Time Series Variation in the Factor Zoo*

Hendrik Bessembinder
W.P. Carey School of Business, Arizona State University

Aaron Burt
Michael F. Price College of Business, University of Oklahoma

Christopher Hrdlicka
Michael G. Foster School of Business, University of Washington

Initial Draft: December 2021

This Draft: August 2022

Abstract

The ability of workhorse three- to six-factor models to explain the cross-section of stock returns varies substantially over time, providing scope for time varying numbers of additional significant factors, which we document. The number of statistically significant factors, as well as the number of principal components obtained from them, varies with the cross-sectional dispersion in individual stock CAPM alphas. The number of factor principal components that optimize out-of-sample Sharpe ratios varies over time and often exceeds twenty, implying that many factors are non-redundant. The number of significant factors is strongly related to numbers of publicly-listed firms and measures of diversity in firm characteristics. On balance, our results suggest that time variation in the number of factors with significant explanatory power reflects the complexity of the economic environment, including changes in investor composition, the types of firms listed, and competitive conditions.

*We are grateful to Turan Bali, Michael Brennan, Stephen Brown, Andrew Chen, Tarun Chordia, Campbell Harvey, Theis Jensen, Bryan Kelly, Scott Murray, Stefan Nagel, Jeffrey Pontiff, and seminar participants at the University of Washington. Author emails are hb@asu.edu, aaronburt@ou.edu, and hrdlicka@uw.edu, respectively.

In the beginning there was chaos. . . Then came the CAPM. . . Then anomalies erupted, and there was chaos again...Fama and French (1993, 1996) brought order once again...Alas, the world again is descending into chaos......I did not say it will be easy! But we must address the factor zoo.

John Cochrane, AFA Presidential Address, 2011

1. Introduction.

The literature has identified hundreds of empirical variables, including firm characteristics and "factors" defined as returns to characteristic-sorted portfolios, that appear to have significant explanatory power for the cross-section of stock returns.[1] However, as the preceding quotation from the former President of the American Finance Association illustrates, there is a widespread perception that finance researchers have collectively identified too many factors. Indeed, foundational asset pricing models such as the CAPM or the consumption-based CAPM imply that a single factor should be sufficient to explain the cross-section of returns if it is measured correctly.

The empirical literature has not resolved whether researchers have identified too many factors. While we do not resolve fully resolve this issue either, we posit and provide empirical evidence that the number of economically relevant factors may vary substantially over time as a function of economic complexity. As Cochrane (2011) observes, essentially all variation in price-to-dividend ratios is attributable to changes in discount rates, i.e., expected returns. Prices, in turn, are determined in the course of market trading, based on the interaction between buy and sell orders. Cochrane (2022, page 31) observes that "the standard models do not produce a hundredth of the observed trading volume." It follows, in our view, that the determinants of expected returns need not be confined to those predicted by the standard models, can vary over time, and depend at least in part on prior trading patterns. The need to

[1] The literature has not always been consistent in usage of the terms "characteristic" and "factor." To be precise, we use the term "characteristic" to refer to a firm-level attribute, such as firm size or profitability, and we use the word "factor" to refer to a return to a long-short portfolio. More specifically, each factor is the return on a portfolio that is long a set of stocks selected with either left or right tail outcomes on a given characteristic, e.g., firms of small size or high profitability, and short a set of stocks with outcomes in the opposite tail, e.g., stocks of large size or low profitability. We do not use the term factor to refer to outcomes obtained by the statistical technique of factor analysis.

be mindful of the possibility of collective data mining and joint hypothesis testing notwithstanding, these considerations support allowing the data to speak on the issues.

More broadly, expected returns can vary for a broad variety of reasons. The risk premium associated with a given factor can vary through time as economic conditions evolve. For example, Bali, Chabi-Yo and Murray (2022) show that expected stock returns depend on the variance risk premium as well as implied volatility spreads. Such variation gives rise to the possibility that more factors will be detectable at times when premia tend to be large. In contrast to the assumptions of representative agent models, investors are diverse in terms of both their sophistication and their investment objectives. Some individual investors may seek to form mean-variance efficient portfolios, while others seek out positive skewness or "lottery" payoffs, and yet others trade in response to comments on discussion boards such as "WallStreetBets."[2] Consistent with the notion that the identity of the marginal investor can differ across stocks, Betermier, Calvet, Knüpfer, and Kvaerner (2021) document that the cross-section of expected stock returns depends in part on the proportion of individual investors that are younger as well as the proportion that are wealthier. Other investors delegate portfolio decisions to professional managers, whose objectives can differ from those of their investors due to agency issues arising from specific compensation plans (e.g., Kashyap, Kovrijnykh, Li and Pavlova, 2021). Further, the trades of professional investors can depend on considerations such as the funding liquidity of their employing firms, and market outcomes have been shown to also depend also on the leverage of financial sector firms.[3] The economic characteristics of newly listed firms can differ from those of existing firms, as shown by Campbell (2001), Fama and French (2004), and Kahle and Stulz (2017). We construct a measure of cross-sectional diversity in the observable characteristics that are collectively known to be related to expected returns, and show that this measure of firm diversity increases systematically with the number of publicly-listed firms.

[2] Recent studies documenting the diversity of individual trading approaches include Barber, Huang, Odean, Schwarz (2021), Chen, Kumar, and Zhang (2021) and Bali, Brown, Murray Tang (2017).

[3] See, for example, Koijen, and Yogo, (2019), He, Kelly and Manela (2017), Tobias, Etula, and Muir (2014) and He and Krishnamurthy (2013).

In addition, the degree of product market competition within industries can vary over time for reasons that are not fully understood, but that can systematically affect investor returns, as documented by Hou and Robinson, (2006), Bustamante, Cecilia and Donangelo, (2017), Van Reenen (2018) and Grullon, Larkin, and Michaely (2019). Further, investors may need time to learn about key parameters in a dynamic economy. Brennan (1998) for example, explores how investors' utility-maximizing portfolio decisions depend on their stock market experience, and Pastor and Veronesi (2009) explore how estimates of firm values evolve as investors learn about firms' growth prospects. We see no reason to dismiss *a priori* the possibilities that multiple factors could be relevant in explaining cross-sectional variation in firm returns, or that the number,identity and premia of such relevant factors could vary over time.

The existing empirical evidence as to whether numerous factors are required to explain the cross-section of returns is mixed. Kelly, Pruitt, and Su (2019) present evidence indicating that a relatively small number of latent factors are sufficient. In contrast, Kozak, Nagel, and Santosh (2020) assert that "models based on present-value identities or q-theory … do not really support the idea that only a few stock characteristics should matter." They incorporate these priors in Bayesian estimation of the number of principal components that explain the pricing of a set of characteristic-sorted equity portfolios and conclude that models relying on a small number of factors are not fully successful in "shrinking the cross-section."

We also conclude that factors beyond the three to six employed in widely used models are required to explain cross-sectional return variation. However, our contributions are distinct. First, we focus attention on the substantial time variation in the number of significant factors, and document that this variation has significant out-of-sample predictive power for portfolio Sharpe Ratios. Second, we show that this variation in the number of economically relevant factors is positively related to a set of variables measuring the complexity of the economy. Third, we focus on the cross-sectional variation in mean returns to individual stocks rather than focusing only on characteristic-sorted portfolios.

We show that the stock-level analysis is more informative as to the number of significant factors. We confirm that widely used three- to six-factor models outperform the CAPM when explaining mean returns to size and book-to-market portfolios, which may contribute to a perception that a few factors are sufficient to explain cross-sectional variation in mean returns. In contrast, we show that the CAPM outperforms these models when explaining the returns to industry portfolios and, in particular, when explaining the individual stocks that comprise the characteristic sorted portfolios. These results imply scope for additional relevant factors beyond those included in the widely used three- to six-factor models.

A central part of our empirical approach is to assess the number of factors with statistically significant alphas based on market-model (CAPM) regressions estimated over rolling sixty-month periods. We document that a large number of factors, exceeding one hundred during some time intervals, have significant explanatory power, and that the number of significant factors varies substantially over time. Further, the number of principal components required to explain variation in the significant factors is also large and is positively correlated with the number of significant factors, implying that these results are not attributable to researchers identifying essentially redundant factors. In addition, we show that the explanatory power of factors beyond the first few is economically meaningful. In particular, out-of-sample Sharpe ratios for portfolios constructed from factor principal components are both economically substantive and are substantially larger when allowing for a greater number of factor principal components. By comparison, Kozak, Nagel and Santosh (2018) find that increasing the number of principal components increases in-sample but not out-of-sample Sharpe ratios. We replicate their result, and show that the divergence of our outcomes from theirs is directly attributable to our allowance for time variation in the set of economically relevant factors.

To assess the scope for multiple factors to explain the cross-section of stock returns, we also estimate firm-specific CAPM alphas on a rolling sixty-month basis. Having done so, we compute the cross-sectional standard deviation of the firm-specific CAPM alpha estimates by month.[4] Time periods with more dispersion in firm-level CAPM alphas indicate greater scope for factors beyond the market to

[4] The cross-sectional standard deviation of alpha estimates is similar to the cross-sectional Gibbons, Ross, and Shanken (1989) test, but does not require estimation of the covariance matrix, which would be impossible in light of the numbers of individual assets.

have explanatory power for expected stock returns. We document that the number of significant factors is positively related to the standard deviation of firm-level CAPM alpha estimates, even while the number of significant factors is not related to the average standard error of alpha estimates or average idiosyncratic volatility. These results support the reasoning that the number of significant factors is related to the extent to which expected stock returns are left unexplained by the CAPM.

Previous literature has used the insignificance for the entire pre- and post-original sample periods of factors to argue that the set of factors that matter is small. In contrast, using subsets of these extended samples, specifically, rolling 60-month windows, we show that many factors are statistically significant in periods both before and after the range of data studied by the authors who originally highlighted the factors' performance. These superficially opposite findings are reconciled by recognizing that the statistical significance of individual factors often varies over time, and a given factor can alternate across periods of significance and insignificance.

We argue this pattern reflects that the number of factors which earn a return premium, or the magnitude of such return premia, vary over time. We acknowledge this variation could be due to two alternative hypotheses. The ebbing and flowing over time of individual factors' statistical significance could simply reflect random noise in a stable economic environment. That is, a factor with a constant, but economically modest, true premium could be associated with significant estimates during some intervals and insignificant estimates during other intervals (e.g., Jensen, Kelly and Pedersen, 2021). . Alternatively, it could be due to data mining as others have argued.

We provide a novel way to differentiate between these possibilities by assessing the extent to which variation in the number of significant factors is related to measures of changes in the economic environment. We document that the number of significant factors is related to a recession indicator variable, interest rates, the percentage of firms that pay dividends, mean institutional ownership rates, and an economic complexity index, and is particularly strongly related to the number of firms that are publicly listed. This positive variation in factor significance with measures of economic complexity support the conclusion that factor premia themselves vary over time.

We document a strong positive relation between the number of significant factors and the number of listed firms. The relation need not arise mechanically. If expected returns were fully explained by a small number of correctly-measured factors, then all that would be required to explain expected returns to additional firms would be estimation of their potentially distinct betas on those factors. We also assess whether this relation could simply arise because a larger number of firms improves statistical power and find that, while the effect is partially attributable to variation in standard errors, larger absolute alpha estimates also contribute significantly. The finding that the number of significant factors is related to the number of listed firms supports the reasoning that newly listed firms systematically differ from existing firms in terms of the risks that are relevant to investors. Indeed, we find that the significance of the number of firms is partially subsumed by our newly constructed measure of the cross-sectional diversity in observable firm characteristics, even though this measure is undoubtably a noisy estimate of cross-sectional variation in risks that are relevant to investors.

One may worry that this variation in the number of factors is merely due to variation in the overlap of the factors. We address this concern by assessing the extent to which factors that are individually significant in terms of explaining the cross-section of returns may actually be redundant of each other. To do so, we form portfolios from the principal components of factor returns and assess the relation between Sharpe ratios and the number of principal components in the portfolio. Redundancy would imply only a small constant number of principal components are needed (e.g., 3 to 6 per many canonical asset pricing models). In contrast, we show that the number of principal components that yields the largest out-of-sample Sharpe ratios both varies over time and often includes more than twenty principal components.

On balance, our findings suggest that a substantial and time varying number of non-redundant factors may be required to price the cross-section of returns as the economy evolves dynamically and new firms are listed. Further, in a dynamic economy a factor can be significant in explaining returns during some periods but not others. This suggests the desirability of a degree of caution in interpreting the results of existing out-of-sample tests, as insignificant out-of-sample outcomes need not imply that the factor was unpriced in the original sample period. However, accommodating such time variation may also provide additional scope for specification searches or other sources of bias, and highlight the importance of developing of econometric methods suitable for such a dynamic environment.

2. Data and Key Variables

a. Data Sources

We rely on two main data sources: monthly returns to individual common stocks and monthly returns to 205 factors derived from cross-sectional characteristics previously documented in the literature. The individual stock returns are obtained from CRSP, and include all stocks listed on the NYSE, AMEX and NASDAQ markets with a share code of 10 or 11 during the period July 1926 to December 2020. The factors are the 161 "clear predictors" and 44 "likely predictors" identified by Chen and Zimmerman (2021).[5] We estimate factor exposures and alphas based on rolling sixty-month regressions. As a consequence, to enter our sample a stock or factor must have 60 prior months of non-missing returns, and our assessment of alpha estimates for stocks and factors begins with estimates obtained for June 1931.[6] We obtain industry, size and book-to-market sorted portfolios along with the market excess returns from Kenneth French's website.

b. Assessing the number of significant factors

We first seek to assess the number of economically relevant factors at each point in time. For each month t and for each factor f, we estimate 60-month rolling CAPM alphas by means of the following regression over the period t-59 to t:

$$R_{ft} = \alpha_{ft} + \beta_{ft} R_{MKTRF,t} + \epsilon_{ft}$$

[5] The authors graciously posted their data to https://www.openassetpricing.com/. We find qualitatively and quantitatively similar results using the factor data of Jensen, Kelly, and Pedersen (2021).
[6] We rely on Hansen-Hodrick standard errors with a bandwidth of sixty for all analyses.

where R_{ft} is the month t return on factor f and $R_{MKTRF,t}$ is the month t value-weighted market excess return, obtained from Kenneth French's website. A positive and significant alpha estimate indicates that the factor has explanatory power for the cross-section of stock returns beyond that which is explained by returns to the overall market. We identify a factor as significant for a given time period if the *t*-statistic for the alpha estimate exceeds positive 3.00, the level recommended by Harvey, Liu and Zhu (2016) to allow for potential effects of multiple hypotheses testing and specification searches in the prior literature.[7] Having done so, we tabulate the number of factors with significant prior-sixty-month CAPM alphas as of each month. The orange solid line in Figure 1 panel A displays time series variation in the number of significant factors. For comparison, we also display with the dotted blue and dashed grey lines the numbers of factors that are significant based on alternative t-statistics hurdles of 1.96 and 4.00, respectively.[8] While we focus on outcomes based on a t-statistic of 3.00, all three measures are highly correlated and support similar conclusions regarding the importance of allowing for time variation in the number of factors.

The literature has noted that anomalous returns to long-short portfolios are often attributable to the short leg, presumably due to higher costs of shorting shares. We assess this issue by displaying on Panel B of Figure 1 the number of significant factors over time based on long-short, long, and short portfolios. The orange solid line counts a factor as significant based on the alpha of the factor's long-short portfolio. We rely on this measure throughout the paper. However, for completeness, we also display with the grey dashed line the number of factors that are significant based on the alpha of the factor's long-only portfolio, and with the blue dotted line the number of significant factors based on the short-only portfolio. The figure shows that long-only factor portfolios have significant alphas more often than short-only portfolios.

[7] Factors are generally constructed by the original authors to have a positive mean return (e.g., the size factor is defined as return the return on a small firm portfolio less the return on a large firm portfolio, not vice versa). As a consequence, fewer than 4% of the alphas we estimate are negative.
[8] In their replication studies, Hou, Xue and Zhang (2020) and Chen and Zimmerman (2021) rely on a t-statistic of 1.06. As noted, Harvey, Liu and Zhu (2016) recommend reliance of t-statistics of 3.0 or greater, while Chordia, Goyal and Saretto (2020) argue for a threshold of 3.78.

c. Measuring stocks' unexplained mean return variation using CAPM alphas

We wish to assess the extent to which cross-sectional variation in mean stock returns allows scope for multiple factors to be relevant. To do so, we estimate alphas from simple market-model regressions of excess firm returns on excess market returns, in each month using data drawn from the prior sixty calendar months. In particular, letting R_{it} denote the month *t* excess return for stock i, we estimate for rolling sixty-month intervals

$$R_{it} = \alpha_{it} + \beta_{it} R_{MKTRF,t} + \epsilon_{it}.$$

Having done so, we compute the cross-sectional standard deviation of the α_i estimates for each month t.[9] The standard deviation or variance of alpha estimates can be thought of as a simpler (unweighted) analog to the Gibbons, Ross, and Shanken (1989) test statistic, except for the focus on deviations from the sample mean alpha estimate rather than deviations from zero. In practice, the information contained in deviations of alpha from zero as opposed to the sample mean is essentially identical.[10] The key advantage to using the standard deviation measure is that the variance-covariance matrix need not be estimated and inverted, which would be impractical in light of the number of individual stocks. The degree of variation across stocks in these alpha estimates provides a measure of the extent to which mean stock returns over the sixty months diverge across stocks in a manner not explained by stocks' betas with respect to the overall market. Time series variation in the degree of cross-sectional variation in alpha estimates, in turn, give indication of the scope for the number of significant factors to vary over time.

While our main focus is on simple market-model alphas, we also consider cross-sectional and time series variation in firm-specific alphas that are estimated with respect to the well-known three- to six-factor models presented by Fama and French (1993), Fama and French (2015), Fama and French

[9] For some tests we rely on value-weighted outcomes, in which case each alpha estimate is weighted by the stock's market capitalization at the beginning of the alpha estimation period, *t-60*. For ease of interpretability, we standardize this variable such that it has mean zero and standard deviation equal to one.

[10] Figure IA-1 in the Internet Appendix displays uncentered and centered measures over time. The two measures are highly correlated; Pearson correlation coefficients are 0.93 and 0.95 for the equal- and value-weighted measures. We rely on the centered measure throughout the paper as it is less noisy than the uncentered measure.

(2018), Carhart (1997), Pastor and Stambaugh (2003), Stambaugh and Yuan (2017), Barillas and Shanken (2018), and Hou, Xue, and Zhang (2015).

3. The Evolution in the Number of Factors Over Time

a. The number of Identified Factors and Unexplained Variation in Mean Stock Returns

Figure 2 displays information regarding the scope for multiple factors to explain returns, and that motivate our subsequent analyses. Panel A of Figure 2 displays the number of factors amongst the 205 studied by Chen and Zimmermann (2021) that were identified in the CRSP data for each of the indicated dates. The dotted blue line displays the factor count starting from the earliest data used in the original studies, while the solid black line includes factors as of the (often earlier) date for which all data necessary to construct the factors is now available.[11] All 205 factors draw on data from 1995 or earlier, and approximately 200 of these factors draw on data from 1991 or earlier. In contrast, only about 10 factors employed data from years prior to 1961 in the studies that originally identified the factors. At present, however, sufficient data is available to implement over fifty factors in data drawn from June 1931 or later, and to implement nearly 120 factors in data drawn from 1961 or later. The key point conveyed by Panel A of Figure 2 is that the literature has identified a substantial number of factors that can be studied even in data from the earlier decades covered by the CRSP dataset.

Panel B of Figure 2 displays information regarding CAPM alpha estimates for factors as well as individual stocks. The dotted orange line displays the number of factors with statistically significant (*t*-statistic greater than 3.00) alpha estimates based on return data for the prior sixty months. The solid blue line displays the cross-sectional standard deviation of estimated individual stock CAPM alphas over the

[11] The single most common reason that we can now construct factors for time periods that were not included in the original studies is that additional accounting data has become available in the intervening years. The second most common reason is that daily data necessary to construct some measures of trading and liquidity for periods prior to 1962 were added to the CRSP data in 2006.

same periods. As noted, we view the cross-sectional variation in CAPM alphas to comprise a useful measure of the amount of variation in mean stock returns that can potentially be explained by pricing factors beyond the overall market, i.e., as a measure of the scope for additional factors to be relevant.

The two curves displayed on Panel B of Figure 2 diverge during the 1930s and 1940s, decades when returns were unusually volatile and for which there is even now insufficient data to implement many factor models. However, since approximately 1950 the number of statistically significant factors and the cross-sectional variability of individual firm market-model alphas appear to move reasonably closely together. We study this relation more rigorously in Section 4 below, focusing in particular on the question of whether the comovement in the number of significant factors and cross-sectional variation in firm-level alphas represents variation in expected returns that require more factors vs. variation in idiosyncratic volatility that could alter the scope for data mining.

As noted, the factors studied here are defined as returns to characteristic-sorted portfolios. As such, the number of potential factors is intrinsically related to the number of firm-level characteristics for which data is available. Figure 3 presents evidence regarding time series variation in the number of characteristics (among the 205 considered) for which characteristic data is available in each month. Panel A of Figure 3 displays outcomes in terms of the numbers of firms, while Panel B displays outcomes in terms of percentages of firms, for which characteristic data is available. Panel A reveals that the numbers of firms for which large numbers of factors are available grew rapidly between the initial years of the sample until the late 1990s. This growth was attributable both to increases in the number of listed firms (demonstrated, for example, by the vertical spike in 1972 when Nasdaq firms were added to the CRSP dataset) and increases in the number of characteristics for which requisite data is available. To distinguish the separable effects of changes in the number of firms and changes in the set of characteristics for which data is available, Panel B of Figure 3 displays for each month the percentage of firms for which the indicated numbers of characteristics are available. The data on Panel B demonstrate that the percentage of firms with a large number of available characteristics has steadily increased over time. The percentage of firms for which 140 or more characteristics are available, for example, has increased from zero in 1970 to approximately 20% in 1980 and 50% in 2015. This implies that the reduction in the numbers of firms with many available characteristics in the years since 2000, as displayed on Panel A of Figure 3, is due to decreases in the number of listed firms. It can also be observed that the percent of firms with large numbers of characteristics decreases during the final few years of the sample. This reflects the recent growth in the number of IPOs, in combination with the fact that prior accounting and return data is necessary to compute some characteristics.

