in average returns disappear far faster than they should: if a firm’s characteristic persists, the persistent part commands neither a premium nor a discount—only the part that eventually disappears does.

Table 5 uses portfolio sorts similar to those reported in Table 4 to measure the difference between persistent and transitory firm characteristics. In this table, we sort stocks into portfolios based on the original predictors (as in Table 4) as well as the permanent and transitory components of these predictors. In Table 5, we report average return differences between the top and bottom deciles over the first year after portfolio formation.

We decompose each stock’s z -score by estimating a cross-sectional regression of today’s z -scores against the ten-year average z -scores:

$$z_{it} = a_t + b_t \times \bar{z}_{it}^{10\text{-year}} + e_{it}. \quad (7)$$

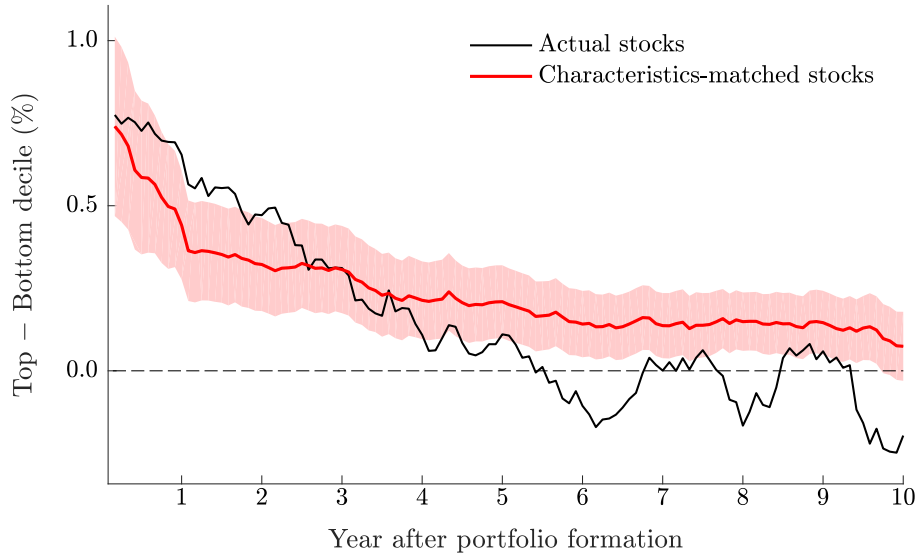


Figure 7: **Returns on characteristics-matched portfolios.** We form decile portfolios by sorting stocks into portfolios by the 34 accounting-based predictors listed in Table 3. We convert each predictor into a z -score by subtracting the cross-sectional average from the predictor and dividing the difference by the cross-sectional standard deviation. Each stock's predictor is the average of its non-missing z -scores. We construct value-weighted portfolios each month and hold these portfolios for up to ten years. The thin black line is the average monthly return difference between the top and bottom deciles in month t following portfolio formation. The thick red line uses returns on characteristics-matched portfolios. Each month we calculate the current portfolio rank of a stock that was originally assigned into the top or bottom decile. If a stock's decile in month t is d , we replace the stock's actual return in month t with the value-weighted return on all stocks that would, as of today, belong to decile d . The red shaded area indicates 95% confidence intervals.

