Noisy Factors

Finance Working Paper N° 920/2023

June 2023

Pat Akey
University of Toronto and ECGI

Adriana Z. Robertson
University of Chicago and ECGI

Mikhail Simutin
University of Toronto

ECGI Working Paper Series in Finance

Noisy Factors

Working Paper N° 920/2023

June 2023

Pat Akey
Adriana Z. Robertson
Mikhail Simutin

We thank Oliver Boguth, Alon Brav, Andrew Chen, Zhi Da, Bob Dittmar, Winston Dou, Jill Fisch, Cam Harvey, Marcel Kahan, Raymond Kan, Jonathan Klick, Tim Kroencke, Genevieve Lakier, Saul Levmore, Marina Niessner, Ella Patelli, Randy Picker, Elizabeth Pollman, Tom Sargent, Gil Segal, Jinfei Sheng, Holger Spamann, Andrea Tamoni, Karaml Todorov, Michael Weber, David Weisbach, Luigi Zingales, and seminar participants at Arizona State Univer- sity, Chapman University, Chicago Quantitative Alliance Spring Conference, Columbia University, ESCP Paris, ESSEC, the European Winter Finance Conference, the Financial Research Association meetings, FIRS, the Foundations of Law and Finance Seminar Series, Harvard University, HEC Paris, Humboldt University Berlin, International Centre for Pension Management Discussion Forum, Iowa State University, Lancaster Univer- sity, London Business School, MFA, NFA, Rice University, SFS Cavalcade North America, Smokey Mountain Finance Conference, Tel Aviv University Finance Conference, Universite Laval, Universite Paris-Dauphine, University of Alberta, University of Bristol, University of British Columbia Summer Conference, University of Chicago, University of Colorado Boulder, University of Dayton, University of Exeter, University of Illinois at Urbana-Champaign, University of Luxembourg, University of Pennsylvania Law/Wharton Finance seminar series, and University of Toronto for helpful comments. Levi Haas provided exceptional research assistance.

Abstract

The Fama-French factors are ubiquitous in empirical finance. We find that factor returns differ substantially depending on when the data were downloaded, and only a small portion of these retroactive changes is explained by revisions to the underlying data. We show that these changes have large effects in two widely-studied contexts: mutual fund performance and cross-sectional equity pricing. Model evaluation tests suggest that more recent vintages do not perform better. Our findings have significant implications for the integrity of finance research and underscore the importance of understanding the provenance of third-party data.

Keywords: Fama French factors, asset pricing, performance evaluation, equities, mutual funds, model fit

JEL Classifications: G10, G12, G14, G20, G31

Pat Akey
Associate Professor of Finance
University of Toronto
105 St George Street
Toronto, Ontario M5S 3E6 M5S1S4, Canada
phone:
e-mail: pat.akey@rotman.utoronto.ca

Adriana Z. Robertson
Donald N. Pritzker Professor of BusinessLaw
University of Chicago
1111 E. 60th Street
Chicago, IL, 60637, USA
phone: 773-702-9494
e-mail: arobertson@uchicago.edu

Mikhail Simutin*
Associate Professor of Finance
University of Toronto
105 St. George Street
Toronto, ON M5S 3E6, Canada
e-mail: mikhail.simutin@rotman.utoronto.ca

*Corresponding Author

Noisy Factors

Pat Akey, Adriana Z. Robertson, and Mikhail Simutin*

June 15, 2023

ABSTRACT

The Fama-French factors are ubiquitous in empirical finance. We find that factor returns differ substantially depending on when the data were downloaded, and only a small portion of these retroactive changes is explained by revisions to the underlying data. We show that these changes have large effects in two widely-studied contexts: mutual fund performance and cross-sectional equity pricing. Model evaluation tests suggest that more recent vintages do not perform better. Our findings have significant implications for the integrity of finance research and underscore the importance of understanding the provenance of third-party data.

*Akey and Simutin are at the University of Toronto. Robertson is at the University of Chicago. Akey and Robertson are also ECGI Research Members. Akey: pat.akey@rotman.utoronto.ca; Robertson: arobertson@uchicago.edu; Simutin: mikhail.simutin@rotman.utoronto.ca. We thank Oliver Boguth, Alon Brav, Andrew Chen, Zhi Da, Bob Dittmar, Winston Dou, Jill Fisch, Cam Harvey, Marcel Kahan, Raymond Kan, Jonathan Klick, Tim Kroencke, Genevieve Lakier, Saul Levmore, Marina Niessner, Ella Patelli, Randy Picker, Elizabeth Pollman, Tom Sargent, Gil Segal, Jinfei Sheng, Holger Spamann, Andrea Tamoni, Karamfil Todorov, Michael Weber, David Weisbach, Luigi Zingales, and seminar participants at Arizona State University, Chapman University, Chicago Quantitative Alliance Spring Conference, Columbia University, ESCP Paris, ESSEC, the European Winter Finance Conference, the Financial Research Association meetings, FIRS, the Foundations of Law and Finance Seminar Series, Harvard University, HEC Paris, Humboldt University Berlin, International Centre for Pension Management Discussion Forum, Iowa State University, Lancaster University, London Business School, MFA, NFA, Rice University, SFS Cavalcade North America, Smokey Mountain Finance Conference, Tel Aviv University Finance Conference, Université Laval, Université Paris-Dauphine, University of Alberta, University of Bristol, University of British Columbia Summer Conference, University of Chicago, University of Colorado Boulder, University of Dayton, University of Exeter, University of Illinois at Urbana-Champaign, University of Luxembourg, University of Pennsylvania Law/Wharton Finance seminar series, and University of Toronto for helpful comments. Levi Haas provided exceptional research assistance.

In one of the most cited papers in financial economics, Fama and French (1993) propose that sensitivities to returns on three long-short portfolios—the excess return on the market, value minus growth stocks (HML), and small minus large stocks (SMB)—explain the cross-section of expected stock returns. This three-factor model has revolutionized financial economics, becoming the go-to model for empirical researchers. In asset pricing, it is used to measure factor-adjusted returns on mutual funds, stocks, and other investments. In corporate finance and accounting, the model is widely used in event studies and cost of capital calculations. It is taught to PhD, MBA and undergraduate students, and is a part of the CFA curriculum. The model has also had a tremendous impact in practice, where it is used to evaluate real and financial investment decisions, as well as in legal settings to establish liability and to estimate damages.

To apply the model, researchers begin with factor returns. While they can construct their own, researchers overwhelmingly rely on factors from Kenneth French's academic website, which are also provided through the Wharton Research Data Service (WRDS). French's website chronicles occasional revisions to the construction methodology. Until late 2022, only the most recent factor vintage was available after each update. In noting that the data change, WRDS explains that "Research Portfolios incorporate any revisions in the historical underlying data, and thus computations that use the most recent vintage ... may differ from computations that use an earlier vintage. The revisions are typically very small and this set is most commonly used in academic studies."[1] Notwithstanding the reassuring tone of the last sentence, we show in this paper that changes to factor returns are frequent, often substantial, and impact conclusions about first-order questions in finance.

We use archived vintages of French's data going back to 2005, selecting the June vintage from each year.[2] We find that the differences in factor returns are substantial even between adjacent vintage years and tend to increase with the length of time between

[1] `https://wrds-www.wharton.upenn.edu/pages/support/manuals-and-overviews/fama-french/fama-french-research-portfolios-and-factors`

[2] A previous version of this paper used archived versions of French's website obtained from the Wayback Machine. In November 2022, French's website was updated to provide vintages of the factors beginning from the start of our initial sample period.

vintages. While there are large changes in all the factors, the revisions are particularly pronounced for HML. For example, comparing monthly HML returns between the 2005 and 2006 vintages, 98% of the observations are different, and the average absolute difference exceeds 2.5% annualized. Means are also affected: the average HML return is 0.57% per year higher in the 2022 vintage than the 2005 vintage, a difference that is both statistically significant and economically large. In making comparisons such as this, we hold the sample period fixed and base the analysis on data common to both factor vintages (in this case, through June 2005).

The bulk of these changes cannot be explained by updates to the underlying raw data. We construct our own versions of the factors by running the same code on archived versions of CRSP, Compustat and the CRSP-Compustat linking file corresponding to the data that would have been available to a researcher on a given date. We compare changes in these fixed-code factors to changes in the archived versions of French's factors. Changes due solely to updates in the underlying CRSP-Compustat data—as measured by changes in the fixed-code factors—explain only a portion of the changes in French's factors. There is substantial variation in this portion over the sample period: In the first half (1926-1964), data updates explain approximately a half of the variation in changes to French's factors. Thereafter, they are unrelated to the changes. We interpret these changes as being the result of revisions to the construction methodology of French's factors.

These methodological changes increase the Sharpe ratio of the HML factor relative to the fixed code factors. For example, using data beginning in 1993—the year of the publication of the Fama-French three-factor model paper—vintage updates to French's HML factor increased its monthly Sharpe Ratio by 0.010, while the fixed code-based Sharpe ratio of HML saw a decline of 0.003. Relative to the 1993-2021 Sharpe ratio of French's HML of 0.037, these changes correspond to an increase of 27% and a decline of 7%, respectively. The opposite was true for SMB. French's SMB factor saw a decline of 30%, while vintage updates had no tangible effect on the Sharpe ratio of the fixed code

SMB factor. The large majority of these changes occurred in updates between 2017 and 2022. Notwithstanding these changes, we find no evidence of consistent changes to the model's performance in pricing risk.

We then turn to the effect of vintage updates in two widely studied settings: actively managed equity mutual funds and the cross-section of stock returns. We begin with mutual funds because their performance and risk—topics of first order interest to academics, practitioners, and individual investors—are routinely quantified using the three-factor model. We find that mutual fund alphas vary dramatically across vintages. For example, switching between 2005 and 2022 vintages causes almost half of annual alpha estimates to change by more than 1%. Moreover, the effects are present across funds with different styles and, if anything, are more pronounced for larger funds. Remarkably, the choice of factor vintage also affects estimates of the average alpha for the industry as a whole: In some years, switching vintages changes the average annual alpha by more than 1%. Mutual fund betas are also affected. For example, while the mean difference in HML loadings across vintages is small, about a third of loadings change by more than 0.1.

Next, we investigate how the choice of factor vintage impacts inferences about the cross-section of stock returns. Estimates of individual stock alphas and betas vary dramatically, both in level and significance, depending on the factor vintage used. Switching between 2005 and 2022 vintages causes more than a quarter (26%) of alphas estimated from three-factor regressions on five years of monthly data to change by more than 1% per year. Estimated loadings on the three factors change by more than 0.1 for between a tenth and a quarter of observations, suggesting important implications for estimates of cost of capital that use betas as inputs. These effects are even more pronounced using shorter (3- or 1-year) estimation periods, and are pervasive across stocks with different characteristics. Unsurprisingly given the results for mutual funds, vintage changes also affect alphas of portfolios sorted on market size, book-to-market ratio, return runups, and other attributes.

In the final set of tests focused on the cross-section of stock returns, we examine performance of "anomalies," or investment strategies thought to generate significant factor-adjusted returns. We obtain returns of 549 high-low anomaly portfolios from four different data sources: the Global-q Data Library (Hou et al., 2020, 2021), the Equity Anomaly Data (Haddad et al., 2020, Giglio et al., 2021), the Open Source Asset Pricing database (Chen and Zimmermann, 2021), and the 100 anomalies from Dong et al. (2022). About 20% of the 53 marginal anomalies in this set—those for which the t-statistic of the unconditional alpha, using at least one of the 18 factor vintages, is between 2.0 and 2.5—lose statistical significance due to changes in factor vintages.

To the extent that the changes to French's factors reduce the noise in approximating the true unobservable factors, updated factor vintages should represent improvements relative to their predecessors. We treat each vintage as a separate "model" and compare their performance using the tests developed by Barillas et al. (2020). The results of these tests, which use factor data as the only input, are inconclusive. While we find nothing to suggest that newer factors perform worse than older ones, we also do not find consistent evidence that they are improving. In the Internet Appendix, we adopt the performance metric from Fama and French (1993) and Fama and French (2015) and use GRS tests (Gibbons et al., 1989) to evaluate the performance across vintage updates using two standard sets of test portfolios. These tests also provide no evidence that the model is improving as a result of the updates. Collectively, our model performance tests suggest that no particular factor vintage dominates the others, nor that the factors are improving over time in their ability to explain returns of stock portfolios.

Taken together, our results suggest that a wide range of commonly studied quantities in finance are sensitive to changes in factors. These changes are substantial, and their impact is far-reaching: estimates of risk and factor-adjusted performance of mutual funds, stocks, characteristic-sorted portfolios, and anomaly portfolios can change significantly when vintages change. Coupled with the ubiquity of the three factors in finance, our results have significant implications for the replicability, robustness, and integrity of finance research. They suggest that some findings may fail to replicate solely because of changes to the factors. While our focus is on the three-factor model, the findings extend to other models that use the market, HML, and SMB factors as inputs, including the four-, five-, and six-factor models that add momentum, profitability, and investment (Carhart, 1997, Fama and French, 2015).

Our results also have substantial real-world implications. Firms commonly use multifactor models such as the three-factor model to estimate their cost of capital (Graham and Harvey, 2001). Because cost of capital estimates can vary significantly across factor vintages, factor updates may contribute to misallocation of capital. Large institutional investors often use the factors for performance evaluation. Finally, the factors are routinely used by expert witnesses in legal settings, both to determine liability and to assess damages.

Based on our findings, we make several recommendations. At a minimum, researchers should facilitate replication by disclosing which factor vintages they use, and evaluate the robustness of their results to different vintages. More fundamentally, our results cast doubt on the reliability of the standard Fama-French factors. While changes to the factors due to data updates are understandable (and perhaps desirable), the fact that methodological changes substantially affect estimates undermines the value of the factors for empirical research. To circumvent this problem, we provide code to generate versions of the factors using a fixed methodology. We invite researchers to use it to construct methodologically transparent, arm's length versions of the Fama-French factors. Alternatively, researchers may consider using diversified index funds and ETFs or other traded portfolios to construct factors, as suggested by Berk and Van Binsbergen (2015).

Researchers who rely on data provided by third parties, including other researchers, should be cognizant of how those data are created and updated. As our findings suggest, using data whose creation is not transparently documented and for which underlying code is not provided can affect inference and replicability. This risk is exacerbated when combined with the risk of academic capture highlighted by Zingales (2013).

Our paper connects to several strands of literature. Our first contribution is to the literature highlighting problems with commonly-used databases in financial economics. For example, Ljungqvist et al. (2009), Patton et al. (2015), Gillan et al. (2018), and Berg et al. (2020) provide evidence that retroactive changes to the I/B/E/S, hedge fund, ExecuComp, and Refinitiv ESG databases, respectively, can change conclusions of research conducted on previous versions of the data.[3] Moreover, researchers have documented that revisions to macroeconomic data are frequent and can impact the conclusions of empirical studies (e.g., Mankiw et al., 1984, Croushore and Stark, 2003, Chang and Li, 2018). In the same spirit, we show that the qualitative and quantitative conclusions of research questions in mutual funds, equity pricing, and corporate finance that rely on the Fama-French factors can change depending on when the factor data used in the analysis were downloaded.

We also add to the ongoing debate over the "replication crisis" in empirical finance. Harvey et al. (2016), Harvey (2017), Chordia et al. (2020) indicate that p-hacking is pervasive in empirical financial economics; Hou et al. (2020) and Linnainmaa and Roberts (2018) suggest that a large number of asset pricing anomalies fail to replicate or are due to data snooping. In contrast, Chen (2020) argues that p-hacking alone cannot explain the large number of asset pricing anomalies that have been identified, and other authors find that many strategies do replicate, although there is evidence that alphas of these trading strategies decay over time (McLean and Pontiff, 2016, Pénasse, 2020, Chen and Zimmermann, 2021, Jensen et al., 2021).[4] Our analyses point to an additional challenge in replicating past studies.

Finally, our study contributes to the growing literature that evaluates the empiri-

[3]For another example, see Spamann (2010), which points out pervasive errors in, and then corrects, a standard dataset. The literature pointing out issues in commonly used financial data is vast. Some of it is summarized on `https://www.kellogg.northwestern.edu/rs/services/computationalconsulting/trainingandreference/database_biases_and_errors.aspx`. See also Evans (2010), Aiken et al. (2013), Karpoff et al. (2014), Heider and Ljungqvist (2015), Schwarz and Potter (2016), Freyberger et al. (2021), and Bryzgalova et al. (2022).

[4]More broadly, recent work has investigated the replicability of research in various fields of economics (e.g., McCullough and Vinod, 2003, McCullough et al., 2006, Glandon, 2010, Chang et al., 2022).

cal practices in financial economics, law, and accounting. A number of recent papers summarize current empirical practices in the field (e.g., Bowen et al., 2017) or provide guidance on best practices (Atanasov and Black, 2016, 2021, Fisch et al., 2017, Fisch and Gelbach, 2021, Harvey et al., 2020, Harvey and Liu, 2021, Heath et al., 2020, Mitton, 2020a,b, Spamann, 2019). Several papers discuss the problem of measurement error in various empirical contexts (e.g., Erickson and Whited, 2000, 2012, Jennings et al., 2020, Pancost and Schaller, 2022). While not noise in the statistical sense, our paper adds to this literature by identifying a previously unappreciated source of variability in empirical results.

I. The (Noisy) Factor Data

We obtain the current and archived versions of the monthly market, value, and size factors from French's website (Fama and French, 1993).[5] These data are widely used by researchers and the current version is also available through WRDS. We likewise obtain archived versions of the five factors of Fama and French (2015). French's website also provides returns on the Carhart (1997) momentum factor, daily factor returns, and a variety of characteristic-sorted and industry portfolios. Because only current versions of this latter set of returns are available on French's website, we obtain historical vintages using the Internet Archive, a non-profit digital library. One of its services, the Wayback Machine, allows users to access archived versions of over 800 billion web pages. We use this service to obtain vintages of the daily factor data and historical vintages of returns of characteristic-sorted and industry portfolios.

For each set of factors or portfolios, we retain a single time series per year. For monthly factors archived on French's website, we keep the vintages with data through the end of June of each year. For portfolios downloaded from the Wayback Machine, we select the one that is closest to the end of June. Vintage availability varies across portfolios, but for the three factors vintages are available for every year between 2005

[5] `https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`

and 2022. While the three factors were available for a number of years prior to that, earlier vintages are not available on French's website nor via the Wayback Machine.

When we compare vintages from different years, we restrict the analysis to the sample period that is common to both vintages. For example, when comparing two vintages containing data through 2005 and 2022, the sample ends in 2005. In all comparisons, we hold everything other than the vintages constant unless otherwise indicated.

A. Factor Differences Across Vintages

We begin by exploring the extent to which the factors vary across vintages. Each panel of Figure 1 compares the earliest vintages of a particular factor to the latest. The solid black line shows the monthly difference in the realized return of the factor between the two vintages. The blue dash-dotted and red dashed lines plot the cumulative returns of each vintage.[6] We also report means and standard deviations of the two vintages and their difference in the top left of each panel.

Panel A presents the results for the market factor. This factor will change across vintages only when the definition of what constitutes the market or the risk-free asset changes, or to the extent that historical stock returns or market capitalizations are revised (perhaps to correct errors in the underlying data). While the average difference between the 2005 and 2022 vintages is small (1 bp per month), the mean *absolute* difference is considerably larger (10 bps). The absolute value of the difference between the two vintages is at least a quarter of a percent in 74 months.

We observe much larger differences across vintages for the HML (value) factor, presented in panel B. The average return in the 2022 vintage is about 10% larger than in the 2005 vintage (45 vs 41 bps), a difference that is both statistically significant (t=2.16) and economically important, producing much larger cumulative returns over the sample. Monthly return differences across the two vintages frequently exceed 1% and are particularly substantial in the beginning (1920-40s) and near the end (1990s-

[6]The results are qualitatively similar using daily data.

2000s) of the sample. The volatility of the difference is large: at 0.67%, the variation in the *difference* between HML factor vintages is more than one sixth (0.67/3.57) of the magnitude of the *total* variation in the HML factor.

We also find non-trivial differences in the SMB (size) factor across vintages. While the mean difference is close to zero, we observe substantial absolute differences (22 bps on average) over time, which are particularly large early and late in the sample. The standard deviation of the difference between the vintages is also large, representing 12% of the standard deviation of either of the vintages.

Turning to the remaining factors, Panel D shows that differences in realizations of the UMD (momentum) factor are particularly large in the first decades of the sample, frequently exceeding 100 bps per month. Panel E shows that the RMW (profitability) factor, whose first vintage dates to 2015, exhibits large absolute differences in returns, particularly since the 1990s. Here, the variation in the differences between the vintages represents more than 18% of the variation in the factor. Finally, Panel F shows that the differences between the earliest (2015) and latest (2022) vintages of the CMA (investment) factor are very small. In the remainder of the paper, we focus our analyses on the three Fama and French (1993) factors.

We present the differences across all pairs of vintages in Table I. The upper triangular entries show the results using monthly data, while the lower triangular use daily data. Each pairwise comparison uses the data that is available in both vintages. As a result, the time series is longer when comparing two later vintages.[7]

Several features stand out from Table I. First, the differences in factor returns across vintages are substantial, even when comparing two vintages that are close in time. For example, the average absolute difference in HML factor returns between the 2005 and 2006 vintages is 21 bps per month. Even between these adjacent vintages, over two thirds of monthly return differences exceed 1% in annualized magnitude, and only 2 percent of the reported returns are identical. Second, the absolute magnitude of the differences tends to increase with the time between vintages. For example, comparing

[7] In some years, vintages from the Wayback Machine are not available for daily data.

the 2005 vintage with the 2022 vintage, 77 percent of HML observations differ by more than 1% annualized, and only 1 percent are the same. Third, the differences tend to be largest for the HML factor, although all three are affected. Finally, differences are large in both monthly and daily data.

B. Changes to Data and Code

We have shown that the returns of the Fama-French factors vary substantially depending on when they were downloaded from French's website. We now investigate how much of the variation in factor returns is due to changes in the raw data used to construct them and how much comes from changes to their construction. Because French's website does not provide the exact code used to construct the Fama-French factors, we write our own code following the description in Davis et al. (2000), which is paralleled on the websites of both French and WRDS.[8] We run this code on historical vintages of raw CRSP and Compustat data to create fixed-code factors. Because the only difference between the vintages of the fixed-code factors is the underlying data, this allows us to isolate the changes to the posted Fama-French factors (hereinafter "the French factors") that can be explained by changes to the data. The earliest vintages of Compustat, CRSP, and the CRSP/Compustat Merged Database Linking Table that we are able to obtain are from 2010. We also obtain the necessary raw CRSP/Compustat data vintages from 2011, 2016, 2020, 2021, and 2022. We document changes in the underlying CRSP, Compustat and linking table data in Section IA.B of the Internet Appendix.

For each of the three factors and each of the six vintages we can compare, the correlation between the French factors and the fixed-code factors is at least 0.99. In each of the vintages, the French version of both the HML and SMB factors have higher mean returns than their fixed-code analogues. For example, for the 2022 vintages (which span the period through the end of 2021), these differences exceed 1 bp monthly

[8] See http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/variable_definitions.html and https://wrds-www.wharton.upenn.edu/pages/support/manuals-and-overviews/fama-french/fama-french-research-portfolios-and-factors.

for both HML and SMB. For HML, the difference between the French factor and the fixed-code factor is positively skewed for all vintages, with an average skewness of 1.13.

We begin by examining the correlation between (a) the changes in the French factors between the 2010 and 2022 vintages and (b) the changes in the fixed-code factors between the 2010 and 2022 vintages. Figure 2 presents the results of this analysis. The black lines of Panels A, C, and E present the time series changes in the French market, HML, and SMB factor returns, while the gray line presents the analogues changes using our fixed-code factors.[9] Two patterns are apparent. First, the variation in the French factors is substantially higher than the variation in the fixed-code factors: for the French factors, the standard deviations of the changes between vintages of the market, HML, and SMB factors are 0.16, 0.60, and 0.39, respectively, compared to 0.07, 0.38, and 0.26 for the fixed-code factors, respectively. Second, the changes in the fixed-code factors are fairly highly correlated with changes in the French factors in the early part of the sample. This correlation is much smaller in the later part of the sample. For example, splitting the sample at the end of 1964 (a common starting point for empirical studies using Compustat data), the correlations between the changes in the French factors and the fixed-code factors are 0.70, 0.65, 0.66 for the market, HML, and SMB factors in the earlier part of the sample but 0.12, -0.03, and -0.01 in the later part of the sample. These correlations imply that changes to the underlying data explain 42–49% of the changes in the first part of the time series, but nearly none of the changes in the latter part of the time series.[10]

[9]We focus our discussion on the differences between the earliest and latest vintages of the French factors and the fixed-code factors available to us. The vintage-by-vintage comparisons are available in the Internet Appendix. See Figures IA.4, IA.5, and IA.6. For each available vintage-by-vintage comparison, we also reproduce the narrative description of the changes in methodology provided on French's website. We summarize the description of changes to the construction of the factors from French's website in Table IA.IV.

[10]The R-squared of a univariate regression is the square of the correlation between the dependent and independent variables. To investigate the source of the changes to the fixed-code factors, we examine how frequently the variables used to construct the factors have changed. We direct interested readers to section IA.B in the Internet Appendix. We also explore the effects of the changes to French's factors on their cumulative returns, presented in Panels B, D, and F of Figure 2, in section IA.C in the Internet Appendix.

C. Methodological Changes and Factor Performance

The results above suggest that changes in the underlying data are responsible for only some of the changes across the vintages of French's factors, and are particularly unimportant in the second half of the sample. Methodological changes must therefore be an important driver of changes across vintages. French's website provides a parsimonious description of certain changes to the factor construction methodology. While this is helpful, the motivation for these changes is not always clear, nor is it obvious how the changes will affect factor performance. Given this ambiguity, we next investigate the effect of methodological changes on factor performance. One hypothesis is that these changes represent efforts to improve the performance of the model. We show in Section IV that more recent factor vintages are not more successful at pricing risk, as measured by standard asset pricing tests. We do, however, find evidence that the changes have led to an improvement in the performance of the value factor.