Figure 4 displays data informative as to the extent to which the use of prominent multi-factor models improves on the CAPM in terms of reducing the cross-sectional standard deviation of individual firm alpha estimates. Large reductions in the variability of firm alpha estimates, if observed, would be indicative that the factors employed in these workhorse models have substantive explanatory power for mean firm returns, implying limited scope for additional factors. The alternative factor models we assess include the Fama and French (1993) 3-factor model (FF3F), the Fama and French (2015) 5-factor model (FF5F), the Fama and French (2018) 6-factor model (FF6F), the FF3F model augmented with the Carhart (1997) momentum factor (FF3F+UMD), this model augmented with the Pastor and Stambaugh (2003) liquidity factor (FF3F+UMD+PSLIQ), the Stambaugh and Yuan (2017) four-factor model (M4), the Barillas and Shanken (2018) 6-factor model (BS6F), and the Hou et al. (2015) q-factor model (Q4).

Panel A of Figure 4 displays the cross-sectional standard deviation of individual firm alpha estimates obtained from rolling 60-month regressions on the various sets of factors, by month. The most noteworthy result is that the multifactor models do not outperform the CAPM in terms of reducing the cross-sectional variability of alpha estimates. The variability of individual stock CAPM alphas, displayed as the solid black line, has been the lowest or among the lowest as compared to the multi-factor models, particularly since about 1961. More specifically, the variability of CAPM alphas is smaller than the variability of alphas from *any* of the six other factor models considered in 65% of the individual months from 1963 to 2020.

That is, the prominent three- to six-factor models typically leave more, not less, unexplained variation in mean individual stock returns, as compared to the CAPM. This reduction in explanatory power is all the more notable because the market return is included as a factor in the three-to-six factor models. The inclusion of factors in addition to the market must necessarily improve fit (as measured by R-squared) in the time series factor regression for each individual stock. However, the inclusion of these additional factors results in stock-specific intercept estimates that are on average further from rather than nearer to the benchmark of zero that is implied by the factor models. The implication is that the non-market factors included in the three-to-six factor models degrade the ability to explain average individual stock returns.

In Panel B of Figure 4 we display cross-sectional standard deviations of alpha estimates that correspond to those on Panel A, except that the rolling sixty-month regressions rely on the thirty industry portfolios identified on Kenneth French’s website. Consistent with the results reported by Ahmed, Bu, and Tsvetanov (2019), the CAPM often performs better than the multi-factor models when explaining cross-sectional variation in mean industry portfolio returns as well. In particular, we find that the cross-sectional standard deviation of CAPM alphas is smaller than any of the factor models in 15% of the individual months since 1963, and is second lowest in another 9% of the individual months.

Kozak, Nagel, and Santosh (2020) observe that the widely-implemented three to six-factor models have mainly been applied to explain the performance of portfolios formed from a small subset of cross-sectional predictive characteristics. Here, we consider the twenty-five size and book-to-market portfolios identified by Fama and French (1993). The results displayed on Panel C show that the multi-factor models virtually always outperform the CAPM when explaining returns to the size and book-to-market portfolios, particularly in recent decades. Specifically, the CAPM is outperformed in terms of cross-sectional standard deviation of alpha estimates by at least one of the factor models in 96.8% of all sample months and 99.6% of months since 1963.

The data displayed in Figures 2 to 4 support several conclusions. First, there is substantial time series variation in the extent to which the CAPM as well as widely-implemented three to six-factor models leave unexplained cross-sectional variation in mean asset returns, thereby leaving scope for additional factors to enter the factor zoo. Further, changes in the numbers of listed firms and in the availability of data necessary to construct factors implies scope for the number of factors that are both economically relevant and observable to econometricians to change through time. In addition, the widely-implemented factor models perform well when explaining mean returns to characteristic-sorted portfolios, in particular the twenty-five size and book-to-market portfolios identified by Fama and French (1993). In contrast, however, when focusing on industry portfolios or, in particular, individual stocks, the widely-implemented factor models do not typically outperform, and in many instances perform worse, than the CAPM in terms of explaining cross-sectional variation in time series mean returns. The fact that the three- to six-factor models outperform the CAPM when explaining mean returns to size and book-to-market portfolios may contribute to a perception that a few factors are sufficient to explain cross-sectional variation in mean returns. In contrast, the fact that these models perform more poorly than the CAPM when explaining individual stock returns implies scope for additional factors beyond those in the widely-used three- to six-factor models.

b. **Are some factors redundant?**

The data displayed on Panel B of Figure 2 shows that as many as 95 factors have statistically significant alpha estimates at certain times during the sample period. Of course, some factors are similar to each other in their construction, and the economic information contained in outcomes on similarly constructed factors could overlap substantially.[12] To assess the extent to which the factors studied here contain overlapping information, we perform principal component analyses to explain variation in the factors.[13] This approach is somewhat similar to that employed by Gu, Kelly, and Xiu (2020) in that they also rely on principal component analysis. However, we focus on the principal components of factors

[12] Jensen, Kelly, and Pedersen (2021) for example, divide factors into thirteen "clusters," such as those based on corporate investment, corporate profitability, return momentum, return skewness, etc.

[13] A simple alternative approach would involve cross-sectional Fama-MacBeth regressions of firm returns on returns to all factors. However, since the number of factors studied here exceeds the number of observations in our rolling sixty-month regressions this approach is infeasible. Lopez-Lira and Roussanov (2021) apply principal component analysis to individual stock returns, and show that portfolios hedged against these components earn high returns relative to their risk. In contrast, we focus on principal components of factor returns because we are interested in the extent to which various factors are redundant.

identified by other researchers while they study principal components of security returns without allowing for time variation.

Figure 5 displays the number of principal components required to explain 95% of the variation across all 205 factors, as well as the number of principal components required to explain 95% of the variation in the statistically significant factors, when each is assessed on a rolling sixty-month basis. Figure 6 displays more granular information, including the number of principal components required to explain 50%, 60%, 75%, 90%, and 95% of the variation in the set of all factors.

This data reveals that, consistent with the results reported by Hou, Xue, and Zhang (2020), approximately fifty to sixty percent of the variation in the factors can be explained by a small number of principal components, ranging at various times from three to eight. However, explaining a larger portion of the variation in the factors requires many more principal components. To explain 95% of the variation requires between 29 and 40 principal components for every rolling sixty-month window from the late 1950s through the end of the sample period.[14] Further, while the incremental explanatory power of additional factors decreases by construction, this does not necessarily imply that the incremental factors are of minor economic significance. Sharpe (1994) shows that estimated *t*-statistics (which here exceed 3.00 for all significant factor alphas) are proportional to Sharpe ratios.[15]

To assess the economic significance of increasing numbers of factors, we measure portfolio Sharpe ratios as the number of principal components is increased, first on an in-sample basis and subsequently on an out-of-sample basis. More specifically, for each month t, we construct optimized portfolios based on the first, first two, first three, etc., up to the first fifty-nine principal components. In each case, portfolio weights are chosen to optimize the portfolio Sharpe ratio. On Figure 7 we display

[14] The number of principal components estimated from monthly data is inherently limited by the fact that only sixty data points are employed for each estimate. When we repeat this procedure using daily data, the total number of principal components is nearly equal to the number of statistically significant factors, suggesting that virtually all of the factors contain significant independent information for the cross-section of stock returns.

[15] As Linnainmaa and Roberts (2018) observe, a non-zero alpha implies that the combination of the right-hand side factors is not mean-variance efficient, and that a portfolio's Sharpe ratio can be improved by adding the left-hand side factor.

Sharpe ratios for portfolios formed from increasing numbers of principal components, in sample. Marginal Sharpe ratios are reflected in the width of the bands displayed on the figure. Panel A plots the Sharpe ratios themselves, while Panel B plots the portfolios' Sharpe ratios as a proportion of the Sharpe ratio of the portfolio comprised of fifty-nine principal components (the maximum possible to consider).

Figure 7 reveals that the Sharpe ratio continues to increase in an economically meaningful manner as the number of factor principal components increases. Even the higher-order principal components provide a non-negligible marginal contribution to a portfolio's Sharpe ratio. In particular, the first twenty principal components rarely contribute more than half of the Sharpe ratio of the portfolio constructed from all available principal components. In addition, the marginal Sharpe ratio contribution of additional principal components exhibit considerable time variation in their relative magnitudes. Overall, the evidence displayed on Figure 7 supports the reasoning that many factors contribute relevant economic information not captured by the other factors.

Table 1 provides additional data that is useful in assessing the extent to which the large number of factors considered in this study contain distinct information. Specifically, we report the results of regressions where the dependent variable during each month is the number of statistically significant factors as measured over the prior sixty months, and the explanatory variables are the number of principal components required to explain 95% of the variation in all factors or 95% of the variation in the statistically significant factors. Columns (1) and (2) pertain to factors measured at the monthly horizon while, to assess robustness, columns (3) and (4) report results for factors measured at the daily horizon.

The central result observed in Table 1 is that there is a strong positive and statistically significant relation between the number of statistically significant factors and the number of principal components in the factors. This result implies that, in those months where more factors have significant CAPM alphas there is also more independent variation in the factors. This result would not be anticipated if researchers had systematically identified new factors that essentially duplicated the information contained in alternative factors. The R-squared statistics for these regressions are quite high, ranging from 0.65 (column 1) to 0.94 (column 2). We conclude that time variation in the number of statistically significant factors is not primarily attributable to increases or decreases in the number of factors that essentially replicate or are replicated by the economic information contained in other factors.

c. **Out-of-sample evidence**

The results reported in the prior section support the conclusions that the number of factors with significant explanatory power for returns is relatively large and varies over time, and that the factors are both economically important and to a substantial extent not redundant of each other. However, the results to this point are all in-sample, as the significance of factors and the maximum implied Sharpe ratios are always assessed within the same sixty-month period. We next assess the extent to which the factors are or are not redundant and improve Sharpe ratios on an out-of-sample basis.

For each month, t, we consider all factors that have non-missing returns for the "in-sample" months $t - 59$ to t. From these, we compute the in-sample eigenvalues and eigenvectors of the standardized factor covariance matrix, sorting the in-sample eigenvectors by decreasing order of their corresponding eigenvalues. For the out-of-sample evaluation we focus on the 36 months from $t+1$ to $t+36$, and multiply these out-of-sample factor returns by the in-sample eigenvectors to create out-of-sample principal components.[16]

We then construct portfolios comprised of increasing numbers of these out-of-sample principal components. More specifically, for each month t, we construct portfolios based on the first, first two, first three, etc., up to the first fifty-nine out-of-sample principal components. Portfolio weights are chosen in each case to maximize the portfolio's *in-sample* Sharpe ratio. We then focus on the returns earned by these portfolios during the subsequent 36-month out-of-sample period. Note that, unlike in-sample

[16] While we report results for a 36-month out-of-sample window, outcomes are similar for both 12- and 60-month windows. We focus on 36 months as a balance between greater noise at short horizons and a potential loss of economic relevance at long horizons attributable to time variation in the economic importance of individual factors.

outcomes, the Sharpe ratios for these out-of-sample returns do not necessarily rise as the number of principal components used to form the portfolios is increased.

Panel A of Figure 8 displays for each month the number of principal components (from 1 to 59) that results in the highest out-of-sample Sharpe ratio. While this figure can only be observed on an ex-post basis, it is informative as to the extent to which factors are redundant on an out-of-sample basis.

Panel B of Figure 8 displays out-of-sample Sharpe ratios. It would not be possible to know *ex ante* how Sharpe ratios vary as a function of the number of principal components included in the out-of-sample portfolios. However, levels of out-of-sample Sharpe ratios and variation in Sharpe ratios as a function of the number of principal components are informative as to whether factors identified in-sample are useful for portfolio formation out-of-sample as well as the extent to which the factors are or are not redundant of each other. More specifically, Panel B displays the Sharpe ratios obtained if the portfolios are formed from the first five principal components, the average Sharpe ratio obtained from all possible numbers of principal components (from one to fifty-nine), the Sharpe ratio obtained from the maximum number of principal components, and the maximum Sharpe ratio obtained for portfolios based on any of one to sixty principal components. Panels C and D of Figure 8 are analogous to Panels A and B, respectively, but display results that are obtained when principal components are formed from only those factors with significant in-sample ($t - 59$ to t) alphas.[17]

The data displayed on Figure 8 indicates that the factors have substantial out-of-sample explanatory power for portfolio returns. The Sharpe ratio for out-of-sample portfolios formed from the first five principal components are virtually always positive, average about 0.5, and reached 1.0 during some portions of the 1980s and 1990s. The average (across various numbers of principal components) Sharpe ratios are generally (particularly since about 1950) larger than those based on five principal components and approached 1.5 during portions of the 1980s and 1990s. The maximum (across any of the numbers of principal components, corresponding to the number of principal components identified on

[17] Some sections of the lines in Panel D are missing due to an insufficient number of significant factors (i.e., only 1 or 0) during those time periods.

Panel A) are always positive, consistently exceeded one from about 1974 to 2005, and reached 2.5 at times. As noted, the maximum Sharpe ratio is greater than the Sharpe ratio for the maximum number of principal components as (unlike a purely in-sample exercise) measured performance need not increase when the portfolio includes additional components.

The data displayed on Panel C of Figure 8 shows that the number of principal components contained in the portfolios with the highest out-of-sample Sharpe ratios is most often as large or nearly as large as the number of factors that were statistically significant in-sample. The implication is that there is almost no redundancy among the statistically significant factors. During approximately the 1980 to 2002 period the number of principal components (drawn from factors significant in sample) that maximized out-of-sample portfolio Sharpe ratios consistently exceeded twenty, and in some months exceeded forty. Maximum Sharpe ratios displayed on Panel D virtually always exceed one half, most often exceed one, and for a substantial number of months exceed two. By comparison, Gu, Kelly, and Xiu (2020) estimate a (non-time varying) out-of-sample Sharpe ratio of 1.35 for a long-short portfolio formed based on their machine learning approach.

The data displayed in Panel D also indicates a strong correlation between out-of-sample Sharpe ratios and the number of factors that are significant in sample. This result also supports that the number of economically significant and non-redundant factors vary over time. However, the fact that factors estimated within sample can be used to improve Sharpe ratios out-of-sample also suggests that time variation in the economically relevant factors is not so rapid as to render the factors useless after the estimation period where their significance is initially assessed.

As noted, the out-of-sample Sharpe ratios displayed on Figure 8 are economically substantive. To obtain insights as to the data structure that could allow for such large out-of-sample Sharpe ratios, we conduct a series of simulations with parameters calibrated to the properties of the return data, as described in detail in Section II of the Appendix. The most important insight obtained based on these simulations is that two key features of the actual data, including (i) maximized out-of-sample Sharpe ratios most often exceed one and (ii) the correlation between the number of factors that are priced in-sample and the number of out-of-sample principal components contained in the highest Sharpe ratio portfolio is high, are only obtained when the correlation across returns on various priced factors is low, more specifically below approximately 0.10. In contrast, high correlations in the returns to the priced factors, implying redundancy, do not allow for these out-of-sample outcomes to be observed.

The results we report on Figure 8 can be contrasted to results obtained by Kozak, Nagel, and Santosh (2018), who employ a procedure that corresponds to our own except that (i) they focus on just thirty portfolios comprising the long and short legs of fifteen arbitrage strategies, while we consider 205 factors, and (ii) they do not allow for time variation in factor premia. They report that while factors beyond the first few principal components contribute substantially to in-sample Sharpe ratios, principal components beyond the first few do not enhance the Sharpe ratio out of sample. They argue that this result is to be expected since even a relatively small number of arbitragers should "be sufficient to ensure that near-arbitrage opportunities—that is, trading strategies that earn extremely high Sharpe ratios do not exist." Yet, we document high out-of-sample Sharpe ratios, to which principal components beyond the first few contribute substantively. We posit that our results may reflect the complexity of an economy where the magnitude of factor premia can vary through time. That is, while competition among arbitragers should reduce large Sharpe ratios, a dynamic economy may require continual investor learning, and it may take a substantial amount of time for arbitragers to obtain reliable estimates of factor premia such that they can be traded on. The arbitrager also faces the risk that a recently identified factor premium may not persist.

We assess whether time variation in factor premia can reconcile the differences in our findings as compared to those of Kozak, Nagel and Santosh (2018). First, we replicate their results. In particular, we compute in- and out-of-sample Sharpe ratios for portfolios formed from the principal components of the same 30 long and short factors that they consider, when the in- and out-of-sample periods are defined based on the first and second half (22 years each) of their sample period.[18] We then assess the effect of

[18] Due to missing data, we identify only 24 principal components.

instead focusing on shorter subsamples, ranging from 36 to 120 months, to define the in- and out-of-sample periods.

Figure 9 displays the findings. Panel A shows the in-sample Sharpe ratios as a function of number of principal components in the portfolio, while Panel B shows the corresponding out-of-sample Sharpe ratios. The bright red lines in both panels confirm the findings of Kozak, Nagel and Santosh (2018). In particular, in-sample Sharpe ratios increase as more principal components are used to form portfolios, while out-of-sample Sharpe ratios increase little or at all beyond the 5th principal component. The additional 6 lines on Figure 9 display average (across time) of portfolio Sharpe ratios for shorter estimation windows. The results show that out-of-sample Sharpe ratios increase to a maximum of nearly 2.5 as the estimation windows are decreased from twenty-two years, as in Kozak, Nagel and Santosh (2018), to 36 months. Further, out-of-sample Sharpe ratios continue to increase beyond the first 5 principal components to approximately twelve. That is, more principal components are relevant out-of-sample as estimation windows become shorter. It can also be observed that out-of-sample Sharpe ratios reach a maximum as the number of principal components increases. That is, out-of-sample Sharpe ratios do not monotonically increase with the number of principal components. On balance, our results support the conclusions of Kozak, Nagel and Santosh (2018) when the research design focuses on long subperiods, but shows that results differ when in-sample estimation and out-of-sample testing are based on shorter intervals. That is, a research design that accommodates time variation in the number of economically relevant factors leads to conclusions that differ from Kozak, Nagel and Santosh (2018).

4. Time series variation means factors have explanatory power outside original sample periods

It has been suggested that most empirical findings related to factors are attributable to specification searches (also referred to as “data snooping” or “p-hacking”) and a failure to incorporate appropriate multiple testing procedures. However, more recent authors, present evidence indicating the large majority of the factor-related findings can indeed be replicated, do not arise from specification searches, and survive adjustment for multiple testing.[19] In addition to the out-of-sample evidence provided above, we contribute to this debate in three ways. First, we assess the extent to which factors have statistically significant explanatory power in subperiods before and after those examined in the studies that originally identified the factors. In doing so, we extend the related results reported by McLean and Pontiff (2016), Linainmaa and Roberts (2018) and Ilmanen, Israel, Moskowitz, Thapar, and Lee (2021) in that we study a substantially larger set of factors.[20] Second, while the related studies have mainly assessed whether factors do or do not have explanatory power for the pre- and post-sample periods as a whole, we assess the extent to which factors' explanatory power changes over time, both within and outside of the authors' original samples. Third, we go on to identify the economic determinants of such time series variation.

Figure 10 displays information regarding the significance of each factor over time. The figure includes one row for each factor, and a column for each sample month.[21] A given row and column contains a dot (either blue or grey) if the *t*-statistic on the alpha estimated in a market-model regression of the factor return on the overall market return over the prior sixty months is greater than 3.00. In addition, each row contains a green dot that denotes the earliest data used in the original study that identified the

[19] Studies that conclude that factor-based evidence is unreliable include Harvey, Liu, and Zhu (2015), Linnainmaa and Roberts (2018), Chordia, Goyal, and Saretto (2020), and Hou, Xue, and Zhang (2020), while the studies indicating that factors do reliably explain returns include Chen (2021), Chen and Zimmerman (2021), and Jensen, Kelly, and Pedersen (2021).

[20] However, replication rates are not directly comparable across our study and theirs, as we focus on a set of 205 factors that were previously verified by Chen and Zimmerman to have significant explanatory power within the authors' original sample periods. In contrast, only 85 of the 97 factors studied by McLean and Pontiff (2016) have an in-sample t-statistic greater than 1.50, and only 32 of the 36 factors studied by Linnainmaa and Roberts (2018) have an in-sample t-statistic greater than 1.96. Ilmanen, Israel, Moskowitz, Thapar, and Lee (2021) study just four factors, but over a 100-year sample period, and in several distinct asset classes. They report little evidence that arbitrage reduces factor premia over time.

[21] For purposes of Figure 6, we follow Chen and Zimmerman (2021) in assigning factors to categories, including (1) "Price", which includes factors mainly constructed from return data, (2) Accounting, which includes factors that rely on financial statement data, (3) Analyst, which rely on analyst estimates, (4) Trading, which use volume and transactional data, (5) 13F, which use institutional holdings data, (6) Options, which use options-related data, and (7) Other, which include hand-collected or other non-standard data. Due to the small numbers of factors, we combine the last two categories as "Other".

factor, a red dot that denotes the latest data used in the original study, and a magenta dot that indicates the earliest date for which we are able to estimate the factor's alpha based on data now available. It is, of course, not possible even now to ascertain if the factor had significant explanatory power for returns for those dates that are earlier than the magenta dots.