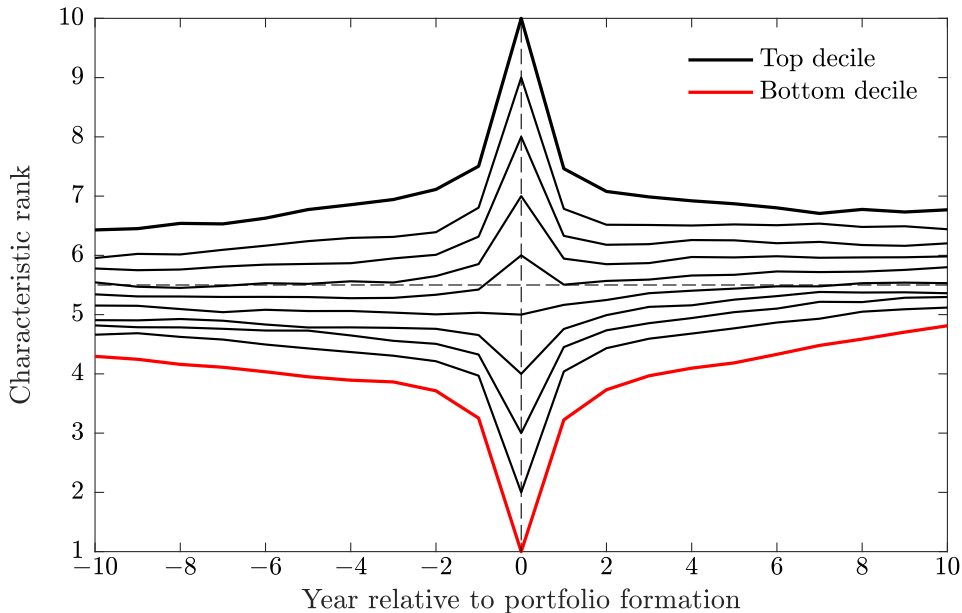


Figure 8: **Changes in firm characteristics before and after portfolio formation.** This figure reports average firm characteristic ranks for 10 years before and after portfolio formation. We assign firms into deciles at the end of each June by a combination of the 34 accounting-based return predictors listed in Table 3. We convert each predictor into a z -score by subtracting the cross-sectional average from the predictor, and then scale by the cross-sectional standard deviation. A firm's decile is determined by the average of its non-missing z -scores.

A firm's persistent component is the fitted value, $\hat{b}_t \times \bar{z}_{it}^{10\text{-year}}$, and the transitory component is the residual, \hat{e}_{it} . The persistent component, as suggested by Figure 6, is economically large. When we compute the z -scores using the accounting-based predictors, the average R^2 is 25%; with return-based predictors, it is 16%.

Figure 8 illustrates the rationale behind this decomposition. This figure is the same as Panel B of Figure 6 except that we also track the evolution of firm characteristics *before* portfolio formation. Top-decile stocks tend to lose their rankings over time, and these stocks are also the ones that gained in rankings before portfolio formation. Those in the bottom decile tend to gain in rankings after portfolio

Table 5: Decomposing predictors into persistent and transitory components

This table reports average returns for high-minus-low strategies that are long in the top decile of stocks and short in the bottom decile of stocks. We assign stocks into deciles alternatively by all 46 return predictors, 34 accounting-based predictors, or 12 return-based predictors (see Table 3). The accounting-based predictors are updated annually at the end of June; the price-based predictors are updated monthly. We convert each predictor into a z -score by subtracting the cross-sectional average from the return and dividing the difference by the cross-sectional standard deviation. Each stock’s predictor is the average of its non-missing z -scores. This table reports the return difference between the value-weighted top and bottom decile portfolios over the first year after portfolio formation. Columns labeled “Original predictor” use unadjusted predictors. Columns labeled “Permanent component” and “Transitory component” decompose characteristics by estimating, for each characteristic, a cross-sectional regression of month- t characteristics against the average characteristics over the prior ten years. The permanent component is the fitted value from this regression, and the transitory component is the residual. Holding-period returns in this table begin in July 1973 (and not in 1963 as in Table 4), and a stock must have at least 10 years of historical z -scores to be included in the decomposition regression.

Return predictors	Original predictor		Permanent component		Transitory component	
	\bar{r}	t -value	\bar{r}	t -value	\bar{r}	t -value
All	0.59	3.52	-0.07	-0.32	0.61	4.79
Accounting-based	0.50	3.69	0.05	0.29	0.47	4.25
Price-based	0.26	1.22	-0.28	-0.93	0.31	1.91

formation, and they are the ones that have fallen in rankings before portfolio formation.¹⁷ The regression in equation (7) decomposes firm characteristics by comparing a stock’s rank at the time of portfolio formation to its average pre-formation rank.