To investigate the effect of the changes in the construction of the factors on their performance, we compute the change in Sharpe ratios between adjacent vintages of each of French's factors. We split the sample in 1993, which corresponds both to the publication of their canonical paper and roughly with the second mass of substantial changes in the returns of HML and SMB. Panel A reports the results using the data from 1993 onward; Panel B reports the results for the 1964–1992 period.

For each factor, and each pair of adjacent vintages, we compute the difference in Sharpe ratios between the two vintages using the sample period common to both. We repeat this process using the fixed code factors. This allows us to isolate changes in the French factors that cannot be explained by changes to the raw data. Because 2010 is the earliest vintage of the fixed code factors we can construct, the first vintage with which we can make this comparison is 2010.

For example, in Panel A, the 2010 factor vintage covers the sample from January 1993 to December 2009. To evaluate how the Sharpe ratio of the French HML factor—computed using post-1993 data—changes between the 2010 and 2011 vintages, we compute two values for its Sharpe ratio, one from each vintage, using the same 1993-2009 sample period. The resulting Sharpe ratios are 0.10625 and 0.10435 for the 2010 and 2011 vintages, respectively, indicating that the changes between the 2010 and 2011 vintages caused the Sharpe ratio of the French HML factor to decrease by 0.0019. This change is represented by the solid gray line in Panel A of Figure 3 decreasing to -0.0019 in 2011. We repeat this process using the fixed-code HML factor and plot the corresponding changes in its Sharpe ratio with the solid black line.[11] The changes in Sharpe ratios of the French and fixed-code market factors are plotted in light blue and dark blue dashed lines, respectively; analogues for SMB are plotted in pink and red.

Between 2010 and 2016, all six of the lines in Panel A of Figure 3 are relatively flat, indicating that the vintage updates have little effect on the cumulative post-publication Sharpe ratio. Beginning in 2017, however, the French HML and SMB factors begin to diverge. French's HML performs substantially better across the remaining updates, although the changes are not monotonic. The fixed-code HML factor, in contrast, remains very flat, drifting slightly down at the end of the sample period. The opposite pattern emerges for SMB: the performance of French's SMB factor deteriorates substantially over the same time period. As with HML, the fixed-code SMB factor shows little change. The fact that the changes to French's HML and SMB factors occur at the same time is unsurprising given the double sort methodology used to construct them. However, such dependency does not necessarily mean that the changes should roughly mirror each other. Vintage updates have negligible effects on the Sharpe ratios of the two versions of the market factor.

A very different pattern emerges in the pre-1993 sample period. All six of the lines in Panel B of Figure 3 are substantially flatter. The Sharpe ratios on both versions of SMB are essentially flat across vintages. The Sharpe ratio on French's market factor increases in 2013 (corresponding to the redefinition of the "market" described on French's website), while the fixed code market factor remains flat. Both versions of the

[11] For the reasons described above, we are only able to compute the fixed code factors for six years during the sample period. While this may mask variability, it will not affect the cumulative changes in Sharpe Ratios.

HML factor deteriorate modestly over the course of the vintage updates.

While not the focus of the analysis in the remainder of the paper, we repeat the analysis using vintages of French's Developed ex-US factors and present the results in Figure 4.[12] Like the US versions of these factors, the changes across vintages of these factors tend to improve the performance of the HML factor while causing the performance of SMB to deteriorate. The 2015 vintage ends in September and the 2016 ends in February, which may explain why all three lines are flat between those years. Data limitations and lack of documentation mean that we cannot create fixed code versions of these factors. However, we note that by construction, the data used to create these factors does not overlap with the data used in the US factors. Accordingly, similar patterns in the two geographically distinct sets of factors are unlikely to be mechanical.

Taken together, these results indicate that the retroactive changes to the factors have led to an apparent improvement in the performance of the value factor. Consequently, a value-based investment strategy may appear more attractive using more recent vintages of French's factors. Many asset managers employ such a strategy, including Dimensional Fund Advisors (DFA), which occasionally refers to French's factor data in its publicly available marketing and educational materials.[13] Both Eugene Fama and Kenneth French have longstanding and clearly disclosed associations with DFA. The relationship between DFA and the French factors is less clear.

For example, the factors have been hosted on French's webpage at Dartmouth College's Tuck School of Business since at least 2001. Its source code states that the images and code are property of Ken French, but that the site was "[d]eveloped by Dimensional Fund Advisors Web Team."[14] There is no specific discussion on the website of

[12]These factors are not archived on French's website, so we rely on data from the Wayback Machine for this analysis. The time series begins in 1990, making it impractical to investigate changes prior to 1993.

[13]Historically, DFA was also associated with a size strategy. Today, DFA offers funds targeting both large- and small-cap stocks, as well as all-cap funds. Its large cap value fund is currently substantially larger—in terms of AUM—than its small cap value fund.

[14]This statement in the source code has been there since at least 2001. Prior drafts of this paper—which did not mention DFA—were sent to French in 2021 and 2022. While he confirmed receipt, he had not provided any substantive reply by the time of this writing.

any relationship between DFA and the factors.

The lack of transparency with respect to both the construction and the provenance of French's factors make it impossible to conclusively establish the reasons for the methodological changes we observe. Rather than speculate, we simply note that this lack of transparency, coupled with the pattern of changes to the factors, may be concerning to researchers who rely on the factors for empirical analysis.

II. Mutual Funds

We now turn to the question of whether the changes to French's factors matter in standard empirical applications. Our first application is the measured performance and risk exposure of equity mutual funds. We therefore examine the extent to which mutual fund alpha and beta estimates are sensitive to switching factor vintages. Factor-adjusted fund performance is an important area of research in both the academic literature and in practice. We begin by studying the pooled sample of mutual funds and then examine heterogeneity across mutual fund characteristics.

A. Individual Funds

Our mutual fund return data are from the February 2022 vintage of the CRSP Survivorship Bias-Free Mutual Fund Database. We use returns from actively managed domestic equity mutual funds from 1980 to 2021. We exclude index, sector, and target date funds and group share classes into funds using the MFLINK dataset. For each mutual fund in the sample, we estimate alphas and three-factor beta loadings annually at the end of every calendar year using each factor vintage. We use one year of monthly return data for our baseline analysis but show robustness to using three- and five-year windows. We choose one year in our main specification for four reasons: (i) microstructure noise is less of a concern for diversified mutual funds than it is for individual stocks; (ii) performance horizons as long as five years are not commonly analyzed in mutual fund settings; (iii) time variation in mutual fund betas can bias long-horizon estimates of performance; and (iv) running five-year regressions may introduce a survivorship bias in the mutual fund sample. We winsorize alphas cross-sectionally at the 1st and 99th percentiles.

We estimate the regressions using each factor vintage, so we obtain 18 sets of regression estimates for each fund × year in the sample. We then compute the difference between the parameter estimates obtained using each pair of vintages. Figure 5 plots histograms and kernel densities of the differences in alphas and betas obtained using the earliest (2005) and the most recent (2022) vintages. Panel A shows that, as expected, average net-of-fees alphas are substantially negative. Their magnitude, however, is sensitive to the choice of factor vintage. Average underperformance is 20 bps per year greater when estimated using the 2022 vintage compared to the 2005 vintage. While this average difference is not statistically significant in the full sample (t=1.38), it exhibits substantial variation over time. Figure 6 shows that the average difference exceeds 1% in some calendar years and is below -1% in others. Put differently, inferences about average yearly performance of the overall active equity mutual fund industry can change quite dramatically due to nothing more than switching factor vintages.

Returning to Panel A of Figure 5, switching between the two factor vintages causes almost half of estimated annual alphas to change by more than 100 bps, and 33% of alphas that are statistically significant using one vintage become insignificant using the other. These results further underscore the extent to which mutual fund performance evaluation is sensitive to factor vintage.

Mutual fund factor loadings are often used to assess the risks that funds are exposed to and to investigate the extent to which funds are following their stated strategies (e.g., Sensoy, 2009). Panel B shows the differences in market beta estimates obtained using the two vintages. While the average difference is close to zero, its standard deviation represents a quarter (0.08/0.24) of the cross-sectional standard deviation in market betas, and 12% of mutual funds have loadings on the market that change by more than 0.1.

Panel C shows that the variation in HML loadings is even larger, consistent with the large differences in HML returns in Figure 1. While the mean difference in HML loadings across vintages is small, the standard deviation of that difference is equivalent to more than a quarter (0.16/0.56) of the cross-sectional variation in the loadings, and 33% of loadings change by more than 0.1. The loadings on SMB are somewhat more stable: 9% change by at least 0.1, and the standard deviation of the difference represents about 16% of the cross-sectional standard deviation.

Next, we investigate the impact of estimation horizon and sample period on these effects. In Table II, we estimate alphas using one, three, and five years of data (Panels A, B, C) and partition the full sample into three subperiods: the 1980s, 1990s, and 2000s.[15] Consistent with the dramatic cross-vintage differences in factor returns in the latter part of the sample (Figure 1), the variation in the effect on mutual fund alphas tends to be largest in the latter part of the sample period. This is significant because the later period is also more frequently used in mutual fund studies. While the alpha estimates in the 1980s display some sensitivity to the choice of factor vintages, the results from the 1990s and 2000s are much more dramatic. Fifty-six percent of one-year alphas in the 1990s—and 49% of those in the 2000s—change by more than 100 bps per year. About a third of alphas in these two later periods lose significance.

While alphas are somewhat less sensitive to changes in factor vintages when estimated using longer windows, they continue to have substantial effects. For example, focusing on the 5-year estimation horizons in the 2000s, the volatility of the difference across vintages is about half as large as it was using the one-year horizon, and the proportion of statistically significant alphas that lose significance is lower (19% vs 33%).

In Table III, we compare the differences in annual mutual fund alphas for each pair of factor vintages. We find that the differences generally increase with the time between vintages, but even adjacent vintages can produce substantially different alpha estimates. For example, 33% of alphas change by over 100 bps when we switch between the 2016 and 2017 vintages, although only 6% of significant alphas lose their significance

[15]The 2000s subsample includes the period from 2000 through 2004.

when we switch between the two most recent vintages (2021 and 2022).

B. Fund Characteristics

Next, we ask whether funds with different characteristics are affected differently by changes in factor vintages. We focus on three characteristics commonly used in the literature: fund size (i.e., assets under management), and exposure to two dimensions of style: size and value. We measure the style tilt of each fund using its lagged loadings on the size and value factors, estimated from three-factor regressions on five years of monthly data. We estimate these loadings using every factor vintage and average SMB (HML) loading across vintages to approximate the fund's exposure to size (value) styles.[16]

We sort funds cross-sectionally into quintiles on the basis of assets under management and exposures to size and value factors. For each quintile, Table IV summarizes the differences in annualized alphas calculated using the 2005 and 2022 factor vintages. Since the average alphas discussed above weigh all funds equally, funds that account for only a small proportion of total AUM may disproportionately drive our results. This concern is unwarranted in our context. Panel A shows that all funds, irrespective AUM, are substantially affected by vintage updates. The proportion of alphas that change by more than 100 bps is stable across quintiles, ranging between 48% and 49%. If anything, the largest funds might be slightly more sensitive to changes in factor vintages: the volatility of the difference in alphas from the two vintages as a share of total cross-sectional volatility increases monotonically with fund size: from 15.5% for the smallest quintile to 19.5% for the largest. The share of fund alphas that lose significance also tends to increase with fund size, from 29% to 39%. This is noteworthy for two reasons. First, by definition, the largest funds represent a disproportionate amount of industry AUM. Second, estimates of the alphas of large funds are commonly thought to be less susceptible to noise than those of smaller funds. This is not the case when it comes to

[16] Alternatively, style can be inferred from objective codes on CRSP. In untabulated results, we confirm that our results are not sensitive to defining size and value styles of the fund using Lipper classifications from CRSP.

the effect of vintage updates.

Panel B of Table IV shows that the effect of factor vintages on alphas is not limited to funds with a particular size exposure. Across size exposure quintiles, the standard deviation of the difference in alphas represents between 14% and 21% of the cross-sectional standard deviation in alphas. Funds that are more tilted towards small stocks (high size factor exposures) are more likely to have alphas that change by more than 100 bps (60%), but the percentage of such observations remains large for funds that are tilted towards large stocks (38%) and those in the middle (47%). Consistent with the evidence in Panel A, it is the large-cap funds that are more likely to see their statistically significant alphas become insignificant (37% compared to 33% and 31% for the middle and top quintiles, respectively).

Finally, Panel C shows that funds with deeper growth or value tilts (those in the Low and High quintiles, respectively) are more likely to experience large changes in alphas when the factor vintage changes (56% and 51%, respectively). This is consistent with the value factor experiencing the most dramatic changes across vintages, in turn causing the alphas of funds with large positive or negative exposure to the factor to change by a large amount. In terms of statistical significance, funds with high value exposure are most affected: 40% and 36% of observations lose statistical significance in the fifth and fourth quintiles, respectively. The mean difference in alphas is monotonically increasing in value exposure: the average alpha for deep value (fifth quintile) funds is 25 bps higher using the 2022 vintage compared to the 2005 vintage. For the deep growth funds, this difference is -75 bps. The opposite pattern appears for size exposure.

Taken together, the results in this section indicate that the choice of factor vintage has a substantial impact on performance evaluation across the mutual fund market. Alphas and betas of individual funds can change dramatically with vintage updates, and large funds are not immune. The effects are present across style tilts, but are particularly pronounced for funds with large tilts to value or growth.

III. Equities

In this section, we turn to the effects of factor vintages on estimates involving individual stocks, standard test portfolios, and published anomaly portfolios. We start by examining the extent to which alpha and beta estimates of individual stocks are sensitive to switching factor vintages. While single-stock estimates are known to be noisy, they are used in a variety of asset pricing, corporate finance and legal applications, including firm valuation and event studies. We then turn to characteristic-sorted and industry portfolios. Even in this context, switching factor vintages can substantially change alpha and beta estimates. Finally, we turn to published long-short anomaly strategies. The impact of the changes in the factors is weakest in this setting. Notwithstanding this, we show that statistical inference about the unconditional alphas of some of these strategies is sensitive to the choice of factor vintage.

A. Individual Stocks

Our tests in this section use the same structure as those in Section II. For each common stock in CRSP listed on the NYSE, Amex, or Nasdaq, we estimate alphas and betas from three-factor regressions at the end of each calendar year using rolling five-year windows. We use monthly return data and require a minimum of 36 monthly observations for a stock $\times$ year to be included in the sample. Alphas are winsorized cross-sectionally at the 1st and 99th percentiles.

As in Section II, we compute the difference between the parameter estimates obtained using each pair of vintages, this time for each stock $\times$ year. Figure 7 plots the histograms and kernel densities for the resulting differences in alpha and beta estimates obtained using the earliest (2005) and the most recent (2022) factor vintages. For convenience, the figure also shows summary statistics in the upper left of each plot.

Panel A shows that the choice of vintages has a large impact on estimated alphas. The average difference in annualized alphas is 29 bps per year, which is both economically meaningful and statistically significant (t=3.8).[17] For more than a quarter of observations, the choice of factor vintage changes the estimated alpha by more than 100 bps.

While individual stock alphas are not widely used, *betas* of individual stocks are used in a variety of settings. Event studies—widely used in the corporate finance literature—commonly rely on individual stock betas to construct abnormal returns (e.g. MacKinlay (1997), Kothari and Warner (2007)). They are also used to estimate the cost of capital for firm valuation, and are therefore of interest in their own right. Panel B of Figure 7 shows that switching between the two vintages causes 12% of estimated market betas to change by more than 0.1. Assuming a market risk premium of about 5% per year, this implies that the choice of factor vintage can generate a difference of 50 basis points per year in the discount rate. The standard deviation of the difference—0.08—is large, amounting to about 12% of the standard deviation of betas estimates.

Consistent with the earlier evidence showing the substantial cross-vintage variation in the HML factor, Panel C of Figure 7 shows that switching between the 2005 and 2022 vintages causes more than a quarter of HML loadings to change by more than 0.1, and the standard deviation of the difference corresponds to 13% of the of the standard deviation of the estimated loadings. The magnitudes for SMB, shown in Panel D, are roughly similar to those for market betas.[18]

B. Characteristic and Industry Sorted Portfolios

We now turn to the question of how differences in factor vintages affect inferences about alphas of commonly studied portfolios. Specifically, we consider value-weighted decile portfolios from Ken French's website that are sorted on 10 attributes: market eq-

[17] A positive mean alpha (using either vintage) reflects the fact that smaller stocks are known to have higher average three-factor alphas. In the pooled mean, all stocks receive the same weight.

[18] We use 2022 CRSP stock return data in these analyses to isolate the impact of the change in French's factors. To ensure that our results are not driven by changes in the underlying CRSP stock return data, we repeat the analyses using contemporaneous CRSP return data and French factor data. Because of data limitations, 2006 is the earliest year for which we can compute contemporaneous alphas and betas. The results are extremely similar to those presented in Figure 7, indicating that changes in CRSP are not driving the results. We direct interested readers to Figure IA.1 in the Internet Appendix. Because of this similarity, we present the results in the remainder of the paper using only 2022 return data.

uity, book-to-market ratio, profitability, investment, accruals, net issuance, momentum, market beta, variance, and residual variance.

Using characteristic-sorted portfolios introduces a second dimension of vintages: not only do factor returns change across vintages, so do the returns of the portfolios themselves. To keep the analysis focused, we compare the earliest vintage of both the factors and the portfolios to the latest vintage of each. As before, we compute alphas at the end of every calendar year using five years of monthly data.

We present the results in Table V. While there is considerable variation across the ten sets of portfolios, vintage updates have a substantial effect on all ten sets of alphas. Portfolios sorted by book-to-market and profitability are most affected, with approximately a third of alpha estimates changing by more than 1% per year (32% and 34%, respectively) and over a half of significant alphas losing significance (52% and 62%, respectively). In untabulated results, we find that some estimates change by as much as 4% per year. For the book-to-market and profitability portfolios, the standard deviation of the difference in alphas amounts to over 50% of the cross-sectional standard deviation of the respective alphas. These are dramatic effects in widely used diversified portfolios.

The effects are also economically meaningful in the other portfolios. For all characteristics, changes in vintages cause alphas of at least 10% of the portfolios to change by at least 100 bps, with a median of 22%; the median the standard deviation of the difference in alphas represents approximately 34% of the cross-sectional standard deviation. These results are all the more striking given that for some portfolios, the earliest available vintages are from 2015, meaning that some comparisons involve comparing vintages from 2015 and 2022.

We also consider unconditional alphas, estimated over the full sample period rather than in five-year windows. Rather than overwhelm the reader with another set of results from a broad set of portfolios, we focus on two commonly used sets of portfolios: 25 portfolios sorted on size and book-to-market, and 17 industry portfolios. For each portfolio, we estimate 216 alphas (18 factor vintages × 12 portfolio vintages), which we round to one tenth of one percent. Figure 8 presents these alphas visually, with the size of the bubble representing the frequency with which a particular rounded alpha estimate occurs within the 216 estimates.

For some portfolios, such as S3V2 (corresponding roughly to mid-cap core stocks), oil, and utilities, alphas exhibit little variation across vintages. For many others, the variation is substantial. For example, alpha estimates for the S1V1 portfolio differ by as much as 1.6% per year, from a low of -10.4% to a high of -8.8%. The estimates for the S2V5 portfolio differ by 1.5% (-0.64% to 0.82%). The effects are even more dramatic in some industry portfolios: alpha estimates for durables range between -3.5% and -0.2%, and the estimates for mines straddle zero, varying between -0.4% and 0.9%.

C. Anomaly Portfolios

Finally, we turn to "anomalies," or investment strategies that have been shown to generate significant factor-adjusted returns. There is a large and active literature studying anomalies, with hundreds of apparent anomalies documented over the last three decades. We obtain returns on 549 long-short anomaly portfolios by pooling data from four sources:[19] (1) the 187 anomaly portfolios from Lu Zhang's Global-q Data Library,[20] constructed by Hou et al. (2020) and used in Hou et al. (2021) to test an augmented version of the q-factor model of Hou et al. (2015); (2) the 207 anomalies from Chen and Zimmermann's Open Source Asset Pricing database,[21] described in Chen and Zimmermann (2021); (3) the 55 portfolios from Serhiy Kozak's Equity Anomaly Data webpage,[22] described in Haddad et al. (2020) and Giglio et al. (2021); and (4) the 100 anomalies from Dong et al. (2022).[23]

[19] We downloaded all four sources in May 2022. Pooling four databases in this way inevitably results in some duplication. In order to minimize the effect of discretionary decision-making, we do not attempt to identify or remove duplicates.

[20] http://global-q.org/testingportfolios.html

[21] https://www.openassetpricing.com/data

[22] https://www.serhiykozak.com/data

[23] We obtain the returns on these 100 anomaly portfolios from the Replication Code on the Journal of Finance website: https://onlinelibrary.wiley.com/doi/abs/10.1111/jofi.13099.

We compute the unconditional alpha of each of these 549 long-short anomaly portfolios using each of the 18 factor vintages. We hold the sample period fixed for each anomaly to restrict attention to the effect of the changes in factor vintages. As a result, each regression uses data through June 2005. We then compute the range of the associated t-statistics for each of these 549 anomalies (i.e., the difference between the largest t-statistic and the smallest t-statistic). This provides a measure of the maximum effect, in sample, of the vintage updates on these anomalies. We present the results in Panel A of Table VI. While there is some variation across the four data sources, the results are largely consistent. On average, t-statistics move by up to about 0.15 with a standard deviation of about 0.08. More important than the mean is the range of the range: in all but two instances, the total impact of the vintage updates on anomaly t-statistics is less than 0.5. This suggests that inferences about anomalies associated with large t-statistics are unlikely to be materially affected. In contrast, the statistical significance of more marginal anomalies—those with a t-statistic of between 2.0 and 2.5—is much more likely to depend on factor vintage.

To investigate this, in Panel B of Table VI we restrict attention to the 53 anomalies (representing 10% of our pooled sample) for which at least one of the computed t-statistics is between 2.0 and 2.5. The largest change in t-statistics in this subsample is somewhat smaller (0.42), although the mean change remains 0.15. Because of the nonlinearity of the t-test, a change of 0.15 can have a substantial effect on the significance level of a marginal anomaly. To zoom in on this point further, in Panel C we present the number of these marginal anomalies whose t-statistic falls below 2 using at least one of the factor vintages. This group represents 11 of the 53 marginal anomalies, or 21%. This proportion varies across the four anomaly data sources, from 14% in the Zhou data to 33% in the Kozak data.

There are several implications from this analysis. First, our results suggest that, retrospectively, anomalies with t-statistics greater than 2.5 are unlikely to be materially affected by changes in the French factors. This provides empirical support for Chen and

Zimmermann (2021)'s characterization of anomalies with t-statistics greater than 2.5 as "clearly significant." Conversely, it suggests that anomalies with t-statistics less than 2.5 are at risk from the factor vintage updates. The 21% of marginal anomalies that we document represents a lower bound: these are the anomalies for which switching factor vintage *alone* can render the anomaly insignificant. The other problems (including data snooping and p-hacking) that are widely discussed in the replication debate operate on top of this effect. This provides additional support for a higher t-statistic cutoff, such as the cutoff of 3.0 proposed by Harvey et al. (2016).

Overall, the results in this section show that changes in factor vintages can have substantial effects on inferences about the risk and return of equities in many contexts. Individual stock alphas and betas can vary dramatically depending on which factor vintage is used to estimate them. So do alphas of standard characteristic-sorted portfolios. To a lesser extent, the choice of factor vintages can also impact inferences about the alphas of anomaly portfolios.

IV. Are The Factors Improving?

As discussed in Section I, the French factors change across vintages because of changes to their construction, and, to a lesser extent, updates to the underlying CRSP and/or Compustat data. Assuming the model is correct, if these changes produce better proxies for the true unobservable factors, more recent factor vintages should represent an "improvement" relative to older vintages. Model performance tests typically compare two competing models, such as the CAPM and the three-factor model, each with its own set of factors. Here, we keep the model fixed and instead compare the performance of different vintages of the Fama-French three-factor model. We formally test the relative performance of each vintage using the "BKRS" test of Barillas et al. (2020). In the Internet Appendix, we repeat this analysis using the classic "GRS" test of Gibbons et al. (1989).[24]

[24] While these are not formal model comparison tests, GRS tests are widely used to rank asset pricing models by relative performance. For example, Fama and French (2015) use their finding that "the five-factor model

Barillas et al. (2020) build on Barillas and Shanken (2017) to develop a formal pairwise model comparison test. When comparing two models, the one whose factors produce a higher squared Sharpe ratio dominates. A particularly attractive feature of these tests is that they require only the factors themselves as inputs, and do not rely on test assets.

Table VII presents the results of BKRS tests under the null hypothesis that the squared Sharpe ratios for each pair of factor vintages are equal. Panel A shows differences in squared Sharpe ratios, and Panel B provides the corresponding p-values. We make a few observations. First, all differences of the squared Sharpe ratios in Panel A are non-negative, suggesting that vintage updates do not cause the model's performance to deteriorate.[25] The differences tend to be smallest close to the diagonal, corresponding to comparisons of immediately adjacent or otherwise close vintages, and largest in the top right corner, which compares some of the most recent and oldest vintages. Second, while the differences in squared Sharpe ratios sometimes increase with the time between vintages, these differences are small and non-monotonic. Third, only three of the p-values in Panel B (out of 153) are less than 5%, all of which are in the first few vintage updates; only 25 p-values are less than 10%, the latest of which is from comparing the 2012 and 2013 vintage update. To the extent that the tests provide weak statistical evidence of improvement, this disappears in the later vintage updates. In particular, the large increases in the French HML factor (and corresponding decreases in the French SMB factor) beginning in 2017, do not coincide with any detectable improvement in model performance.[26]

Overall, we find no consistent evidence that one factor vintage is preferable to any other. Any improvements in French's factors across vintages are statistically and eco-

produces lower GRS statistics than the original three-factor model" to establish the former's superiority over the latter. We direct interested readers to section IA.D in the Internet Appendix.

[25]In Table IA.I of the Internet Appendix, we find no significant difference between squared Sharpe ratios of the fixed-code factors and French's factors.