Two points can be observed visually on Figure 10. First, factors often display statistically significant explanatory power in data drawn from months both before and after the data used in the original study that identified the factor. Panel A of Table 2 reports on the extent, indicating that over three quarters (77%) of the factors have significant explanatory power during at least one sixty-month interval prior to the range of dates used in the original studies, and a remarkable 93% have significant explanatory power during at least one sixty-month interval after the range of dates used in the original studies, when significance is assessed by a t-statistic on the alpha estimate of 1.96 or greater. If statistical significance is defined based on a larger t-statistic (lower p-value) the proportion of factors that are significant outside of the original sample period declines, but remains large. For example, applying the t-statistic of 3.00 (p-value is .003 or less) used elsewhere in this paper, 54% of factors are significant during at least one earlier sixty-month interval and 69% are significant during at least one subsequent sixty-month interval (as compared to the original study sample period).. This out-of-sample evidence supports the reasoning that the factors' success in the original studies did not simply reflect data mining or specification searches.

The second observation that can be gleaned from Figure 10 is that the statistical significance of individual factor alphas varies over time; in many cases a given factor is significant for periods spanning multiple years, loses significance for a time, and then regains significance. Panel B of Table 2 reports on the distribution of the number of non-overlapping periods, or "spells" of significance, for various t-statistic cutoffs. For example, relying on a t-statistic cutoff of 3.00, the median number of significance spells is 6 per factor, while the mean is 7.7. Panel C of Table 2 reports on the distribution of the duration of such significance spells. Once again based on a t-statistic cutoff of 3.00, the median length of a significance spell is 13 months, while the mean length is 22 months.

A pattern whereby statistical significance for individual factors ebbs and flows over time could simply reflect random noise in a stable economic environment. That is, a factor with a constant premium equal to zero or an economically modest level could be associated with significant estimates during some intervals and insignificant estimates during other intervals. Alternatively, the pattern could reflect that the number of factors that earn a return premium, or the magnitude of such return premia, vary over time. In the following sections we present evidence that helps to differentiate between these possibilities by assessing the extent to which variation in the number of significant factors is or is not related to measures of changes in the economic environment.

5. Variation in Significant Factors and Firm Alphas

The empirical results reported in the prior sections demonstrate that (i) a substantive but time-varying number of factors have explanatory power that is both statistically and economically significant for the cross-section of stock returns during certain time periods, (ii) most of the factors are significant in periods before and after the time intervals studied by the authors who originally identified them and are useful in out-of-sample portfolio selection, (iii) the factors are generally not redundant of each other, in that the number of principal components required to explain their variation is substantial, both within and out-of-sample, (iv) the extent to which the simple CAPM explains the cross-section of mean returns to individual stocks varies substantially over time, and (v) popular three- to six-factor models most often underperform the CAPM in explaining the cross-section of mean returns to individual stocks, and, therefore, do not substantively diminish the potential role of additional factors in explaining returns.

As noted, we view cross-sectional variability in CAPM alpha estimates as being informative as to the scope for multiple factors to explain average returns. However, it has been documented, e.g., by Goyal and Santa Clara (2003), that aggregate idiosyncratic volatility, i.e., the cross-sectional average standard deviation of residuals from market-model regressions, varies over time. An alternative hypothesis is that the standard deviation of *ex post* alphas varies over time because of variation in aggregate idiosyncratic risk, not because of variation across stocks in *ex ante* expected returns.

To guide our analysis, we provide in section I of the Appendix to this paper an assessment of the statistical determinants of such variation. Let μ_g denote the vector of expected return premia associated with non-market factors and let Σ_γ denote the cross-sectional covariance matrix of firm exposures to the non-market factor returns. Expression (7) in section I of Appendix verifies that the cross-sectional variance of market-model alphas depends both on the cross-sectional average idiosyncratic volatility, denoted, $\overline{\sigma_\varepsilon^2}$, and the weighted sum of squared non-market factor return premia, captured by the term $\mu_g' \Sigma_\gamma \mu_g$. This formalizes the intuition that the cross-sectional volatility of market-model alpha estimates can increase either because non-market factors have become more important (either through larger premia or greater cross-sectional dispersion in factor exposures) or because of increases in average idiosyncratic volatility.

We seek to assess the effects of non-factor premia while controlling for the effect of idiosyncratic volatility. To do so, we consider also the cross-sectional average of the standard errors of the firm-specific CAPM alpha estimates. Let μ_γ denote the vector of cross-sectional mean non-market factor exposures and Σ_g denote the covariance matrix of non-market factor returns. Expression (9) of section I of the Appendix shows that the cross-sectional average standard error depends on both $\mu_\gamma' \Sigma_g \mu_\gamma$ and $\overline{\sigma_\varepsilon^2}$. Comparing equations (7) and (9) it can be observed that the cross-sectional variance of alpha estimates contains variation attributable to risk premia associated with non-market factors while the average standard error of the alpha estimates does not. Further, the cross-sectional mean exposures to non-market factors, which appear in expression (9), are empirically small. The cross-sectional mean non-market factor exposure is .006, while the median is -.015, implying that the $\mu_\gamma' \Sigma_g \mu_\gamma$ term in (9) is relatively unimportant. As a consequence, idiosyncratic return volatility is relatively more important in explaining the average standard error of the alpha estimates, while return premia associated with non-market factors are relatively more important in explaining cross-sectional variation in market-model alpha estimates. Thus, when both the volatility and the average standard error of alpha estimates are included as explanatory variables in the same regression, the coefficient on the former primarily reflects the effect of non-market return factor premia, while the coefficient on the latter primarily reflects average idiosyncratic risk.

Table 3 reports the results of time series regressions where the dependent variable in each month is the number of statistically significant factors, and the explanatory variables include the cross-sectional standard deviation of firm-specific CAPM alpha estimates. If, as we hypothesize, greater cross-sectional variation in CAPM alphas indicates greater variation in *expected* returns across stocks, and that variation is attributable to return premiums associated with factors other than the market, then we should observe a positive coefficient on this variable, with or without inclusion of the residual volatility variables. We include as control variables either the cross-sectional mean standard error of the alpha estimates, or to assess robustness, the cross-sectional average market-model idiosyncratic volatility, σ_{ε}^2, itself.

Columns (1) to (3) are based on estimation with equal weighting of each observation, while to assess robustness we report in Columns (4) to (6) corresponding results when each observation is weighted by the firm's market capitalization at the beginning of the sixty-month estimation period. The results reported in Table 3 indicate positive and significant coefficient estimates on the standard deviation of alpha estimates, with or without inclusion of the control variables, and by either weighting method. While the coefficient estimates on the control variables are negative, suggesting that higher idiosyncratic volatility tends to reduce the number of statistically significant factors due to reduced statistical power, these estimates are not themselves significant. It is also of interest to observe that the inclusion of the control variables substantially increases the magnitude of the coefficient estimates on the cross-sectional standard deviation of the alpha estimates. That is, these results show that the number of significant factors is significantly explained by the cross-sectional standard deviation of firm-specific alphas, but not by aggregate idiosyncratic volatility. These results support the reasoning that greater cross-sectional variation in firm-specific alpha estimates results from greater dispersion in expected returns attributable to return premia associated with non-market factors which in turn allows for the empirical estimation of a larger number of such factors.

We next assess the extent to which the observed positive relation between the number of significant factors and the cross-sectional variation in CAPM alphas is attributable to firms of varying size. To do so, we assign each firm to one of three size buckets; large stocks are those with month *t* market capitalization greater than the median for NYSE-listed firms, mid-size stocks are those with market capitalization between the 20th percentile and the median for NYSE-listed firms, and the remaining stocks are assigned as small capitalization. Recall that the sample estimate of the cross-sectional variance of firm-specific alphas is obtained as the sum across stocks of squared deviations of alpha estimates from the cross sectional mean estimate (divided by one less than the cross-sectional sample size). We decompose this sum into the portions attributable to large, medium and small firms.

On Table 4 we report the results obtained when we regress the number of significant factors on the sum of squared deviations across all firms (columns 1 and 3), and for sums computed within size groups (columns 2 and 4). Columns (1) and (2) report results for the full sample, while columns (3) and (4) pertain to the more recent 1968 to 2020 period during which accounting data allows for the construction of additional factors.

The coefficient estimates in columns (1) and (3) are positive and significant, indicating that inference is unaltered when focusing on the cross-sectional variance of alpha estimates (which can be cleanly decomposed into the sum of firm-size contributions) rather than the standard deviation. The results in columns (2) and (4) indicate that the positive relation between the number of significant factors and cross-sectional variability in firm alphas is primarily attributable to small firms. While a positive and significant coefficient is estimated for small firms both for the full sample and for the 1968 to 2020 sample, the coefficient estimates for the mid- and large-capitalization samples are either insignificant or, in the case of large firms during the full sample, negative and marginally significant. This result is broadly consistent with the Hou, Xue, and Zhang (2020) and Pontiff (2006) result that "anomalous" returns are most prevalent in small capitalization stocks, where arbitrage trades are risky and costly. Some researchers have argued that empirical regularities that are strongest in smaller stocks are uninteresting. We recognize that documenting differential expected returns for stocks that are relatively illiquid and/or of high risk may not provide actionable information to hedge funds or other arbitrageurs. However, in our view the information contained in small stocks returns is not inherently less useful than the information contained in large stock returns if the objective is to understand the economic determinants of cross-sectional variation in equilibrium expected returns. Further, in contrast to some of the Hou, Xue, and Zhang (2020) findings, the relation between the number of significant factors and cross-sectional variation in alphas remains significant in value-weighted regressions, as demonstrated in Table 3.

6. Variation in Significant Factors and Economic Complexity

The relevance of a given factor in terms of explaining the cross-section of stock returns can depend on the volatility of the factors, the variation across stocks in sensitivities of firm returns to factor outcomes, as well as the magnitude of the risk premia associated with the factor. Each of these variables potentially change when new firms enter the market, when the identity of the marginal investor changes, or as the economy evolves through the business cycle and over time.

a. The role of the number of listed firms.

We first focus on relations between the number of factors with statistically significant CAPM alphas and the number of firms traded in the U.S. markets. We reason that large increases or decreases in the number of publicly traded firms are likely to be accompanied by shifts in the types of firms available for public investment. Indeed, Fama and French (2004) show that the characteristics of firms newly listed on major U.S. stock markets varies substantially over time. Multiple and varied risk factors may be necessary to explain patterns in the returns of varying firm types. To assess robustness, we use three measures of the number of firms; the number listed as of month t, the number listed as of month *t-60,* and the number continuously listed from time *t-60* to *t* (so that alpha can be estimated).

Results for the full 1931 to 2020 sample, as well as for the more recent 1968 to 2020 period, are reported in Table 5, and indicate that the number of significant factors is positively related to each measure of the number of firms. The results in columns (1) to (3) for the full sample and columns (5) to (7) for the later subperiod indicate that the positive relation is statistically significant for any of the three measures considered individually. When all three measures are included in multiple regressions, the results in column (4) for the full sample as well as column (8) for the most recent subsample indicate that the number of firms listed at date *t* has significant explanatory power, while the other two measures do not, so we focus on this measure of the number of firms going forward.

The positive and significant coefficient estimates reported in Table 5 support the reasoning that a larger number of factors are required to explain cross-sectional variation in mean returns when more firms are listed. This result need not arise mechanically. As a simple example, suppose the CAPM determined expected returns for all stocks. The addition of new stocks with unique characteristics would only require estimation of their potentially distinct market betas. The empirical fact that more factors have significant explanatory power at times when more firms are listed is consistent with the reasoning that the firms that enter and depart the CRSP database differ from other firms in that they are exposed to differing sources of priced risk, rather than simply having differential exposures to a fixed set of priced systematic risks.

A simple alternative explanation for the observed positive relation between the number of statistically significant factors and the number of listed firms is that a larger cross-sectional sample size improves statistical power, such that estimated return premia of given economic magnitudes are more likely to become statistically significant. We demonstrate in section 1 of the appendix that this possibility arises, in particular, when the non-market factors are not directly observable, and consistent with actual practice, the empirical analyses are implemented based factors created from returns to portfolios sorted based on observable firm characteristics.

To distinguish between these possibilities, we conduct cross-sectional regressions of the number of statistically significant factors on (i) the cross-factor average standard error of the alpha estimates, (ii) the cross-factor average absolute alpha estimate, and (iii) the number of firms. The first four columns of Table 6 report results for the full 1931-2020 sample, while columns (5) to (8) contain corresponding results for the more recent 1968 to 2020 period.

As would be anticipated, the result reported in columns (1) and (5) of Table 6 indicate that fewer factors are statistically significant in periods where alpha standard errors are larger. As we show in section 1 of the appendix, one determinant of these standard errors is the number of firms in the sample. In columns (2) and (6) of Table 6 we report results that are obtained when the number of firms is included in the regression along with the average standard error. We continue to estimate negative coefficients on the mean standard error, even while we estimate positive coefficients on the number of firms. These results indicate that the positive relation between the number of statistically significant factors and the number of firms is not solely attributable to the effect of the number of firms on the standard errors.

Of course, the number of significant factors depends on the magnitude of the factor alpha estimates as well. In columns (3) and (7) of Table 6 we report results obtained when the explanatory variables include the mean absolute alpha as well as the mean standard error of the alpha estimates, while in columns (4) and (8) we report results when the number of sample firms is included as the third explanatory variable. These results confirm that, while a larger mean alpha is, as expected, associated with more statistically significant factors, the number of firms continues to have a significantly positive effect as well. We conclude that the number of listed firms has explanatory power for the number of significant factors that is distinct from the improvement in statistical power associated with a larger sample size, and that the number of firms contributes explanatory beyond any direct effect on the mean alpha estimate.

b. Economic Complexity and Observable Diversity

The results reported in the prior section show that the number of significant factors is strongly related to the number of publicly-traded firms. We next assess the extent to which the number of significant factors is related to aspects of economic complexity, and to the diversity of observable firm characteristics. We include several variables that plausibly proxy for various aspects of economic complexity. To facilitate interpretation, we standardize each of the following variables relative to its own time series. Thus, regression coefficients are interpreted as a response to a one-standard deviation change in that variable.

We conjecture that the business cycle will be relevant, both because of potential variation in the magnitude of risk premia and due to changes in firm types, with diversity increasing during periods of economic expansions characterized by high rates of firm entry and decreasing during recessions as firms exit. To capture these effects, we rely on an indicator variable equal to 1 for recession months, as defined by the National Bureau of Economic Research, and the unemployment rate reported by the US Bureau of Labor Statistics. We also consider two interest rate series, the Fed Funds rate (which begins in 1954) and the 10-year treasury note yield (which begins in 1964). Interest rates potentially capture the effects of monetary policy and funding conditions. The unemployment rate, fed funds rate and treasury yields are all obtained from the Federal Reserve Economic Data (FRED) website.

Fama and French (2001) suggest that the disappearance of dividend-paying firms reflects the changing characteristics of publicly traded firms. Thus, to further measure variation in firm types, we compute the proportion of dividend-paying common stocks as the number of firms paying at least one cash dividend in the previous 12-months relative to the total number of common stocks. Variation in firm characteristics such as the propensity to pay dividends could arise as firms respond to demand from different investor types. Further, the preferences of the marginal investor who effectively sets prices for specific stocks can depend on whether the investor is an individual or an institution.[22] To potentially capture the impact of changes in the composition of the investor base we measure the proportion of each firm's shares outstanding held by 13-F institutions in the Thomson-Reuters database.

We also consider the possibility that the number of significant factors may be related to market liquidity. To the extent that factor premia arise because investors are unable to profitably trade to eliminate mispricing, we should observe that more factor premia are significant when markets are less

[22] Lewis and Santosh (2021), for example, show that an asset pricing model where betas are defined relative to the portfolios held by active institutional investors performs better than the standard CAPM where betas are defined relative to aggregate market holdings.

liquid. To assess this possibility, we compute, on a monthly basis, the average across stocks of the Amihud (2002) illiquidity measure.

As a proxy for general economic complexity, we use the Economic Complexity Index constructed by Simoes and Hidalgo (2011), which is a measure of "the relative knowledge intensity of an economy." In addition, we include a measure of product market competition. Specifically, we compute for each month the cross-industry average of Herfindahl-Hirschman indices constructed for each three-digit NAICS industry, based on firm and industry sales. The Herfindahl index is identical in its construction to the Simpson (1949) measure, which is typically applied in the ecology literature to measure the diversity of species in a given environment. In our setting, the Herfindahl index, which is increasing in market concentration, is a measure of the inverse of diversity of firms within an industry. In the extreme, a 100% market share for a single firm would imply no firm diversity in the industry, while equal market shares for all firms would imply maximal within industry firm diversity.

To measure diversity in firm characteristics, we first standardize each of the 205 characteristics within each firm, such that the mean is zero and standard deviation is one. Having done so we compute, for each month *t*, the standard deviation of characteristics across all firms with available month *t* data. Finally, we sum these standard deviations across characteristics for each month.[23] The result can be interpreted as a measure of average cross-sectional dispersion in those characteristics observable for the available sample of firms in each month. Note that no relation necessarily exists between the number of firms and cross-sectional dispersion in characteristics; if newly listed firms were predominantly similar to the typical existing firm in terms of observable characteristics the diversity measure would decline rather than increase as more firms listed. An increase in the cumulative dispersion in characteristics across firms indicates, in contrast, that the underlying firms themselves are becoming increasingly differentiated.

[23] Results are virtually identical if we focus on the mean rather than the sum across characteristics.

Figure 11 displays the average number of characteristics that can be computed, delineated by the number of months since the firm initially appears in the database. In the first few months, less than twenty characteristics can be computed, on average. Thirty-six months after listing, approximately one hundred characteristics can be computed. This rapid growth reflects that many characteristics require prior accounting statement data (which is often sparsely collected at the beginning of a firm's public life), as well as prior return history.[24] However, the fact that a given characteristic cannot yet be computed by an econometrician need not imply that market participants are unaware of the characteristic. To accommodate the "burn in" period between the addition of a firm and the time when characteristics become observable we focus on cross-sectional variation in characteristics in month t+36 to measure firm characteristic diversity as of month t.

Table 7 reports the results of regressions of the number of statistically significant factors on these measures of economic complexity. Columns (1) to (11) report results of univariate regressions for each variable in turn, while columns (12) to (14) report multivariate outcomes. We omit mean institutional ownership from two of the multivariate specifications because data is available only from 1980 onward, and we omit the economic complexity index from the specification reported in column (12) because the data is not available after 2017 or before 1964. Panel A of Table 7 pertains to the full 1931 to 2020 sample while Panel B reports results for the recent 1968 to 2020 subperiod. Column (1) of Table 7 reproduces results obtained when the number of listed firms as of month t is the explanatory variable for comparison to results in other columns.

The univariate results reported in Table 7 show that the number of statistically significant factors is related to macroeconomic conditions, decreasing during recessions and increasing during periods of higher interest rates, based both on the Federal Funds rate and the Treasury-bond rate (though the former is not statistically significant during the more recent subsample). The unemployment rate, in contrast,

[24] The remaining upward drift observable on Figure 11 is attributable to growth in the number of variables contained in the Compustat database over time.

does not have significant explanatory power. It is, however, noteworthy that the macroeconomic variables have much less explanatory power for the number of significant factors as compared to the number of listed firms. Focusing on the full sample results in Panel A of Table 7, the R-squared statistics for the statistically significant macroeconomic variables vary from 0.02 for the recession indicator to 0.18 for the Treasury bond rate, as compared to 0.63 for the number of listed firms.

The coefficient estimates reported in column (6) of Table 7 indicate that the number of statistically significant factors is strongly negatively related to the percentage of firms that pay dividends, with a full-sample r-squared of 0.28. This result is consistent with the reasoning that the listing of non-dividend paying firms, which tend also to be younger and less familiar to investors, is associated with an increase in the number of significant factors, and more broadly with the notion that more factors are required to explain returns when listed firms are more diverse. The coefficient estimates reported in column (7) of Table 7 indicate that the number of statistically significant factors is also strongly negatively related to mean institutional ownership, with an R-squared statistic equal to 0.46. If institutions invest with a differing objective function as compared to individuals (due, for example, to agency issues or heterogeneity across individual investors) then changes in institutional ownership can effectively alter the identity and objective of the marginal stock market investor. The negative coefficient estimates reported on Table 7 imply that increased institutional ownership reduces the number of significant factors, potentially because it effectively reduces variation in the identity of the marginal investor. The coefficient estimate on the economic complexity index in column (8) is positive for both the full sample and for the more recent subsample, but is statistically significant only in the latter case. The coefficient estimate on the average industry HHI in column (9) is negative and highly significant both for the full sample and for the more recent subperiod. The full sample r-squared statistic for the average HHI is 0.44, being one of the highest r-squared among the single explanatory variables, with the exception of except the mean institutional ownership and number of publicly traded firms.

The coefficient estimate for the average Amihud illiquidity in the cross-section of stocks shown in column (10) of Panel A is negative and highly statistically significant for the full sample, and results in a r-squared of 0.05. This result indicates that an increase in market-wide illiquidity is associated with a decreased number of significant factors, and is at odds with the reasoning that factor premia arise because illiquidity impedes trading on mispricing. In the more recent subsample, the coefficient estimate on the Amihud measure in column (10) of Panel B is positive, but insignificant. These results do not support the reasoning that significant factor premia arise because illiquidity impedes trading on mispricing.

The coefficient estimate for the diversity of firm characteristics (column 11) is positive and statistically significant, both for the full sample and for the more recent subsample. The full-sample r-squared statistic is 0.42. This result implies that more factors are significant during those periods when there is greater cross-sectional variability in the firm characteristics that are observable to econometricians.

Columns 12 to 14 present results for multivariate specifications. The unemployment rate remains significant in all specifications, and the recession indicator, the 10-year yield and the mean Amihud illiquidity remain significant in Column 12 to 13. The proportion of firms paying dividends, and the ECI are insignificant, even in the more recent subperiod. Notably, the number of publicly listed firms is no longer significant in the multivariate setting, while the coefficient on the average HHI remains significant when all explanatory variables are included in column 14. The negative coefficient estimate on the average HHI implies that less competitive product markets are associated with fewer significant factors. This result suggests that firms' declining market power may be associated with increases in the number of priced risks. An alternative interpretation arises in light of the fact that the HHI is also equivalent to the Simpson (1949) measure of diversity. In particular, greater within-industry diversity may arise when firms' product offerings are only slightly differentiated and sales are correspondingly spread out across firms. This result is suggestive that explaining the number of significant asset pricing factors may also require better understanding of the industrial organization of industries, including the economic forces that allow for greater or lesser industry sales concentration.