Table 5 shows the average return on a strategy that assigns stocks into deciles based on the transitory component is approximately as high as that on a strategy based on the original signal. When we use all 46 predictors, the average return difference between the top and bottom deciles based on the original z -score is 59 basis points (t -value = 3.52). The average return difference based on the transitory component of the z -score is 61 basis points with a t -value of 4.79. The average return on a strategy that assigns stocks into deciles based on the permanent component, by contrast, is close to zero. Using

¹⁷The near-symmetry of Figure 8 is not a coincidence. The post-formation pattern, for example, shows high-decile stocks turn into lower-decile stocks over time. Reading this post-formation pattern backwards therefore implies low-decile stocks were, on average, higher-decile stocks in the past. That is, the post-formation pattern implies the pre-formation pattern. The figure is not perfectly symmetrical, because the observations at the beginning of the sample lack pre-formation data, and those toward the end lack post-formation data.

all 46 predictors, this average return difference between the top and bottom deciles is -7 basis points (t -value = -0.32).

8 Conclusions

We show that cross-sectional differences in stocks' expected returns converge to zero in five years. Using different combinations of 46 return predictors, we are unable to identify any differences in long-term returns beyond this point. We devise a characteristics-free bootstrapping procedure that measures our ability to detect persistent cross-sectional differences in expected returns, and find no evidence of such differences. If differences in expected returns persisted, past average stock returns would positively and significantly predict the cross section of stock returns. In the data, this sign is negative.

Our results are consistent with firms' risks changing over time. We show that in production-based asset pricing models such as Berk, Green, and Naik (1999) and Zhang (2005), firms' expected returns converge toward the mean the same way they do in the actual data. The convergence in expected returns through discount rate shocks can generate long-term reversals, that is, the negative association between past and future returns. If expected returns were constant, past returns would measure differences in the *levels* of expected returns. However, when expected returns vary across firms, and these differences rapidly converge to zero, past returns measure the direction of discount rate shocks. Long-term reversals are the inevitable consequence of the lack of differences in long-term discount rates; they are thus not necessarily an indicator of the markets overreacting to new information (De Bondt and Thaler 1985).

If, as our results suggest, firms' expected returns rapidly converge to the mean, analysts should use about the same discount rates to value stocks. To illustrate the importance of the changes in expected returns to the valuation of cash flows, consider the Gordon growth model, $P/D = 1/(r - g)$, where P is the stock price, D is the dividend next year, r is the discount rate, and g is the dividend growth

rate. Suppose two firms exist that will each pay a $D = \$1$ dividend next year and that these dividends thereafter increase at the rate of $g = 2\%$. If one firm's discount rate is fixed at 12% and the other's at 4% , the stock prices of the two firms are $\$10$ and $\$50$, respectively. However, if both firms' discount rates change to 8% in five years, their values are $\$14.2$ and $\$19.8$ instead. These valuations are the same as if the market discounted all of the first firm's dividends at a rate of 9% and those of the second firm at a rate of 7% . The amount of cross-sectional dispersion in *short-term* discount rates far exceeds that in firm valuations.

The relevance of the short-term differences in expected returns depends on the lifespan of a firm's projects. If the average project generates long-term cash flows, the role of initial differences in discount rates will be small. Dechow, Sloan, and Soliman (2004) estimate that the equity duration of publicly traded U.S. firms is on average 15.1 years. This estimate together with ours suggests the effective long-term discount rates of most firms must be close to each other.

The finding that differences in unconditional expected returns are small goes against the conventional wisdom that discount rates should vary across firms and even across projects within the same firm. For example, corporate finance textbooks allow for considerable variation in the cost of equity, and rarely consider the idea that firms' discount rates would converge over time.¹⁸

Readers suggest our results are surprising in light of Proposition II of Modigliani and Miller (1958): if two firms' *unlevered* rates of return on equity are the same, but their debt-to-equity ratios differ, their expected levered rates of returns differ as well. The assumption of similar unlevered rates of return on equity is reasonable. High rates of returns attract competition, which can be expected to drive the rates of return on investment toward equality (Stigler 1963; Fama and French 2000). Debt-to-equity ratios