[26]In addition to pairwise tests, we also consider multiple-model comparison tests of Barillas et al. (2020). The null hypothesis of these tests is that none of the other models is superior to the benchmark model. Given that no vintage consistently dominates another economically or statistically in pairwise tests, it is not surprising that no single model emerges as dominant in multiple-model comparison tests. We omit these tests for brevity. We thank Raymond Kan for offering the code for the pairwise and multiple-model tests.

nomically small. We do note, however, that we find no evidence that updates to the factors lead to *worse* model performance.

V. Implications and Recommendations

Until the code is made public, we do not believe that academic finance can justify the continued use of French's factors.To be clear, nothing in this analysis speaks to the validity of the three- (or five-) factor *model*, only to this particular source of factor *data*. Moreover, we do not believe that there is a viable econometric or statistical solution that can salvage French's factors. The evidence does not support the conclusion that the "noise" we document is, or can be reasonably approximated by, classical measurement error. Because the changes appear to be the result of intentional modifications to the code, it is unrealistic to assume that we can predict what changed might look like going forward.

We have shown that the changes to French's factors have large effects on the economic magnitudes and statistical significance of empirical research in finance. Our results have obvious implications for discussions about the state of replicability in financial economics. This is particularly troubling because of the source of the "noise." Even more troubling, the impact of these retroactive changes is not confined to academic research. The Fama-French model is frequently taught as a "gold standard" to undergraduate and MBA students, and standard textbooks explain that the factor data can be obtained from French's website. The model is used to evaluate the performance of mutual funds and therefore affects allocation of investment capital and career outcomes of managers. Firms use the model to calculate their cost of capital in capital budgeting decisions. Single-firm event studies are commonly used by expert witnesses to determine liability and damages in securities litigation, often relying on French's data. These changes across vintage have therefore contaminated analyses in practice as well as in the literature.

The most obvious implication of our findings is that empirical results obtained using one factor vintage may fail to replicate using a different vintage. Researchers attempting replications should be aware that results relying upon the French factors may fail to replicate solely because of revisions to the factors. More generally, our results raise questions about the continued use of intermediate datasets that are not accompanied by the code used to generate them. While changes due to data updates are understandable (and perhaps desirable), methodological changes pose a greater challenge.

As an alternative to the French factors, we make the code used to generate the fixed-code factors discussed in Section I.B freely available.[27] Researchers are invited to download this code and use it—in conjunction with the standard CRSP and Compustat data—to construct the three standard Fama-French factors. We believe that these factors have several appealing qualities. First, they are convenient and easy to use. Second, they are at arm's length from the research question at issue, thereby alleviating concerns that the researcher might strategically choose, or create, factors that strengthen the author's empirical claims. These are, of course, the two most appealing qualities of the French factors. But our factors have two additional qualities that distinguish them. First, because the code is publicly available, all methodological choices are completely transparent. Second, we commit to not changing the construction methodology. To the extent that changes or updates to the methodology may become desirable at some point in the future, we commit to keeping archived versions of all past distributions available in the data archive.

Alternatively, our results also provide support for the approach taken by Berk and Van Binsbergen (2015) to evaluating mutual fund performance. Unlike factors that consist of hypothetical portfolios, factors based on returns on traded assets like index funds and ETFs are much less likely to be subject to retroactive changes.

[27] http://www.law.uchicago.edu/law-finance/code/NoisyFactors

VI. Conclusion

The returns on French's factors—among the most ubiquitous inputs in empirical finance—are subject to large retroactive changes. These changes are due to updates to the raw data and to methodological changes in the factor construction. The retroactive changes have large effects on estimates that rely upon the factors. We show this in several contexts, in all cases holding the sample period fixed to restrict attention to the effects of the changes in the factors. Annual alphas of almost half of actively managed mutual funds change by more than 1%. In the cross-section of equities, changes in factor vintages substantially effect estimated alphas and factor loadings of both individual stocks and portfolios. Some "anomaly" portfolios are also affected. Our findings have significant implications for the replicability and robustness of finance research and have a direct bearing on a variety of industry and legal contexts.

References

Aiken, Adam L, Christopher P Clifford, and Jesse Ellis, 2013, Out of the dark: Hedge fund reporting biases and commercial databases, *The Review of Financial Studies* 26, 208–243.

Atanasov, Vladimir, and Bernard Black, 2021, The trouble with instruments: The need for pretreatment balance in shock-based instrumental variable designs, *Management Science* 67, 1270–1302.

Atanasov, Vladimir A, and Bernard S Black, 2016, Shock-based causal inference in corporate finance and accounting research, *Critical Finance Review* 5, 207–304.

Barillas, Francisco, Raymond Kan, Cesare Robotti, and Jay Shanken, 2020, Model comparison with sharpe ratios, *Journal of Financial and Quantitative Analysis* 55, 1840–1874.

Barillas, Francisco, and Jay Shanken, 2017, Which alpha?, *The Review of Financial Studies* 30, 1316–1338.

Berg, Florian, Kornelia Fabisik, and Zacharias Sautner, 2020, Rewriting history ii: The (un) predictable past of esg ratings, *European Corporate Governance Institute Working Paper* .

Berk, Jonathan B, and Jules H Van Binsbergen, 2015, Measuring skill in the mutual fund industry, *Journal of Financial Economics* 118, 1–20.

Bowen, Donald E, Laurent Frésard, and Jérôme P Taillard, 2017, What's your identification strategy? innovation in corporate finance research, *Management Science* 63, 2529–2548.

Bryzgalova, Svetlana, Sven Lerner, Martin Lettau, and Markus Pelger, 2022, Missing financial data, *Available at SSRN 4106794* .

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

Chang, Andrew C, and Phillip Li, 2018, Measurement error in macroeconomic data and economics research: Data revisions, gross domestic product, and gross domestic income, *Economic Inquiry* 56, 1846–1869.

Chang, Andrew C, Phillip Li, et al., 2022, Is economics research replicable? sixty published papers from thirteen journals say 'often not', *Critical Finance Review* 11, 185–206.

Chen, Andrew Y, 2020, The limits of p-hacking: Some thought experiments, *The Journal of Finance* .

Chen, Andrew Y, and Tom Zimmermann, 2021, Open source cross-sectional asset pricing, *Critical Finance Review* .

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

Croushore, Dean, and Tom Stark, 2003, A real-time data set for macroeconomists: Does the data vintage matter?, *Review of Economics and Statistics* 85, 605–617.

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

Dimensional Fund Advisors, 2013, Putting financial science to work for you, Technical report.

Dong, Xi, Yan Li, David E Rapach, and Guofu Zhou, 2022, Anomalies and the expected market return, *The Journal of Finance* 77, 639–681.

Erickson, Timothy, and Toni M Whited, 2000, Measurement error and the relationship between investment and q, *Journal of political economy* 108, 1027–1057.

Erickson, Timothy, and Toni M Whited, 2012, Treating measurement error in tobin's q, *The Review of Financial Studies* 25, 1286–1329.

Evans, Richard B, 2010, Mutual fund incubation, *The Journal of Finance* 65, 1581–1611.

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, 2015, A five-factor asset pricing model, *Journal of financial economics* 116, 1–22.

Fisch, Jill E, and Jonah B Gelbach, 2021, Power and statistical significance in securities fraud litigation, *Harvard Business Law Review* 11, 55.

Fisch, Jill E, Jonah B Gelbach, and Jonathan Klick, 2017, The logic and limits of event studies in securities fraud litigation, *Tex. L. Rev.* 96, 553.

Freyberger, Joachim, Björn Höppner, Andreas Neuhierl, and Michael Weber, 2021, Missing data in asset pricing panels, *Available at SSRN* .

Gibbons, Michael R, Stephen A Ross, and Jay Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica* 1121–1152.

Giglio, Stefano, Bryan T Kelly, and Serhiy Kozak, 2021, Equity term structures without dividend strips data, *Available at SSRN 3533486* .

Gillan, Stuart L, Jay C Hartzell, Andrew Koch, and Laura T Starks, 2018, Getting the incentives right: Backfilling and biases in executive compensation data, *Review of Financial Studies* 31, 1460–1498.

Glandon, Philip, 2010, Report on the american economic review data availability compliance project, *Vanderbilt University* 1–12.

Graham, John R, and Campbell R Harvey, 2001, The theory and practice of corporate finance: Evidence from the field, *Journal of financial economics* 60, 187–243.

Haddad, Valentin, Serhiy Kozak, and Shrihari Santosh, 2020, Factor timing, *The Review of Financial Studies* 33, 1980–2018.

Harvey, Campbell R, 2017, Presidential address: The scientific outlook in financial economics, *The Journal of Finance* 72, 1399–1440.

Harvey, Campbell R, and Yan Liu, 2021, Lucky factors, *Journal of Financial Economics* .

Harvey, Campbell R, Yan Liu, and Alessio Saretto, 2020, An evaluation of alternative multiple testing methods for finance applications, *The Review of Asset Pricing Studies* 10, 199–248.

Harvey, Campbell R, Yan Liu, and Heqing Zhu, 2016, ... and the cross-section of expected returns, *The Review of Financial Studies* 29, 5–68.

Heath, Davidson, Matthew Ringgenberg, Mehrdad Samadi, and Ingrid M Werner, 2020, Reusing natural experiments, *Fisher College of Business Working Paper* 021.

Heider, Florian, and Alexander Ljungqvist, 2015, As certain as debt and taxes: Estimating the tax sensitivity of leverage from state tax changes, *Journal of financial economics* 118, 684–712.

Hou, Kewei, Haitao Mo, Chen Xue, and Lu Zhang, 2021, An augmented q-factor model with expected growth, *Review of Finance* 25, 1–41.

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

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

Jennings, Jared N, Jung Min Kim, Joshua A Lee, and Daniel J Taylor, 2020, Measurement error and bias in causal models in accounting research, *Available at SSRN 3731197* .

Jensen, Theis Ingerslev, Bryan T Kelly, and Lasse Heje Pedersen, 2021, Is there a replication crisis in finance?, Technical report, National Bureau of Economic Research.

Karpoff, Jonathan M, Allison Koester, D Scott Lee, and Gerald S Martin, 2014, Database challenges in financial misconduct research, *University of Washington working paper* .

Kothari, Sagar P, and Jerold B Warner, 2007, Econometrics of event studies, in *Handbook of empirical corporate finance*, 3–36 (Elsevier).

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

Ljungqvist, Alexander, Christopher Malloy, and Felicia Marston, 2009, Rewriting history, *The Journal of Finance* 64, 1935–1960.

MacKinlay, A Craig, 1997, Event studies in economics and finance, *Journal of economic literature* 35, 13–39.

Mankiw, N Gregory, David E Runkle, and Matthew D Shapiro, 1984, Are preliminary announcements of the money stock rational forecasts?, *Journal of Monetary economics* 14, 15–27.

McCullough, Bruce D, Kerry Anne McGeary, and Teresa D Harrison, 2006, Lessons from the jmcb archive, *Journal of Money, Credit and Banking* 1093–1107.

McCullough, Bruce D, and Hrishikesh D Vinod, 2003, Verifying the solution from a nonlinear solver: A case study, *American Economic Review* 93, 873–892.

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

Mitton, Todd, 2020a, Economic significance in corporate finance, *Available at SSRN 3667830* .

Mitton, Todd, 2020b, Methodological variation in empirical corporate finance, *Available at SSRN 3304875* .

Pancost, N. Aaron, and Garrett Schaller, 2022, Measuring measurement error, Working Paper, University of Texas at Austin.

Patton, Andrew J, Tarun Ramadorai, and Michael Streatfield, 2015, Change you can believe in? hedge fund data revisions, *The Journal of Finance* 70, 963–999.

Pénasse, Julien, 2020, Understanding alpha decay, *Available at SSRN 2953614* .

Schwarz, Christopher G, and Mark E Potter, 2016, Revisiting mutual fund portfolio disclosure, *The Review of Financial Studies* 29, 3519–3544.

Sensoy, Berk A, 2009, Performance evaluation and self-designated benchmark indexes in the mutual fund industry, *Journal of Financial Economics* 92, 25–39.

Spamann, Holger, 2010, The "antidirector rights index" revisited, *The Review of Financial Studies* 23, 467–486.

Spamann, Holger, 2019, On inference when using state corporate laws for identification, *Harvard Law School John M. Olin Center Discussion Paper* .

Zingales, Luigi, 2013, Preventing economists' capture, in David Carpenter, and David Moss, eds., *Preventing Regulatory Capture: Special Interest Influence and How to Limit it*.

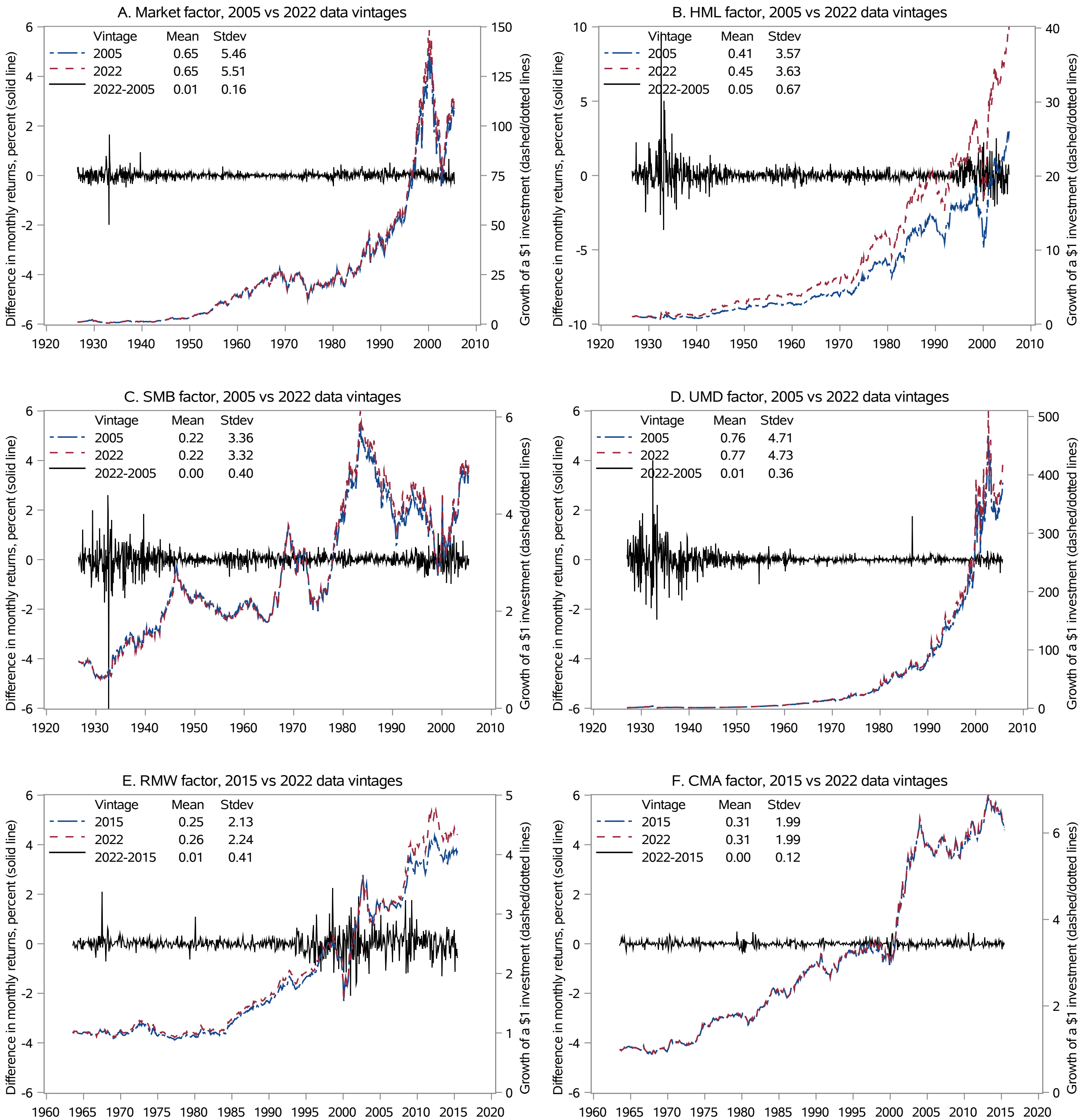


Figure 1. Factor returns from different vintages

This figure plots differences between monthly returns of factors from two vintages (the solid black line). It also shows cumulative returns of factors from the two vintages (the dashed and dash-dotted lines). The top left of each panel reports means and standard deviations of factor returns, in percent per month.

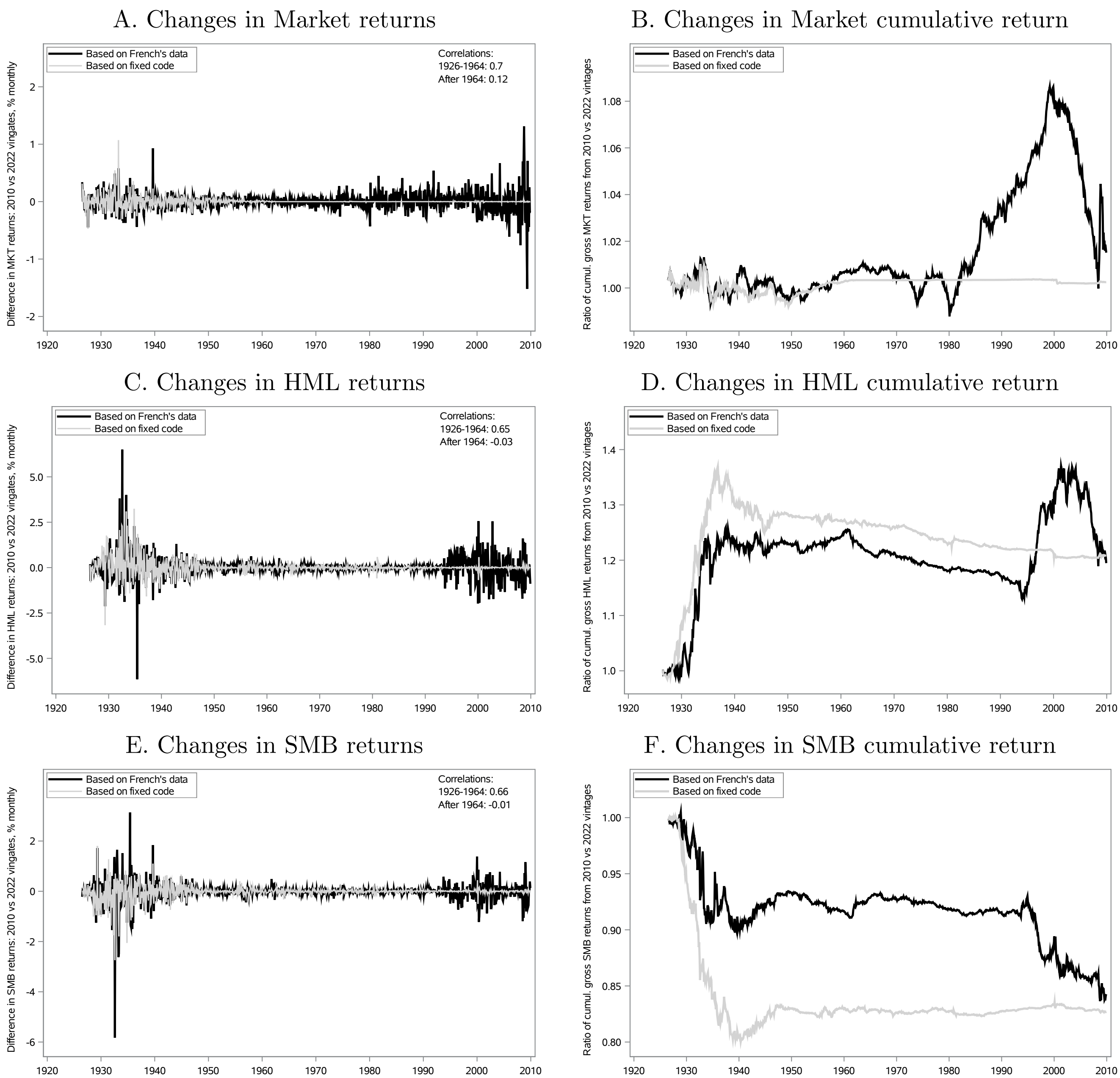


Figure 2. Changes in the French factors and the fixed-code factors

This figure plots the differences in returns in the fixed-code factors and French factors, along with the changes in cumulative returns of the fixed-code factors and the French factors. Panels A, C, and E present the changes in factor returns, while panels B, D, and F present the changes in cumulative returns. Panels A and B examine the market factor, panels C and D examine the HML factor, while panels E and F present the SMB factor. Data used for the fixed-code factors comes CRSP and Compustat.

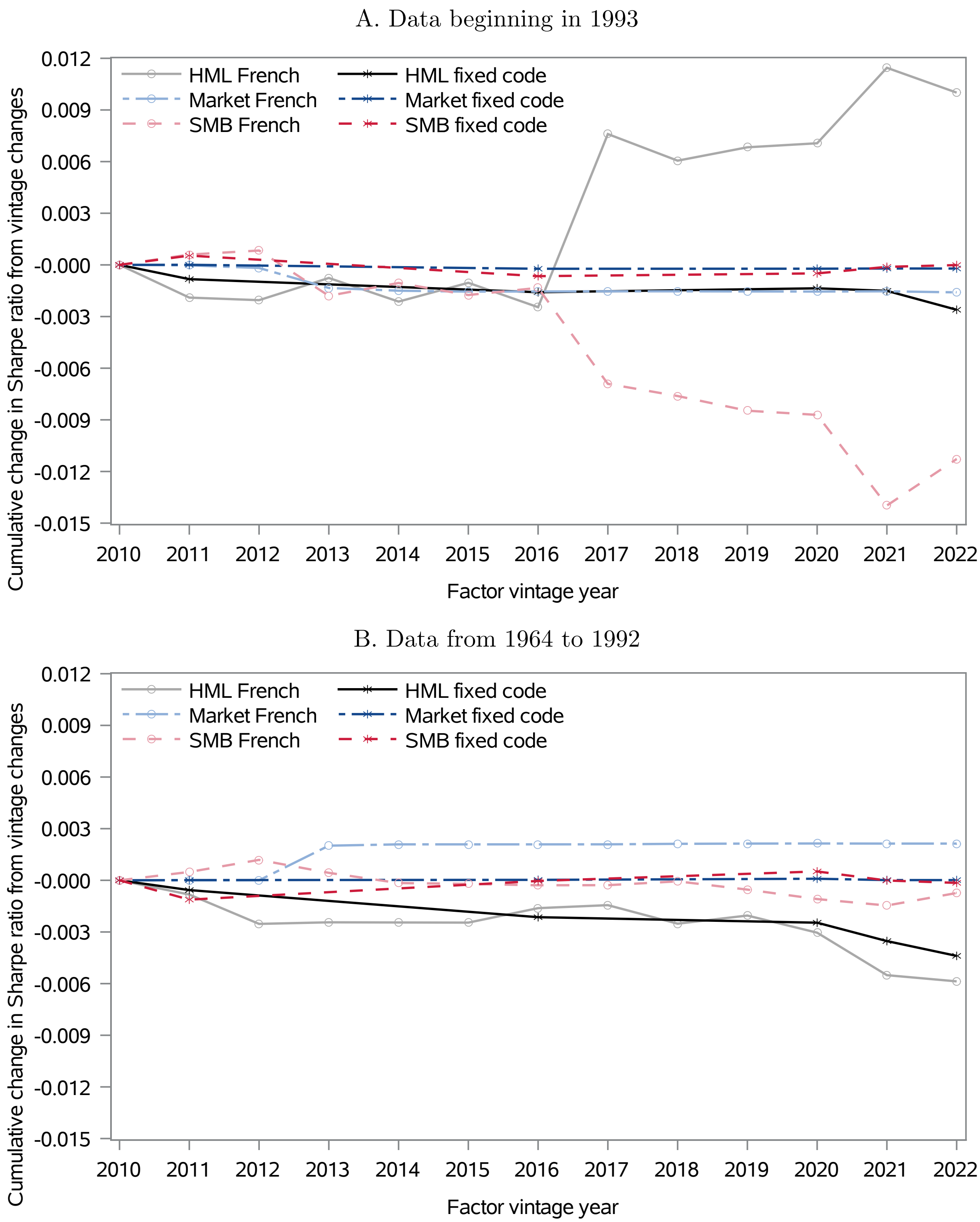


Figure 3. Cumulative changes in Sharpe ratios of factors across vintages

This figure plots cumulative changes in Sharpe ratios of the three factors that arise due to updating factor vintages. For each factor and each adjacent pair of vintages, Sharpe ratios are calculated using data common to both vintages. The differences in the two Sharpe ratio estimates are then cumulated over time. Panel A uses data from 1993 through 2021; Panel B uses data from 1964 through 1992.

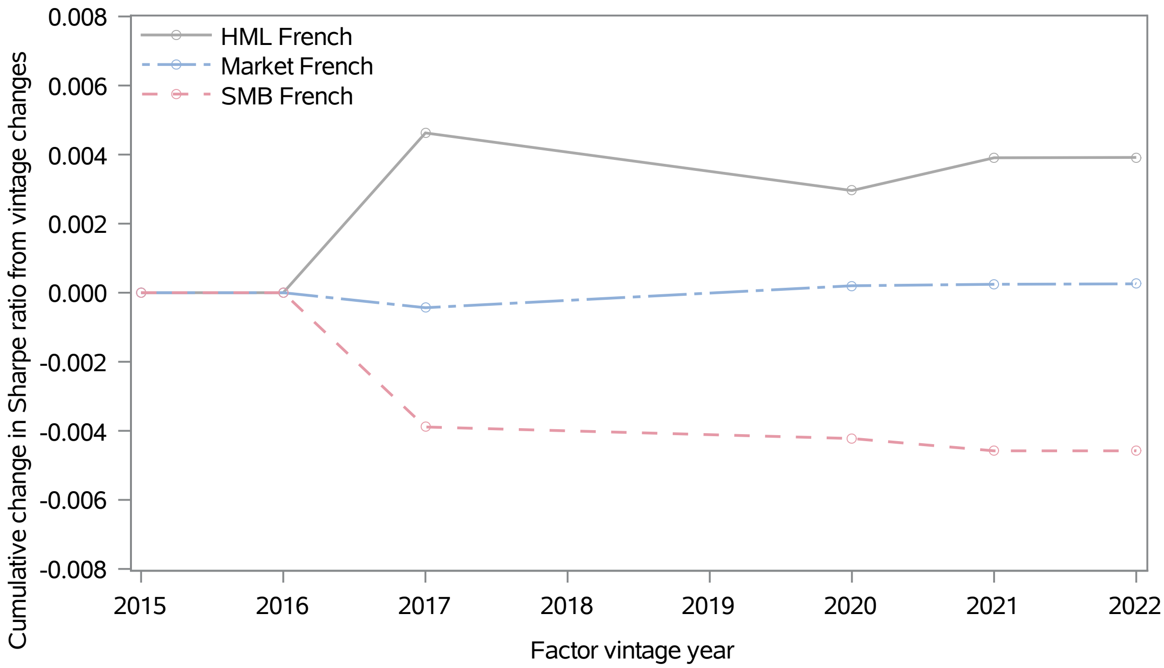


Figure 4. Cumulative changes in Sharpe ratios of Ex-US factors across vintages This figure plots cumulative changes in Sharpe ratios of the Developed Ex-US three factors that arise due to updating factor vintages. Prior to 2020, they were called "Global ex US" factors, but included data from the same set of countries as the present "Developed ex US" factors. For each factor and each adjacent pair of vintages, Sharpe ratios are calculated using data common to both vintages. The differences in the two Sharpe ratio estimates are then cumulated over time. The figure uses data beginning in 1993.