The diversity of firm characteristics remains significant in each specification. That is, even after allowing for the explanatory power of macroeconomic variables such as the unemployment rate and interest rates, changes in institutional ownership, and the strong effects of industry HHI, cross-sectional variation in firm characteristics has explanatory power for the number of significant factors. In combination, the results here provide strong support for the notion that time variation in the number of significant factors is not random, but rather is linked to variation in macroeconomic conditions, intra-industry concentration, and observable diversity in firms' characteristics.

7. Conclusions

The reasoning that only a few factors should be necessary to explain the cross section of mean returns is attractive in part because parsimony is desirable. So, should the fact that the literature shows that hundreds of empirically observable factors have explanatory power for the cross-section of stock returns be viewed as a collective failure? We think not, if the reason is that financial markets and the broader economy are complex and dynamic. The characteristics of the firms that are available for investment can change through time as existing firms evolve and new firms are listed or delisted. Investors are diverse in terms of their investment horizon and objectives. Some investors trade on their own account, while others rely on professional managers whose strategies can be affected by agency issues related to their compensation. Return premia have been shown to depend on intermediaries' funding liquidity, leverage, and balance sheets, as well as on the state of the economy. The identity of the marginal investor can differ across stocks, and in any given stock can vary through time. In short, it is unclear that return premiums in actual capital markets need to be governed by only a few factors.

We present a number of empirical findings relevant to these issues, showing that a large number of factors, exceeding one hundred during some time intervals, have significant explanatory power, and that the number of significant factors varies substantially over time. Further, the number of principal components required to explain variation in the significant factors is also large and is positively correlated with the number of significant factors, both in sample and out-of-sample, implying that the results do not simply arise because various researchers identify factors that are redundant of each other.

To assess the scope for multiple factors to explain the cross-section of stock returns we estimate firm-specific CAPM alphas on a rolling sixty-month basis. Having done so, we compute the cross-sectional standard deviation of the firm-specific CAPM alpha estimates by month. We document that the number of significant factors is positively related to both the mean and standard deviation of factor CAPM alpha estimates, even while the number of significant factors is not related to the average standard error of alpha estimates or average idiosyncratic volatility. These results support the reasoning that the number of significant factors is related to the extent to which expected stock returns are left unexplained by the CAPM, and do not support the alternative hypothesis that more significant factors are identified because higher volatility allows for more effective data mining or specification searches.

We assess the extent to which widely-used three- to six-factor models do a better job of explaining the cross-section of returns as compared to the CAPM. While these models outperform the CAPM in terms of explaining returns to characteristic-sorted (size and market-to-book) portfolios, they do not reliably outperform for industry portfolios, and most often perform worse than the CAPM for the cross-section of stock returns. The last result is noteworthy in part because the three- to six-factor models all include the market factor, implying that the non-market factors included degrade the ability of the models to explain mean returns to individual stocks. To the extent that the perception that only a few factors should matter for stocks in general is based on the performance of three- to six-factor models in explaining returns to characteristic-sorted portfolios, the perception is misplaced.

We also provide evidence that the number of significant factors varies through time with measures of economic complexity and firm diversity. In particular, the number of significant factors is related to a recession indicator variable, interest rates, the percentage of firms that pay dividends, mean institutional ownership rates, and an economic complexity index, and is particularly strongly related to the number of firms that are publicly listed, cross-sectional variation in observable firm characteristics, and industry-level sales concentration. The finding with respect to the number of firms supports the reasoning that newly listed firms systematically differ from existing firms in terms of systematic risks relevant to investors. The finding with respect to concentration ratios is suggestive that diversity of firms within industries and by extension the industrial organization of product markets may be a source of systematic variation in firm profitability and returns that is relevant to investors. Finally, the finding with respect to diversity of firm characteristics suggests that more factors are relevant when firms themselves are more distinct.

On balance, our findings suggest that the collection of asset pricing factors documented in the literature do not simply reflect collective data snooping. To the contrary, they suggest that multiple and time-varying factors may be required to price the cross-section of returns as the economy continues to evolve dynamically and new firms are listed. Further, in a dynamic economy a factor can be significant in explaining returns during some periods but not others. This suggests the desirability of a degree of caution in interpreting the results of existing out-of-sample tests, as insignificant out-of-sample outcomes need not imply that the factor was unpriced in the original sample period, and the need for the development of econometric methods for out-of-sample tests suitable to the dynamic environment.

REFERENCES

Adrian, Tobias, Erkko Etula, and Tyler Muir, 2014, Financial Intermediaries and the Cross-Section of Asset Returns, *Journal of Finance*, 69, 2557-2596.

Ahmed, Shamim, Ziwen Bu, and Daniel Tsvetanov, 2019, Best of the best: a comparison of factor models, *Journal of Financial and Quantitative Analysis* 54, 1713-1758.

Amihud, Yakov, 2002, Illiquidity and stock returns: cross-section and time-series effects, *Journal of Financial Markets*, 5, 31-56.

Bali, Turan, Stephen Brown, Scott Murray, and Yi Tang, 2017, A Lottery-Demand-Based Explanation of the Beta Anomaly, *Journal of Financial and Quantitative Analysis,* 52, 2369-2397.

Barber, Brad M. and Huang, Xing and Odean, Terrance and Schwarz, Christopher, 2021, Attention Induced Trading and Returns: Evidence from Robinhood Users, *Journal of Finance*, Forthcoming.

Barillas, Francisco, and Jay Shanken, 2018, Comparing asset pricing models, *Journal of Finance* 73, 715-754.

Betermier, Sebastien, Laurent Calvet, Samuli Knüpfer, and Jens Kvaerner, 2021, What do the portfolios of individual investors reveal about the cross-section of equity returns?, working paper, download at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3795690.

Brennan, Michael, 1998. The role of learning in dynamic portfolio decisions, *European Finance Review*, 1, 295–306.

Bustamante, M Cecilia, and Andres Donangelo, 2017, Product market competition and industry returns, *Review of Financial Studies* 30, 4216-4266.

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

Chen, Andrew Y, 2021, The limits of p-hacking: Some thought experiments, *Journal of Finance* 76, 2447-2480.

Chen, Andrew Y., and Tom Zimmermann, Open source cross sectional asset pricing, *Critical Finance Review,* Forthcoming.

Chen, Yao, Alok Kumar, and Chendi Zhang, 2021, Searching for Gambles: Investor Attention, Gambling Sentiment, and Stock Market Outcomes, *Journal of Financial and Quantitative Analysis*, 56, 2010-2038.

Chordia, Tarun, Amit Goyal, and Alessio Saretto, 2020, Anomalies and false rejections, *Review of Financial Studies* 33, 2134-2179.

Cochrane, John H, 2011, Presidential address: Discount rates, *Journal of Finance* 66, 10471108.

Fama, Eugene F, 1998, Determining the number of priced state variables in the ICAPM, *Journal of Financial and Quantitative Analysis* 33, 217-231.

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

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

Fama, Eugene F, and Kenneth R French, 2004, New lists: Fundamentals and survival rates, *Journal of Financial Economics* 73, 229-269.

Fama, Eugene F, and Kenneth R French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1-22.

Fama, Eugene F, and Kenneth R French, 2018, Choosing factors, *Journal of Financial Economics* 128, 234-252.

Gibbons, Michael, Stephen Ross, and Jay Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica*, 57, 1121-52.

Goodwin, Thomas H, 1998, The information ratio, *Financial Analysts Journal* 54, 34-43.

Goyal, Amit, and Pedro Santa-Clara, 2003, Idiosyncratic risk matters!, *Journal of Finance* 58, 975-1007.

Grullon, Gustavo, Yelena Larkin, and Roni Michaely, 2019, Are US industries becoming more concentrated?, *Review of Finance* 23, 697-743.

Harvey, Campbell R., Yan Liu, and Heqing Zhu, 2016, ... and the cross-section of expected returns, *Review of Financial Studies* 26, 5-68.

He, Zhiguo, Bryan Kelly, and Asaf Manela, 2017, Intermediary asset pricing: New evidence from many asset classes. *Journal of Financial Economics* 126, 1-35.

He, Zhiguo and Arvind Krishnamurthy, 2013, Intermediary Asset Pricing, *American Economic Review*, 103, 732-770.

Gu, Shihao, Bryan Kelly, and Dacheng Xiu, 2020, Empirical Asset Pricing via Machine Learning, *Review of Financial Studies*, 33, 2223-2273.

Hou, Kewei, and David T Robinson, 2006, Industry concentration and average stock returns, *Journal of Finance* 61, 1927-1956.

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

Hou, Kewei, Chen Xue, and Lu Zhang, 2020, Replicating anomalies, *Review of Financial Studies* 33, 2019-2133.

Ilmanen, Antti, Ronen Israel, Tobias J Moskowitz, Ashwin K Thapar, and Rachel Lee, 2021, How do factor premia vary over time? A century of evidence, *Journal of Investment Management* 19, 15–57.

Jensen, Theis, Bryan T Kelly, and Lasse Heje Pedersen, 2021, Is there a replication crisis in finance?, NBER Working Paper .

Kahle, Kathleen M, and Rene M Stulz, 2017, Is the US public corporation in trouble?, *Journal of Economic Perspectives* 31, 67-88.

Kashyap, Anil, Natalia Kovrijnykh, Jian Li, and Anna Pavlova, 2021, The benchmark inclusion subsidy, *Journal of Financial Economics*, 142, 756-774.

Kelly, Bryan, Seth. Pruitt, and Yinan Su, 2019, Characteristics are covariances: A unified model of risk and return, *Journal of Financial Economics*, 134, 501-524

Koijen, Ralph SJ, and Motohiro Yogo, 2019, A demand system approach to asset pricing, *Journal of Political Economy* 127, 1475-1515.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2018, Interpreting factor models, *Journal of Finance*, 133, 1183-1223.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2020, Shrinking the cross-section, *Journal of Financial Economics*, 135, 271-292.

Lewis, Ryan, and Shrihari Santosh, 2012, Investor betas, working paper, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3739424.

Linnainmaa, Juhani T, and Michael R Roberts, 2018, The history of the cross-section of stock returns, *Review of Financial Studies* 31, 2606-2649.

Lopez-Lira, Alejandro, and Nikolai Roussanov, 2021, Do common factors really explain the cross-section of returns? Working paper, https://papers.ssrn.com/sol3/Papers.cfm?abstract_id=3628120.

McLean, R David, and Jeffrey Pontiff, 2016, Does academic research destroy stock return predictability?, *Journal of Finance* 71, 5-32.

Pastor, Lubos, and Robert F Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 111, 642-685.

Pastor, Lubos, and Pietro Veronesi, 2009, Learning in financial markets, *Annual Review of Financial Economics*, 1, 361-381.

Pontiff, Jeffrey, 2006, "Costly Arbitrage and the Myth of Idiosyncratic Risk,", *Journal and Accounting and Economics*, Vol. 42, 35-52.

Sharpe, William F, 1994, The Sharpe Ratio, *Journal of Portfolio Management* 21, 49-58

Simoes, Alexander James Gaspar, and Cesar A Hidalgo, 2011, The economic complexity observatory: An analytical tool for understanding the dynamics of economic development, in *Workshops at the twenty-fifth AAAI conference on artificial intelligence*.

Simpson, E., 1949, Measurement of Diversity, *Nature*, 163, 688-688.

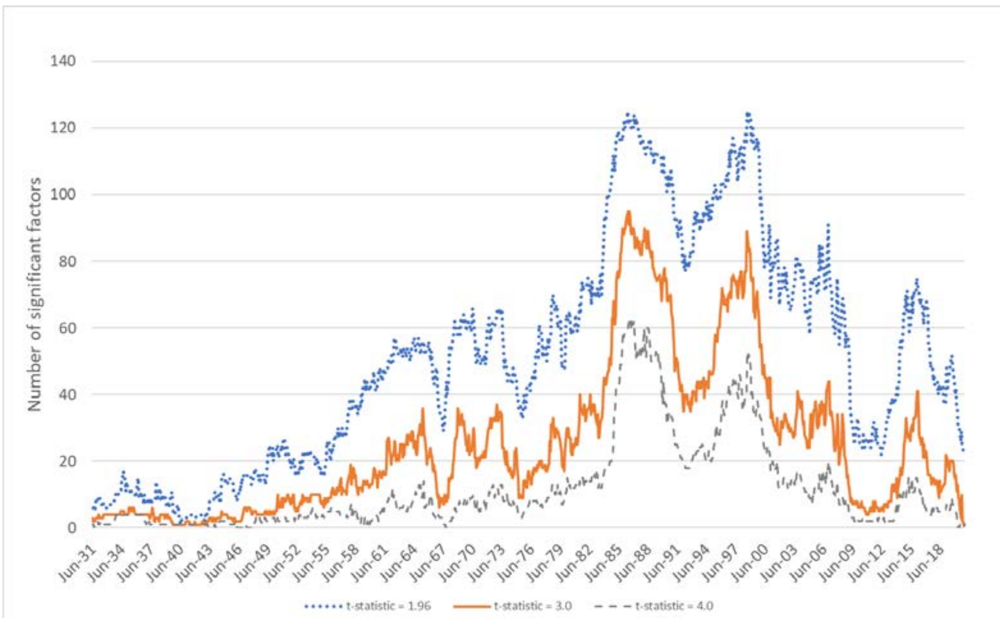
Stambaugh, Robert F, and Yu Yuan, 2017, Mispricing factors, *Review of Financial Studies* 30, 1270-1315.

Stulz, Rene M, 2018, The shrinking universe of public firms: Facts, causes, and consequences, *NBER Reporter* 12-15.

Van Reenen, John, 2018, Increasing Difference Between Firms: Market Power and the Macro Economy, Changing Market Structures and Implications for Monetary Policy, *Kansas City Federal Reserve: Jackson Hole Symposium*, 19-65.

Figure 1. Time series variation in the number of significant factors. This figure shows the time series variation in the number of significant factors based on the alphas obtained from the sample of factor returns. For each factor at each month t, we regress each of the three portfolios' monthly returns from $t - 59$ to t on the market's monthly excess returns to obtain each portfolio's CAPM alpha. To be included, portfolios must have 60 non-missing returns over the alpha estimation period. For each month t, we count the number of significant factors based on each of the three portfolios' alphas. A factor is significant at month t if the t-statistic of its CAPM alpha on a given portfolio exceeds defined thresholds. Panel A shows the number of significant factors for a t-statistic cutoff of 1.96 (dotted blue line), 3.00 (solid orange line) and 4.00 (dashed grey line).. Panel B shows the time series variation in the number of significant factors based on the alphas of 1) the long-short portfolio of the factor (solid orange line), 2) the long-only portfolio of the factor (dashed grey line), and 3) the short-only portfolio of the factor (dotted blue line).

Panel A: Number of significant factors as determined by varying t-statistic cutoffs



Panel B: Number of significant factors arising from long and short portfolios

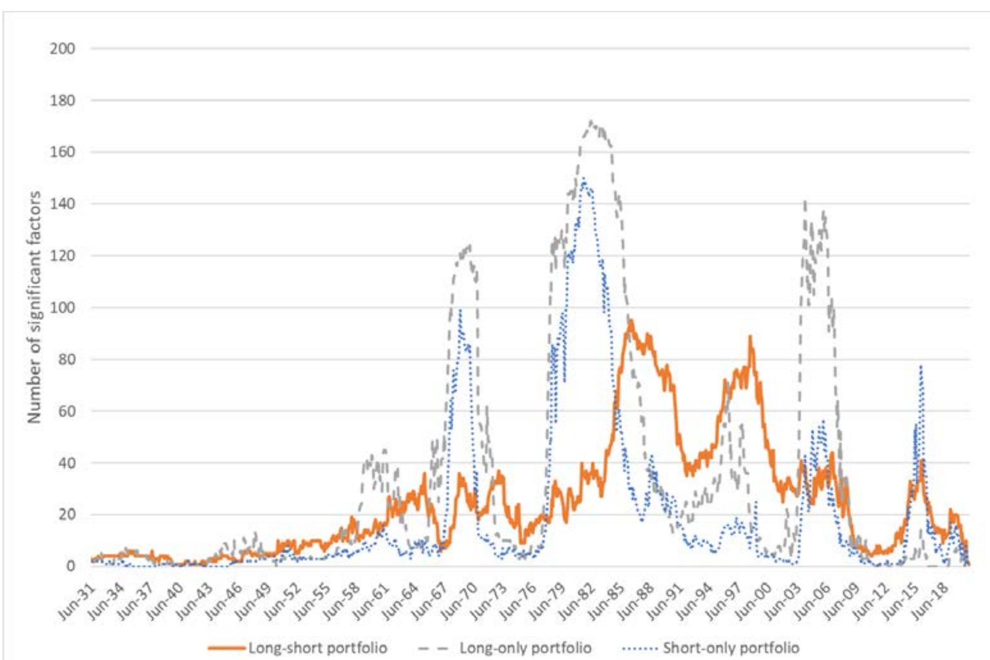
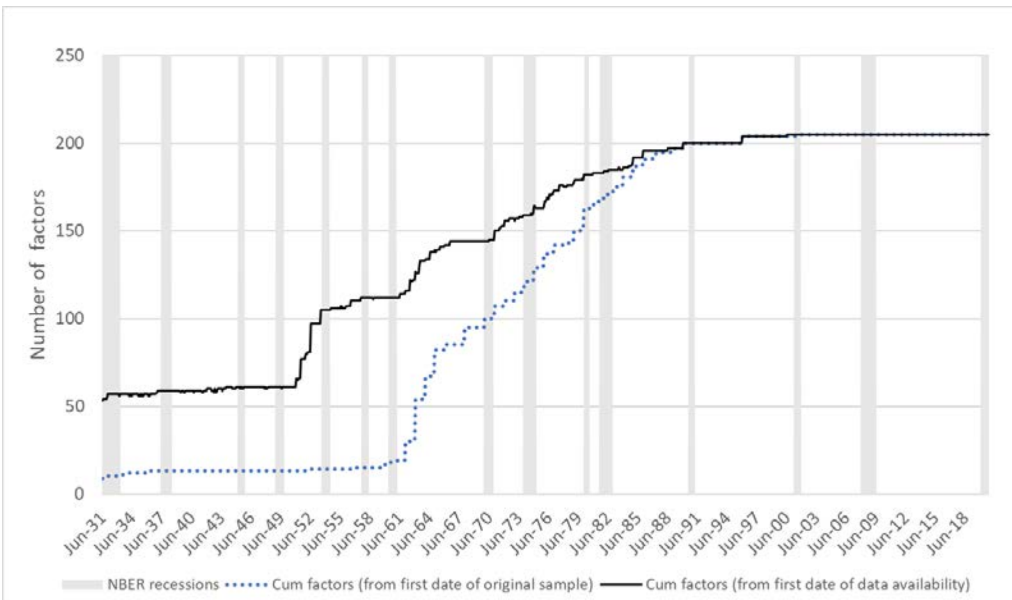


Figure 2. Time series variation in number of and significance of factors. This figure shows the cumulative factors over time as documented in the finance literature. The sample of factors comes from the set of "clear" and "likely" predictors provided by Chen and Zimmermann (Forthcoming) from 1931 to 2020. Panel A shows the cumulative number of factors over time computed in two ways. The solid black line is incremented at the date of each factor's first available return given the data available today. The dashed blue line is incremented at the date of each factor's first available return based on the time period of the data used in the original paper's sample. Panel B relates the number of significant factors to the variation in the standard deviation of stock-level CAPM alphas. For each month *t*, we regress each factor's (stock's) monthly returns from *t – 59* to *t* on the market's monthly excess returns to obtain each factor's (stock's) CAPM alpha. To be included, factors (stocks) must have 60 non-missing returns over the alpha estimation period. A factor is counted significant at month *t* if the *t*-statistic of its CAPM alpha exceeds 3.00. The solid blue line shows the standard deviation of all stock-level CAPM alphas computed at month *t*. The dashed orange line shows the number of significant factors at month *t*. The grey vertical bars represent periods of NBER-defined recessions.