¹⁸Welch (2017) is a rare exception to this tradition. His view reflects that of Levi and Welch (2017), who recommend that the "cost-of-capital estimates should be shrunk far more than is common practice." Blume (1971) is the first to report that betas revert to the mean. Cohen, Polk, and Vuolteenaho (2009) also document that value stocks' betas increase and growth stocks' betas decrease after portfolio formation. See Levi and Welch for a review of the cost-of-capital prescriptions of corporate finance textbooks and academic literature.

are also known to differ considerably across firms. Although leverage ratios display convergence toward the mean, economically significant differences in leverage persist for long periods of time (Lemmon, Roberts, and Zender 2008).

Why do the data then not support the prediction that expected *levered* rates of return differ across firms?¹⁹ The idea that equity risk is increasing in leverage relies on the assumption that markets are frictionless, which leads firms' investment and financing decisions to be independent of one another. George and Hwang (2010) argue financial distress costs affect firms' capital structure decisions, inducing low-leverage firms to assume greater exposures to systematic risk than high-leverage firms. This neutralizes the mechanical effect leverage has on equity risk. Johnson et al. (2011) draw similar conclusions from a generalized version of the George and Hwang (2010) model. The puzzle can thus be explained by rational models, albeit ones with market frictions.

Our results also have asset pricing implications. Reliably extracting the risk-return relationship from the data may be difficult even when one exists. Suppose, for example, the true data-generating process is a conditional CAPM. Our results would then imply firms' betas must change rapidly to match the changes in discount rates. Lin and Zhang (2013) make an analogous argument by simulating data from Zhang's (2005) model and showing the resulting data appear to support the characteristics-based model of Daniel and Titman (1997) and not the risk-based model. Our results on the changes in average returns, betas, and characteristics point toward the same conclusion. To give asset pricing models a fair chance, econometric methods for testing them must accommodate rapid changes in firm risks.

¹⁹For example, Linnainmaa and Roberts (2018, Table 5) document that leverage is not a significant predictor of returns.

REFERENCES

- Asness, C., A. Frazzini, and L. Pedersen (2019). Quality minus junk. *Review of Accounting Studies* 24(1), 34–112.
- Asness, C. S., R. B. Porter, and R. L. Stevens (2000). Predicting stock returns using industry-relative firm characteristics. AQR Capital Management working paper.
- Berk, J. B., R. C. Green, and V. Naik (1999). Optimal investment, growth options, and security returns. *Journal of Finance* 54(5), 1553–1607.
- van Binsbergen, J. H. and C. C. Opp (2019). Real anomalies. *Journal of Finance*. Forthcoming.
- Black, F. (1986). Noise. *Journal of Finance* 41(3), 528–543.
- Blume, M. E. (1971). On the assessment of risk. *Journal of Finance* 26(1), 1–10.
- Christiano, L. J. and M. Eichenbaum (1992). Current real-business-cycle theories and aggregate labor-market fluctuations. *American Economic Review* 82(3), 430–450.
- Cochrane, J. H. (1991). Production-based asset pricing and the link between stock returns and economic fluctuations. *Journal of Finance* 46(1), 209–237.
- Cochrane, J. H. (2011). Presidential address: Discount rates. *Journal of Finance* 66(4), 1047–1108.
- Cohen, R., C. Polk, and T. Vuolteenaho (2009). The price is (almost) right. *Journal of Finance* 64(6), 2739–2782.
- Cohen, R. B. and C. Polk (1996). An investigation of the impact of industry factors in asset-pricing tests. Working Paper, London School of Economics.
- Conrad, J. and G. Kaul (1998). An anatomy of trading strategies. *Review of Financial Studies* 11, 489–519.
- Cooper, I. (2006). Asset pricing implications of nonconvex adjustment costs and irreversibility of investment. *Journal of Finance* 61(1), 139–170.
- Daniel, K. and S. Titman (1997). Evidence on the characteristics of cross sectional variation in stock returns. *Journal of Finance* 52(1), 1–33.
- De Bondt, W. F. M. and R. H. Thaler (1985). Does the stock market overreact? *Journal of Finance* 40(3), 379–395.