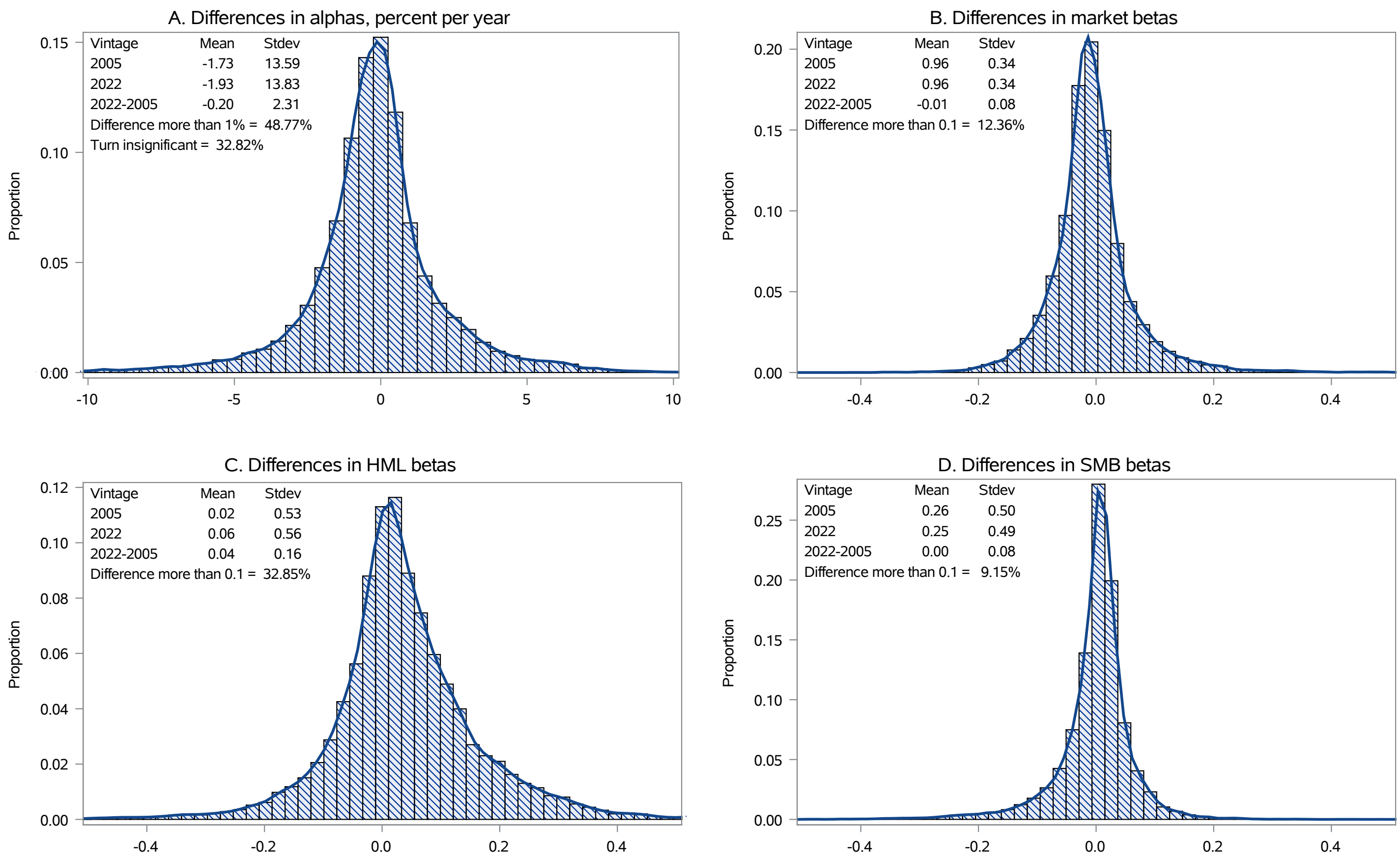


Figure 5. Differences in mutual fund alphas and betas: 2005 vs 2022 factor vintages

This figure plots histograms and kernel densities of differences in alphas (Panel A) and betas (B, C, D) of individual mutual funds estimated using 2005 and 2022 factor vintages. Alphas, in percent per year, and betas are estimated every calendar year using three-factor regressions on monthly data. Top left of each panel reports means and standard deviations of estimates from the two vintages and of their differences. It also shows the fraction of observations with absolute differences above a certain threshold and in Panel A the proportion of alphas that are significant in one vintage but not in the other.

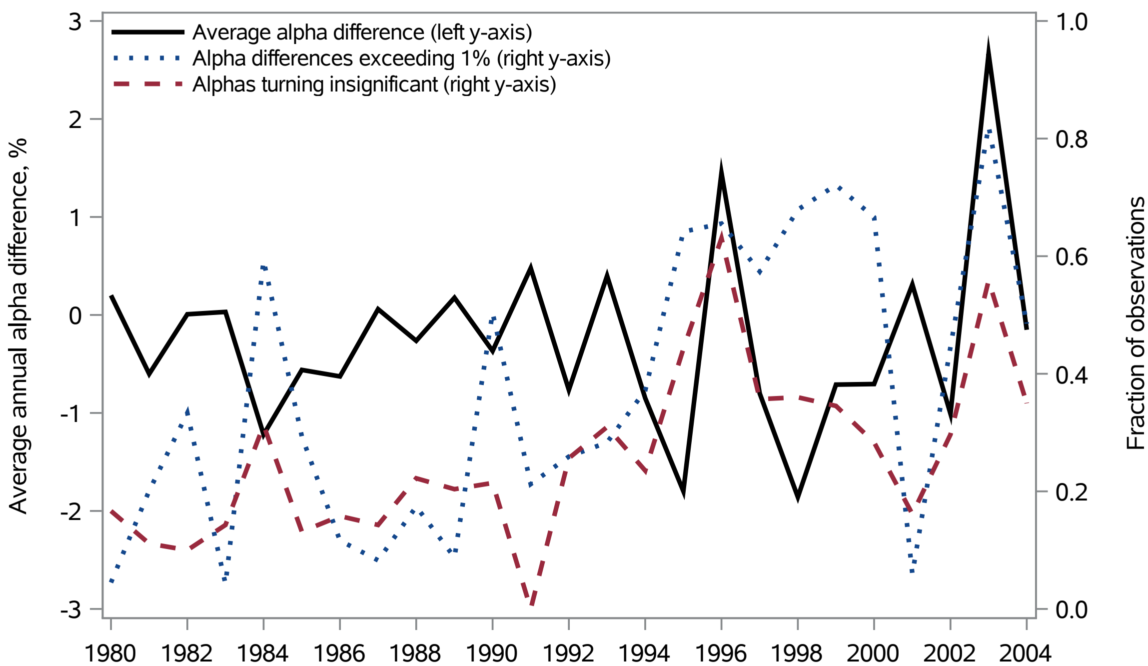


Figure 6. Evolution of differences in mutual fund alphas: 2005 vs 2022 factor vintages

This figure plots the time series of statistics of estimates of mutual fund alphas obtained using 2005 and 2022 factor vintages. Alphas are estimated every calendar year using three-factor regressions on monthly data. The solid line shows average alpha, in percent per year. The dotted line plots the fraction of funds whose alphas from the two vintages differ by more than 1% annualized. The dashed line indicates the proportion of alphas that are significant in one vintage but not in the other.

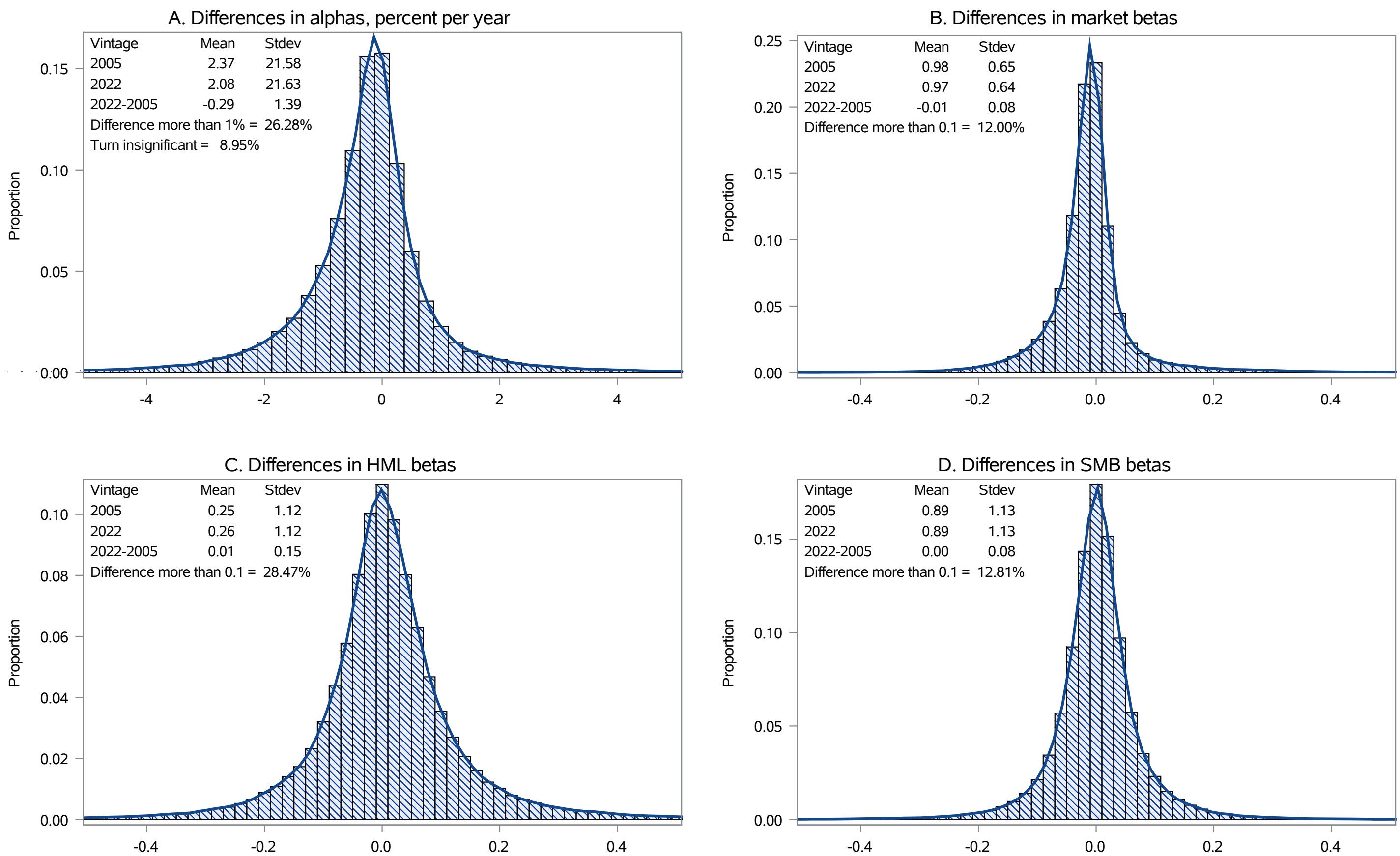


Figure 7. Differences in stock-level alphas and betas: 2005 vs 2022 factor vintages

This figure plots histograms and kernel densities of differences in alphas (Panel A) and betas (B, C, D) of individual stocks estimated using 2005 and 2022 factor vintages. Alphas, in percent per year, and betas are estimated at the end of every calendar year using three-factor regressions on five years of monthly data. Top left of each panel reports means and standard deviations of estimates from the two vintages and of their differences. It also shows the fraction of observations with absolute differences above a certain threshold and in Panel A the proportion of alphas that are significant in one vintage but not in the other.

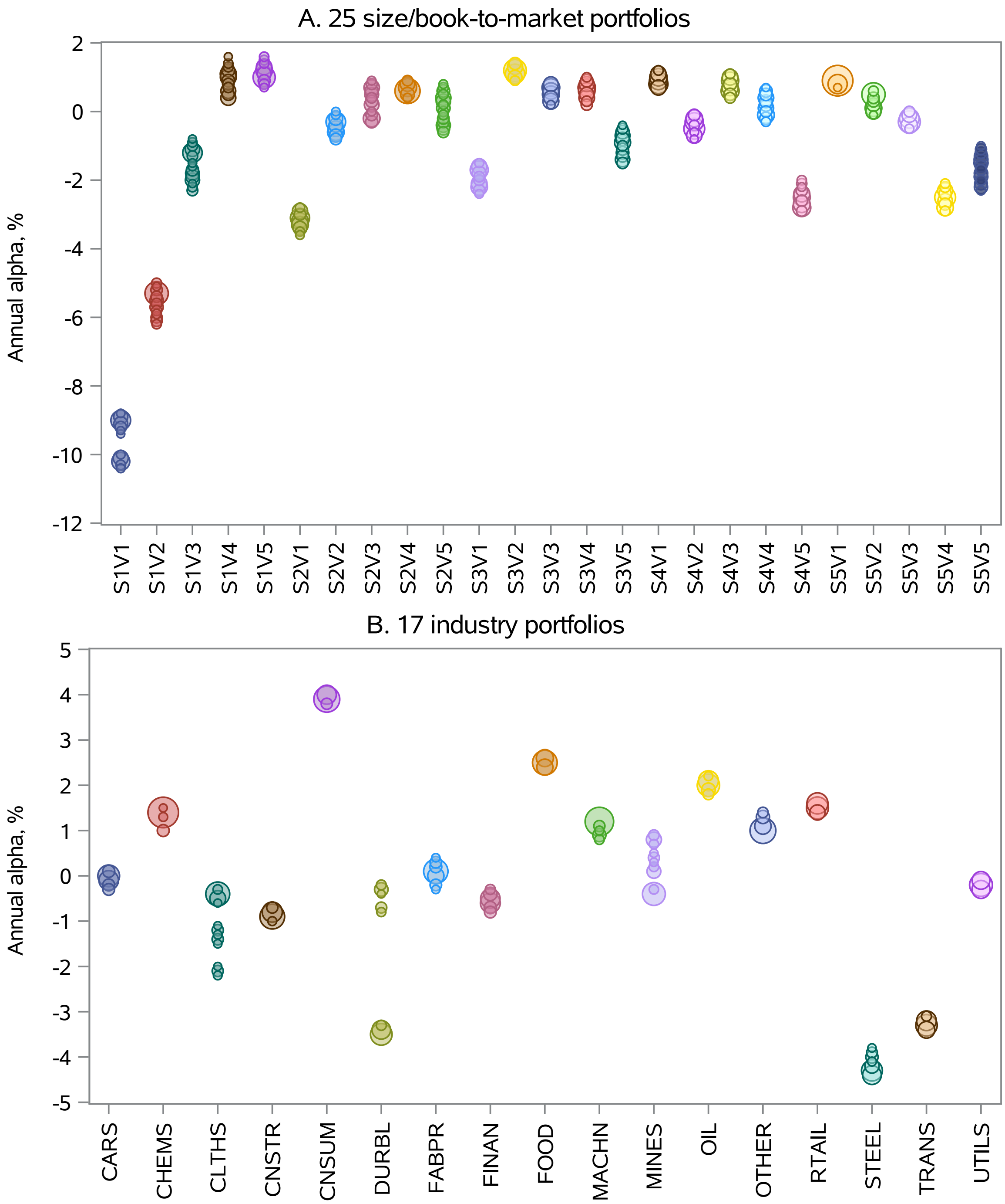


Figure 8. Alphas of 25 size/book-to-market and 17 industry portfolios estimated using different factor and portfolio vintages

This figure plots unconditional three-factor alphas, in percent per year, of 25 size and book-to-market sorted portfolios (Panel A) and industry portfolios (B). Alphas are estimated for each of 18 factor vintages and 12 portfolio vintages using the sample common to all vintages: 07/1926-06/2005 in Panel A and 07/1926-05/2005 in Panel B. Alphas are rounded to the nearest 0.1%, and the size of the bubbles represents the relative frequency of estimates.

Table I
Differences in returns of factors across vintages

This table reports statistics for differences in returns of market (Panel A), HML (B), and SMB (C) factors from different factor vintages. Upper and lower triangular entries reflect the results using monthly and daily data, respectively. Mean $|\text{Diff}|$ is the average absolute difference in factor returns, in percent per month. SD Diff is the standard deviation of the difference, in percent per month. $|\text{Diff}| > 1\%/\text{yr}$ is the proportion of observations where the absolute difference exceeds 1% per year, which translates into 1%/12 in monthly data and $1\%/(12 \times 21)$ in daily data. The row labeled Not same shows the proportion of factor return observations that is different in the two compared vintages. When comparing vintages, all data common to both vintages is used.

		Vintage 1																	
Vintage 2	Variable	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
A. Market factor																			
2005	Mean $\lvert$Diff$\rvert$		0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.09	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
	SD Diff		0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.15	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16
	$\lvert$Diff$\rvert$ > 1%/yr		0.10	0.11	0.11	0.11	0.11	0.11	0.11	0.39	0.41	0.42	0.42	0.42	0.43	0.43	0.43	0.43	0.43
	Not same		0.67	0.68	0.67	0.67	0.68	0.68	0.68	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
2006	Mean $\lvert$Diff$\rvert$	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
	SD Diff	0.01		0.01	0.02	0.02	0.02	0.02	0.02	0.12	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13
	$\lvert$Diff$\rvert$ > 1%/yr	0.00		0.00	0.01	0.01	0.01	0.01	0.01	0.35	0.38	0.39	0.39	0.39	0.39	0.39	0.39	0.39	0.39
	Not same	0.00		0.10	0.12	0.12	0.14	0.14	0.16	0.94	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
2007	Mean $\lvert$Diff$\rvert$	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
	SD Diff	0.01	0.02		0.02	0.02	0.02	0.02	0.02	0.12	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13
	$\lvert$Diff$\rvert$ > 1%/yr	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.35	0.38	0.39	0.39	0.39	0.39	0.39	0.39	0.39	0.39
	Not same	0.00	0.00		0.04	0.05	0.07	0.07	0.09	0.94	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
2008	Mean $\lvert$Diff$\rvert$					0.00	0.00	0.00	0.00	0.09	0.09	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
	SD Diff					0.00	0.00	0.00	0.00	0.13	0.13	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14
	$\lvert$Diff$\rvert$ > 1%/yr					0.00	0.00	0.00	0.00	0.35	0.39	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40
	Not same					0.01	0.03	0.04	0.06	0.95	0.96	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
2009	Mean $\lvert$Diff$\rvert$						0.00	0.00	0.00	0.09	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
	SD Diff						0.00	0.00	0.00	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16
	$\lvert$Diff$\rvert$ > 1%/yr						0.00	0.00	0.00	0.36	0.39	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40
	Not same						0.03	0.04	0.06	0.95	0.96	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
2010	Mean $\lvert$Diff$\rvert$	0.01	0.01	0.01				0.00	0.00	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
	SD Diff	0.05	0.05	0.08				0.00	0.00	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16
	$\lvert$Diff$\rvert$ > 1%/yr	0.02	0.02	0.02				0.00	0.00	0.37	0.40	0.41	0.41	0.41	0.41	0.41	0.41	0.41	0.41
	Not same	0.02	0.02	0.02				0.00	0.04	0.95	0.96	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
2011	Mean $\lvert$Diff$\rvert$								0.00	0.10	0.10	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11
	SD Diff								0.00	0.16	0.16	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17
	$\lvert$Diff$\rvert$ > 1%/yr								0.00	0.37	0.40	0.41	0.41	0.41	0.41	0.41	0.41	0.41	0.41
	Not same								0.04	0.95	0.96	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
2012	Mean $\lvert$Diff$\rvert$	0.01	0.01	0.01			0.00			0.10	0.10	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11
	SD Diff	0.06	0.06	0.08			0.03			0.16	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17
	$\lvert$Diff$\rvert$ > 1%/yr	0.04	0.04	0.04			0.02			0.38	0.41	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42
	Not same	0.04	0.04	0.04			0.02			0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
2013	Mean $\lvert$Diff$\rvert$										0.02	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
	SD Diff										0.05	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07
	$\lvert$Diff$\rvert$ > 1%/yr										0.07	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08
	Not same										0.51	0.53	0.53	0.54	0.54	0.54	0.54	0.55	0.54
2014	Mean $\lvert$Diff$\rvert$	0.45	0.45	0.46			0.55		0.57			0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	SD Diff	0.66	0.67	0.69			0.90		0.92			0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
	$\lvert$Diff$\rvert$ > 1%/yr	0.80	0.80	0.80			0.81		0.82			0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
	Not same	0.80	0.80	0.80			0.81		0.82			0.25	0.25	0.27	0.28	0.29	0.29	0.30	0.29
2015	Mean $\lvert$Diff$\rvert$	0.45	0.45	0.46			0.55		0.57		0.06		0.00	0.00	0.00	0.00	0.00	0.00	0.00
	SD Diff	0.66	0.67	0.69			0.90		0.92		0.28		0.00	0.00	0.00	0.00	0.00	0.00	0.00
	$\lvert$Diff$\rvert$ > 1%/yr	0.80	0.80	0.80			0.81		0.82		0.14		0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Not same	0.80	0.80	0.80			0.81		0.82		0.14		0.00	0.02	0.03	0.04	0.04	0.05	0.05
2016	Mean $\lvert$Diff$\rvert$	0.45	0.45	0.46			0.55		0.57		0.06	0.00		0.00	0.00	0.00	0.00	0.00	0.00
	SD Diff	0.66	0.67	0.69			0.90		0.92		0.28	0.00		0.00	0.00	0.00	0.00	0.00	0.00
	$\lvert$Diff$\rvert$ > 1%/yr	0.80	0.80	0.80			0.81		0.82		0.14	0.00		0.00	0.00	0.00	0.00	0.00	0.00
	Not same	0.80	0.80	0.80			0.81		0.82		0.14	0.00		0.02	0.04	0.04	0.04	0.05	0.05
2017	Mean $\lvert$Diff$\rvert$	0.45	0.45	0.46			0.55		0.57		0.06	0.00	0.00		0.00	0.00	0.00	0.00	0.00
	SD Diff	0.66	0.67	0.69			0.90		0.92		0.28	0.01	0.01		0.00	0.00	0.00	0.00	0.00
	$\lvert$Diff$\rvert$ > 1%/yr	0.80	0.80	0.80			0.81		0.82		0.14	0.00	0.00		0.00	0.00	0.00	0.00	0.00
	Not same	0.80	0.80	0.80			0.81		0.82		0.14	0.00	0.00		0.01	0.02	0.02	0.03	0.03
2018	Mean $\lvert$Diff$\rvert$															0.00	0.00	0.00	0.00
	SD Diff															0.00	0.00	0.00	0.00
	$\lvert$Diff$\rvert$ > 1%/yr															0.00	0.00	0.00	0.00
	Not same															0.01	0.01	0.02	0.02
2019	Mean $\lvert$Diff$\rvert$																0.00	0.00	0.00
	SD Diff																0.00	0.00	0.00
	$\lvert$Diff$\rvert$ > 1%/yr																0.00	0.00	0.00
	Not same																0.00	0.01	0.02
2020	Mean $\lvert$Diff$\rvert$	0.45	0.45	0.46			0.55		0.57		0.06	0.00	0.00	0.00				0.00	0.00
	SD Diff	0.66	0.67	0.69			0.90		0.92		0.28	0.02	0.02	0.02				0.00	0.00
	$\lvert$Diff$\rvert$ > 1%/yr	0.80	0.80	0.80			0.81		0.82		0.15	0.01	0.01	0.01				0.00	0.00
	Not same	0.80	0.80	0.80			0.81		0.82		0.15	0.01	0.01	0.01				0.01	0.02
2021	Mean $\lvert$Diff$\rvert$	0.45	0.45	0.46			0.55		0.57		0.06	0.00	0.00	0.00			0.00		0.00
	SD Diff	0.66	0.67	0.69			0.90		0.92		0.28	0.02	0.02	0.02			0.01		0.00
	$\lvert$Diff$\rvert$ > 1%/yr	0.80	0.80	0.80			0.81		0.82		0.15	0.01	0.01	0.01			0.00		0.00
	Not same	0.80	0.80	0.80			0.81		0.82		0.15	0.01	0.01	0.01			0.00		0.01
2022	Mean $\lvert$Diff$\rvert$	0.45	0.45	0.46			0.55		0.57		0.06	0.00	0.00	0.00			0.00	0.00	
	SD Diff	0.66	0.67	0.69			0.90		0.92		0.28	0.02	0.02	0.02			0.01	0.01	
	$\lvert$Diff$\rvert$ > 1%/yr	0.80	0.80	0.80			0.81		0.82		0.15	0.01	0.01	0.01			0.00	0.00	
	Not same	0.80	0.80	0.80			0.81		0.82		0.15	0.01	0.01	0.01			0.00	0.00	