Panel A: Cumulative factors over time



Panel B: Number of significant factors relative to stock-level alpha dispersion

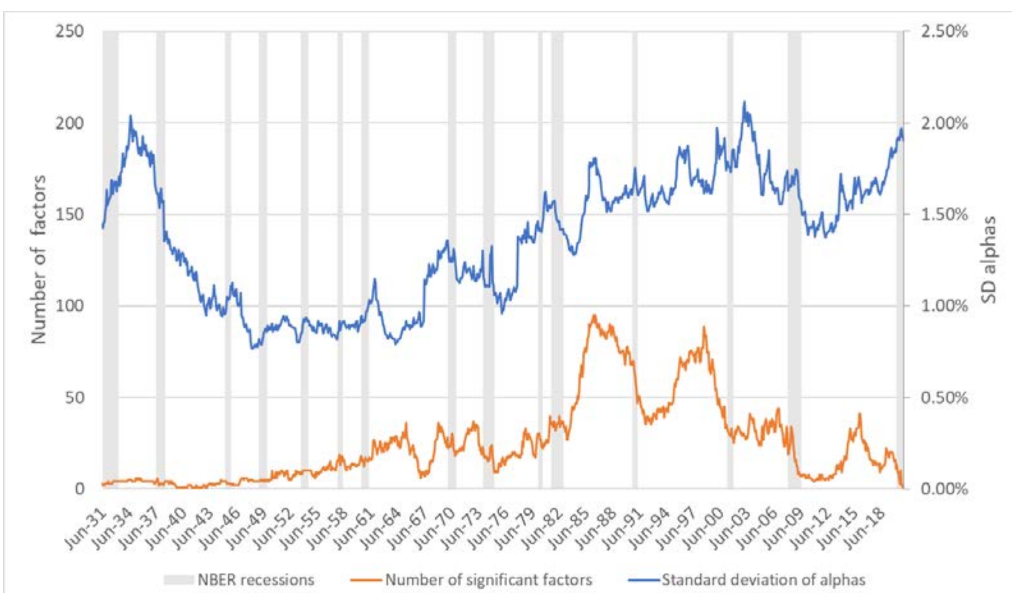
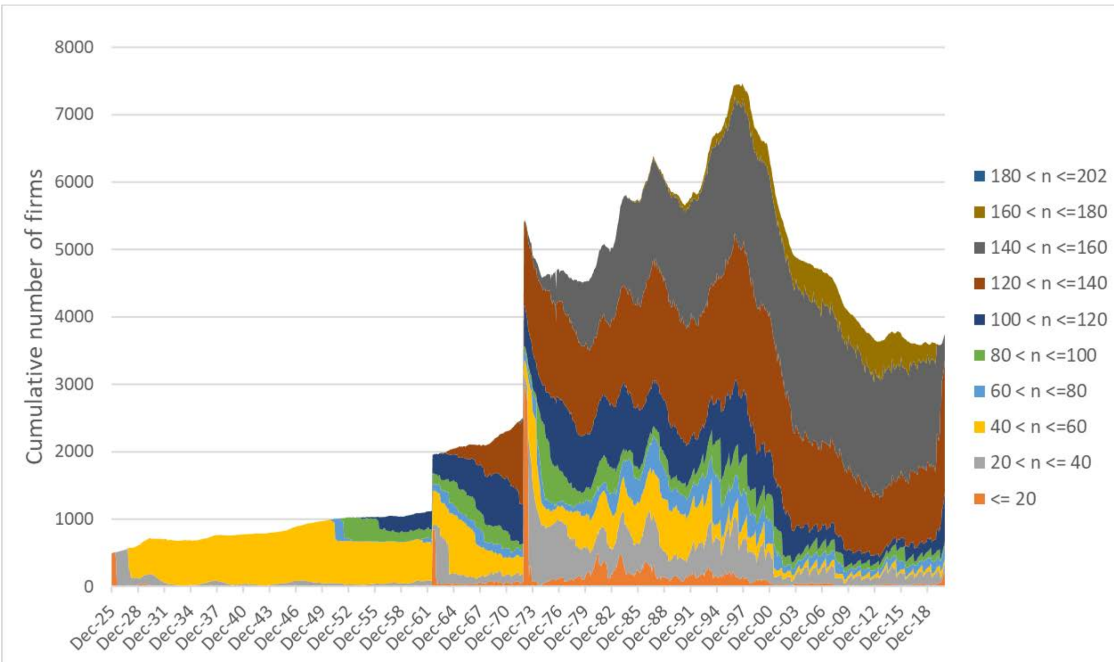


Figure 3. Number of firm-level non-missing characteristics over time. This figure shows the number of firms grouped by the total number of non-missing characteristics for each firm at each month in the sample. For each firm at each date, we compute the number of non-missing characteristics in the original sample of cross-sectional characteristics provided by Chen and Zimmerman (2021). Panel A shows the cumulative number of firms as the number of characteristics increases across groupings for the full sample from 1925-2020. Panel B shows the cumulative percentage of firms at each date that fall in each grouping for the subsample of years 1963-2020.

Panel A: Cumulative number of firms grouped by number of non-missing characteristics (1925-2020)



Panel B: Percentage of firms grouped by number of non-missing characteristics (1963-2020)

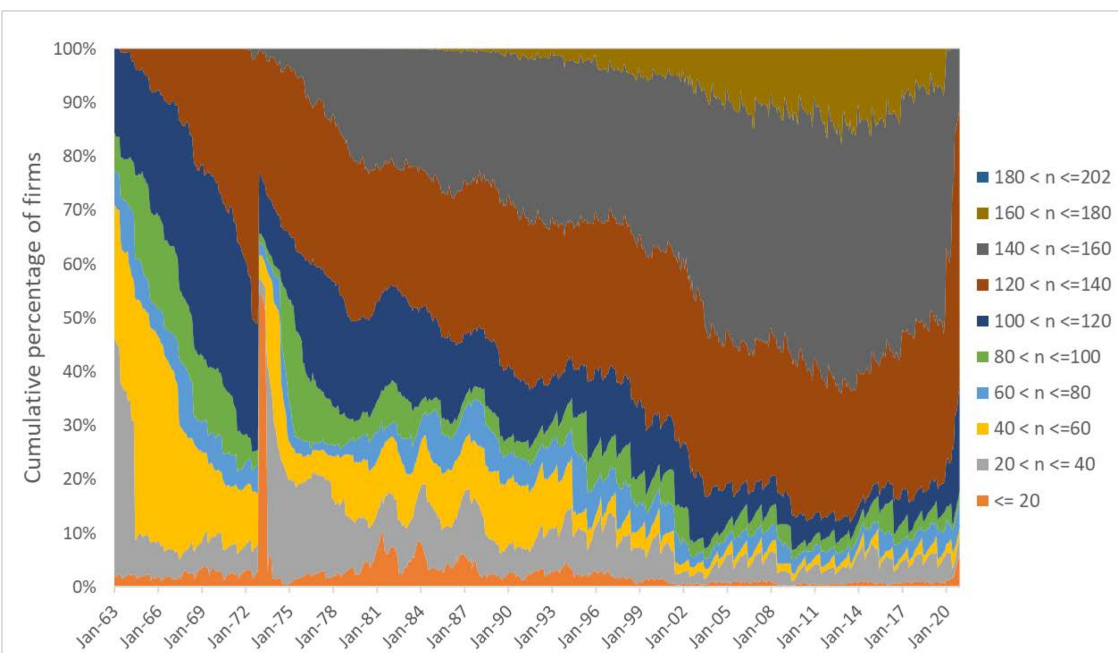


Figure 4. Time series variation in alphas of various asset pricing models This figure shows the time series variation in the standard deviation of alphas obtained from various asset pricing models of different sets of test assets. For each month *t*, a test asset's excess monthly returns from *t – 59* to *t* are regressed on factors of various asset pricing models to obtain an alpha relative to that asset pricing model. To be included, a test asset is required to have 60 non-missing returns over the estimation period. At each month *t*, we plot the standard deviation of all alphas obtained from a specific asset pricing. We use four samples of test assets when computing the alphas. Panel A consists of all stocks in the CRSP universe. Panel B consists of the set of 205 factors from Chen and Zimmermann (Forthcoming). Panel C consists of the 25 size and book-to-market portfolios. Panel D consists of the Fama-French 30 industry portfolios. The asset pricing models include the Capital Asset Pricing Model (CAPM), the Fama and French (1993) 3-factor model (FF3F), the Fama and French (2015) 5-factor model (FF5F), the Fama and French (2018) 6-factor model (FF6F), the FF3F model augmented with Carhart (1997) momentum factor (FF3F+UMD), the FF3F+UMD model augmented with momentum and Pastor and Stambaugh (2003) liquidity factor (FF3F+UMD+PSLIQ), the Stambaugh and Yuan (2017) factor model (M4), the Barillas and Shanken (2018) 6-factor model (BS6F) and the Hou et al. (2015) q-factor model (Q4). The standard deviation of alphas is in percent per month. The grey vertical bars represent periods of NBER-defined recessions.

Panel A: CRSP universe of common stocks

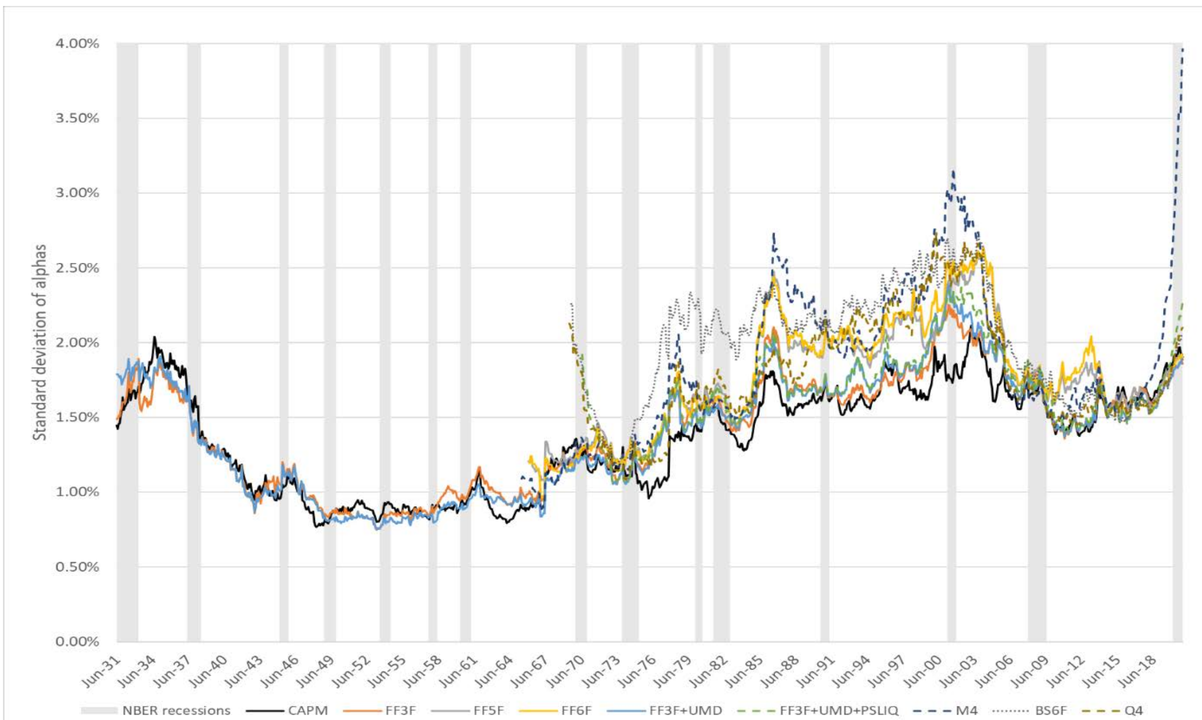
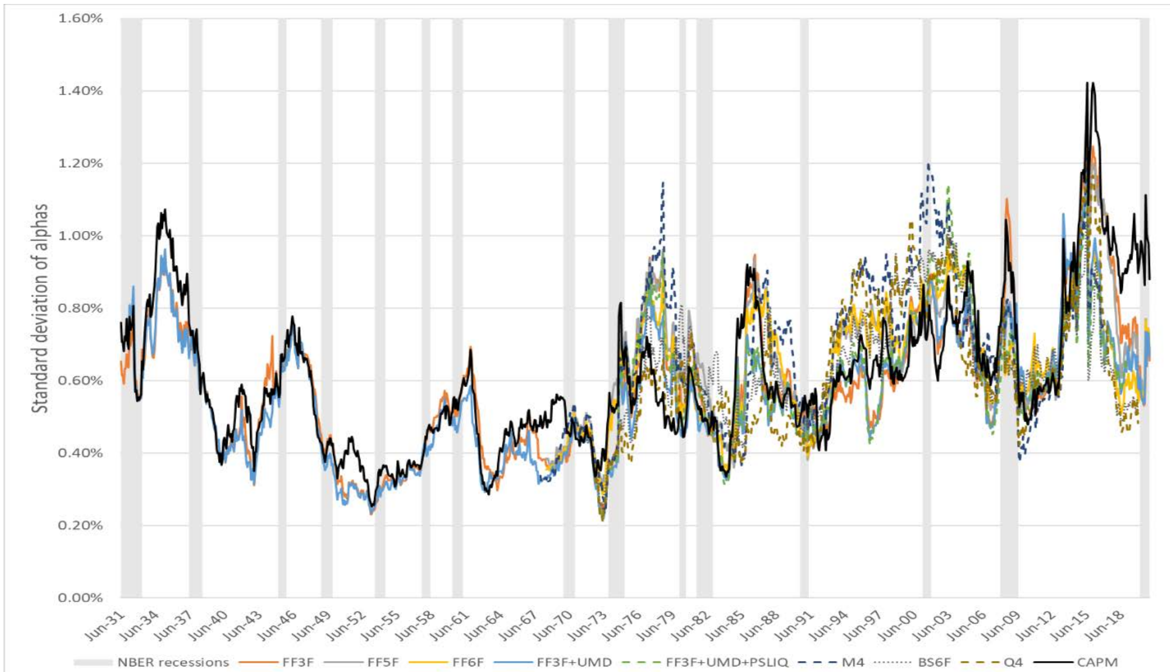


Figure 4 continued.

Panel B: 30 Fama-French industry portfolios



Panel C: 25 size and book-to-market portfolios

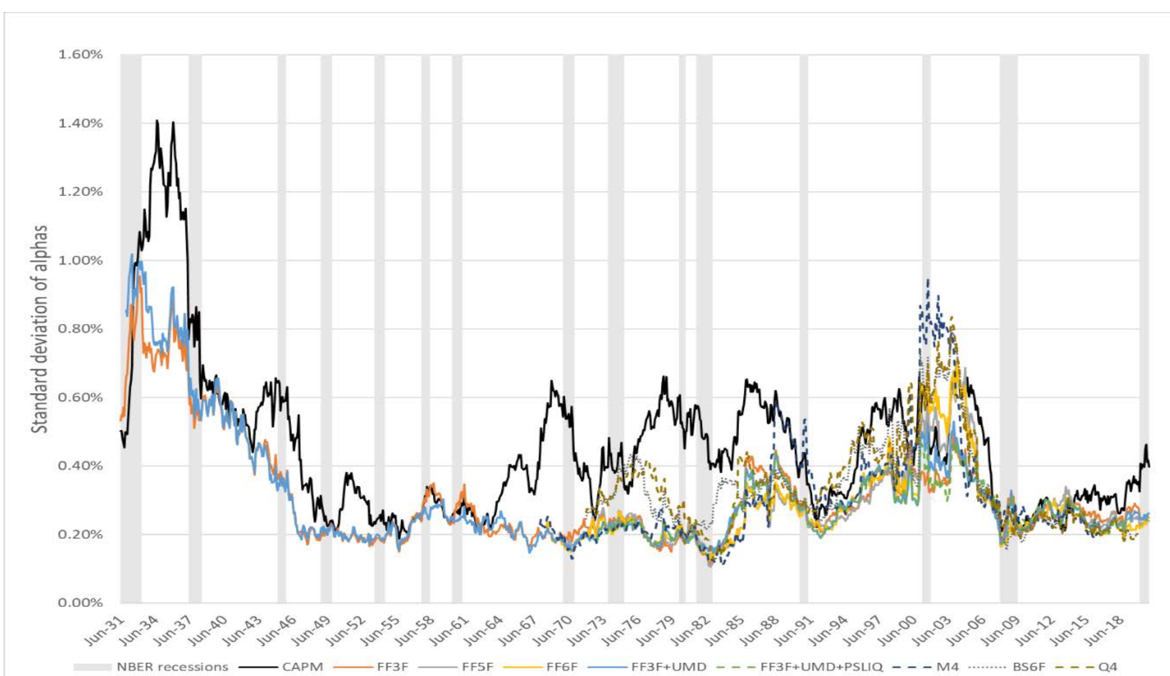


Figure 5. Time series variation in principal components of factors This figure shows the variation in the number of significant principal components in the sample of 205 factors across time. For each month *t*, we regress each factor's monthly returns from *t – 59* to *t* on the market's monthly excess returns to obtain each factor's CAPM alpha. A factor is significant at month *t* if the *t*-statistic of its CAPM alpha exceeds 3.00. To be included, factors must have 60 non-missing returns over the alpha estimation period. The orange dashed line shows the number of significant factors at each date. We also compute the number of significant principal components at each month *t* by counting the number of principal components required to explain 95% of the cumulative variation of a set of factor returns from *t – 59* to *t*. A factor must have 60 non-missing returns to be included in the principal component analysis at month *t*. We compute the number of principal components based on two rolling samples of factors. The black solid line shows the number of significant principal components for the sample of all factors. The blue dotted line shows the number of significant principal components for the sample of factors which have a significant CAPM alpha over the previous 60 months. The grey vertical bars represent periods of NBER-defined recessions.

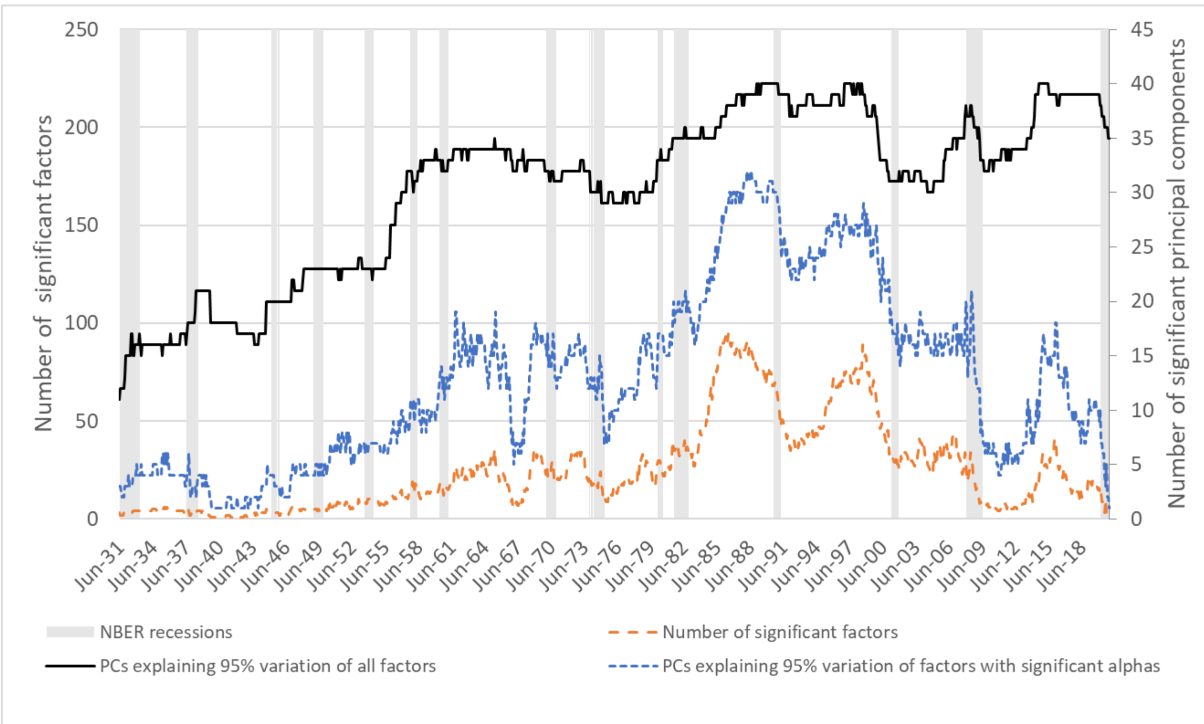


Figure 6. Cumulative variation explained by principal components of all factor returns. This figure shows the cumulative variation explained by increasing numbers of principal components for the set of all factor returns. At each month t, we compute the number of significant components of the factors' monthly returns from *t – 59* to *t*. Factors must have 60 non-missing returns over the analysis window to be included in the sample at a given month *t*. The figure shows the cumulative number of principal components at each date that make up different percentages of variation. The grey vertical bars represent periods of NBER-defined recessions.

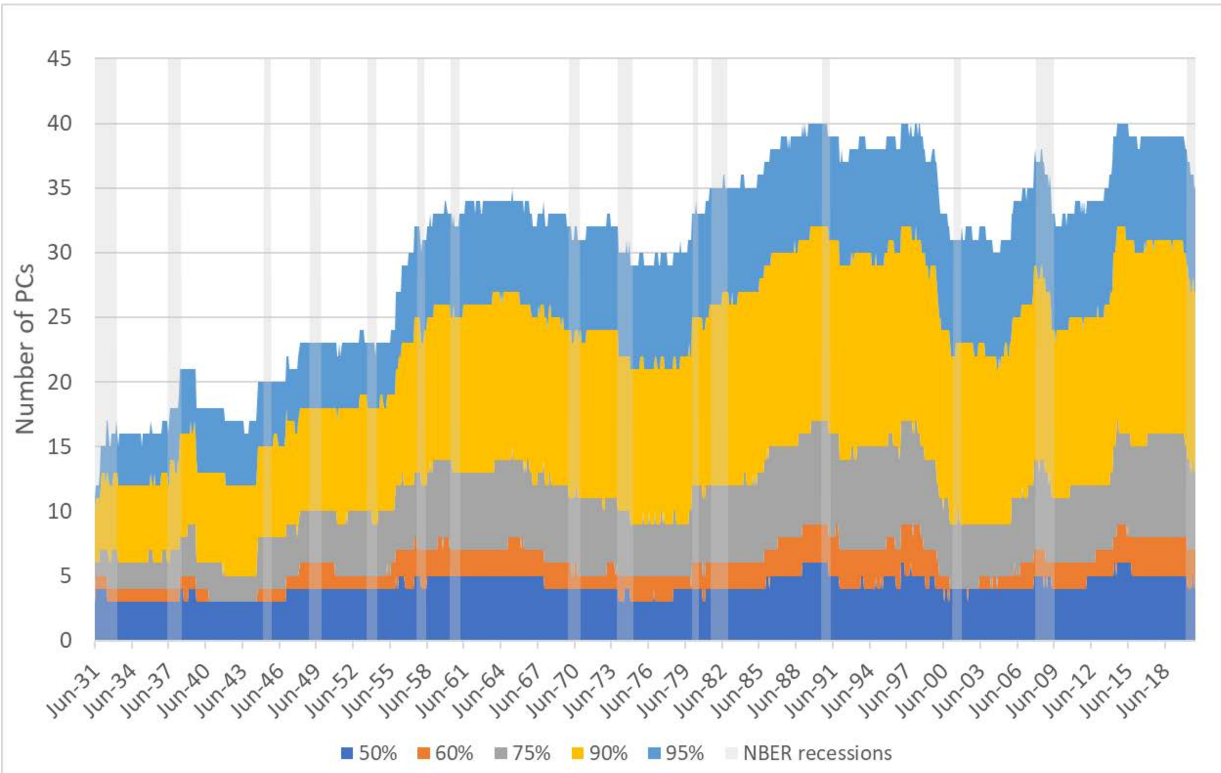


Figure 7. Marginal Sharpe ratio contribution of additional principal components. This figure shows the marginal Sharpe ratio contributions that additional principal components add to an optimal tangency portfolio. From 1968-2020, at each month t, we compute the principal components of the factors' monthly returns from $t - 59$ to t. Factors must have 60 non-missing returns over the analysis window to be included in the sample at a given month t. We construct up to 59 portfolios by incrementally adding a principal component to the portfolio. The weights in each portfolio are chosen to optimize the portfolio's Sharpe ratio. Panel A shows the cumulative Sharpe ratio obtained after including additional principal components in, while Panel B shows each Sharpe ratios as a proportion of the Sharpe ratio obtained in the portfolio which includes all 59 principal components.