- Dechow, P. M., R. G. Sloan, and M. T. Soliman (2004). Implied equity duration: A new measure of equity risk. *Review of Accounting Studies* 9(2–3), 197–228.
- DeMiguel, V., L. Garlappi, and R. Uppal (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies* 22(5), 1915–1953.
- Fama, E. F. and K. R. French (2000). Forecasting profitability and earnings. *Journal of Business* 73(2), 161–175.
- George, T. J. and C.-Y. Hwang (2010). A resolution of the distress risk and leverage puzzles in the cross section of stock returns. *Journal of Financial Economics* 96(1), 56–79.
- Gomes, J., L. Kogan, and L. Zhang (2003). Equilibrium cross section of returns. *Journal of Political Economy* 111(4), 693–732.
- Goyal, A. and S. Wahal (2015). Is momentum an echo? *Journal of Financial and Quantitative Analysis* 50(6), 1237–1267.
- Gu, L., D. Hackbarth, and T. Johnson (2017). Inflexibility and stock returns. *Review of Financial Studies* 31(1), 278–321.
- Hackbarth, D. and T. Johnson (2015). Real options and risk dynamics. *Review of Economic Studies* 82(4), 1449–1482.
- Harvey, C. R., Y. Liu, and H. Zhu (2016). . . . And the cross-section of expected returns. *Review of Financial Studies* 29(1), 5–68.
- Heston, S. L. and R. Sadka (2008). Seasonality in the cross-section of stock returns. *Journal of Financial Economics* 87(2), 418–445.
- Higham, N. J. (2002). Computing the nearest correlation matrix—a problem from finance. *IMA Journal of Numerical Analysis* 22(3), 329–343.
- Hou, K., C. Xue, and L. Zhang (2015). Digesting anomalies: An investment approach. *Review of Financial Studies* 28(3), 650–705.
- Huberman, G. and S. Kandel (1987). Mean-variance spanning. *Journal of Finance* 42(4), 873–888.
- Jegadeesh, N. (1990). Evidence of predictable behavior of security returns. *Journal of Finance* 45(3), 881–898.

- Jegadeesh, N. and S. Titman (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance* 48(1), 65–91.
- Johnson, T. C., T. Chebonenko, I. Cunha, F. D’Almeida, and X. Spencer (2011). Endogenous leverage and expected stock returns. *Finance Research Letters* 8(3), 132–145.
- Kan, R. and G. Zhou (2007). Portfolio choice with parameter uncertainty. *Journal of Financial and Quantitative Analysis* 42(3), 621–656.
- Kydland, F. E. and E. C. Prescott (1982). Time to build and aggregate fluctuations. *Econometrica* 50(6), 1345–1370.
- Lemmon, M. L., M. R. Roberts, and J. F. Zender (2008). Back to the beginning: Persistence and the cross-section of corporate capital structure. *Journal of Finance* 63(4), 1575–1608.
- Levi, Y. and I. Welch (2017). Best practice for cost-of-capital estimates. *Journal of Financial and Quantitative Analysis* 52(2), 427–463.
- Lewellen, J. (2002). Momentum and autocorrelation in stock returns. *Review of Financial Studies* 15(2), 533–563.
- Lin, X. and L. Zhang (2013). The investment manifesto. *Journal of Monetary Economics* 60(3), 351–366.
- Linnainmaa, J. T. and M. R. Roberts (2018). The history of the cross section of stock returns. *Review of Financial Studies* 31(7), 2606–2649.
- Lo, A. W. and A. C. MacKinlay (1990). When are contrarian profits due to stock market overreaction? *Review of Financial Studies* 3(2), 175–205.
- MacKinlay, A. G. and L. Pastor (2000). Asset pricing models: Implications for expected returns and portfolio selection. *Review of Financial Studies* 13(4), 883–916.
- Martin, I. and C. Wagner (2019). What is the expected return on a stock? *Journal of Finance*. Forthcoming.
- McLean, R. D. and J. Pontiff (2016). Does academic research destroy stock return predictability? *Journal of Finance* 71(1), 5–32.
- Modigliani, F. and M. H. Miller (1958). The cost of capital, corporation finance and the theory of investment. *American Economic Review* 48(3), 261–297.