		Vintage 1																	
Vintage 2	Variable	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
B. HML factor																			
2005	Mean |Diff|		0.21	0.23	0.23	0.22	0.23	0.23	0.23	0.28	0.29	0.31	0.31	0.36	0.36	0.36	0.36	0.36	0.37
	SD Diff		0.36	0.38	0.40	0.40	0.40	0.40	0.40	0.46	0.49	0.58	0.59	0.65	0.65	0.65	0.65	0.65	0.67
	|Diff| > 1%/yr		0.68	0.70	0.70	0.69	0.69	0.69	0.71	0.73	0.73	0.74	0.73	0.76	0.77	0.77	0.76	0.76	0.77
	Not same		0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.99	0.99	0.98	0.98	0.99	0.99	0.98	0.99
2006	Mean |Diff|	0.13		0.04	0.06	0.07	0.07	0.07	0.07	0.16	0.17	0.20	0.20	0.27	0.27	0.27	0.27	0.30	0.30
	SD Diff	0.23		0.12	0.18	0.18	0.18	0.18	0.18	0.32	0.36	0.44	0.44	0.51	0.51	0.51	0.51	0.54	0.56
	|Diff| > 1%/yr	0.47		0.10	0.16	0.21	0.23	0.22	0.23	0.40	0.42	0.46	0.47	0.56	0.59	0.58	0.58	0.62	0.61
	Not same	0.47		0.51	0.78	0.83	0.83	0.81	0.82	0.94	0.96	0.95	0.96	0.97	0.97	0.97	0.97	0.97	0.97
2007	Mean |Diff|	0.17	0.10		0.04	0.05	0.05	0.05	0.05	0.16	0.17	0.20	0.20	0.27	0.28	0.28	0.28	0.31	0.31
	SD Diff	0.31	0.25		0.15	0.16	0.16	0.16	0.16	0.32	0.37	0.45	0.45	0.51	0.52	0.52	0.52	0.56	0.57
	|Diff| > 1%/yr	0.54	0.34		0.10	0.16	0.17	0.16	0.18	0.40	0.41	0.45	0.46	0.55	0.58	0.58	0.58	0.62	0.62
	Not same	0.54	0.34		0.71	0.77	0.77	0.74	0.75	0.92	0.95	0.96	0.96	0.97	0.96	0.97	0.97	0.97	0.97
2008	Mean |Diff|					0.03	0.03	0.02	0.03	0.14	0.16	0.19	0.19	0.27	0.28	0.28	0.28	0.31	0.31
	SD Diff					0.04	0.05	0.04	0.05	0.30	0.35	0.44	0.44	0.51	0.54	0.55	0.55	0.58	0.60
	|Diff| > 1%/yr					0.07	0.07	0.07	0.09	0.34	0.34	0.40	0.42	0.52	0.56	0.57	0.57	0.60	0.60
	Not same					0.58	0.59	0.50	0.51	0.88	0.95	0.94	0.95	0.96	0.97	0.97	0.97	0.97	0.96
2009	Mean |Diff|						0.01	0.03	0.03	0.15	0.16	0.19	0.20	0.28	0.29	0.29	0.29	0.31	0.32
	SD Diff						0.03	0.05	0.05	0.30	0.35	0.44	0.44	0.52	0.55	0.56	0.56	0.59	0.60
	|Diff| > 1%/yr						0.03	0.07	0.09	0.40	0.39	0.44	0.46	0.55	0.58	0.59	0.59	0.62	0.62
	Not same						0.41	0.59	0.59	0.90	0.94	0.95	0.96	0.97	0.97	0.96	0.96	0.97	0.98
2010	Mean |Diff|	0.29	0.26	0.26				0.02	0.03	0.14	0.16	0.19	0.19	0.28	0.29	0.29	0.29	0.31	0.32
	SD Diff	0.44	0.39	0.41				0.04	0.05	0.30	0.35	0.43	0.43	0.52	0.55	0.55	0.55	0.58	0.60
	|Diff| > 1%/yr	0.73	0.70	0.69				0.04	0.07	0.37	0.36	0.41	0.43	0.54	0.58	0.58	0.58	0.61	0.62
	Not same	0.73	0.70	0.69				0.57	0.59	0.91	0.94	0.94	0.94	0.96	0.97	0.96	0.97	0.97	0.98
2011	Mean |Diff|								0.02	0.14	0.15	0.19	0.19	0.27	0.28	0.29	0.29	0.31	0.32
	SD Diff								0.03	0.29	0.35	0.43	0.43	0.52	0.55	0.55	0.55	0.58	0.60
	|Diff| > 1%/yr								0.03	0.34	0.34	0.38	0.41	0.52	0.55	0.56	0.56	0.61	0.61
	Not same								0.47	0.89	0.95	0.96	0.95	0.96	0.97	0.96	0.96	0.97	0.97
2012	Mean |Diff|	0.30	0.27	0.27			0.21			0.13	0.15	0.18	0.18	0.27	0.28	0.29	0.29	0.31	0.32
	SD Diff	0.42	0.39	0.41			0.32			0.29	0.34	0.43	0.43	0.52	0.55	0.55	0.55	0.59	0.60
	|Diff| > 1%/yr	0.75	0.73	0.71			0.63			0.30	0.31	0.36	0.38	0.51	0.54	0.54	0.54	0.60	0.62
	Not same	0.75	0.73	0.71			0.63			0.89	0.94	0.95	0.95	0.96	0.95	0.96	0.96	0.97	0.98
2013	Mean |Diff|										0.07	0.12	0.12	0.22	0.24	0.24	0.24	0.27	0.28
	SD Diff										0.21	0.34	0.34	0.45	0.48	0.48	0.48	0.52	0.54
	|Diff| > 1%/yr										0.18	0.26	0.29	0.44	0.47	0.48	0.49	0.55	0.60
	Not same										0.76	0.80	0.82	0.89	0.95	0.95	0.95	0.97	0.98
2014	Mean |Diff|	0.31	0.28	0.29			0.29		0.21			0.10	0.11	0.21	0.22	0.23	0.23	0.26	0.27
	SD Diff	0.45	0.42	0.45			0.44		0.32			0.27	0.28	0.41	0.43	0.43	0.43	0.48	0.50
	|Diff| > 1%/yr	0.76	0.73	0.72			0.73		0.63			0.22	0.24	0.42	0.45	0.47	0.47	0.54	0.59
	Not same	0.76	0.73	0.72			0.73		0.63			0.60	0.76	0.85	0.93	0.93	0.94	0.96	0.97
2015	Mean |Diff|	0.39	0.38	0.39			0.40		0.35		0.69		0.01	0.12	0.14	0.14	0.14	0.18	0.21
	SD Diff	0.58	0.59	0.60			0.67		0.60		1.68		0.03	0.30	0.33	0.33	0.33	0.39	0.41
	|Diff| > 1%/yr	0.79	0.79	0.78			0.79		0.73		0.70		0.02	0.21	0.26	0.28	0.28	0.35	0.47
	Not same	0.79	0.79	0.78			0.79		0.73		0.70		0.35	0.46	0.66	0.66	0.68	0.72	0.92
2016	Mean |Diff|	0.39	0.38	0.39			0.40		0.35		0.69	0.00		0.12	0.14	0.14	0.15	0.18	0.21
	SD Diff	0.58	0.59	0.60			0.67		0.60		1.68	0.01		0.31	0.34	0.34	0.34	0.40	0.41
	|Diff| > 1%/yr	0.79	0.79	0.78			0.79		0.73		0.70	0.00		0.22	0.27	0.29	0.29	0.37	0.47
	Not same	0.79	0.79	0.78			0.79		0.73		0.70	0.00		0.43	0.66	0.66	0.69	0.72	0.93
2017	Mean |Diff|	0.87	0.86	0.88			1.05		1.04		1.10	0.53	0.53		0.03	0.04	0.04	0.11	0.14
	SD Diff	1.66	1.65	1.66			2.17		2.18		2.28	1.57	1.57		0.13	0.13	0.14	0.25	0.27
	|Diff| > 1%/yr	0.84	0.83	0.83			0.84		0.80		0.76	0.29	0.30		0.08	0.10	0.11	0.32	0.45
	Not same	0.84	0.83	0.83			0.84		0.80		0.76	0.29	0.30		0.63	0.60	0.63	0.73	0.93
2018	Mean |Diff|															0.02	0.02	0.10	0.13
	SD Diff															0.05	0.06	0.22	0.24
	|Diff| > 1%/yr															0.03	0.03	0.29	0.43
	Not same															0.40	0.49	0.70	0.92
2019	Mean |Diff|																0.01	0.09	0.13
	SD Diff																0.03	0.21	0.24
	|Diff| > 1%/yr																0.01	0.27	0.42
	Not same																0.23	0.69	0.92
2020	Mean |Diff|	0.88	0.87	0.89			1.07		1.06		1.15	0.61	0.62	0.14				0.09	0.13
	SD Diff	1.67	1.66	1.67			2.18		2.19		2.30	1.60	1.61	0.35				0.21	0.24
	|Diff| > 1%/yr	0.85	0.84	0.84			0.84		0.82		0.82	0.48	0.48	0.37				0.27	0.43
	Not same	0.85	0.84	0.84			0.84		0.82		0.82	0.48	0.48	0.37				0.67	0.92
2021	Mean |Diff|	1.11	1.11	1.13			1.31		1.32		1.31	0.79	0.79	0.48			0.43		0.07
	SD Diff	2.07	2.05	2.06			2.56		2.60		2.52	1.90	1.90	1.07			1.02		0.13
	|Diff| > 1%/yr	0.87	0.87	0.87			0.87		0.85		0.86	0.57	0.57	0.56			0.51		0.27
	Not same	0.87	0.87	0.87			0.87		0.85		0.86	0.57	0.57	0.56			0.51		0.91
2022	Mean |Diff|	1.11	1.11	1.13			1.32		1.36		1.35	0.97	0.97	0.66			0.64	0.35	
	SD Diff	2.07	2.06	2.07			2.58		2.61		2.54	2.00	2.01	1.23			1.21	0.69	
	|Diff| > 1%/yr	0.87	0.87	0.87			0.88		0.88		0.88	0.81	0.81	0.80			0.79	0.71	
	Not same	0.87	0.87	0.87			0.88		0.88		0.88	0.81	0.81	0.80			0.79	0.71	

Vintage 2	Variable	Vintage 1																	
		2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
C. SMB factor																			
2005	Mean $\lvert$Diff$\rvert$		0.14	0.15	0.15	0.15	0.15	0.15	0.15	0.18	0.19	0.20	0.20	0.22	0.22	0.22	0.22	0.22	0.22
	SD Diff		0.24	0.24	0.26	0.26	0.26	0.26	0.26	0.31	0.35	0.37	0.37	0.38	0.38	0.38	0.38	0.38	0.40
	$\lvert$Diff$\rvert$ > 1%/yr		0.53	0.52	0.54	0.54	0.54	0.54	0.56	0.60	0.59	0.60	0.60	0.64	0.65	0.64	0.64	0.66	0.65
	Not same		0.97	0.97	0.98	0.97	0.98	0.98	0.97	0.97	0.98	0.97	0.97	0.97	0.97	0.97	0.97	0.98	0.98
2006	Mean $\lvert$Diff$\rvert$	0.09		0.03	0.04	0.05	0.05	0.05	0.05	0.11	0.13	0.14	0.14	0.16	0.16	0.16	0.16	0.17	0.18
	SD Diff	0.18		0.08	0.12	0.12	0.12	0.12	0.12	0.25	0.30	0.32	0.32	0.33	0.33	0.33	0.33	0.34	0.37
	$\lvert$Diff$\rvert$ > 1%/yr	0.36		0.08	0.12	0.14	0.15	0.15	0.17	0.30	0.32	0.32	0.32	0.40	0.42	0.43	0.43	0.44	0.47
	Not same	0.36		0.53	0.80	0.84	0.84	0.80	0.81	0.90	0.94	0.94	0.95	0.95	0.96	0.95	0.95	0.96	0.97
2007	Mean $\lvert$Diff$\rvert$	0.13	0.08		0.03	0.04	0.04	0.04	0.04	0.11	0.12	0.14	0.14	0.16	0.16	0.16	0.16	0.17	0.18
	SD Diff	0.23	0.18		0.10	0.11	0.11	0.11	0.11	0.24	0.30	0.32	0.32	0.33	0.33	0.33	0.33	0.34	0.37
	$\lvert$Diff$\rvert$ > 1%/yr	0.45	0.31		0.07	0.09	0.10	0.09	0.11	0.29	0.31	0.32	0.33	0.40	0.41	0.42	0.42	0.45	0.47
	Not same	0.45	0.31		0.72	0.77	0.76	0.75	0.76	0.90	0.93	0.94	0.94	0.94	0.95	0.95	0.95	0.96	0.96
2008	Mean $\lvert$Diff$\rvert$					0.02	0.02	0.01	0.02	0.09	0.11	0.12	0.13	0.15	0.16	0.16	0.16	0.17	0.18
	SD Diff					0.03	0.03	0.03	0.04	0.22	0.28	0.31	0.31	0.32	0.34	0.34	0.34	0.35	0.38
	$\lvert$Diff$\rvert$ > 1%/yr					0.02	0.03	0.02	0.04	0.25	0.26	0.27	0.28	0.37	0.39	0.40	0.40	0.42	0.45
	Not same					0.57	0.58	0.47	0.48	0.82	0.92	0.91	0.91	0.93	0.94	0.94	0.94	0.94	0.97
2009	Mean $\lvert$Diff$\rvert$						0.01	0.02	0.02	0.10	0.12	0.13	0.13	0.16	0.17	0.17	0.17	0.18	0.19
	SD Diff						0.02	0.03	0.04	0.22	0.28	0.31	0.31	0.33	0.35	0.35	0.35	0.36	0.39
	$\lvert$Diff$\rvert$ > 1%/yr						0.02	0.03	0.05	0.28	0.29	0.30	0.31	0.40	0.42	0.42	0.43	0.45	0.47
	Not same						0.37	0.56	0.58	0.88	0.93	0.93	0.94	0.94	0.95	0.96	0.96	0.94	0.97
2010	Mean $\lvert$Diff$\rvert$	0.21	0.19	0.18				0.02	0.02	0.10	0.11	0.13	0.13	0.16	0.16	0.17	0.17	0.18	0.19
	SD Diff	0.32	0.28	0.28				0.03	0.03	0.22	0.28	0.30	0.30	0.33	0.34	0.34	0.35	0.36	0.38
	$\lvert$Diff$\rvert$ > 1%/yr	0.65	0.62	0.60				0.01	0.03	0.27	0.28	0.30	0.31	0.39	0.41	0.41	0.42	0.44	0.48
	Not same	0.65	0.62	0.60				0.56	0.59	0.88	0.93	0.93	0.94	0.94	0.95	0.94	0.94	0.95	0.97
2011	Mean $\lvert$Diff$\rvert$								0.01	0.09	0.11	0.12	0.12	0.16	0.16	0.16	0.16	0.18	0.18
	SD Diff								0.02	0.21	0.28	0.30	0.30	0.33	0.34	0.34	0.34	0.36	0.38
	$\lvert$Diff$\rvert$ > 1%/yr								0.01	0.24	0.25	0.28	0.28	0.38	0.40	0.41	0.41	0.43	0.46
	Not same								0.47	0.86	0.93	0.92	0.92	0.93	0.94	0.95	0.95	0.96	0.97
2012	Mean $\lvert$Diff$\rvert$	0.22	0.20	0.19			0.16			0.09	0.11	0.12	0.12	0.15	0.16	0.16	0.16	0.17	0.18
	SD Diff	0.33	0.29	0.29			0.24			0.21	0.28	0.30	0.30	0.32	0.34	0.34	0.34	0.36	0.38
	$\lvert$Diff$\rvert$ > 1%/yr	0.67	0.64	0.63			0.57			0.23	0.25	0.27	0.27	0.38	0.39	0.41	0.41	0.42	0.46
	Not same	0.67	0.64	0.63			0.57			0.81	0.90	0.90	0.89	0.92	0.91	0.92	0.92	0.93	0.96
2013	Mean $\lvert$Diff$\rvert$										0.05	0.08	0.08	0.12	0.12	0.13	0.13	0.14	0.15
	SD Diff										0.13	0.19	0.19	0.23	0.25	0.25	0.25	0.27	0.29
	$\lvert$Diff$\rvert$ > 1%/yr										0.13	0.19	0.20	0.34	0.35	0.37	0.37	0.38	0.43
	Not same										0.77	0.81	0.81	0.87	0.91	0.92	0.92	0.93	0.95
2014	Mean $\lvert$Diff$\rvert$	0.22	0.20	0.19			0.21		0.15			0.07	0.07	0.11	0.12	0.12	0.12	0.13	0.15
	SD Diff	0.32	0.29	0.30			0.31		0.24			0.18	0.18	0.22	0.24	0.24	0.24	0.26	0.27
	$\lvert$Diff$\rvert$ > 1%/yr	0.67	0.65	0.62			0.65		0.53			0.17	0.18	0.31	0.33	0.34	0.34	0.36	0.41
	Not same	0.67	0.65	0.62			0.65		0.53			0.58	0.69	0.79	0.89	0.90	0.89	0.92	0.95
2015	Mean $\lvert$Diff$\rvert$	0.38	0.37	0.37			0.43		0.40		0.59		0.01	0.05	0.06	0.06	0.06	0.08	0.10
	SD Diff	0.68	0.70	0.69			0.79		0.77		1.47		0.02	0.13	0.15	0.15	0.16	0.19	0.20
	$\lvert$Diff$\rvert$ > 1%/yr	0.76	0.75	0.75			0.78		0.73		0.71		0.01	0.14	0.16	0.18	0.18	0.21	0.32
	Not same	0.76	0.75	0.75			0.78		0.73		0.71		0.29	0.41	0.61	0.61	0.65	0.68	0.92
2016	Mean $\lvert$Diff$\rvert$	0.38	0.37	0.37			0.43		0.40		0.59	0.00		0.05	0.06	0.06	0.06	0.08	0.10
	SD Diff	0.68	0.70	0.69			0.79		0.77		1.47	0.01		0.13	0.15	0.15	0.16	0.19	0.20
	$\lvert$Diff$\rvert$ > 1%/yr	0.76	0.75	0.75			0.78		0.73		0.71	0.00		0.15	0.17	0.18	0.19	0.21	0.32
	Not same	0.76	0.75	0.75			0.78		0.73		0.71	0.00		0.39	0.60	0.60	0.63	0.68	0.92
2017	Mean $\lvert$Diff$\rvert$	0.53	0.52	0.52			0.62		0.59		0.70	0.22	0.22		0.02	0.02	0.02	0.06	0.08
	SD Diff	0.95	0.94	0.93			1.16		1.12		1.59	0.72	0.72		0.08	0.09	0.09	0.14	0.16
	$\lvert$Diff$\rvert$ > 1%/yr	0.80	0.79	0.78			0.81		0.77		0.74	0.25	0.25		0.03	0.05	0.05	0.17	0.29
	Not same	0.80	0.79	0.78			0.81		0.77		0.74	0.25	0.25		0.56	0.54	0.59	0.66	0.91
2018	Mean $\lvert$Diff$\rvert$															0.01	0.01	0.05	0.08
	SD Diff															0.04	0.04	0.12	0.14
	$\lvert$Diff$\rvert$ > 1%/yr															0.02	0.03	0.14	0.27
	Not same															0.38	0.47	0.66	0.91
2019	Mean $\lvert$Diff$\rvert$																0.01	0.04	0.07
	SD Diff																0.02	0.11	0.13
	$\lvert$Diff$\rvert$ > 1%/yr																0.01	0.13	0.26
	Not same																0.23	0.61	0.90
2020	Mean $\lvert$Diff$\rvert$	0.54	0.53	0.53			0.64		0.60		0.73	0.27	0.27	0.09				0.04	0.07
	SD Diff	0.95	0.95	0.94			1.17		1.13		1.61	0.76	0.76	0.23				0.11	0.13
	$\lvert$Diff$\rvert$ > 1%/yr	0.80	0.80	0.79			0.81		0.78		0.78	0.39	0.39	0.28				0.12	0.25
	Not same	0.80	0.80	0.79			0.81		0.78		0.78	0.39	0.39	0.28				0.57	0.90
2021	Mean $\lvert$Diff$\rvert$	0.63	0.62	0.61			0.73		0.70		0.79	0.35	0.35	0.23			0.19		0.05
	SD Diff	1.09	1.07	1.07			1.33		1.30		1.68	0.94	0.94	0.57			0.52		0.09
	$\lvert$Diff$\rvert$ > 1%/yr	0.82	0.81	0.81			0.83		0.80		0.80	0.45	0.45	0.43			0.35		0.17
	Not same	0.82	0.81	0.81			0.83		0.80		0.80	0.45	0.45	0.43			0.35		0.88
2022	Mean $\lvert$Diff$\rvert$	0.64	0.62	0.63			0.75		0.74		0.83	0.49	0.49	0.38			0.36	0.25	
	SD Diff	1.10	1.08	1.07			1.33		1.31		1.70	1.02	1.02	0.71			0.68	0.47	
	$\lvert$Diff$\rvert$ > 1%/yr	0.82	0.82	0.82			0.84		0.84		0.83	0.73	0.73	0.71			0.69	0.64	
	Not same	0.82	0.82	0.82			0.84		0.84		0.83	0.73	0.73	0.71			0.69	0.64	

Table II
Mutual fund alphas estimated using 2005 and 2022 factor vintages: Varying sample periods and estimation horizons

This table reports statistics for alphas of individual mutual funds estimated using 2005 and 2022 factor vintages. Three-factor alphas are estimated at the end of every calendar year using one, three, or five years of monthly data (Panels A, B, and C, respectively). The columns show results in subperiods and for the full sample. Reported are annualized means and standard deviations of alphas from the two vintages, as well as of the difference in alphas. |Difference| > 1% indicates the proportion of estimated alpha differences that exceed 1% in magnitude. The row labeled Lose significance shows the proportion of alphas that are significant in one vintage but not in the other.

	1980s	1990s	2000s	1980-2004
A. 1-year estimation horizon				
Mean 2005	1.56	-1.86	-2.41	-1.73
Mean 2022	1.30	-2.49	-2.18	-1.93
Mean difference	-0.25	-0.63	0.23	-0.20
SD 2005	9.69	15.79	11.85	13.59
SD 2022	9.79	16.32	11.73	13.83
SD difference	0.87	2.50	2.28	2.31
|Diff| > 1%	0.18	0.56	0.49	0.49
Lose significance	0.18	0.36	0.33	0.33
B. 3-year estimation horizon				
Mean 2005	1.62	-1.81	-0.65	-0.89
Mean 2022	1.31	-2.07	-0.52	-0.97
Mean difference	-0.32	-0.26	0.13	-0.08
SD 2005	6.67	7.77	8.13	7.90
SD 2022	6.68	7.76	8.09	7.88
SD difference	0.38	0.79	0.59	0.70
|Diff| > 1%	0.04	0.15	0.09	0.11
Lose significance	0.14	0.23	0.15	0.18
C. 5-year estimation horizon				
Mean 2005	1.50	-1.19	-0.42	-0.51
Mean 2022	1.15	-1.42	-0.44	-0.66
Mean difference	-0.35	-0.22	-0.03	-0.14
SD 2005	5.60	6.11	6.51	6.30
SD 2022	5.58	6.12	6.55	6.33
SD difference	0.31	0.47	0.56	0.52
|Diff| > 1%	0.03	0.05	0.08	0.06
Lose significance	0.13	0.17	0.19	0.17

Table III
Mutual fund alphas estimated using different factor vintages

This table reports statistics for alphas of individual mutual funds estimated using different factor vintages. Three-factor alphas are estimated in every calendar year using monthly data. Reported are annualized means and standard deviations of alphas from the two vintages (Panel A), as well as of the difference in alphas (B). |Diff| > 1% indicates the proportion of estimated alpha differences that exceed 1% in magnitude. The row labeled Lose signif shows the proportion of alphas that are significant in one vintage but not in the other.