Panel A: Cumulative Sharpe ratios of portfolios obtained from additional principal components

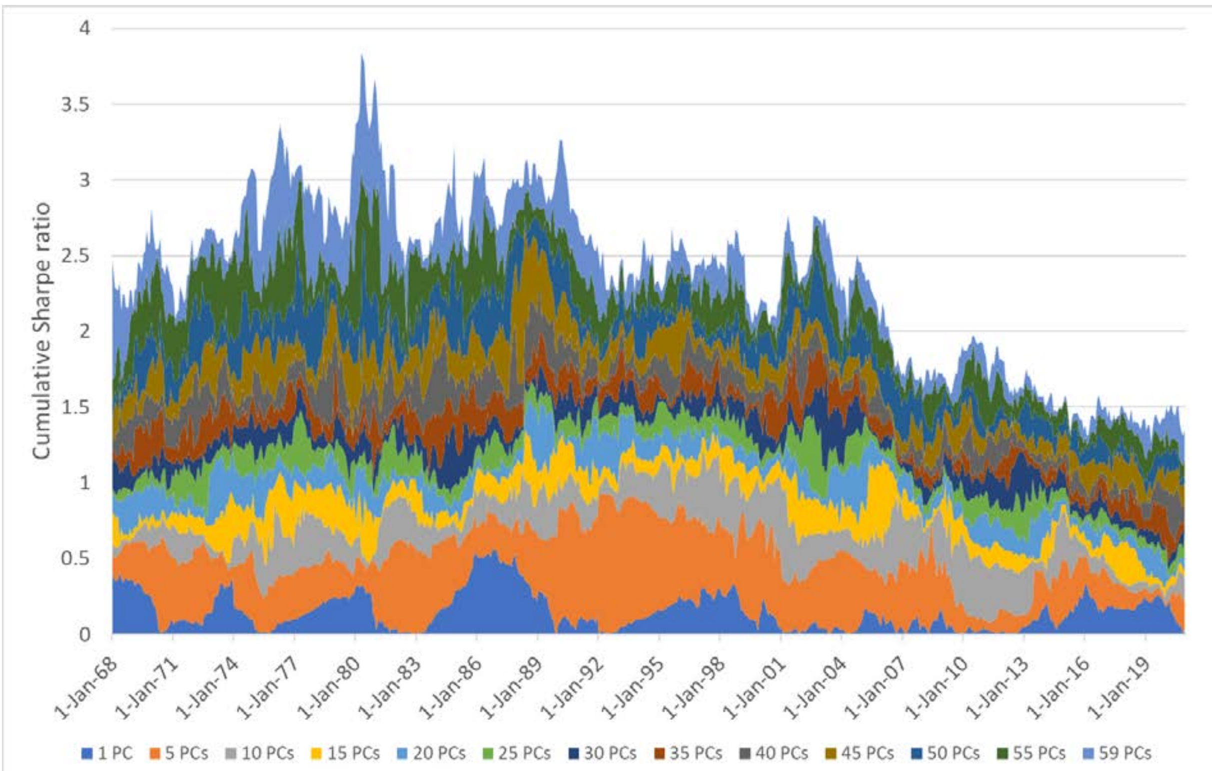


Figure 7 continued.

Panel B: Portfolio Sharpe ratio as a proportion of 60-principal component portfolio Sharpe ratio

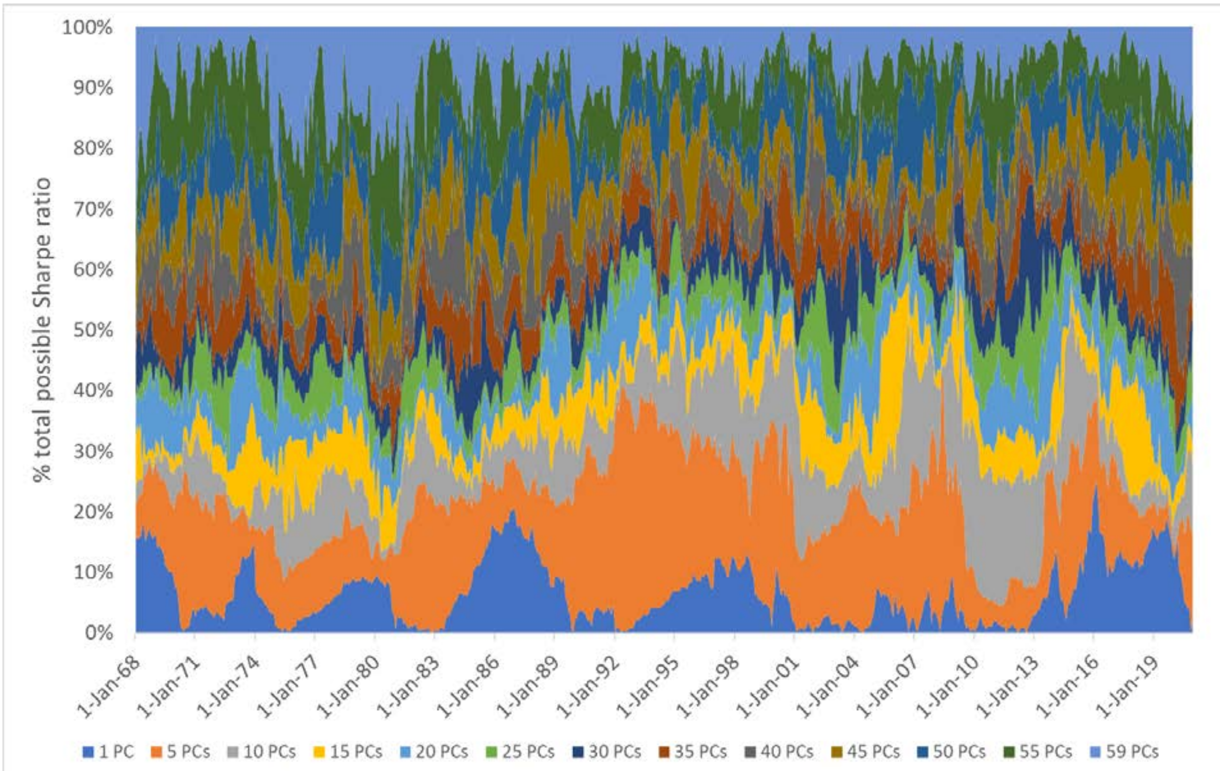


Figure 8 – Number of PCs and Sharpe ratio of portfolios constructed from out-of-sample principal components. These figures show the number of PCs and the maximum Sharpe ratio for portfolios formed from varying number of principal components. The PCs are calculated from the factor returns from t-59 to t. The portfolios are formed using the out-of-sample data from t+1 to t+36. See section 3.C for the full methodology. Panels A and C compare the number of principal components that form a portfolio with the highest Sharpe ratio. The orange dots in panels A and C are the number of principal components that form a portfolio with the maximum Sharpe ratio and the blue lines are the number of significant factors at month *t*. Panels B and D show the Sharpe ratios of portfolios formed from 4 different sets of PCs: 1) the first 5 PCs, 2) the average Sharpe ratio across all portfolios formed by increasing numbers of PCs, 3) the maximum number of principal components, and 4) the number of PCs that form the portfolio with the maximum Sharpe ratio. Panels A and B show results for PCs formed from all factors. Panels C and D shows results based on the subset of significant factors.

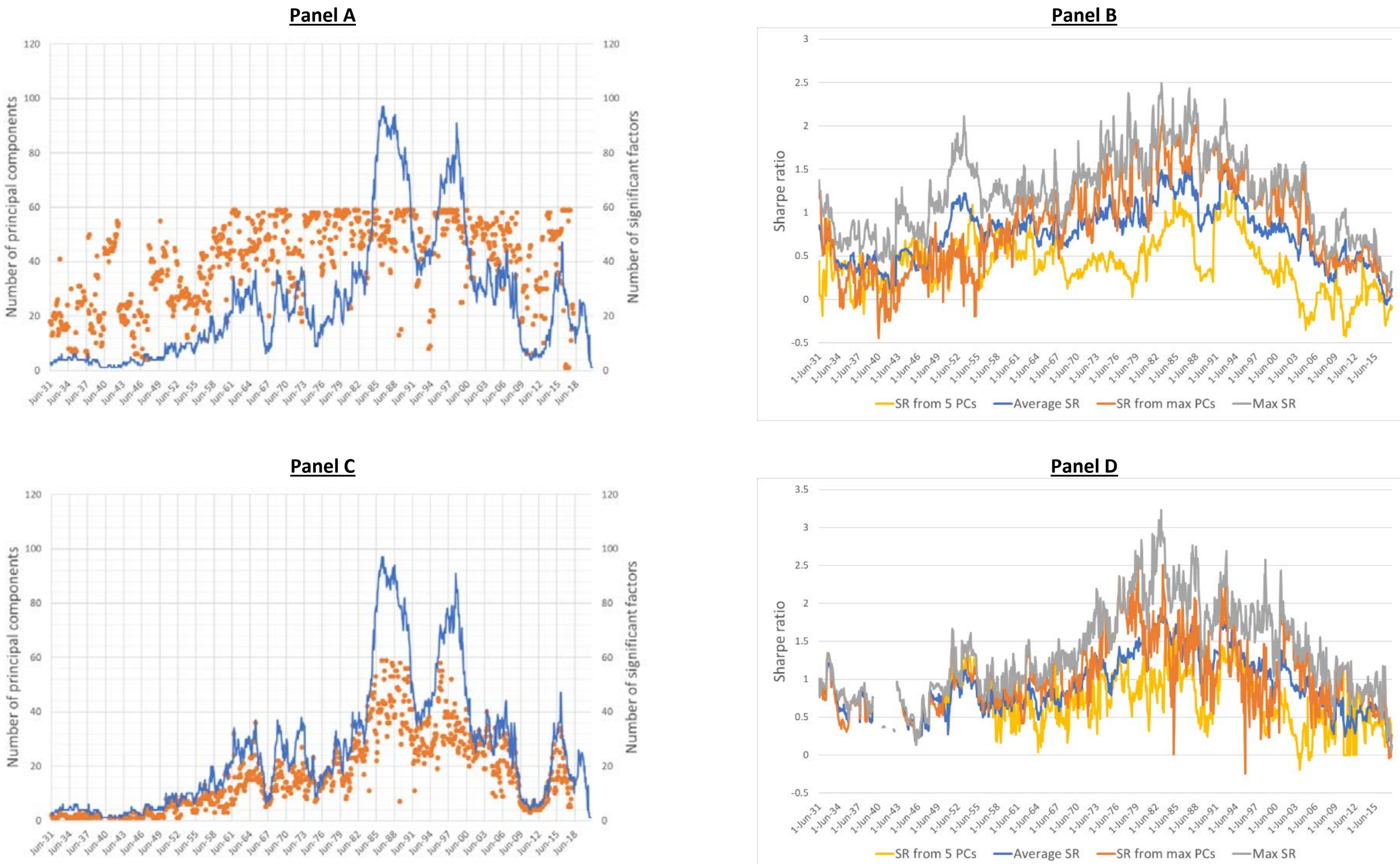
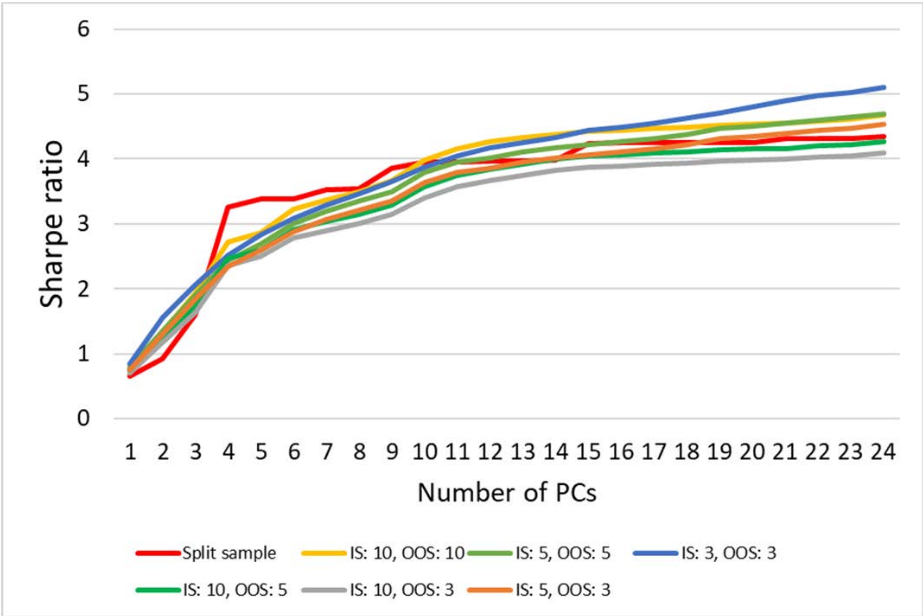


Figure 9. Effect of principal components' estimation and forecasting horizons on out-of-sample Sharpe ratios. This figure shows the average Sharpe ratios of portfolios formed from increasing numbers of principal components for different estimation windows. Panel A shows the rolling average of in-sample Sharpe ratios obtained from forming optimal portfolios consisting of different numbers of principal components. The principal components are computed for a sample of the short and long legs of 12 factors (see Kozak, Nagel and Santosh, 2018 for a description) for a given estimation window from $t - k$ to t, where k can be ~22, 10, 5 or 3 years of daily returns. Principal components are computed on a rolling monthly basis. Panel B shows the rolling averages of out-of-sample Sharpe ratios obtained from portfolios of principal components for a given forecasting window ranging from $t + k$ to t, where k can be ~22, 10, 5, or 3 years of daily returns. The out-of-sample portfolios are constructed using the in-sample optimal weights. The "split sample" lines in both panels confirm the findings of Kozak, Nagel and Santosh (2018).

<u>Panel A</u>: In-sample Sharpe ratios.



<u>Panel B</u>: Out-of-sample Sharpe ratios.

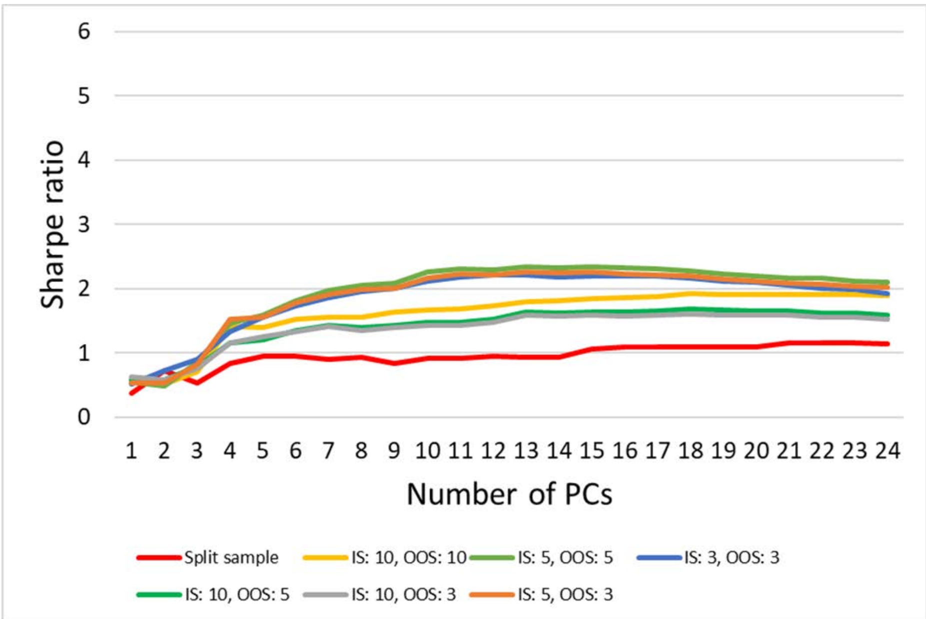


Figure 10. Time series of factor significance This figure shows whether a factor at a specific date has a statistically significant CAPM alpha over the preceding 60 months. For each month *t*, we regress each factor's monthly returns from *t* – *59* to *t* on the market's monthly excess returns to obtain each factor's CAPM alpha. A factor is significant at month *t* if the *t*-statistic of its CAPM alpha exceeds a specific threshold. Factors must have 60 non-missing returns over the alpha estimation period. Each horizontal series represents a different factor with the blue or grey dots signifying a month in which the factor is significant. A green dot denotes the earliest data used in the original study that identified the factor. A red dot denotes the latest data used in the original study. A magenta dot indicates the earliest date for which we are able to estimate the factor's alpha based on data now available. The left vertical axis lists the category of factors according to Chen and Zimmermann (2021). Categories are assigned primarily based on the data source underlying the characteristic used to form the factor. The grey vertical bars represent periods of NBER-defined recessions.

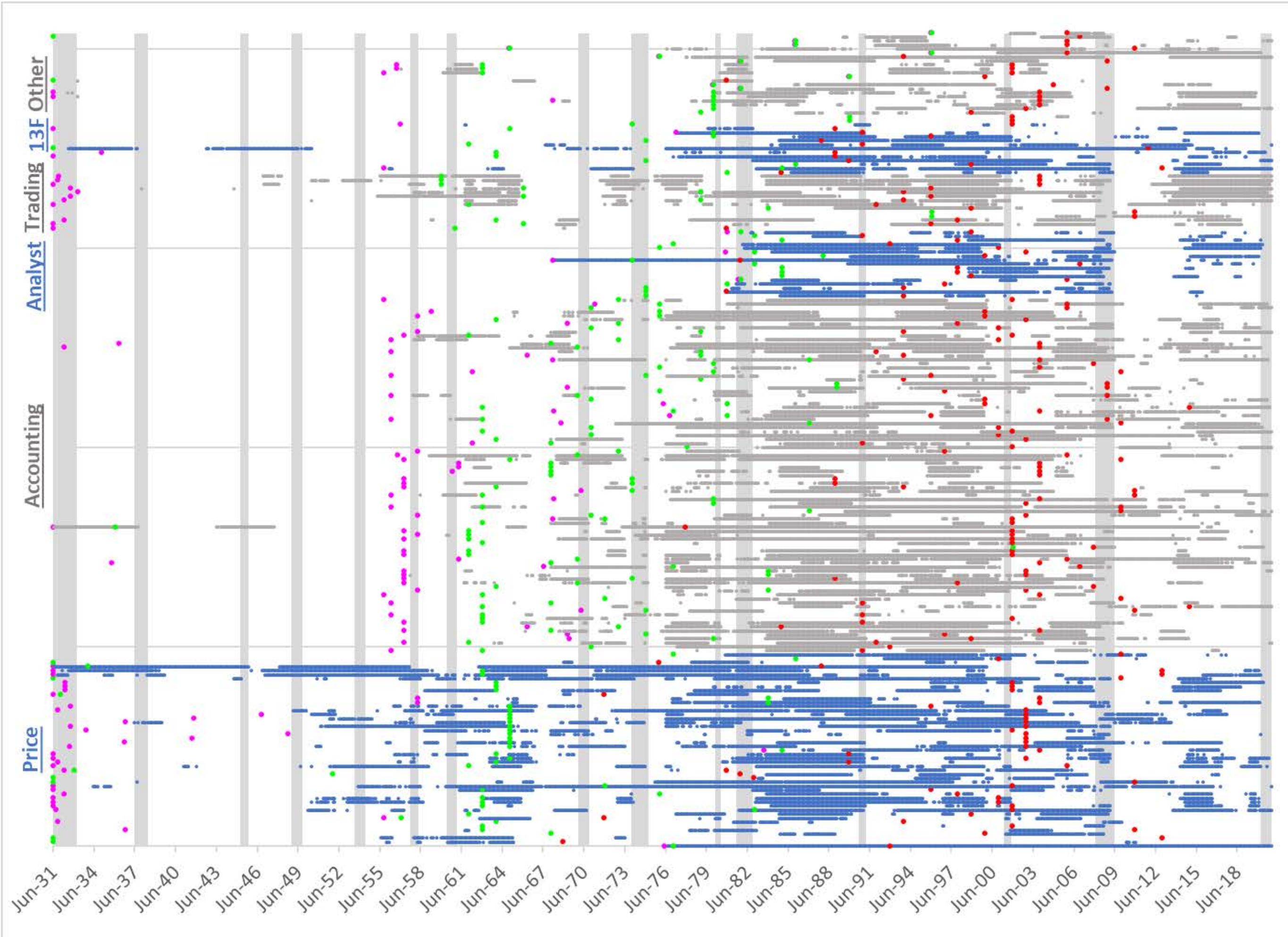


Figure 11. Cumulative number of non-missing cross-sectional characteristics over a firm's lifecycle.
This figure shows the average number of cross-sectional characteristics available for each firm in a given month since the firm first appears in the cross-sectional characteristics dataset of Chen and Zimmerman (2021). The first month the firm appears is indexed at zero. The blue line is for firms that first appeared at any time during the sample. The orange line is the set of firms that first appeared after January 1963.