- Novy-Marx, R. (2013). The other side of value: The gross profitability premium. *Journal of Financial Economics* 108(1), 1–28.
- Shumway, T. (1997). The delisting bias in CRSP data. *Journal of Finance* 52(1), 327–340.
- Shumway, T. and V. A. Warther (1999). The delisting bias in CRSP's Nasdaq data and its implications for the size effect. *Journal of Finance* 54(6), 2361–2379.
- Stigler, G. J. (1963). Competition and the rate of return. In *Capital and Rates of Return in Manufacturing Industries*, Chapter 3, pp. 54–71. Princeton University Press.
- Welch, I. (2017). *Corporate Finance, 4th edition*. United States: Lulu.
- Zhang, L. (2005). The value premium. *Journal of Finance* 60(1), 67–103.

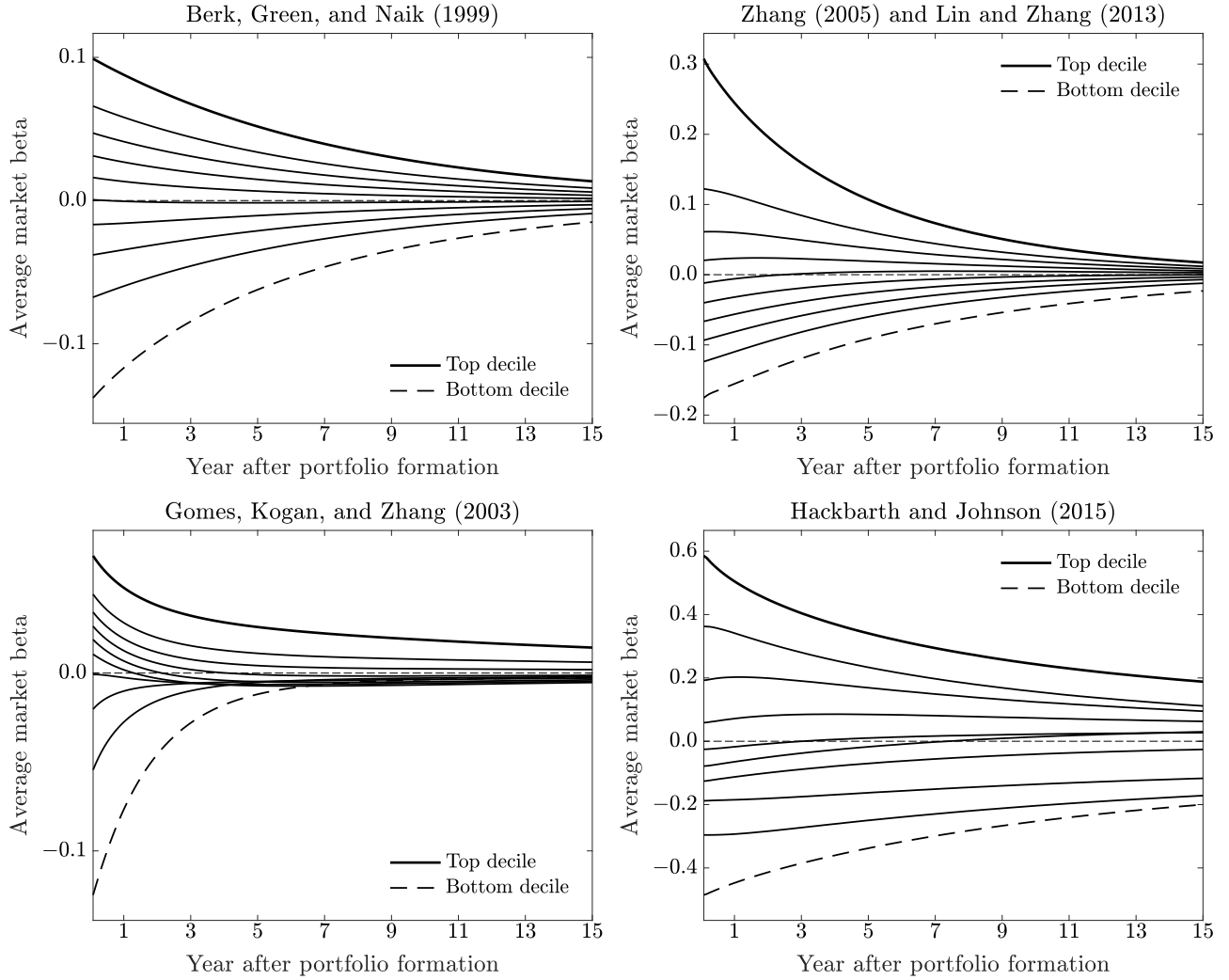


Figure A1: **Average market betas on stocks sorted by expected returns.** We simulate 1,000 months of return data from the four models described in Section 2. We run these simulations using the same parameters as those used in the original studies. We discard the first 400 months and then begin ranking stocks into deciles based on expected returns. We report average cross-sectionally demeaned market betas for these deciles over the next 15 years after portfolio formation.