		Vintage 1																	
Vintage 2	Variable	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
A. Moments of alphas in different factor vintages																			
	Mean	-1.73	-1.84	-1.83	-1.84	-1.86	-1.86	-1.85	-1.85	-1.92	-1.90	-1.91	-1.88	-2.06	-2.05	-2.05	-2.06	-1.90	-1.93
	SD	13.6	13.5	13.5	13.5	13.5	13.5	13.5	13.5	13.6	13.6	13.6	13.5	13.8	13.8	13.8	13.8	13.8	13.8
B. Statistics for differences in alphas between factor vintages																			
2006	Mean	-0.11																	
	SD	1.22																	
	Lose signif	0.13																	
	|Diff| > 1%	0.18																	
2007	Mean	-0.10	0.00																
	SD	1.22	0.16																
	Lose signif	0.13	0.02																
	|Diff| > 1%	0.19	0.00																
2008	Mean	-0.11	-0.01	-0.01															
	SD	1.28	0.17	0.16															
	Lose signif	0.15	0.03	0.02															
	|Diff| > 1%	0.19	0.00	0.00															
2009	Mean	-0.13	-0.03	-0.03	-0.03														
	SD	1.18	0.31	0.31	0.30														
	Lose signif	0.13	0.04	0.04	0.04														
	|Diff| > 1%	0.18	0.02	0.02	0.01														
2010	Mean	-0.13	-0.03	-0.03	-0.02	0.00													
	SD	1.18	0.34	0.35	0.31	0.16													
	Lose signif	0.13	0.05	0.05	0.04	0.02													
	|Diff| > 1%	0.18	0.03	0.02	0.02	0.00													
2011	Mean	-0.12	-0.02	-0.02	-0.02	0.01	0.01												
	SD	1.29	0.27	0.29	0.22	0.26	0.22												
	Lose signif	0.15	0.04	0.04	0.03	0.04	0.03												
	|Diff| > 1%	0.19	0.01	0.01	0.01	0.01	0.01												
2012	Mean	-0.12	-0.02	-0.02	-0.02	0.00	0.00	0.01											
	SD	1.29	0.29	0.31	0.26	0.28	0.25	0.14											
	Lose signif	0.15	0.05	0.05	0.04	0.05	0.05	0.03											
	|Diff| > 1%	0.20	0.02	0.01	0.01	0.01	0.01	0.00											
2013	Mean	-0.19	-0.05	-0.10	0.00	0.05	0.16	0.23	0.16										
	SD	1.56	0.77	0.79	0.79	0.89	1.02	1.00	1.01										
	Lose signif	0.22	0.18	0.20	0.20	0.21	0.21	0.23	0.24										
	|Diff| > 1%	0.34	0.18	0.19	0.21	0.25	0.27	0.30	0.31										
2014	Mean	-0.17	-0.03	-0.08	0.01	0.07	0.18	0.24	0.17	0.01									
	SD	1.56	0.78	0.79	0.80	0.90	1.03	1.01	1.02	0.26									
	Lose signif	0.24	0.19	0.22	0.21	0.22	0.23	0.24	0.25	0.04									
	|Diff| > 1%	0.36	0.18	0.19	0.21	0.25	0.28	0.31	0.31	0.01									
2015	Mean	-0.18	-0.04	-0.09	0.01	0.06	0.17	0.23	0.18	0.02	0.01								
	SD	1.55	0.78	0.79	0.81	0.91	1.04	1.02	1.03	0.27	0.14								
	Lose signif	0.24	0.19	0.21	0.21	0.22	0.23	0.24	0.25	0.05	0.03								
	|Diff| > 1%	0.36	0.18	0.19	0.21	0.25	0.27	0.30	0.31	0.01	0.00								
2016	Mean	-0.15	-0.02	-0.06	0.03	0.07	0.17	0.24	0.18	0.02	0.01	0.00							
	SD	1.52	0.79	0.79	0.79	0.92	1.03	1.00	1.02	0.34	0.20	0.19							
	Lose signif	0.24	0.19	0.21	0.21	0.22	0.22	0.24	0.25	0.06	0.03	0.03							
	|Diff| > 1%	0.36	0.17	0.18	0.19	0.25	0.27	0.30	0.30	0.02	0.00	0.01							
2017	Mean	-0.33	-0.27	-0.32	-0.31	-0.30	-0.23	-0.15	-0.22	-0.35	-0.24	-0.25	-0.25						
	SD	2.40	1.77	1.74	1.89	2.04	1.99	1.94	1.94	1.65	1.79	1.77	1.71						
	Lose signif	0.29	0.26	0.29	0.29	0.30	0.31	0.32	0.32	0.22	0.24	0.23	0.23						
	|Diff| > 1%	0.47	0.41	0.42	0.44	0.46	0.46	0.47	0.47	0.34	0.35	0.34	0.33						
2018	Mean	-0.32	-0.26	-0.31	-0.30	-0.29	-0.22	-0.14	-0.21	-0.34	-0.24	-0.24	-0.24	0.01					
	SD	2.40	1.78	1.76	1.91	2.06	2.01	1.95	1.95	1.67	1.81	1.78	1.72	0.15					
	Lose signif	0.29	0.26	0.29	0.30	0.31	0.31	0.32	0.32	0.22	0.25	0.23	0.23	0.03					
	|Diff| > 1%	0.47	0.41	0.42	0.44	0.47	0.47	0.48	0.48	0.34	0.36	0.35	0.34	0.00					
2019	Mean	-0.32	-0.27	-0.32	-0.31	-0.30	-0.23	-0.15	-0.22	-0.34	-0.24	-0.24	-0.24	0.01	0.00				
	SD	2.38	1.77	1.75	1.90	2.04	1.99	1.94	1.95	1.66	1.83	1.80	1.74	0.15	0.14				
	Lose signif	0.30	0.26	0.29	0.30	0.31	0.31	0.32	0.32	0.23	0.25	0.24	0.23	0.04	0.02				
	|Diff| > 1%	0.47	0.40	0.42	0.44	0.46	0.46	0.47	0.47	0.34	0.36	0.35	0.33	0.00	0.00				
2020	Mean	-0.33	-0.27	-0.32	-0.31	-0.30	-0.23	-0.15	-0.22	-0.34	-0.23	-0.23	-0.24	0.01	0.00	0.00			
	SD	2.38	1.76	1.74	1.88	2.03	1.99	1.94	1.94	1.65	1.82	1.79	1.73	0.16	0.15	0.07			
	Lose signif	0.30	0.26	0.29	0.30	0.31	0.31	0.32	0.32	0.22	0.25	0.23	0.23	0.03	0.02	0.01			
	|Diff| > 1%	0.46	0.40	0.42	0.44	0.46	0.46	0.47	0.47	0.34	0.36	0.35	0.34	0.00	0.00	0.00			
2021	Mean	-0.17	-0.07	-0.15	-0.33	-0.33	-0.22	-0.14	-0.21	-0.33	-0.21	-0.22	-0.22	0.03	0.03	0.01	0.01		
	SD	2.27	2.16	2.12	3.06	3.05	2.97	2.93	2.91	2.70	2.78	2.74	2.63	1.62	1.58	1.55	1.52		
	Lose signif	0.33	0.33	0.35	0.37	0.37	0.37	0.38	0.38	0.29	0.31	0.29	0.28	0.16	0.16	0.16	0.15		
	|Diff| > 1%	0.48	0.47	0.48	0.51	0.52	0.52	0.54	0.53	0.40	0.42	0.40	0.39	0.20	0.19	0.18	0.18		
2022	Mean	-0.20	-0.11	-0.16	-0.32	-0.31	-0.19	-0.11	-0.19	-0.31	-0.19	-0.19	-0.20	0.05	0.04	0.03	0.03	0.02	
	SD	2.31	2.16	2.12	2.92	2.93	2.86	2.82	2.80	2.57	2.71	2.67	2.57	1.54	1.51	1.47	1.45	0.40	
	Lose signif	0.33	0.33	0.35	0.37	0.37	0.37	0.38	0.38	0.29	0.31	0.28	0.28	0.16	0.16	0.16	0.16	0.06	
	|Diff| > 1%	0.49	0.48	0.49	0.52	0.53	0.53	0.55	0.54	0.40	0.42	0.41	0.40	0.22	0.22	0.21	0.20	0.03	

Table IV
Mutual fund alphas estimated using 2005 and 2022 factor vintages: Varying fund characteristics

This table reports statistics for alphas of individual mutual funds estimated using 2005 and 2022 factor vintages. Three-factor alphas are estimated in every calendar year using monthly data. Funds are grouped into quintiles on the basis of characteristics shows in panel headings using most recent characteristic available prior to the alpha estimation window. Reported are annualized means and standard deviations of alphas from the two vintages, as well as of the difference in alphas. |Difference| > 1% indicates the proportion of estimated alpha differences that exceed 1% in magnitude. The row labeled Lose significance shows the proportion of alphas that are significant in one vintage but not in the other.

	Low	Q2	Med	Q4	High
A. Fund size					
Mean 2005	-0.39	-1.72	-2.10	-2.16	-2.32
Mean 2022	-0.58	-1.86	-2.31	-2.40	-2.55
Mean difference	-0.19	-0.14	-0.21	-0.24	-0.23
SD 2005	15.12	14.28	13.59	12.98	11.70
SD 2022	15.38	14.63	13.83	13.16	11.81
SD difference	2.36	2.32	2.33	2.26	2.30
|Difference| > 1%	0.49	0.49	0.49	0.49	0.48
Lose significance	0.29	0.32	0.32	0.33	0.39
B. Size factor exposure					
Mean 2005	-1.10	-2.15	-3.66	-2.07	-1.19
Mean 2022	-1.01	-2.22	-3.79	-2.27	-1.70
Mean difference	0.09	-0.06	-0.13	-0.20	-0.51
SD 2005	9.82	11.24	14.32	15.17	13.85
SD 2022	9.97	11.47	14.70	15.71	13.50
SD difference	1.42	1.77	2.29	2.62	2.88
|Difference| > 1%	0.38	0.40	0.47	0.56	0.60
Lose significance	0.37	0.36	0.33	0.32	0.31
C. Value factor exposure					
Mean 2005	-0.84	-2.41	-2.43	-2.12	-2.35
Mean 2022	-1.59	-2.76	-2.52	-2.02	-2.10
Mean difference	-0.75	-0.35	-0.09	0.11	0.25
SD 2005	14.72	12.20	12.13	12.72	13.35
SD 2022	14.40	12.30	12.45	13.11	13.95
SD difference	2.65	2.05	2.03	2.04	2.38
|Difference| > 1%	0.56	0.46	0.43	0.46	0.51
Lose significance	0.32	0.31	0.30	0.36	0.40

Table V
Stock portfolio alphas estimated using different factor vintages

This table reports statistics for alphas of characteristic-sorted value-weighted decile portfolios from Kenneth French's website. The earliest vintage of both the factors and the portfolios ('vintage 1') is compared to the latest vintage of each ('vintage 2'). The earliest vintages of factors and portfolios sorted on market equity and book-to-market ratio is 2005. For runup portfolios, the earliest vintage is 2007, and for all other portfolios it is 2015. The latest vintage is 2022 for factors and all portfolios. Three-factor alphas are estimated at the end of every calendar year using five years of monthly data. Reported are annualized means and standard deviations of alphas from the two vintages, as well as of the difference in alphas. |Difference| > 1% indicates the proportion of estimated alpha differences that exceed 1% in magnitude. The row labeled Lose significance shows the proportion of alphas that are significant in one vintage but not in the other.

	Mkt equity	BM ratio	Runup	Profitability	Investment
Mean vintage 1	0.16	0.03	-0.50	-0.26	0.68
Mean vintage 2	-0.07	-0.11	-0.68	-0.34	0.51
Mean difference	-0.23	-0.15	-0.18	-0.08	-0.17
SD vintage 1	2.09	2.94	5.43	3.12	2.87
SD vintage 2	2.03	3.05	5.54	3.11	2.65
SD difference	0.89	1.59	1.16	1.72	0.95
|Difference| > 1%	0.11	0.32	0.12	0.34	0.23
Lose significance	0.49	0.52	0.13	0.62	0.33

	Accruals	Beta	Issuance	Variance	Res. variance
Mean vintage 1	0.67	0.43	-0.24	0.00	-0.14
Mean vintage 2	0.54	0.28	-0.29	-0.17	-0.29
Mean difference	-0.13	-0.15	-0.05	-0.17	-0.16
SD vintage 1	3.41	2.80	4.00	4.78	4.86
SD vintage 2	3.15	2.87	3.80	4.84	4.93
SD difference	1.28	0.76	1.34	0.67	1.05
|Difference| > 1%	0.25	0.12	0.23	0.10	0.20
Lose significance	0.44	0.19	0.27	0.08	0.22

Table VI
***T*-statistics of anomaly strategies**

This table reports the results of tests examining the t-statistics of long-short anomaly strategies across different vintages of the French factors. We calculate the t-statistic for each strategy using each vintage of the Fama-French factors and compute the range of the t-statistics. We present the average, standard deviation, maximum and minimum range for the anomalies, along with the number of strategies. Panel A presents these statistics for the entire sample, while panel B presents these statistics for the "marginal" anomalies (those that have at least one t-statistic that is between 2 and 2.5 (inclusive) across the various vintages of the Fama-French factors). Panel C presents the number and proportion of "marginal" anomalies that lose statistical significance using at least one vintage of the Fama-French factors. The 549 long-short strategies are from Hou, Xue, and Zhang (2020), Haddad, Kozak, and Santosh (2020), Chen and Zimmermann (2021), and Dong, Li, Rapach, Zhou (2022).

A. All anomalies

Data source	N	Mean	Std Dev	Minimum	Maximum
Hou, Xue, Zhang (2020)	187	0.141	0.072	0.031	0.352
Haddad, Kozak, and Santosh (2020)	55	0.161	0.089	0.038	0.417
Chen and Zimmermann (2021)	207	0.142	0.080	0.030	0.565
Dong, Li, Rapach, Zhou (2022)	100	0.171	0.078	0.024	0.340
Combined	549	0.149	0.079	0.024	0.565

B. Marginal anomalies (at least one $2 \leq t \leq 2.5$)

Data source	N	Mean	Std Dev	Minimum	Maximum
Hou, Xue, Zhang (2020)	26	0.122	0.071	0.031	0.292
Haddad, Kozak, and Santosh (2020)	6	0.234	0.135	0.040	0.417
Chen and Zimmermann (2021)	14	0.147	0.076	0.064	0.265
Dong, Li, Rapach, Zhou (2022)	7	0.202	0.069	0.107	0.328
Combined	53	0.152	0.088	0.031	0.417

C. Marginal anomalies losing significance

Data source	Anomalies	Lose significance	Prop. lose significance
Hou, Xue, Zhang (2020)	26	5	0.192
Haddad, Kozak, and Santosh (2020)	6	2	0.333
Chen and Zimmermann (2021)	14	3	0.214
Dong, Li, Rapach, Zhou (2022)	7	1	0.143
Combined	53	11	0.208

Table VII
Tests of equality of squared Sharpe ratios

This table reports results of pairwise tests of equality of the squared Sharpe ratios of the three-factor model with different factor vintages. Panel A reports the difference between the bias-adjusted sample squared Sharpe ratios of the models based on vintages showing in columns and rows. Panel B shows the associated p-values.

	Vintage 1																
Vintage 2	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
A. Differences in squared Sharpe ratio																	
2005	0.000	0.000	0.000	0.001	0.001	0.000	0.000	0.001	0.001	0.001	0.001	0.002	0.002	0.002	0.001	0.002	0.001
2006		0.000	0.000	0.001	0.001	0.001	0.000	0.001	0.001	0.001	0.001	0.002	0.002	0.002	0.002	0.002	0.001
2007			0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.001	0.001	0.002	0.002	0.001	0.001	0.002	0.001
2008				0.000	0.000	0.000	0.000	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
2009					0.000	0.000	0.000	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
2010						0.000	0.000	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
2011							0.000	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
2012								0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
2013									0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2014										0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000
2015											0.000	0.000	0.000	0.000	0.000	0.000	0.000
2016												0.000	0.000	0.000	0.000	0.000	0.000
2017													0.000	0.000	0.000	0.000	0.000
2018														0.000	0.000	0.000	0.000
2019															0.000	0.000	0.000
2020																0.000	0.000
2021																	0.000
B. p-values																	
2005	0.921	0.940	0.651	0.485	0.510	0.579	0.653	0.109	0.160	0.159	0.156	0.111	0.145	0.162	0.170	0.139	0.198
2006		0.483	0.126	0.056	0.062	0.078	0.122	0.027	0.062	0.064	0.062	0.056	0.068	0.078	0.081	0.093	0.115
2007			0.166	0.059	0.070	0.102	0.184	0.032	0.079	0.085	0.082	0.069	0.087	0.100	0.105	0.113	0.147
2008				0.047	0.093	0.279	0.946	0.065	0.162	0.159	0.154	0.113	0.161	0.184	0.194	0.192	0.257
2009					0.544	0.253	0.125	0.117	0.266	0.244	0.238	0.160	0.227	0.259	0.272	0.255	0.348
2010						0.294	0.130	0.110	0.252	0.232	0.226	0.154	0.218	0.249	0.261	0.246	0.334
2011							0.172	0.091	0.214	0.201	0.195	0.138	0.193	0.221	0.232	0.225	0.302
2012								0.074	0.177	0.169	0.164	0.122	0.169	0.193	0.203	0.201	0.267
2013									0.520	0.923	0.941	0.568	0.785	0.875	0.908	0.760	0.979
2014										0.749	0.733	0.372	0.563	0.648	0.681	0.569	0.819
2015											0.825	0.335	0.612	0.738	0.788	0.631	0.962
2016												0.358	0.634	0.760	0.809	0.650	0.977
2017													0.368	0.250	0.223	0.729	0.378
2018														0.352	0.317	0.891	0.598
2019															0.657	0.729	0.738
2020																0.670	0.794
2021																	0.276

Internet Appendix to

Noisy Factors

Intended for online publication.

IA.A. Additional Tables and Figures

Table IA.I
Tests of equality of squared Sharpe ratios: fixed-code vs French factors

This table reports results of pairwise tests of equality of the squared Sharpe ratios of the fixed-code and the French factors. All tests use data common to all vintages.

Vintage	Difference in squared Sharpe ratio	p-value of difference
2010	0.000	0.568
2011	0.000	0.548
2016	0.000	0.547
2020	0.000	0.762
2021	0.000	0.783
2022	0.000	0.959

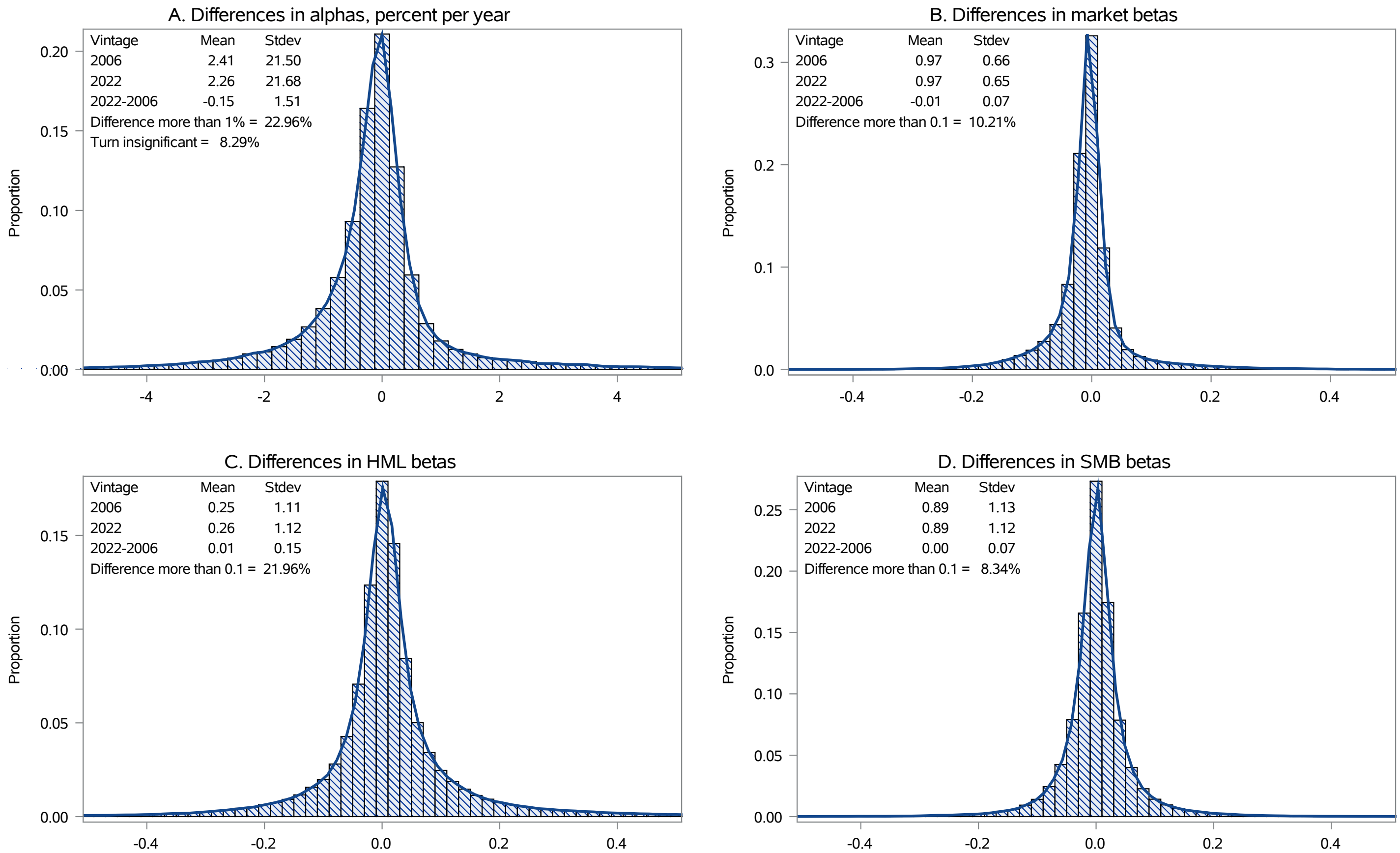


Figure IA.1. Differences in contemporaneous stock-level alphas and betas: 2006 vs 2022 factor vintages

This figure plots histograms and kernel densities of differences in alphas (Panel A) and betas (B, C, D) of individual stocks estimated using 2006 and 2022 factor vintages of factors and CRSP data. Specifically, the 2006 (2022) estimation uses information available to an empiricist at the time: the factors downloaded in 2006 (2022) and CRSP data through the end of 2005 (2021). Alphas, in percent per year, and betas are estimated at the end of every calendar year using three-factor regressions on five years of monthly data. Top left of each panel reports means and standard deviations of estimates from the two vintages and of their differences. It also shows the fraction of observations with absolute differences above a certain threshold and in Panel A the proportion of alphas that are significant in one vintage but not in the other.

IA.B. Which Variables Are the Noisiest?

While changes to the underlying data explain only a small part of the changes in the French factors, our fixed-code factors show that data changes have non-trivial effects on monthly factor returns, especially in the first decades of the sample.

We contacted WRDS to inquire about the possibility of getting a larger time series of historical data vintages. We received the following in response: "We do not have archived versions of the data available on WRDS. It is very rare for historical data to be changed in both the CRSP and Compustat products. You should be able to run a query and receive the same historical data." Compustat offers the 'Snapshot' dataset, which can be used to construct historical vintages of the Compustat fundamentals file, but the data starts in 1970s. No such historical dataset vintages are available for CRSP or for the linking table. Accordingly, we rely on archived data from previous projects.

We examine how frequently the variables used to construct the factors have changed and present the results in Figure IA.2. We calculate the percent of observations that are different between the 2010 and 2022 vintages of Compustat (for book equity, panel A), CRSP (for market equity and returns in panels B and C, respectively), and the CRSP/Compustat Merged Database Linking Table (for CRSP/Compustat linkages, panel D) in each year. There are very few changes to book equity values except for the last handful of years in the sample. These could be due to accounting revisions or restatements. CRSP's data cleanup led to many revisions to the market value of equity prior to the 1950s. After that, there are very few changes to market equity except for the last year in the sample, which may also reflect corrections. There are also very few revisions to monthly returns. In contrast, there are substantial revisions to the CRSP/Compustat linking table between 2010 and 2022. On average more than 5% of linkage observations are not the same between vintages in each year.[28]

The analysis in this section has shown that there are frequent, retroactive revisions to the French factors. These revisions can be large in magnitude and materially change the cumulative returns of the HML and SMB factors. Data updates can explain a substantial portion of these revisions in the first half of the sample but cannot explain the revisions in the more commonly studied post-1964 half of the sample.

[28]To the best of our knowledge, neither CRSP nor Compustat provides any indication that the linking table has changed over time, nor are historical vintages made available. Many of these changes affect records that at some point were classified as having been confirmed by research.

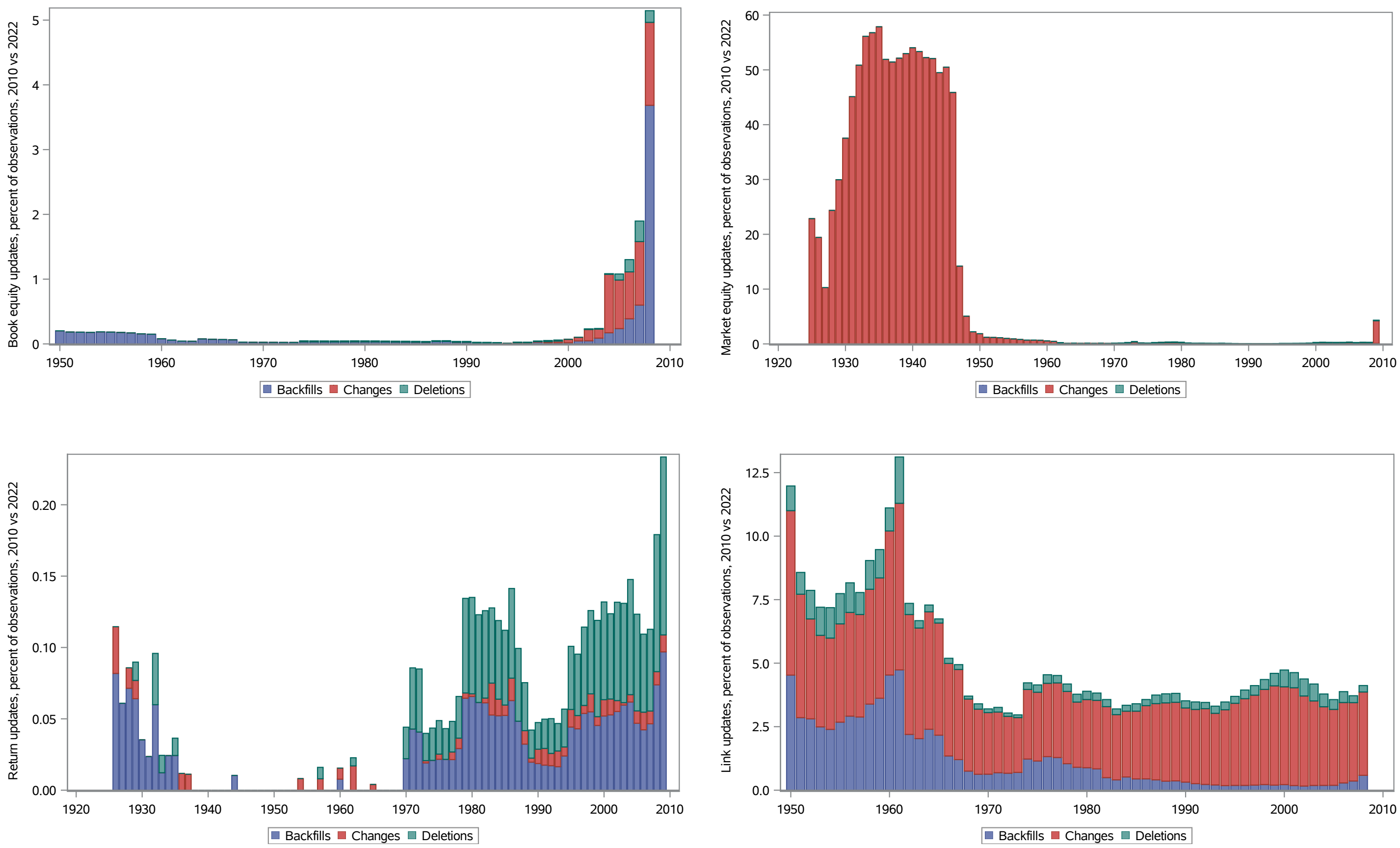


Figure IA.2. Revisions to Compustat and CRSP data

This figure examines the proportion of observations in a given year that was not the same in the 2010 and 2022 versions of CRSP, Compustat or the Compustat-CRSP Linking Table. Panel A presents changes to book equity. Panel B presents changes to market equity. Panel C presents changes to equity returns. Panel D presents the proportion of firm observations that are associated with a different entity in 2010 and 2022 by the CRSP-Compustat Linking Table.