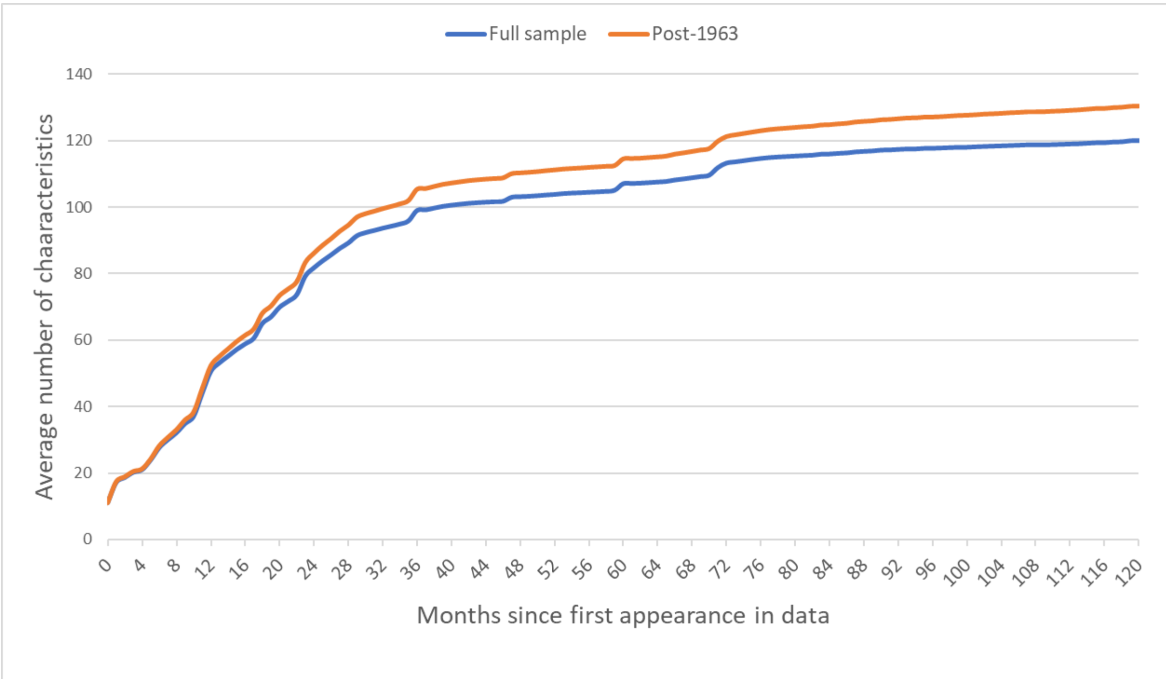


Table 1. Correlation of number of significant factors and significant principal components This table shows the results from regressing the number of significant factors on the number of principal components obtained using those factors. For each month *t*, we regress each factor's monthly returns from $t-59$ to t on the market's monthly excess returns to obtain each factor's CAPM alpha. A factor is significant at month *t* if the *t*-statistic of its CAPM alpha exceeds 3.00. Factors must have 60 non-missing returns over the alpha estimation period. To compute the number of significant principal components at each date *t*, we count the number of principal components required to explain 95% of the cumulative variation of a set of factor returns from $t-59$ to t. We compute the principal components from four samples of factor returns: 1) monthly returns of all factors, 2) monthly returns of significant factors, 3) daily returns of all factors, and 4) daily returns of significant factors. We standardize each independent variable by subtracting the mean of that variable over the time series and dividing that difference by the variable's standard deviation over the time series. Hansen-Hodrick standard errors with a bandwidth of 60 are in parentheses. ***, **, * denote statistical significance at the 1%, 5%, and 10% levels.

	Dep var: Number of significant factors			
	(1)	(2)	(3)	(4)
PCs monthly factors	3.62***			
	(0.63)			
PCs monthly sign. factors		3.98***		
		(0.17)		
PCs daily factors			1.04***	
			(0.13)	
PCs daily sign. factors				1.55***
				(0.14)
constant	-58.51***	0.15	-16.34***	17.36***
	(14.89)	(2.48)	(5.24)	(4.91)
R-squared	0.65	0.94	0.75	0.83
N	1075	1075	1075	1075

Table 2. Summary statistics of factor significance spells across various thresholds of significance For each month *t*, we regress each factor's monthly returns from *t – 59* to *t* on the market's monthly excess returns to obtain each factor's CAPM alpha. A factor is significant at month *t* if the *t*-statistic of its CAPM alpha exceeds one of the various thresholds listed in the table. Factors must have 60 non-missing returns over the alpha estimation period. A significance spell for a given factor is the number of months (i.e., spell length) the factor is continuously significant. Panel A shows the proportion of factors that exhibit at least one significance spell before (after) the sample period of the original paper to identify the factor. Panel B provides summary statistics on the number of significance spells for the cross-section of factors. Panel C computes each factor's average length of a spell and shows summary statistics of this measure for the cross-section of factors. The exceptions are that "Abs min" and "Abs max" show the absolute minimum and maximum spell length of all factors. p-values are calculated from Hansen-Hodrick standard errors with a bandwidth of 60.

Panel A: Proportion of factors with at least one significance spell

		% significant:	
t-statistic	p-value	before original sample	after original sample
1.96	0.0500	77.2	92.6
2.00	0.0455	77.2	92.1
2.50	0.0124	66.9	82.3
2.58	0.0100	65.4	80.3
3.00	0.0027	54.3	68.5
3.50	0.0005	44.1	50.7
4.00	0.0001	23.6	36.5

Panel B: Number of significance spells per factor

t-statistic	p-value	Mean	SD	Median	Min	Max
1.96	0.0500	11.9	6.8	11	0	35
2.00	0.0455	11.7	6.9	11	0	37
2.50	0.0124	9.8	6.5	9	0	33
2.58	0.0100	9.5	6.4	9	0	31
3.00	0.0027	7.7	5.8	6	0	23
3.50	0.0005	5.2	4.6	4	0	21
4.00	0.0001	3.5	4.2	2	0	19

Panel C: Average length of significance spell

t-statistic	p-value	Mean	SD	Median	Min	Max	Abs Min	Abs Max
1.96	0.0500	32.3	52.9	20.0	1.4	535	1	624
2.00	0.0455	31.9	52.6	20.0	1.2	535	1	624
2.50	0.0124	22.1	25.7	15.1	1	233	1	572
2.58	0.0100	21.8	26.7	15.4	1	233	1	571
3.00	0.0027	21.8	45.5	12.7	1	523	1	523
3.50	0.0005	18.3	30.3	11.4	1	260	1	427
4.00	0.0001	20.1	29.6	12.2	1	233	1	415

Table 3. Correlation of the number of significant factors and the standard deviation of stock-level alphas This table shows the results of regressing the number of significant factors in each period on the standard deviation of CAPM alphas for all common stocks in the CRSP universe. For each month *t*, we regress each factor's monthly returns from *t – 59* to *t* on the market's monthly excess returns to obtain each factor's CAPM alpha. A factor is significant at month *t* if the *t*-statistic of its CAPM alpha is greater than 3.00. The dependent variable is a count of the number of significant factors at each month *t*. Stock-level CAPM alphas are obtained by regressing a stock's returns from *t – 59* to *t* on the market return. The standard deviation is either equal-weighted or value-weighted (using each stock's market capitalization at *t – 60)*. We control for the mean standard error and the mean residual volatility of the stock-level CAPM alphas. We standardize each independent variable by subtracting the mean of that variable over the time series and dividing that difference by the variable's standard deviation over the time series. To be included, stocks and factors must have 60 non-missing returns over the alpha estimation period. Hansen-Hodrick standard errors using a bandwidth of 60 are in parentheses. ***, **, * denote statistical significance at the 1%, 5%, and 10% levels.

	Dep var: Number of significant factors					
	(1)	(2)	(3)	(4)	(5)	(6)
Equal-weighted (%):						
Standard deviation of alphas	9.80**	22.69*	22.22*			
	(4.69)	(11.79)	(11.37)			
Mean standard error of alphas		-14.47				
		(10.44)				
Mean residual volatility			-14.14			
			(9.93)			
Value-weighted (%):						
Standard deviation of alphas				9.39**	12.57**	12.61**
				(4.68)	(5.62)	(5.54)
Mean standard error of alphas					-4.67	
					(4.05)	
Mean residual volatility						-4.85
						(3.92)
R-squared	0.18	0.26	0.27	0.17	0.19	0.19
N	1075	1075	1075	1075	1075	1075

Table 4. Comovement of the number of significant factors with size-based components of the cross-sectional dispersion of stock-level alphas This table shows the results of regressing the number of significant factors at each period on various size-based components that make up the cross-sectional variance of stocks' CAPM alphas sorted by firm size. For each month *t*, we regress each factor's monthly returns from *t – 59* to *t* on the market's monthly excess returns to obtain each factor's CAPM alpha. A factor is significant at month *t* if the *t*-statistic of its CAPM alpha is greater than 3.00. The dependent variable is a count of the number of significant factors at each month *t*. We also compute each stock's CAPM alpha at each month *t* by regressing each stock's monthly returns from *t – 59* to *t* on the market's monthly excess returns to obtain a stock's CAPM alpha. We then compute the sum of squared deviations divided by the total number of firms (minus 1) in the full sample at each month for the samples of 1) all firms, 2) large-cap firms, 3) mid-cap firms and 4) small-cap firms. Note that the measure for the sample of all firms is equivalent to the variance while each size-based component is the variation that contributes to the full sample's variance. Large-caps are stocks with a market cap above the 50th NYSE percentile. Mid-caps are stocks with a market cap between the 20th and 50th NYSE percentile. Small-caps are stocks with a market cap less than the 20th NYSE percentile. We standardize each of the four independent variables by subtracting the mean cross-sectional dispersion of the full sample of firms across all dates from each variable and then dividing by the standard deviation of the full sample across all firms. Panel A covers the sample of factors from 1931-2020. Panel B covers the sample of factors from 1968-2020. To be included, stocks and factors must have 60 non-missing returns over the alpha estimation period. Hansen-Hodrick standard errors are in parentheses. ***, **, * denote statistical significance at the 1%, 5%, and 10% levels.

	Dep. var: Number of significant factors			
	1931-2020		1968-2020	
	(1)	(2)	(3)	(4)
Sample:				
All firms	9.48**		9.57*	
	(4.71)		(5.28)	
Large-cap firms		-7.14*		-15.17
		(4.26)		(9.25)
Mid-cap firms		0.36		6.70
		(4.06)		(7.75)
Small-cap firms		12.76***		8.67**
		(2.83)		(3.66)
R-squared	0.17	0.42	0.09	0.18
N	1075	1075	636	636

Table 5. Comovement of the number of significant factors with the number of public firms This table shows the results of regressing the number of significant factors in each period on the number of public firms at each date. For each month *t*, we regress each factor's monthly returns from *t – 59* to *t* on the market's monthly excess returns to obtain each factor's CAPM alpha. A factor is significant at month *t* if the *t*-statistic of its CAPM alpha is greater than 3.00. The dependent variable is a count of the number of significant factors at each month *t*. The number of public firms consists of all common stocks trading on the NYSE, NASDAQ or Amex and is computed in three ways: 1) all public firms at month *t*, 2) all public firms at month *t – 59*, and 3) all public firms with a non-missing alpha over the period *t – 59* to *t*. To be included, stocks and factors must have 60 non-missing returns over the alpha estimation period. Panel A covers the sample of factors from 1931-2020. Panel B covers the sample of factors from 1968-2020. We standardize each independent variable by subtracting the mean of that variable over the time series and dividing that difference by the variable's standard deviation over the time series. Hansen-Hodrick standard errors are in parentheses. ***, **, * denote statistical significance at the 1%, 5%, and 10% levels.

	Dep var: Number of significant factors							
	1931-2020				1968-2020			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Total number of firms:								
At time t	18.24***			29.89**	26.74***			32.46**
	(3.08)			(13.76)	(7.97)			(12.99)
At time t-60		15.42***		5.98		13.86***		2.85
		(3.92)		(9.28)		(4.88)		(9.51)
with 60-month alpha			15.69***	-18.23			17.01***	-11.10
			(3.70)	(11.37)			(5.67)	(11.58)
R-squared	0.63	0.45	0.47	0.67	0.50	0.17	0.19	0.52
N	1075	1075	1075	1075	636	636	636	636

Table 6. Disentangling noise from power This table shows the results of regressing the number of significant factors in each period on the cross-sectional mean of the absolute value of all factors' CAPM alphas, the mean standard error of those alphas and the number of public firms. For each month *t*, we regress each factor's monthly returns from *t – 59* to *t* on the market's monthly excess returns to obtain the factor's CAPM alpha and its corresponding standard error. The dependent variable is a count of the number of significant factors at each month *t*. A factor is significant at month *t* if the *t*-statistic of its CAPM alpha is greater than 3.00. To be included, factors must have 60 non-missing returns over the alpha estimation period. The number of public firms is a count of all common stocks outstanding at *t*. We standardize each independent variable by subtracting the mean of that variable over the full time series and dividing that difference by the variable's standard deviation over the time series. Hansen-Hodrick standard errors with a bandwidth of 60 are in parentheses. ***, **, * denote statistical significance at the 1%, 5%, and 10% levels.

	Dependent var: Number of significant factors							
	1931-2020				1968-2020			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean standard error of factor alphas	-8.69**	-2.69	-30.01***	-19.39***	-17.01	-17.61**	-37.64***	-34.61***
	(3.48)	(1.98)	(6.98)	(6.72)	(13.42)	(8.22)	(4.03)	(4.13)
Number of public firms		17.31***		11.16***		26.91***		7.97***
		(2.90)		(1.67)		(7.26)		(1.67)
Mean factor absolute alpha			27.08***	18.51***			37.26***	31.47***
			(7.75)	(6.63)			(3.80)	(4.35)
R-squared	0.14	0.64	0.67	0.83	0.09	0.60	0.90	0.92
N	1075	1075	1075	1075	636	636	636	636

Table 7. Comovement of the number of significant factors and economy and firm characteristics This table shows the results of regressing the number of significant factors in each period on various economic measures at each month. For each month *t*, we regress each factor's monthly returns from *t – 59* to *t* on the market's monthly excess returns to obtain the factor's CAPM alpha and its corresponding standard error. A factor is significant at month *t* if the *t*-statistic of its CAPM alpha is greater than 3.00. The dependent variable is a count of the number of significant factors at each month *t*. The number of public firms is a count of all common stocks at *t* traded on the NYSE, NASDAQ or Amex at month *t*. The NBER recession indicator is an indicator equal to one if the month is classified as an NBER recession and zero otherwise. The unemployment rate is the number of unemployed as a percentage of the labor force as provided by the U.S. bureau of labor statistics. The federal funds rate is the established rate by the Federal Reserve at month *t*. The 10-year treasury bond yield is the market yield on U.S. treasury securities at a 10-year constant maturity. The percent of dividend-paying firms is the total number of common stocks which have paid a dividend in the previous 12 months divided by the number of firms at month *t*. The mean institutional ownership is the fraction of a firm's shares outstanding held by 13-f firms. The economic complexity index is a measure of economic complexity used from Simoes and Hidalgo (2011). Diversity of firm characteristics is a measure of diversity in the cross-sectional characteristics across firms. See Appendix Table A1 for a complete description of the measures. Panel A covers the sample of factors from 1931-2020. Panel B covers the sample of factors from 1968-2020. Hansen-Hodrick standard errors are in parentheses. ***, **, * denote statistical significance at the 1%, 5%, and 10% levels.

Panel A: Sample of factors from 1931-2020

	Dep var: Number of significant factors													
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Number of public firms	18.24***											5.85	-8.13	-26.36
	(3.08)											(6.41)	(9.01)	(20.29)
NBER recession indicator		-9.41**										-9.90**	-7.34*	1.38
		(4.66)										(4.17)	(3.91)	(3.86)
Unemployment rate			0.11									-11.16***	-8.71***	-8.33***
			(4.38)									(1.77)	(2.72)	(1.76)
Fed funds rate				8.95*								-5.62**	-2.33	-0.34
				(5.18)								(2.46)	(2.56)	(4.19)
10-year T-Bond yield					9.64*							15.46***	9.41***	1.33
					(5.23)							(4.48)	(3.19)	(4.31)
% dividend-paying firms						-12.12***						-1.32	2.55	-14.77
						(4.47)						(8.98)	(7.32)	(14.09)
Mean institutional ownership							-16.81***							-13.79*
							(4.46)							(7.13)
Economic complexity index								5.59					2.66	-1.63
								(4.72)					(2.75)	(2.74)
Mean industry HHI									-15.06***				-29.53	-40.18*
									(4.29)				(19.83)	(20.76)
Mean Amihud illiquidity										-5.30***		12.97***	9.17**	-3.41
										(1.90)		(4.33)	(4.08)	(5.76)
Diversity of firm characteristics											14.76***		9.42**	18.57***
											(4.76)		(3.81)	(7.05)
R-squared	0.63	0.02	0.00	0.16	0.18	0.28	0.46	0.06	0.44	0.05	0.42	0.67	0.73	0.83
N	1075	1075	876	798	708	1075	483	648	852	1075	1039	672	648	456

Table 7 continued.

Panel B: Sample of factors from 1968-2020

	Dep var: Number of significant factors													
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Number of public firms	26.74***											6.19	-16.54**	-26.36
	(7.97)											(4.93)	(7.60)	(20.29)
NBER recession indicator		-11.90**										-7.25**	-2.62	1.38
		(5.58)										(3.48)	(1.93)	(3.86)
Unemployment rate			-5.30									-11.25***	-8.94***	-8.33***
			(4.22)									(2.24)	(2.75)	(1.76)
Fed funds rate				7.58								-5.62*	-1.95	-0.34
				(5.36)								(2.90)	(2.97)	(4.19)
10-year T-Bond yield					9.22*							16.90***	7.35***	1.33
					(5.31)							(3.95)	(2.58)	(4.31)
% dividend-paying firms						-12.09*						-4.35	1.86	-14.77
						(6.21)						(7.59)	(5.03)	(14.09)
Mean institutional ownership							-16.81***							-13.79*
							(4.46)							(7.13)
Economic complexity index								6.79					0.88	-1.63
								(4.20)					(2.63)	(2.74)
Mean industry HHI									-42.55***				-52.22***	-40.18*
									(8.44)				(17.88)	(20.76)
Mean Amihud illiquidity										8.25		9.20***	1.40	-3.41
										(11.77)		(3.30)	(1.93)	(5.76)
Diversity of firm characteristics											17.23***		8.55***	18.57***
											(5.71)		(2.97)	(7.05)
R-squared	0.50	0.03	0.05	0.12	0.17	0.11	0.46	0.09	0.64	0.01	0.38	0.71	0.80	0.83
N	636	636	636	636	636	636	483	600	636	636	600	600	600	456

Appendix

I. Theory and Intuition

Consider an APT-type return process with the market factor, f, and a vector g of additional non-market factors. The exposure vectors are β and γ.

$$R_{i,t} = \beta_{i,t} f_t + \gamma_{i,t}' g_t + \epsilon_{i,t} \tag{1}$$

Parameter Definitions. We assume for simplicity the market and non-market factors are orthogonal, but this is not necessary for the main points. The idiosyncratic returns ϵ are also orthogonal to all the factors. There are two sets of parameters: those for the factors and those for the distribution of stocks in the cross-section. Though formally we will allow them to be time varying, we will assume for a given sample window that the parameters are fixed to simplify the notation. Further, we will assume the cross-sectional stock level exposures are constant over the sample windows. Also, we will assume all the factors and idiosyncratic shocks are independent across time.

We will then compare the key statistics of interest as a function of these parameters, which one can think of as changing across sample windows. We can generalize to the parameters changing within the sample windows. Doing so leaves the main conclusions, except for in perverse situations.

All parameters are assumed to be normally distributed with mean denoted by μ and variances denoted by Σ (unless a scalar as in the case of idiosyncratic shocks, in which case we will denote the standard deviation as σ). The idiosyncratic shocks are mean zero by construction. The cross-section distribution of the idiosyncratic shock variance must be non-normal. We take no stand on that distribution beyond it having a defined mean: $\bar{\sigma}_t$ There are a finite number of stocks N in the cross-section.

Statistics of interest. In this section we derive how the key statistics used in the empirical section depend upon the parameters of the return generating process. For all statistics we focus on the expected value of these statistics given the parameters. This expectation is across both time series draws in a particular sample period and across cross-sectional draws of stocks.

Consider the regression of individual stocks over a 60-month sample window on the market factor f:

$$R_{i,t} = \alpha_{i,T} + \beta_{i,T} f_t + \theta_{i,t} \quad \text{for } t \in \{T-59, \ldots, T\} \tag{2}$$

We compute the cross-sectional variance of the alpha estimates and the average of the standard variances. (We focus on variances rather than standard deviations for tractability.) For these calculations, we ignore the variation in the market factor that is not removed by the regressions, as this is not the main focus of the paper.

We then focus on regressions of non-market factor returns on the market factor:

$$g[k]_t = a + b \quad f_t + \eta_{k,t} \quad \text{for } t \in \{T-59, \ldots, T\}, \tag{3}$$

where k denotes the k^{th} non-market factor. We compute the expected probability of the factor having an intercept greater than a statistical cutoff t^*. Because the market and non-market factors are orthogonal, we again will ignore the limited variation introduced by random variation in the market factor.

A. Cross-sectional variance of stock level alphas

The estimated alpha from the stock-level regression, Equation (2), (ignoring the market factor contribution) is

$$\hat{\alpha}_{i,T} = \frac{1}{T} \sum_{\tau=T-59}^{T} \left(\epsilon_{i,t} + \gamma'_{i,t} g_t \right) \tag{4}$$

We suppress the time subscripts on the parameters since we are assuming the parameters are constant through the estimation window. The cross-sectional variance of these alphas is

$$\sigma^2_{\alpha,T,CS} = \frac{1}{N-1} \sum_{i \epsilon N} (\hat{\alpha}_{i,T} - \bar{\bar{\alpha}}_T)^2 \tag{5}$$

where

$$\bar{\hat{\alpha}}_T = \frac{1}{N}\sum_{i\in N}\hat{\alpha}_{i,T} \tag{6}$$

Plugging Equations (6) and (4) into Equation (5) and taking its expectation gives

$$E\left[\sigma^2_{\alpha,T,CS}\right] = \left(\frac{N-1}{N}\right)\left[\frac{\bar{\sigma}^2_{\epsilon}}{60} + \mu_g'\Sigma_\gamma\mu_g + \frac{1}{60}h\left(\Sigma_\gamma,\Sigma_g\right)\right] \tag{7}$$

where h is a symmetric increasing function of the two variances.