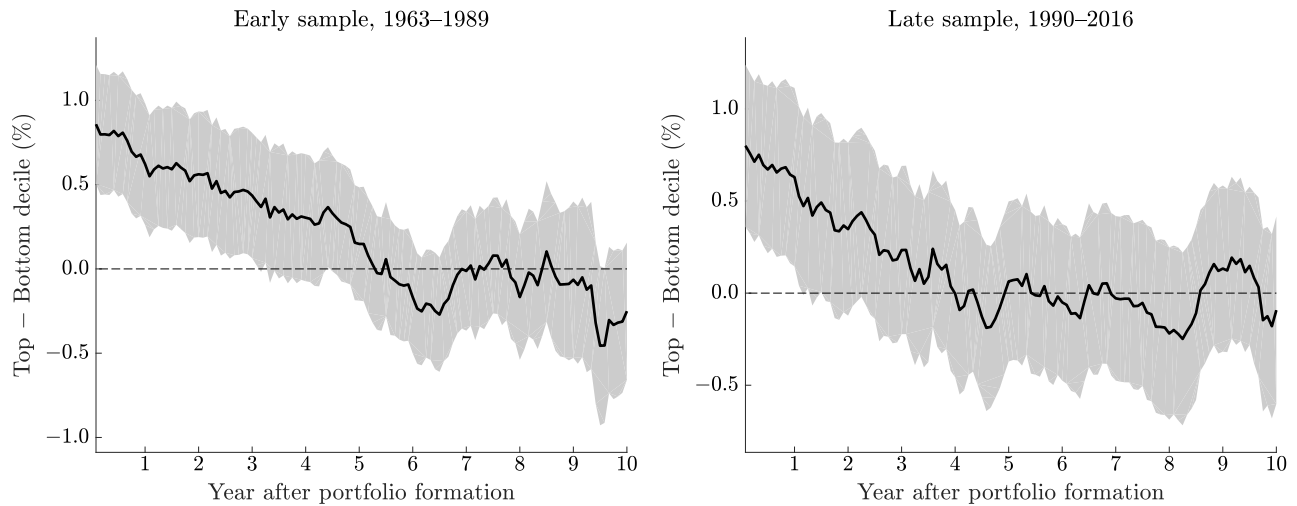


Figure A2: Average monthly returns on long-short strategies in months 1–120 after portfolio formation: Subsamples We form decile portfolios at the end of each month by sorting stocks into portfolios by different combinations of the 34 accounting-based return predictors listed in Table 3. We convert each predictor into a z -score by subtracting the cross-sectional average from the predictor and dividing the difference by the cross-sectional standard deviation. Each stock’s predictor is the average of its non-missing z -scores. We construct value-weighted portfolios each month and hold these portfolios for up to ten years. Panel A reports average returns using the 1963–1989 sample; Panel B uses the 1990–2016 sample. The average return in year k is its year k return (“forward rate”), not its average return from today to year k . The shaded areas indicate 95% confidence intervals.