IA.C. Cumulative Changes to Factor Returns

In this section, we discuss the changes in cumulative returns for the French factors (the market, HML and SMB, respectively) and the fixed code factors. The black lines in Panels B, D, and F of Figure 2 present the changes in cumulative returns for the French factors (the market, HML and SMB, respectively). The gray line presents the same information for the fixed-code factors. A value of one for a particular year indicates that the cumulative return (from the beginning of the sample to that date) on the 2022 vintage of the factor is the same as the cumulative return on the 2010 vintage of the factor. This can be interpreted as the cumulative return on a long-short portfolio where the hypothetical investor buys the 2022 vintage and sells the 2010 vintage of the factor at the beginning of the sample period. For the market, both the black and the gray lines are relatively flat in the first part of the sample, indicating that neither changes to the data (captured by the fixed-code factors in gray) nor changes to the methodology used to construct the French factors had a substantial impact on the cumulative profitability of the market factor. From the early 1980s through the late 1990s, however, the black line rises sharply, indicating that the cumulative return on the 2022 vintage of the French factor rose relative to the 2010 vintage: by the end of 1999, the cumulative return on the long-short portfolio is 8%. It then drops even more sharply between 2000 and 2007, and oscillates more to end the sample with a cumulative return of 2%. In contrast, the fixed-code factor exhibits very little systematic variation in any subsample.

Panel D shows that the 2022 vintage of HML had considerably higher cumulative returns than the 2010 vintage. This is true for both the fixed-code factors and the French factors. Both the black and the gray lines rise sharply in the beginning of the sample. This likely reflects the changes to CRSP—which were released in January 2015—resulting from CRSP's review of, and corrections to, its shares outstanding data from 1926–1945. The increase in cumulative return over this time period was higher for the fixed-code factors than for the French factors. Both of these increases in cumulative returns deteriorate for the next several decades at roughly the same rate, and, in the case of the fixed-code factors, this deterioration continues until the end of the sample. In contrast, the French version of the 2022 vintage increases sharply again (relative to the 2010 vintage) between 1994 to 2004 before falling in the last few years of the sample.

Finally, we turn to SMB in panel F. The cumulative returns on both the fixed-code and the French SMB factors have deteriorated substantially between the 2010 vintages and the 2022 vintages, although that deterioration occurred at different rates and over parts of the sample. For the fixed-code SMB factor, all of the deterioration occurred in the first 15 years or so of the sample: the 2022 vintage of the fixed-code SMB factor declined by 20% relative to the 2010 vintage by the early 1940s before rebounding slightly through the 1940s. It remained at roughly that level through the end of the sample. Accordingly, the changes to the fixed-code factors are likely due to the same CRSP data cleanup that affected HML during that period. The CRSP cleanup appears to have had a much smaller effect on the French SMB factor, as the deterioration in the 1940s was about half as large. It then remained roughly flat until the mid-1990 before declining again through the end of the sample.

IA.D. GRS Tests

To implement the GRS tests, we treat each factor vintage as a "model," and compare the performance of each such model against the others. We use the sample common to all vintages, which spans July 1926 to June 2005. We consider two sets of test portfolios, the first of which is the widely used 25 value-weighted portfolios sorted on size and book-to-market. As we noted in Section III.B, portfolio vintages also undergo changes. We consider all available factor and portfolio vintages, resulting in 216 GRS tests (18 factor vintages $\times$ 12 portfolio vintages).

Panel A of Table IA.II summarizes the F-statistics from the 216 GRS tests. Lower values indicate superior "models." Rows and columns of the table correspond to different vintages of portfolios and factors, respectively, and the highlighted cells show statistics from tests that use contemporaneous portfolio and factor vintages. There are several striking features of the results. First, there is substantial variation in the test statistics across vintages, with the largest value (3.45) exceeding the smallest (2.46) by over 40%. In other words, "model" performance is highly unstable, varying substantially due solely to changes in vintages.

Second, the model does not systematically perform better when the factors and the portfolios are from the same vintage. That is, there is no evidence that the F-statistics are systematically lower along the highlighted diagonal than in off-diagonal cells. This is noteworthy given that these portfolios and factors are presumably formed using same versions of CRSP and Compustat data.

Third, the F-statistics tend to decline as we move down the diagonal of highlighted cells, with the lowest value appearing in the 2021 vintages. In other words, the changes do tend to result in better model performance, albeit not monotonically. Curiously, these improvements seem to be due to changes in *portfolio*—rather than factor—vintages: test statistics are fairly stable as we move across rows (perhaps even increasing slightly) but tend to decline as we move down the columns.

We repeat the analysis using the 17 value-weighted industry portfolios. The resulting F-statistics are summarized in Panel B of Table IA.II. Again, we see wide variation in F-statistics across vintages: from a low of 3.39 to a high of 4.41. One important difference, however, is in the analysis of contemporaneous vintages of factors and portfolios: here, the F-statistics increase from 3.46 (in 2005) to 4.25 (in 2022). While vintage updates appear to have resulted in better model performance when the 25 size and book-to-market portfolios are used as test assets in Panel A, the updates led to a deterioration in performance when using industry portfolio.

Overall, the results in Table IA.II are inconclusive. There is no systematic evidence that the factor updates represent an improvement in the model's performance in pricing test portfolios. The fluctuations, however, can significantly affect the interpretation of standard asset pricing tests. For example, suppose a researcher ran the tests using the 2006 factor and portfolio vintages. She would find that the 3-factor model performs about the "same"—in terms of the F-statistics—in pricing the 25 size and book-to-market portfolios as it does in pricing the 17 industry portfolios (3.32 versus 3.49). Were she to try to replicate this result in 2022, however, she would find a very different picture: using this vintage, the model performs markedly better for the size and book-to-market portfolios than the industry portfolios (F-statistics of 2.55 versus 4.25).

To investigate the effect of the changes to the factor construction methodology on the performance of the factors, we repeat the analysis using the fixed-code factors. While the fixed-code factors produce slightly larger F-statistics—indicating slightly worse performance—when pricing the 25 size and book-to-market portfolios, they perform somewhat better when pricing the 17 industry portfolios. We interpret this as evidence that the methodological changes are not improving the overall performance of the factors. We present the results in Table IA.III.

Table IA.II
GRS F-test statistics from different vintages of factors and portfolios

This table reports F-statistics from the GRS tests using different vintages of the three factors and portfolios. All tests use data common to all vintages: 07/1926-08/2005 for portfolios sorted on size and book-to-market, and 07/1926-05/2005 for industry portfolios. Highlighted cells indicate contemporaneous vintages of factors and portfolios.

Portfolio	Factor vintage																	
vintage	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
B. 25 portfolios sorted on size and book-to-market ratio																		
2005	2.89	2.90	2.88	2.90	2.91	2.91	2.91	2.90	2.95	2.91	2.93	2.94	2.97	2.96	2.96	2.95	2.95	2.98
2006	3.32	3.32	3.29	3.32	3.35	3.35	3.34	3.33	3.39	3.35	3.39	3.39	3.45	3.42	3.41	3.41	3.43	3.44
2007	3.20	3.20	3.18	3.22	3.25	3.25	3.23	3.22	3.29	3.25	3.30	3.30	3.36	3.33	3.32	3.31	3.33	3.35
2010	3.11	3.08	3.06	3.09	3.11	3.11	3.10	3.09	3.15	3.10	3.13	3.13	3.19	3.17	3.16	3.16	3.17	3.18
2012	3.11	3.09	3.07	3.10	3.13	3.12	3.11	3.10	3.17	3.12	3.16	3.16	3.21	3.19	3.18	3.18	3.20	3.21
2014	2.85	2.80	2.80	2.80	2.80	2.80	2.80	2.80	2.82	2.79	2.79	2.79	2.82	2.80	2.80	2.80	2.81	2.82
2015	2.77	2.66	2.66	2.65	2.65	2.65	2.64	2.64	2.69	2.65	2.62	2.62	2.65	2.63	2.63	2.63	2.65	2.65
2016	2.77	2.66	2.66	2.65	2.65	2.65	2.65	2.64	2.69	2.65	2.62	2.62	2.65	2.63	2.63	2.63	2.66	2.65
2017	2.87	2.77	2.79	2.76	2.75	2.75	2.75	2.75	2.73	2.71	2.68	2.68	2.71	2.70	2.70	2.70	2.73	2.72
2020	2.84	2.74	2.75	2.73	2.72	2.72	2.72	2.72	2.72	2.69	2.66	2.66	2.69	2.68	2.68	2.68	2.71	2.70
2021	2.63	2.55	2.57	2.54	2.53	2.53	2.53	2.53	2.50	2.48	2.46	2.46	2.47	2.46	2.46	2.46	2.49	2.48
2022	2.71	2.63	2.65	2.62	2.60	2.60	2.60	2.61	2.57	2.55	2.53	2.53	2.53	2.53	2.53	2.53	2.56	2.55
B. 17 Industry portfolios																		
2005	3.46	3.58	3.58	3.62	3.64	3.64	3.64	3.63	3.54	3.48	3.53	3.53	3.59	3.57	3.56	3.55	3.56	3.57
2006	3.39	3.49	3.49	3.53	3.55	3.55	3.55	3.54	3.47	3.41	3.45	3.46	3.51	3.49	3.48	3.48	3.48	3.48
2007	3.57	3.69	3.68	3.73	3.75	3.75	3.74	3.73	3.66	3.59	3.64	3.64	3.70	3.68	3.67	3.67	3.67	3.67
2010	4.23	4.35	4.35	4.39	4.41	4.41	4.41	4.40	4.34	4.28	4.35	4.35	4.40	4.39	4.38	4.37	4.37	4.38
2012	4.21	4.33	4.34	4.38	4.40	4.40	4.40	4.38	4.33	4.26	4.34	4.34	4.39	4.38	4.37	4.36	4.36	4.37
2014	4.05	4.17	4.17	4.21	4.23	4.23	4.23	4.21	4.17	4.12	4.20	4.20	4.25	4.23	4.22	4.21	4.22	4.22
2015	4.03	4.13	4.13	4.17	4.19	4.19	4.19	4.18	4.14	4.10	4.18	4.19	4.23	4.22	4.21	4.20	4.19	4.20
2016	4.03	4.13	4.13	4.17	4.19	4.19	4.19	4.18	4.14	4.10	4.18	4.19	4.23	4.22	4.21	4.20	4.19	4.20
2017	4.03	4.13	4.13	4.17	4.19	4.19	4.19	4.18	4.14	4.10	4.18	4.19	4.23	4.22	4.21	4.20	4.19	4.20
2020	4.03	4.13	4.13	4.17	4.19	4.19	4.19	4.18	4.15	4.11	4.19	4.19	4.24	4.22	4.21	4.20	4.20	4.20
2021	4.03	4.13	4.13	4.17	4.19	4.19	4.19	4.18	4.15	4.11	4.19	4.19	4.24	4.22	4.21	4.20	4.20	4.20
2022	4.07	4.17	4.17	4.21	4.23	4.23	4.23	4.22	4.19	4.15	4.23	4.23	4.28	4.26	4.25	4.24	4.24	4.25

Table IA.III
Differences in GRS *F*-statistics: fixed-code vs French factors

This table reports the differences in GRS F-statistics between the fixed-code and French factors. Each estimation is performed on the same vintage of test portfolios and risk factors. Positive values indicate that the F-statistics of the fixed-code factors are higher in a given vintage and indicate that the French factors explain more of the variation in the test asset returns. Negative values indicate that the F-statistics of the fixed-code factors are lower in a given vintage and indicate that the fixed-code factors explain more of the variation in the test asset returns. All tests use data common to all vintages.

Vintage	Size and book-to-market portfolios	Industry portfolios
2010	0.00	-0.26
2011	0.00	-0.28
2016	0.15	-0.16
2020	0.12	-0.19
2021	0.10	-0.18
2022	0.10	-0.18

IA.E. Portfolios of Random Stocks

Many settings in empirical finance involve analyzing portfolios rather individual stocks. While we have already shown the effect of vintage updates on mutual funds and characteristic-sorted portfolios, in this section we investigate their impact on portfolios of randomly selected stocks. At the beginning of each calendar year t, we split the cross-section of stocks into portfolios containing N random stocks. We hold these portfolios for five years without rebalancing, and calculate alphas of each portfolio during the t through $t+4$ holding period by regressing monthly value-weighted portfolio returns in excess of the risk-free rate on the three factors from different vintages.

We consider portfolios of $N = 10$, 30, 50, and 100 stocks, and repeat the process of creating random portfolios 1, 3, 5, or 10 times, respectively, each year. That is, when forming portfolios of 100 stocks at the beginning of year t, we randomly split the cross-section into portfolios containing 100 stocks each and then repeat this process 10 times. In a year when the cross section contains 3,000 stocks this procedure produces $(3,000/100) \times 10 = 300$ random portfolios.

Figure IA.3 shows histograms and kernel densities of the differences in alphas computed using the 2005 and 2022 factor vintages for the resulting portfolios. Changes in the factors meaningfully affect alphas even in these random positions. While the percentage of alphas that change by more than 100 bps falls relative to what we observed for individual stocks (26%), it remains substantial at 15%, 9%, 7%, and 4% for portfolios containing 10, 30, 50, and 100 stocks, respectively.

At the same time, the proportion of alphas that lose significance actually increases relative to the single-stock setting (9%), rising to over 15% for the 100-stock portfolios.

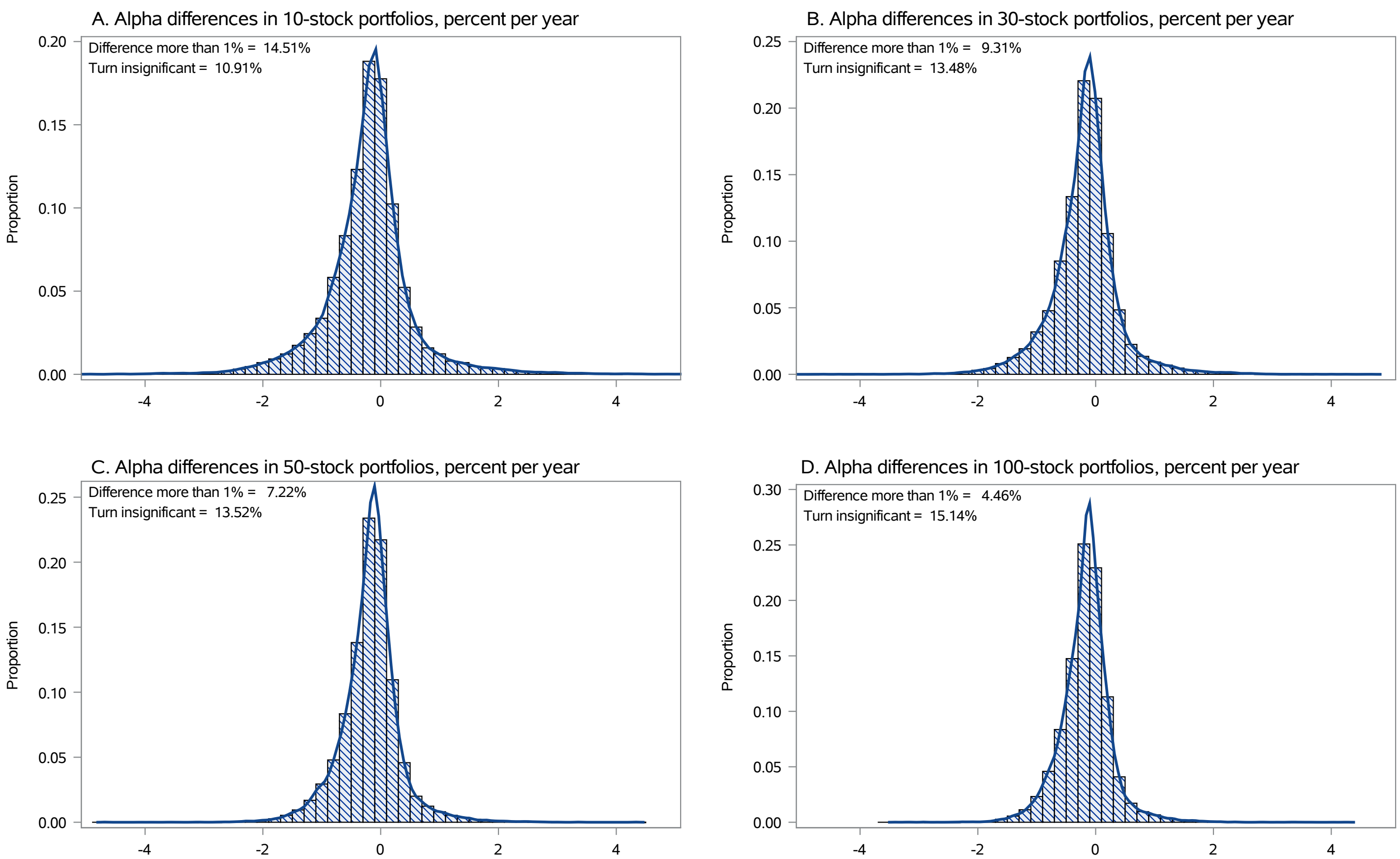


Figure IA.3. Differences in alphas of random stock portfolios: 2005 vs 2022 factor vintages
This figure plots histograms and kernel densities of differences in alphas estimated using 2005 and 2022 factor vintages for portfolios of randomly chosen stocks. At the beginning of every five-year period, random portfolios are created to contain between 10 (Panel A) and 100 (D) stocks. Alphas, in percent per year, are estimated for each portfolio and factor vintage using regressions of five years of monthly value-weighted portfolio returns in excess of the risk-free rate on the three factors. Top left of each panel shows the fraction of observations with absolute differences above a certain threshold and the proportion of alphas that are significant in one vintage but not in the other.

IA.F. Methodological Changes

In this section of the Internet Appendix, we provide the full text of the disclosed methodological changes from French's website. We then present vintage-by-vintage changes in French's factors as well as the fixed code factors for each of the three factors for each data vintage that we have.

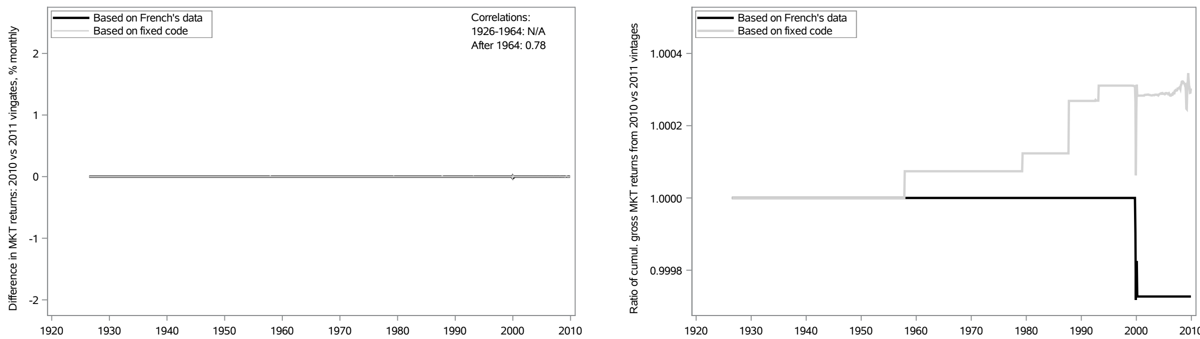
Table IA.IV

Description of Changes to Construction of the Fama-French Three Factors

This table summarizes the changes to the construction of the Fama-French monthly three factors as described on French's website. The last three columns indicate whether the described change is expected to affect monthly returns of the factors.

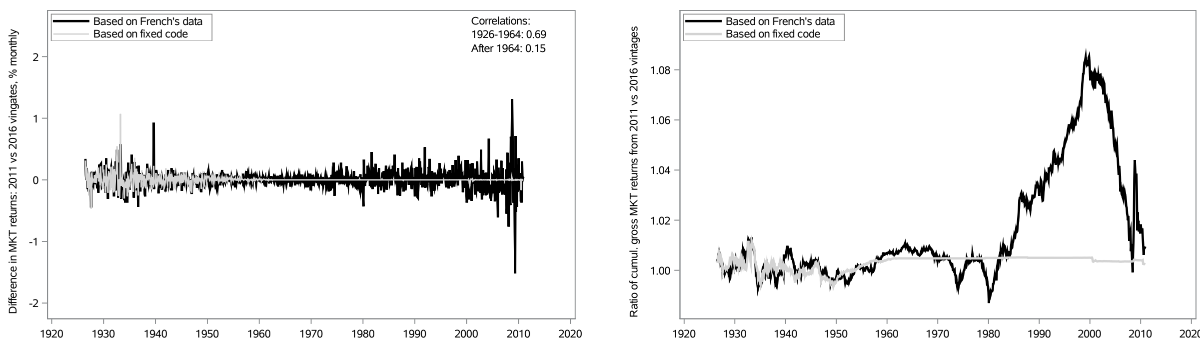
	Monthly factor returns affected		
Description of changes to factor construction	Market	HML	SMB
In October 2012, we revised the market return used to measure Rm-Rf in the US. It is now the value-weight return of all CRSP firms incorporated in the US and listed on the NYSE, AMEX, or NASDAQ that have (i) a CRSP share code of 10 or 11 at the beginning of month t, (ii) good shares and price data at the beginning of t, and (iii) good return data for t. Previously we used the CRSP NYSE/AMEX/NASDAQ Value-Weighted Market Index as the proxy for the market return. The set of firms in the new series is more consistent with the universe used to compute the other US returns.	Yes	No	No
In January 2015, CRSP completed an extensive review of their shares outstanding data for 1925-1946. The file they released in January 2015 (with data through December 2014) incorporates over 4000 changes that affect 400 Permnos. As a result, many of the returns we report for 1925-1946 change in our January 2015 update and some of the changes are large. Please see Changes in CRSP Data for descriptions of data changes by CRSP affecting the data series above.	Yes	Yes	Yes
In May 2015, we made two changes in the way we compute daily portfolio returns so the process is closer to the way we compute monthly portfolio returns. In daily files produced in May 2015 or thereafter, stocks are dropped from a portfolio immediately after their CRSP delist date; in files produced before May 2015, those stocks are held until the portfolio is reconstituted, at the end of June. Also, in daily files produced before May 2015 we exclude a stock from portfolios during any period in which it is missing prices for more than 10 consecutive trading days; in daily files produced in May 2015 and thereafter, we exclude a stock if there is no price for more than 200 consecutive trading days.	No	No	No
Because of changes in the treatment of deferred taxes described in FASB 109, files produced after August 2016 no longer add Deferred Taxes and Investment Tax Credit to BE for fiscal years ending in 1993 or later.	No	Yes	Yes
In August 2018, we have revised the method for computing Operating Profitability. We now include minority interest in the denominator, so the operating profitability ratio used to form portfolios in June of year t is annual revenues minus cost of goods sold, interest expense, and selling, general, and administrative expense divided by the sum of book equity and minority interest for the last fiscal year ending in t-1.	No	No	No
In August 2020, we removed the adjustment to book equity related to FASB Statement No. 106, Employers' Accounting for Postretirement Benefits Other Than Pensions, which was issued in 1990. This adjustment affects portfolios formed on book-to-market equity and portfolios formed on profitability, which is defined as operating income before depreciation and amortization minus interest expense scaled by book equity.	No	Yes	Yes
In September 2021, we transitioned from using our proprietary links between CRSP and Compustat data to those provided by CRSP after examining their consistency. We also updated the eligible universe through time to apply time-sensitive evaluation of stocks on criteria such as whether they are investment funds.	Maybe	Yes	Yes

Panel A: 2010–2011



Description of changes to factor construction over time period: None

Panel B: 2011–2016



Description of changes to factor construction over time period:

In October 2012, we revised the market return used to measure Rm-Rf in the US. It is now the value-weight return of all CRSP firms incorporated in the US and listed on the NYSE, AMEX, or NASDAQ that have (i) a CRSP share code of 10 or 11 at the beginning of month t, (ii) good shares and price data at the beginning of t, and (iii) good return data for t. Previously we used the CRSP NYSE/AMEX/NASDAQ Value-Weighted Market Index as the proxy for the market return. The set of firms in the new series is more consistent with the universe used to compute the other US returns.

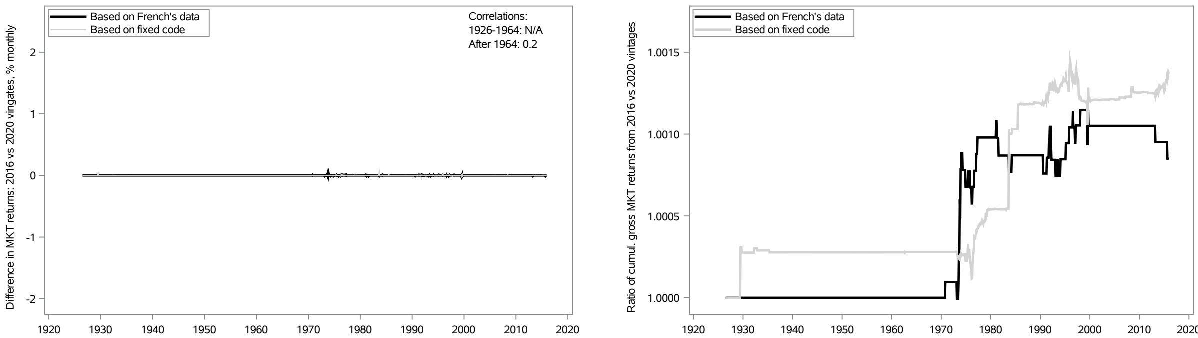
In January 2015, CRSP completed an extensive review of their shares outstanding data for 1925-1946. The file they released in January 2015 (with data through December 2014) incorporates over 4000 changes that affect 400 Permnos. As a result, many of the returns we report for 1925-1946 change in our January 2015 update and some of the changes are large. Please see Changes in CRSP Data for descriptions of data changes by CRSP affecting the data series above.