Thus, the expected cross-sectional variance is determined primarily by the non-market factor premia μ_g and the cross-sectional exposure to the factors Σ_γ. It is less so determined by the average idiosyncratic risk and the combined variation in the exposure to the factors and factor volatility.

In the empirical results we control for these additional two effects by including the average standard error in of the stock level alphas. We derive this next.

B. Average standard variance of the alphas

The standard variance of the stock-level alphas from Equation (2), ignoring the contribution from the modeled factors realizations, is

$$\left[SE\left(\alpha_{i,T}\right)\right]^2 = \left(\frac{1}{60-2}\right)\left(\frac{1}{60}\right)\sum_{\tau=T-59}^{T}\left(\epsilon_{i,T} + \gamma_{i,\tau}' - \frac{1}{60}\sum_{s=T-59}^{T}\left\{\epsilon_{i,s} + \gamma_{i,s}'g_s\right\}\right)^2 \tag{8}$$

Taking expectations and then averaging across firms in the cross-section yields

$$Ave\left(E\left[SE\left(\alpha_{i,T}\right)^2\right]\right) = \left(\frac{1}{60-2}\right)\left(\frac{59}{60}\right)\left[\bar{\sigma}^2_\epsilon + \mu_\gamma'\Sigma_g\mu_\gamma + h\left(\Sigma_\gamma,\Sigma_g\right)\right] \tag{9}$$

Thus, this average standard variance is more heavily dependent upon the idiosyncratic risk and the combined volatility of the factor and factor exposure than the cross-sectional standard deviation, and comprises a good control variable for these components. It also depends upon the interaction of the average factor exposure and the volatility of the factor, which under the assumption that average factor exposures outside the market are close to zero, will contribute a negligible amount. Hence, this makes a good control to remove the parts of the cross-sectional standard deviation of alpha that are of less interest.

C. Probability that unmodeled factors are significant

In the empirical section we measure the statistical (and economic significance) of the unmodeled factors by looking at the t-values of the intercept of the unmodeled factors on the modeled factors. In particular, we calculate the number of significant factors as those with t-values above a cutoff t^*. To understand how the likelihood of a factor being significant is a function of the parameters, we compute the expected power. We begin by considering the case where the factor is directly observable and then generalize to the case where the factor is formed from a long-short portfolio from characteristic-based sorts.

C.1 Power of observable factor

This case corresponds directly to the regression in Equation (3). Again, we ignore the contribution of the random realization of the market factor. The expected power for factor *k* is a standard calculation:

$$1 - \Phi\left(t^* - \frac{\sqrt{60}\mu_g[k]}{\sigma_{\eta_k}}\right) \tag{10}$$

where Φ is the CDF of a standard normal distribution. As would be anticipated, power increases in the factor premium and decreases in the volatility of the factor, σ_{η_k}. Thus, other things equal there will be a positive correlation between the cross-sectional variance of alphas from individual stock regressions and the number of significant factors.

C.2. Power of unobservable factor

The positive relation between the number of listed firms and the number of significant factors is potentially attributable in part to the increased precision of the factor measurement with more firms. To assess this issue, let the non-market factors be unobserved, but there be an observable firm level characteristic $C[k]_{i,t}$ that is jointly normal with firm i's factor exposure $\gamma[k]_{i,t}$. Let the variance-covariance matrix for this joint distribution be $\Sigma_{C,t}$ For conciseness, we assume that the characteristics and exposures are fixed over a sample window. Let us form the factor $G[k]_{i,t}$during $t \in \{T-59, \ldots, T\}$

from a long-short portfolio. The long and short ends are formed by sorting all stocks at $T - 60$ into Q quintiles based on the characteristic $C[k]_{i,T-60}$

The factor $G[k]$ so formed will be a function of the number of firms and the number of quintiles. Our interest lies in the expected mean and the expected variance of the factor, as these affect the power to detect its significance. For a finite number of firms, the breakpoints and hence, average exposure, to $g[k]$ will vary from sample to sample based on the random realizations of the characteristics and factor exposures. The amount of variation in these will decrease in the number of firms, however these exposures will not vary in expectation with the number of firms. Hence the contribution to the expected mean and variance of the factor $G[k]$ will not vary from the exposure to $g[k]$ as the number of firms change.

Nevertheless, the total expected variance of $G[k]$ will vary with the number firms due to amount of idiosyncratic risk that diversifies away in the long and short end of the portfolio. The number of firms in each long and short portfolio is $\frac{N}{Q}$. Under the assumption of idiosyncratic risk being independent of everything else, the expected variance of the idiosyncratic risk in each of these portfolios is $\frac{Q}{N}\bar{\sigma}_{\epsilon}^{2}$. Because this risk is additive across the long and short portfolio the total variance of factor $G[k]$ due to this diversified idiosyncratic risk is

$$\frac{2Q}{N}\bar{\sigma}_{\epsilon}^{2}. \tag{11}$$

Thus, this component decreases in the number of firms. Since the variance of the factor decreases as we increase the number of firms, the power to detect a factor as significant increases with the number of firms, as shown in (10). Thus, either increasing the factor premium of g or the number of firms (via reduced volatility) can drive the number of significant factors detected. To distinguish whether the relation we find between the number of firms and the number of significant factors is driven by only the later or both parameters, we can take advantage of the standard error of the intercept estimates in the regression

$$G[k]_t = a + b' f_t + \eta_{k,t} \text{ for } t \in \{T-59, \dots, T\} \tag{12}$$

which is also driven by this idiosyncratic risk in the long-short portfolio (Equation (11)). Thus, we can regress the number of significant factors on the average intercept estimates (alphas) and the average intercept standard errors. The prediction is that we will observe a positive coefficient on the former and a negative coefficient on the later. If both are true, then the relation between the number of significant factors and the number of firms is not simply attributable to increased precision in the observability of the unmodeled factors from an increasing number of firms.

II. Simulation

To validate our findings regarding the magnitude of out-of-sample Sharpe ratios formed from portfolios of principal components and those portfolios' relevance in identifying non-redundant factors. We simulate a time series of returns for a total of 205 factors for 96 periods t = 1 to t = 96. All factor returns are assumed to follow a normal distribution. The number of priced factors varies from n = 1 to n = 100 (as in the actual data), with the remaining factors being unpriced. Priced factors have a positive mean return, μ_p, and standard deviation, σ_p. Unpriced factors have a zero mean return and standard deviation, σ_u. We begin our analysis by assuming all factors to be independent, that is the cross-correlations among factors is 0.

After constructing the simulated factor returns for 96 time periods and 205 factors, we compute the principal component factor loadings for the set of simulated factor returns for the in-sample period of the first 60 months, $t - 59$ to t. We construct the out-of-sample principal components by applying the in-sample factor loadings to the out-of-sample simulated data for the final 36 months, i.e., from $t + 1$ to $t + 36$. We then construct in-sample and out-of-sample portfolios with an increasing number of principal components included in each portfolio and measure the Sharpe ratios. More specifically, we construct optimized portfolios based on the first, first two, first three, etc., up to the first sixty principal components. In the in-sample case, portfolio weights are chosen to optimize the portfolio Sharpe ratio. For the out-of-sample portfolios, we use the optimal portfolio weights from the in-sample portfolios. We repeat this simulation 10,000 times resulting in 59 portfolios from 10,000 draws of simulated data. We average the Sharpe ratio within portfolios of the same number of principal components across all simulations, resulting in 59 portfolios and their average Sharpe ratios. We report the maximum Sharpe ratio of the 59 portfolios and the number of principal components that are used to form this Sharpe ratio.

To capture the time variation in the number of significant factors and the Sharpe ratios of those factors, we calibrate the mean of the priced factors, μ_p, and the standard deviations, σ_p, of the priced factors by using the values measured in the actual data during a high Sharpe ratio regime (e.g., around

1986) or during a low Sharpe ratio regime (e.g., around 1967 or 2009 – See figure 8 Panels B and D). We also vary the standard deviation of the unpriced factors, σ_u, between the high and low values observed in the data for each of these two regimes. Table A2 presents the numerical values of these parameters. More specifically, μ_p varies between 0.04% and 0.16%, σ_p is calibrated to be either 1.8% or 3.2%, and σ_u is calibrated to be either 2.5% or 3.6%.

The baseline scenario is calibrated to match the parameters measured in the data during a high SR regime with a low standard deviation of unpriced factors. The blue solid line in Figure A1 Panel A presents the maximum portfolio Sharpe ratios of the baseline scenario as a function of the number of priced factors. We find that increasing the number of priced factors while holding other parameters constant increases the out-of-sample Sharpe ratios. This increase occurs at a decreasing rate as more priced factors are likely to be partially redundant with earlier factors. Importantly, at 50 priced factors, the simulation matches the Sharpe ratios of approximately 2.5 and 0.7 measured in the data and shown in Figure 8 Panel B for the high and low regime periods. As expected, the change in the unpriced standard deviation (as shown by the dotted lines), has little effect on the Sharpe ratios, suggesting the out-of-sample Sharpe ratios are a good measure of redundancy.

Panel B of Figure A1 shows the maximum out-of-sample portfolio Sharpe ratios as a function of μ_p for 1, 50 and 100 priced factors. For each case, we vary σ_p between 0.18 and 0.32, and σ_u between 0.25 and 0.36. We find that an increase in the mean of the priced factors results in monotonically increasing Sharpe ratios, again confirming that the out-of-sample Sharpe ratios is a reliable measure reflecting the number of priced factors. The figure also reveals that as the unpriced factor standard deviations increase, the out-of-sample Sharpe ratios decrease. This is at This variation in the parameters allows us to match the out-of-sample Sharpe ratios found in the actual data.

Thus far, we have assumed independence between all factors, both unpriced and priced. We now assess the extent to which this assumption has on the out-of-sample Sharpe ratios. We focus on correlations between same-type (i.e., priced/priced or unpriced/unpriced) factors and cross-correlations between different-type (i.e., priced/unpriced) factors. In Table A2 Panel C, we plot the maximum Sharpe ratios for all correlations varying from 0.00 to 0.20 and as a function of the number of priced factors. We present results for a calibration with the cross-correlations zeroed out and the cross-correlation positive and equal to the same-type correlations. Our first finding is that the maximum out-of-sample Sharpe ratio is decreasing as the mean correlation among factors is increasing. As the correlation increases, the priced factors become increasingly redundant resulting in lower Sharpe ratios. This suggests that the data must have a high number of priced factors as we require very low correlations among priced factors to match the data. We also observe that for correlations greater than 0.05, positive cross correlation results in increased out-of-sample Sharpe ratios. Essentially, positive cross-correlations allow for free hedges from the unpriced factors. This is not economically sensible. If the correlations are sufficiently low, the free hedges become small and contribute very little to the diversification effects in the portfolio and, thereby, the out-of-sample Sharpe ratio.

Finally, we explore how the correlation affects the number of principal components that yields the portfolio with the maximum Sharpe ratio as a function of the number of priced factors and mean correlation between same-type factors. Panel D of Figure A2 shows for the factor independence case, the number of priced factors is only weakly positively correlated with the number of principal components generating the portfolio with the maximum Sharpe ratio. As the correlation among same-type factors increases, however, we see a strong negative correlation with the number of principal components that make up the out-of-sample portfolio with the highest Sharpe ratio. In essence, the same-type correlations increase the redundancy among priced factors, thereby limiting the number of principal components that provide marginal increases to the Sharpe ratio. This result is the analog of the hump-shaped pattern of the out-of-sample Sharpe ratios found in Panel B of Figure 9.

Overall, the simulation provides evidence of limited redundancy in factor data and confirms the magnitudes of out-of-sample Sharpe ratios obtained from the data.

<u>Figure A1</u> – **Simulation of Sharpe ratios obtained from portfolios of out-of-sample principal components.** These figures show the maximum Sharpe ratios of portfolios consisting of out-of-sample principal components constructed using simulated data. The simulation assumes 205 factors. See section XXX for details of the simulation. Panel A shows how the maximum Sharpe ratios vary with the number of priced factors under parameters obtained from the factor returns during periods of high and low Sharpe ratios, and high and low standard deviations of unpriced factors. Panel B shows how the maximum Sharpe ratios vary with the mean return of the priced factors, while also varying the number of priced factors and the standard deviations of the priced and unpriced factors. Panel C shows how the Sharpe ratios vary with the number of priced factors and different correlations of the factors. The solid lines in Panel C represent the case when zero cross-correlation exists between the priced and unpriced factors, while the dashed lines have those correlations set to the same mean as the correlations. Panel D shows the number of principal components in the portfolio that yields the maximum Sharpe ratio for the case of zero cross-correlations between priced and unpriced factors.

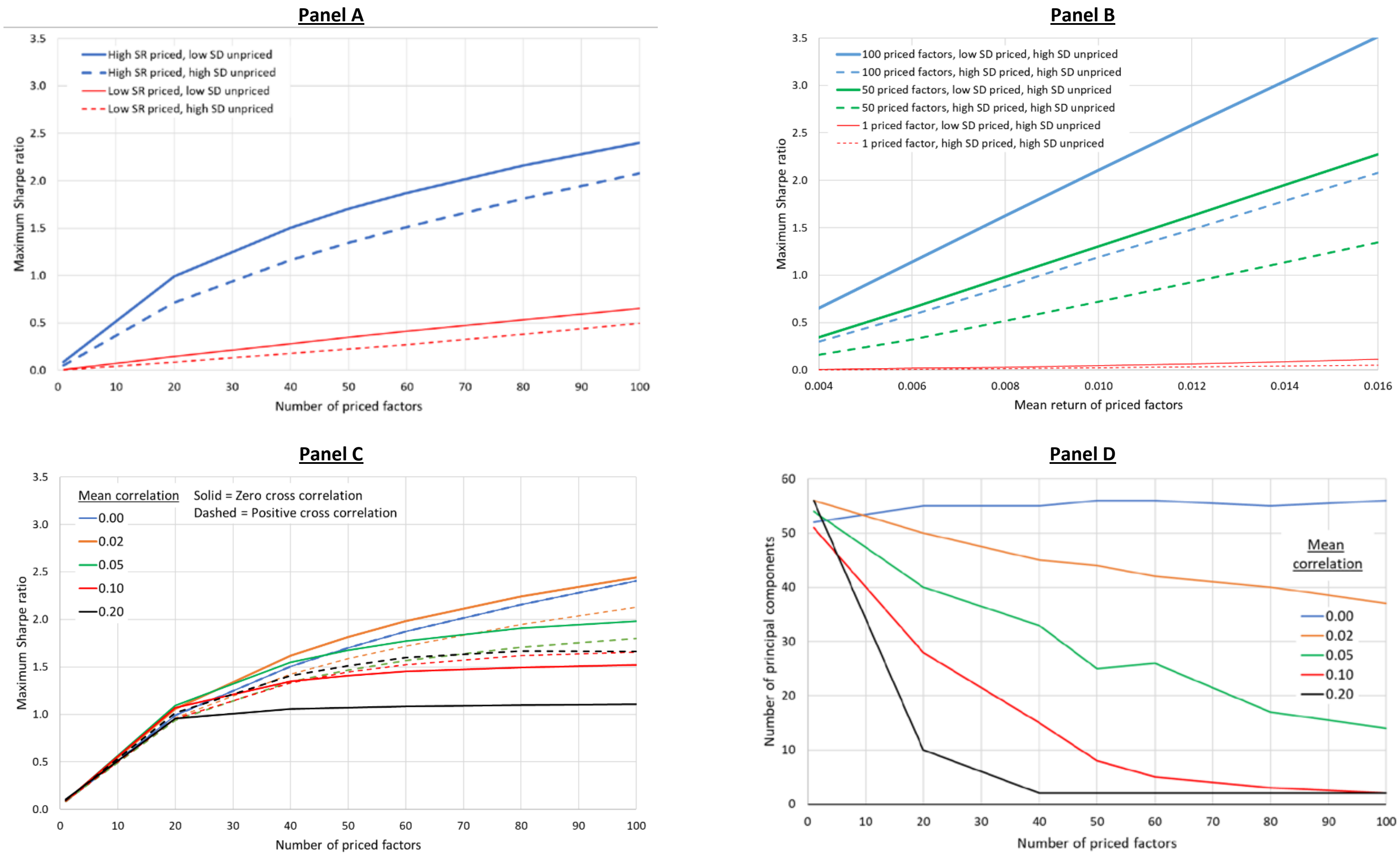


Table A1. Variable definitions This table summarizes the various variables we use throughout the analysis. The variables are listed in order of appearance in the paper.

Variable name	Description
Number of significant factors	The total number of significant factors at each month *t*. At each month *t*, we regress each factor's returns from *t-60* to *t-1* on the market's excess returns over the same period to obtain the factor's CAPM alpha. A factor is considered significant if the *t*-statistic of its CAPM alpha is greater than 3.00. To be included, the factor must have zero non-missing returns over the 60-month period. Factors come from Chen and Zimmerman (2021) and are categorized as "clear" or "likely" predictors.
Standard deviation of stocks' alphas	The equal-weighted (value-weighted) cross-sectional standard deviation of stocks' alphas at month t. At each month *t*, we regress each factor's returns from *t-60* to *t-1* on the market's excess returns to obtain the stock's CAPM alpha. To be included, the stock must have zero non-missing returns over the 60-month period. Stocks are all common stocks (CRSP share codes 10 or 11) listed on the NYSE, AMEX and NASDAQ. The value-weighted cross-sectional standard deviation is weighted by each stock's market capitalization at *t-61*. Units are expressed as percentage points. This measure is standardized across the sample for ease of interpretability.
NBER recession	A recession indicator equal to 1 if the economy at month t was in a recession as defined by the National Bureau of Economic Research. Data can be obtained here: https://fred.stlouisfed.org/series/USREC
25 Size and Book-to-Market portfolios	Monthly equal-weighted returns from 25 size and book-to-market portfolios provided on Ken French's website: https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html
30 Fama-French industry portfolios	Monthly equal-weighted returns from Fama-French 30 industry portfolios provided on Ken French's website: https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html
Number of significant principal components	The number of significant components at month *t* required to explain 95% of the variation in factor returns from *t-60* to *t-1*. The set of factors may be either all factors or only significant factors during the time period. Returns may be at either the monthly or daily frequency.
Significance spell of factor	The number of consecutive months for which a factor remains significant.

Mean standard error of alphas	The equal-weighted average standard error of CAPM alphas for all stocks at month *t*. The value-weighted mean standard error is weighted by each stock's market capitalization at *t-61*. This measure is standardized across the sample for ease of interpretability.
Mean residual volatility	The average residual volatility across stocks obtained from the 60-month stock-level CAPM regressions at each month *t*. This measure is standardized across the sample for ease of interpretability.
Number of public firms	The total number of CRSP common stocks listed on the NYSE, AMEX or NASDAQ at time *t*. This measure is standard across the sample for ease of interpretability.
Number of public firms at beginning of estimation period	The total number of CRSP common stocks listed on the NYSE, AMEX or NASDAQ at time *t-60*. This measure is standard across the sample for ease of interpretability.
Number of public firms with alpha at *t*	The total number of CRSP common stocks at month *t* listed on the NYSE, AMEX or NASDAQ for which an alpha can be calculated (i.e., stock has zero non-missing returns from *t-60* to *t-1*). This measure is standard across the sample for ease of interpretability.
Mean absolute alpha	The equal-weighted average of the absolute value of alpha for all factors at month *t*. This measure is standardized across the sample for ease of interpretability.
Mean standard error of alphas	The equal-weighted average standard error of alphas for all factors at month *t*. This measure is standardized across the sample for ease of interpretability.
Unemployment rate	The percentage of the labor force unemployed at *t* as determined by the US Bureau of Labor Statistics. This measure is standardized across the sample for ease of interpretability. Data can be obtained here: https://fred.stlouisfed.org/series/UNRATE
Fed funds rate	The federal funds rate at the end of each month *t*. This measure is standardized across the sample for ease of interpretability. Data can be obtained here: https://fred.stlouisfed.org/series/FEDFUNDS

10-year Treasury Bond yield	The 10-year Treasury bond yield at the end of each month *t*. This measure is standardized across the sample for ease of interpretability. Data can be obtained here: https://fred.stlouisfed.org/series/DGS10
% of dividend-paying firms	The total number of common stocks which pay a dividend divided by the total number of common stocks at each month *t*. A stock is defined as paying a dividend if at least one dividend was paid over the previous year. This measure is standardized across the sample for ease of interpretability.
Mean institutional ownership	The average institutional ownership across stocks at month *t*. For each stock at month *t*, the percentage of institutional ownership is determined by the total number of shares held by institutions divided by the total number of shares outstanding. Institutional shareholdings are obtained from Thomson-Reuters 13-F database. This measure is standardized across the sample for ease of interpretability.
Economic complexity index	An annualized measure of economic complexity based on the complexity of trade activities within the United States. Each month *t* uses the measure from December of the most previous year. This measure is standardized across the sample for ease of interpretability. Data can be obtained here: https://oec.world/en/rankings/legacy/eci
Average 3-digit industry HHI	The equal-weighted average of the Herfindahl-Hirschman Index (HHI) across industries at each monthly *t*. The HHI is computed based on firm sales. Industries are identified using 3-digit SIC codes. This measure is standardized across the sample for ease of interpretability.
Mean Amihud Illiquidity	For each stock in each month, we compute the Amihud (2002) illiquidity measure using daily data. We require at least 10 trading days in a month. We then average this measure across all stocks in that month.
Diversity of firm characteristics	For each firm, we standardize each of the 205 cross-sectional characteristics across the entire time series of that predictor within that firm. We then compute the cross-sectional average at month *t* of each standardized predictor across all firms. We then sum all the averaged standardized characteristics available at each month. Finally, we move the measure 36 months back in time to account for the delayed introduction of characteristics during the first 3 years from which a firm first appears in the data. The final measure is standardized across the sample for ease of interpretability.