Figure A3: **Average monthly returns long-short strategies in months 1–120 after portfolio formation: Condition on survival** We form decile portfolios at the end of each month by sorting stocks into portfolios by different combinations of the 34 accounting-based return predictors listed in Table 3. We convert each predictor into a z -score by subtracting the cross-sectional average from the predictor and dividing the difference by the cross-sectional standard deviation. Each stock’s predictor is the average of its non-missing z -scores. We construct value-weighted portfolios each month and hold these portfolios for up to ten years. This table differs from Panel A of Figure 3 in that we condition on firm survival: each month the sample includes only those firms that will go on to survive for at least 10 years. The average return in year k is its year k return (“forward rate”), not its average return from today to year k . The shaded areas indicate 95% confidence intervals.

Table A1: Measuring differences in average long-term returns: Equal-weighted portfolios

This table is an alternative version of Table 4 with equal- instead of value-weighted portfolios. We form decile portfolios at the end of each month by sorting stocks into portfolios by different combinations of the 46 return predictors listed in Table 3. Row “All” uses all predictors; “Accounting-based” uses the 34 predictors that use some income statement or balance sheet information; “Return-based” uses the 12 predictors that use price, return, or volume information; and “Ex-post long-lived” use the 21 predictors in Table 3 that, when used in isolation, persist for at least a year and a month. We convert each predictor into a z -score by subtracting the cross-sectional average from the predictor and dividing the difference by the cross-sectional standard deviation. Each stock’s predictor is the average of its non-missing z -scores. We construct equal-weighted portfolios each month and hold these portfolios for up to ten years. Panel A reports average monthly returns and t -values for equal-weighted high-minus-low portfolios for different horizons following portfolio formation; Panel B reports monthly CAPM alphas and t -values. The holding periods range from one month after portfolio formation to year ten after portfolio formation. Average returns are computed using the Jegadeesh and Titman (1993) method to avoid the use of overlapping return observations.

Panel A: Average returns

Return predictor	Horizon								
	1 month	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	Year 8
Average monthly long-short returns									
All	1.53	0.65	0.24	0.15	0.09	0.03	0.08	0.03	−0.03
Accounting-based	1.05	0.87	0.60	0.46	0.29	0.18	0.06	0.07	0.06
Price-based	1.65	0.18	−0.27	−0.18	−0.08	−0.19	−0.01	−0.12	−0.19
Ex-post long-lived	1.12	0.89	0.66	0.51	0.33	0.21	0.07	0.06	0.06
t-values									
All	8.39	3.82	1.50	1.00	0.68	0.25	0.67	0.26	−0.25
Accounting-based	7.00	6.02	4.30	3.37	2.30	1.52	0.56	0.64	0.55
Price-based	8.20	1.04	−1.73	−1.24	−0.62	−1.65	−0.07	−1.05	−1.72
Ex-post long-lived	6.80	5.64	4.39	3.46	2.38	1.60	0.61	0.55	0.50

Panel B: CAPM alphas

Return predictor	Horizon								
	1 month	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	Year 8
CAPM alphas									
All	1.78	0.89	0.45	0.35	0.28	0.20	0.23	0.18	0.12
Accounting-based	1.24	1.05	0.77	0.62	0.45	0.33	0.19	0.20	0.18
Price-based	1.89	0.40	-0.10	-0.01	0.07	-0.06	0.12	0.00	-0.07
Ex-post long-lived	1.32	1.09	0.84	0.68	0.50	0.36	0.20	0.19	0.17
<i>t</i>-values									
All	11.17	5.98	3.14	2.62	2.27	1.74	2.17	1.66	1.13
Accounting-based	9.18	8.16	6.12	5.08	3.98	3.14	1.94	2.08	1.95
Price-based	10.67	2.56	-0.70	-0.10	0.63	-0.58	1.13	-0.01	-0.67
Ex-post long-lived	8.84	7.64	6.15	5.03	3.85	3.00	1.83	1.77	1.67