In May 2015, we made two changes in the way we compute daily portfolio returns so the process is closer to the way we compute monthly portfolio returns. In daily files produced in May 2015 or thereafter, stocks are dropped from a portfolio immediately after their CRSP delist date; in files produced before May 2015, those stocks are held until the portfolio is reconstituted, at the end of June. Also, in daily files produced before May 2015 we exclude a stock from portfolios during any period in which it is missing prices for more than 10 consecutive trading days; in daily files produced in May 2015 and thereafter, we exclude a stock if there is no price for more than 200 consecutive trading days.

Figure IA.4. Vintage-by-vintage changes in the market factor

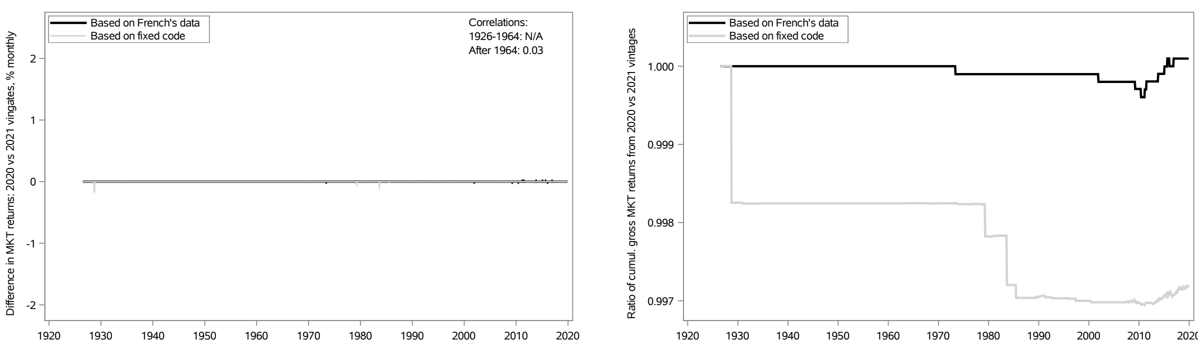
This figure plots the differences in returns in the fixed-code market factor and French market factor in the left panels, along with the changes in cumulative returns of the fixed-code factor and the French factor in the right panels for various changes in vintages. We copy the text of any changes to the methodology of any factor that are described on Ken French's website over the corresponding time period.

Panel C: 2016–2020



Description of changes to factor construction over time period:
In August 2018, we have revised the method for computing Operating Profitability. We now include minority interest in the denominator, so the operating profitability ratio used to form portfolios in June of year t is annual revenues minus cost of goods sold, interest expense, and selling, general, and administrative expense divided by the sum of book equity and minority interest for the last fiscal year ending in t-1.
Because of changes in the treatment of deferred taxes described in FASB 109, files produced after August 2016 no longer add Deferred Taxes and Investment Tax Credit to BE for fiscal years ending in 1993 or later.

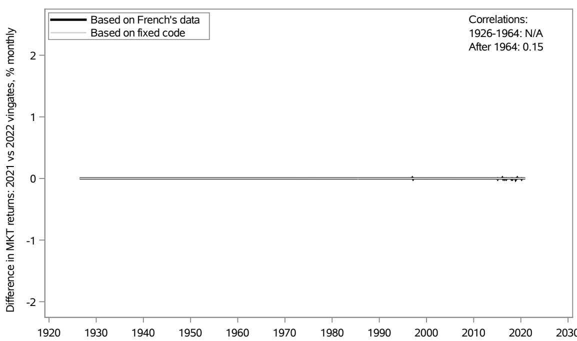
Panel D: 2020–2021

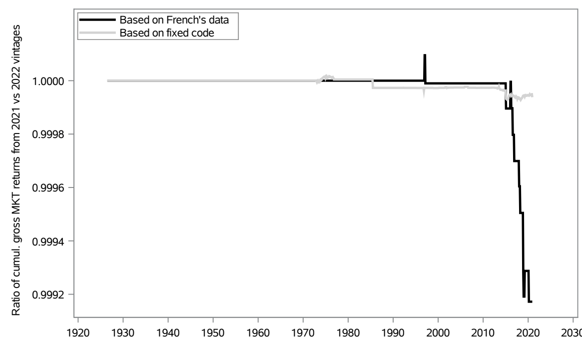


Description of changes to factor construction over time period:
In August 2020, we removed the adjustment to book equity related to FASB Statement No. 106, Employers' Accounting for Postretirement Benefits Other Than Pensions, which was issued in 1990. This adjustment affects portfolios formed on book-to-market equity and portfolios formed on profitability, which is defined as operating income before depreciation and amortization minus interest expense scaled by book equity.

Figure IA.4. Vintage-by-vintage changes in the market factor
This figure continues from the previous page.

Panel E: 2021–2022

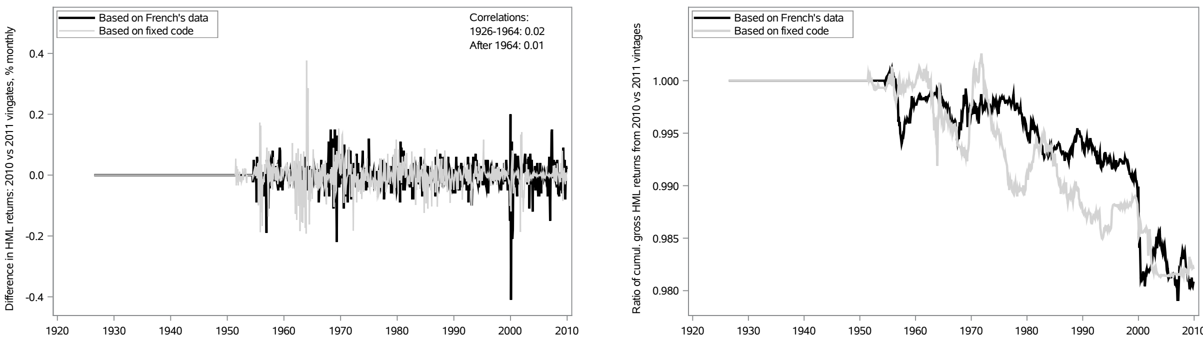




Description of changes to factor construction over time period:
In September 2021, we transitioned from using our proprietary links between CRSP and Compustat data to those provided by CRSP after examining their consistency. We also updated the eligible universe through time to apply time-sensitive evaluation of stocks on criteria such as whether they are investment funds.

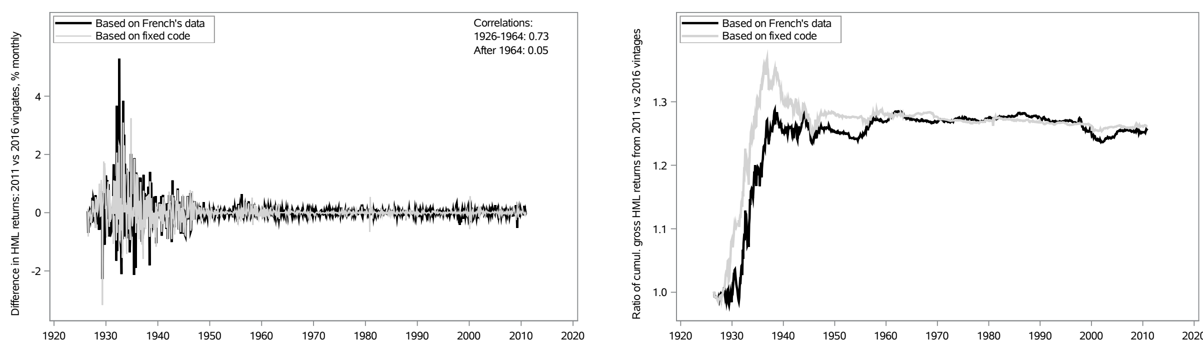
Figure IA.4. Vintage-by-vintage changes in the market factor
This figure continues from the previous page.

Panel A: 2010–2011



Description of changes to factor construction over time period: None

Panel B: 2011–2016



Description of changes to factor construction over time period:

In October 2012, we revised the market return used to measure Rm-Rf in the US. It is now the value-weight return of all CRSP firms incorporated in the US and listed on the NYSE, AMEX, or NASDAQ that have (i) a CRSP share code of 10 or 11 at the beginning of month t, (ii) good shares and price data at the beginning of t, and (iii) good return data for t. Previously we used the CRSP NYSE/AMEX/NASDAQ Value-Weighted Market Index as the proxy for the market return. The set of firms in the new series is more consistent with the universe used to compute the other US returns.

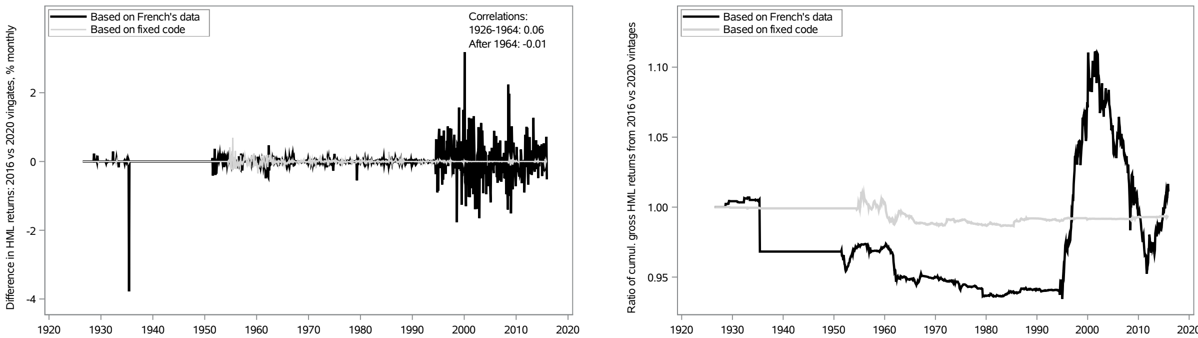
In January 2015, CRSP completed an extensive review of their shares outstanding data for 1925-1946. The file they released in January 2015 (with data through December 2014) incorporates over 4000 changes that affect 400 Permnos. As a result, many of the returns we report for 1925-1946 change in our January 2015 update and some of the changes are large. Please see Changes in CRSP Data for descriptions of data changes by CRSP affecting the data series above.

In May 2015, we made two changes in the way we compute daily portfolio returns so the process is closer to the way we compute monthly portfolio returns. In daily files produced in May 2015 or thereafter, stocks are dropped from a portfolio immediately after their CRSP delist date; in files produced before May 2015, those stocks are held until the portfolio is reconstituted, at the end of June. Also, in daily files produced before May 2015 we exclude a stock from portfolios during any period in which it is missing prices for more than 10 consecutive trading days; in daily files produced in May 2015 and thereafter, we exclude a stock if there is no price for more than 200 consecutive trading days.

Figure IA.5. Vintage-by-vintage changes in the HML factor

This figure plots the differences in returns in the fixed-code HML factor and French HML factor in the left panels, along with the changes in cumulative returns of the fixed-code factor and the French factor in the right panels for various changes in vintages. We copy the text of any changes to the methodology of any factor that are described on Ken French's website over the corresponding time period.

Panel C: 2016–2020

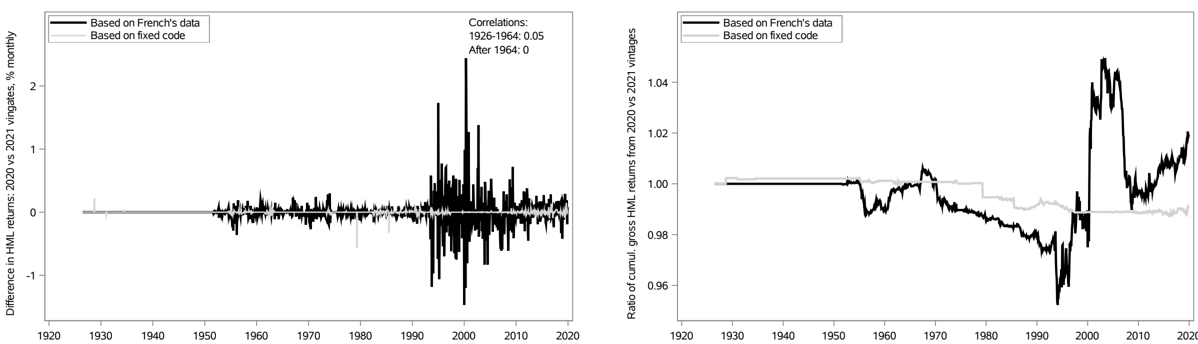


Description of changes to factor construction over time period:

In August 2018, we have revised the method for computing Operating Profitability. We now include minority interest in the denominator, so the operating profitability ratio used to form portfolios in June of year t is annual revenues minus cost of goods sold, interest expense, and selling, general, and administrative expense divided by the sum of book equity and minority interest for the last fiscal year ending in t-1.

Because of changes in the treatment of deferred taxes described in FASB 109, files produced after August 2016 no longer add Deferred Taxes and Investment Tax Credit to BE for fiscal years ending in 1993 or later.

Panel D: 2020–2021



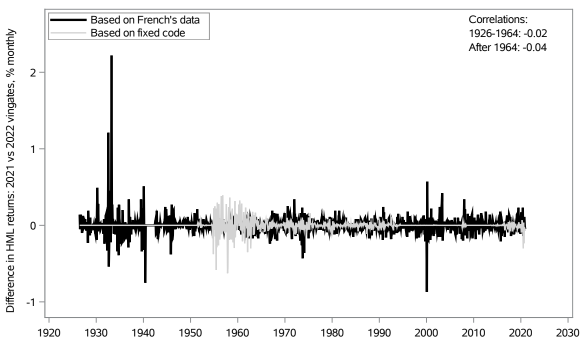
Description of changes to factor construction over time period:

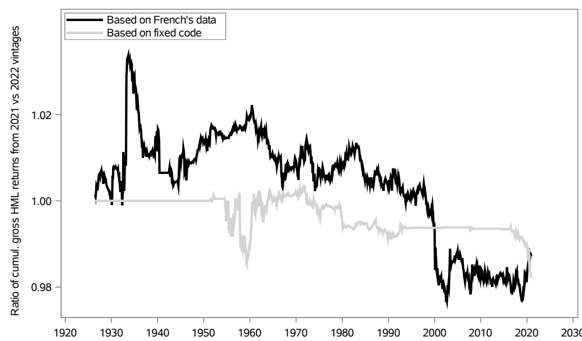
In August 2020, we removed the adjustment to book equity related to FASB Statement No. 106, Employers' Accounting for Postretirement Benefits Other Than Pensions, which was issued in 1990. This adjustment affects portfolios formed on book-to-market equity and portfolios formed on profitability, which is defined as operating income before depreciation and amortization minus interest expense scaled by book equity.

Figure IA.5. Vintage-by-vintage changes in the HML factor

TThis figure continues from the previous page.

Panel E: 2021–2022





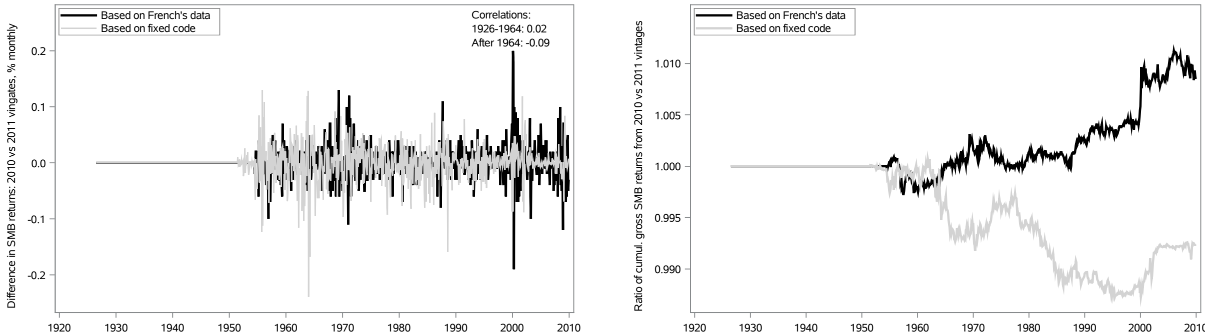
Description of changes to factor construction over time period:

In September 2021, we transitioned from using our proprietary links between CRSP and Compustat data to those provided by CRSP after examining their consistency. We also updated the eligible universe through time to apply time-sensitive evaluation of stocks on criteria such as whether they are investment funds.

Figure IA.5. Vintage-by-vintage changes in the HML factor

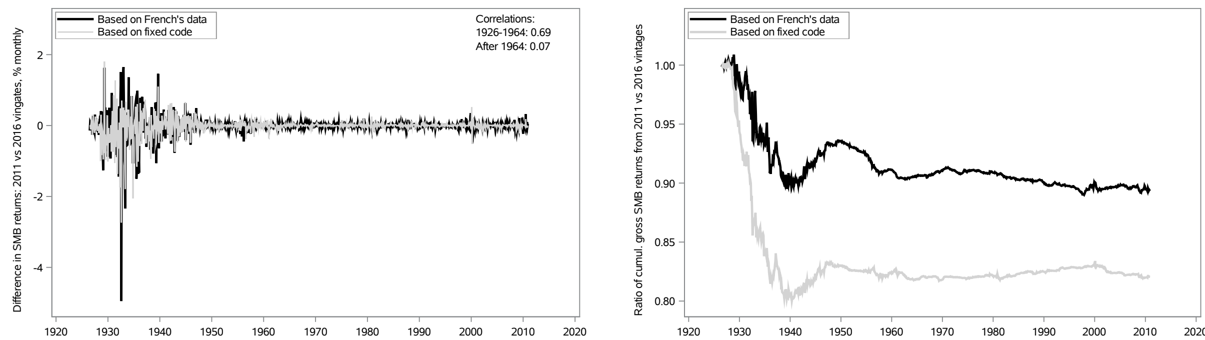
This figure continues from the previous page.

Panel A: 2010–2011



Description of changes to factor construction over time period: None

Panel B: 2011–2016



Description of changes to factor construction over time period:

In October 2012, we revised the market return used to measure Rm-Rf in the US. It is now the value-weight return of all CRSP firms incorporated in the US and listed on the NYSE, AMEX, or NASDAQ that have (i) a CRSP share code of 10 or 11 at the beginning of month t, (ii) good shares and price data at the beginning of t, and (iii) good return data for t. Previously we used the CRSP NYSE/AMEX/NASDAQ Value-Weighted Market Index as the proxy for the market return. The set of firms in the new series is more consistent with the universe used to compute the other US returns.

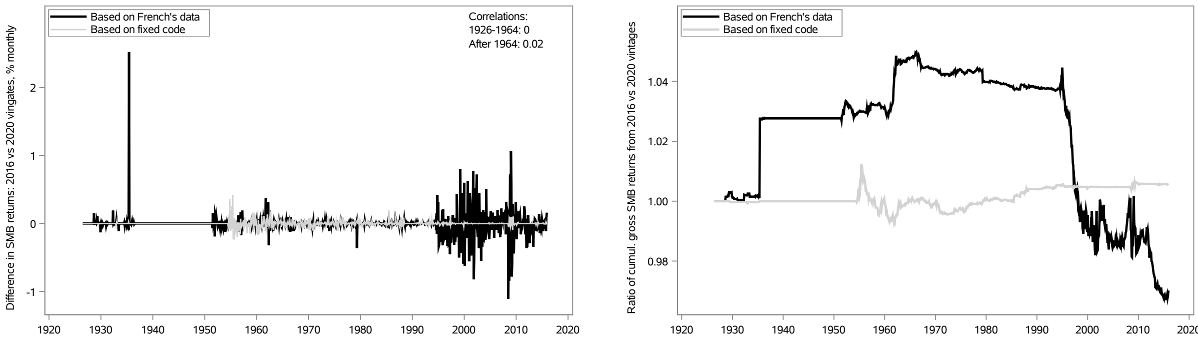
In January 2015, CRSP completed an extensive review of their shares outstanding data for 1925-1946. The file they released in January 2015 (with data through December 2014) incorporates over 4000 changes that affect 400 Permnos. As a result, many of the returns we report for 1925-1946 change in our January 2015 update and some of the changes are large. Please see Changes in CRSP Data for descriptions of data changes by CRSP affecting the data series above.

In May 2015, we made two changes in the way we compute daily portfolio returns so the process is closer to the way we compute monthly portfolio returns. In daily files produced in May 2015 or thereafter, stocks are dropped from a portfolio immediately after their CRSP delist date; in files produced before May 2015, those stocks are held until the portfolio is reconstituted, at the end of June. Also, in daily files produced before May 2015 we exclude a stock from portfolios during any period in which it is missing prices for more than 10 consecutive trading days; in daily files produced in May 2015 and thereafter, we exclude a stock if there is no price for more than 200 consecutive trading days.

Figure IA.6. Vintage-by-vintage changes in the SMB factor

This figure plots the differences in returns in the fixed-code SMB factor and French SMB factor in the left panels, along with the changes in cumulative returns of the fixed-code factor and the French factor in the right panels for various changes in vintages. We copy the text of any changes to the methodology of any factor that are described on Ken French's website over the corresponding time period.

Panel C: 2016–2020

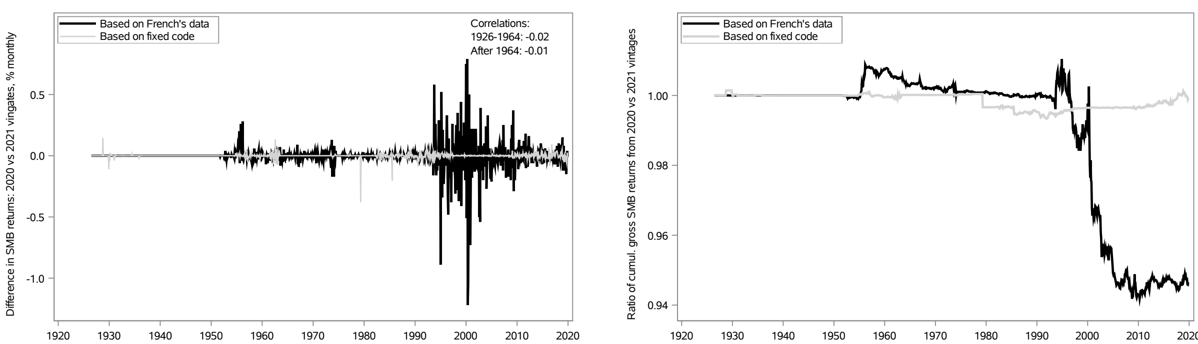


Description of changes to factor construction over time period:

In August 2018, we have revised the method for computing Operating Profitability. We now include minority interest in the denominator, so the operating profitability ratio used to form portfolios in June of year t is annual revenues minus cost of goods sold, interest expense, and selling, general, and administrative expense divided by the sum of book equity and minority interest for the last fiscal year ending in t-1.

Because of changes in the treatment of deferred taxes described in FASB 109, files produced after August 2016 no longer add Deferred Taxes and Investment Tax Credit to BE for fiscal years ending in 1993 or later.

Panel D: 2020–2021



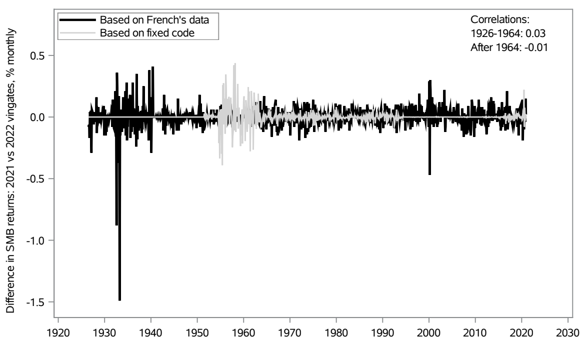
Description of changes to factor construction over time period:

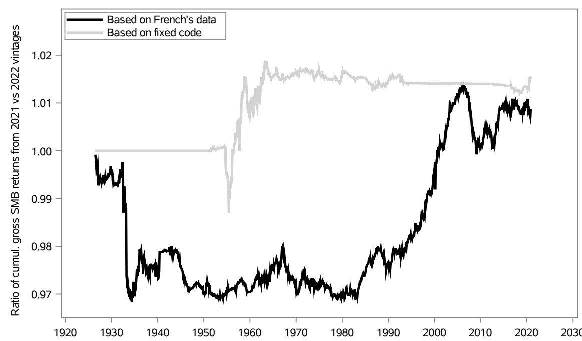
In August 2020, we removed the adjustment to book equity related to FASB Statement No. 106, Employers' Accounting for Postretirement Benefits Other Than Pensions, which was issued in 1990. This adjustment affects portfolios formed on book-to-market equity and portfolios formed on profitability, which is defined as operating income before depreciation and amortization minus interest expense scaled by book equity.

Figure IA.6. Vintage-by-vintage changes in the SMB factor

This figure continues from the previous page.

Panel E: 2021–2022





Description of changes to factor construction over time period:
In September 2021, we transitioned from using our proprietary links between CRSP and Compustat data to those provided by CRSP after examining their consistency. We also updated the eligible universe through time to apply time-sensitive evaluation of stocks on criteria such as whether they are investment funds.

Figure IA.6. Vintage-by-vintage changes in the SMB factor

This figure continues from the previous page.

about ECGI

The European Corporate Governance Institute has been established to improve *corporate governance through fostering independent scientific research and related activities.*

The ECGI will produce and disseminate high quality research while remaining close to the concerns and interests of corporate, financial and public policy makers. It will draw on the expertise of scholars from numerous countries and bring together a critical mass of expertise and interest to bear on this important subject.

The views expressed in this working paper are those of the authors, not those of the ECGI or its members.

Electronic Access to the Working Paper Series

The full set of ECGI working papers can be accessed through the Institute's Web-site (www.ecgi.global/content/working-papers) or SSRN:

Finance Paper Series	http://www.ssrn.com/link/ECGI-Fin.html
Law Paper Series	http://www.ssrn.com/link/ECGI-Law.html