Vanguard

Can active fixed income managers be cloned using style funds? A practical testing method

Vanguard Research

February 2021

Inna Zorina, CFA, and Douglas M. Grim, CFA

- Developments in research, technology, and product strategies have given investors new ways to evaluate active fixed income managers. Investors can now more accurately assess whether a manager's excess returns represent positive alpha or exposure to well-known styles (or factors), such as term or credit, that can now be purchased directly and cheaply.

- Some may be content with what they find. Others, however, may be surprised to learn that what they have been getting from a manager is just a fairly consistent tilt toward one or more of these factors, when their objective for hiring the manager may have been solely to generate positive alpha.

- When traditional active managers are not adding unique value or are charging too much, replacing them with low-cost style funds may give the end investor a way to generate excess return with greater transparency, more risk control, and lower implementation costs.

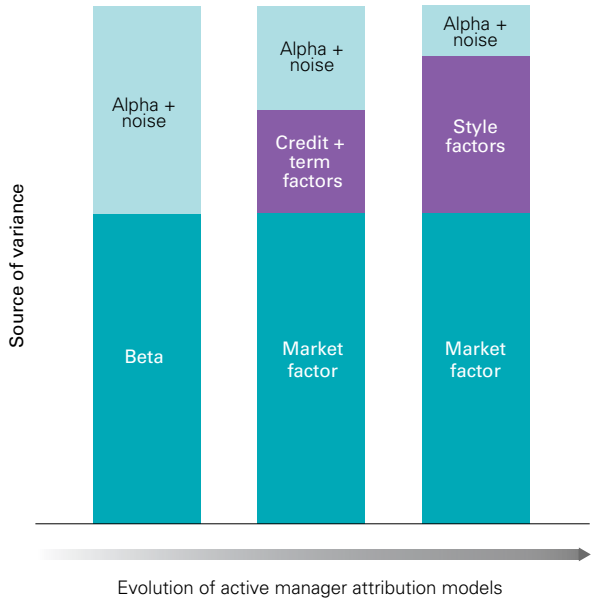
- In this paper, we present an accessible method to test to what extent the returns of an active fixed income manager can be replicated using style funds. This framework has a number of practical applications, including replacing an active fixed income manager that is not delivering net alpha above investable factor exposure and undertaking due diligence on current and prospective managers.

Introduction

Performance attribution, including attribution through a factor lens, has been a standard part of top-tier active manager due diligence for decades. **Figure 1** illustrates the evolution of factor-based evaluation in assessing active fixed income managers. In some cases, investors may discover that the returns they have been getting come from a consistent tilt toward one or more well-known factors (i.e. styles), such as term and credit, when their sole objective in hiring the manager may have been to generate positive, manager-specific alpha.[1]

The continued development of low-cost investment strategies to harvest fixed income factor returns has effectively raised the bar for many managers, who need to produce unique, positive net returns to justify their higher fees. We aim to empower the average professional investor by providing a simple but powerful testing framework that can reveal the extent to which a manager's returns can be cloned using widely available, lower-cost, more transparent style (factor) products.

Figure 1. Attribution models have evolved to distinguish factor exposure from true alpha contribution



Notes: This is a hypothetical scenario for illustrative purposes only. "Noise" refers to the fact that in any period, some degree of statistical randomness will affect results.

Source: Vanguard, 2020.

Notes on risk

All investing is subject to risk, including possible loss of principal. *Past performance does not guarantee future results.* When interest rates rise, the price of a bond or bond fund will decline. Bonds are subject to credit risk and inflation risk. Credit risk is the risk that a bond issuer will fail to make timely payments of interest and principal. Inflation risk is the possibility that increases in the cost of living will decrease or eliminate the returns of an investment. Because high-yield bonds are considered speculative, investors should be prepared to assume a substantially greater level of credit risk than with other types of bonds. There is no guarantee that any particular asset allocation or mix of funds will meet your investment objectives or provide you with a given level of income. *The performance of an index is not an exact representation of any particular investment, as you cannot invest directly in an index.* In a diversified portfolio, gains from some investments may help offset losses from others. However, diversification does not ensure a profit or protect against a loss. Investments in securities issued by non-U.S. companies are subject to risks including country/ regional risk and currency risk. These risks are especially high in emerging markets. While U.S. Treasury or government agency securities provide substantial protection against credit risk, they do not protect investors against price changes due to changing interest rates. Unlike stocks and bonds, U.S. Treasury bills are guaranteed as to the timely payment of principal and interest.

1 Similar to Zorina, Scholz, and Grim (2020), our definition of alpha is value added or subtracted by an active manager through difficult-to-replicate security selection or sector/factor timing.

Factor tilts in the fixed income markets

Although limited in comparison to the plethora of research covering equity factors, the fixed income literature suggests that active bond returns are predominantly driven by consistent tilts toward certain factors. Research shows that the majority of active returns for fixed income managers can be explained by overweights to the credit factor, which can potentially have significant implications for investor portfolios.[2]

The distinct lack of consideration for investable products in the academic research is the major drawback of the current state of analyses for practitioners. While examining factors in a theoretical manner produces robust and insightful results, most fixed income factor models do not consider actual investability. This makes it challenging for investors to decipher whether an active manager is generating returns that cannot be produced through alternative means.

Our paper attempts to fill this gap by building a bridge between existing academic literature and practical implementation of the concept. Our goal is to determine to what extent a manager's performance can be replicated simply by systematically tilting toward well-known fixed income factors. If the manager's performance is almost entirely explained by such factor tilts, the investor will have to determine whether the same results could be achieved in a more transparent, risk-controlled, lower-cost way using widely available off-the-shelf fixed income factor products.

Investors would benefit from using this framework to perform due diligence on current and prospective active fixed income managers to understand if the manager employs a static credit or other tilt(s) that might have implications for their overall portfolio diversification.[3]

In this paper, we apply the framework developed in Zorina, Scholz, and Grim (2020) to analyze factor exposures of active fixed income managers. While we discuss all elements of the framework, the original paper, which evaluated active equity managers, provides more in-depth explanation of the approach.

Evaluating a manager's factor profile

Methods of assessing the factor profile of an active manager date back at least to the early 1990s (for example, Sharpe, 1992, Fama and French, 1993, and Blake, Elton, and Gruber, 1993). Since that time, many techniques have been proposed in the academic literature. At the highest level, there are two general approaches: returns-based style analysis (RBSA) and holdings-based style analysis (HBSA).[4] Both methods are useful but have well-documented limitations, and neither stands out as the clearly superior approach.

In this paper, we focus on a specific RBSA approach, which has several benefits for our practical goals. It requires fewer inputs and is relatively easy to perform and customize (Christopherson and Sabin, 1999). It aims to match the month-to-month historical performance of a manager with that of a buy-and-hold set of factors. It attempts to reflect the typical behavior of the manager and can make it easy to measure corresponding factor weights and differences in performance. Last, the investor can decide which factors to include in the analysis and which strategy to use to represent each factor's performance.[5]

Selecting a suitable approach

We build on existing portfolio construction and factor research (such as Sharpe, 1992, and Zorina, Scholz, and Grim, 2020) by using an RBSA method that employs long-only, investable factor proxies. This creates a fair alternative investment option to a long-only manager or multimanager portfolio that is subject to real-world fees and transaction costs.

To make the factor portfolio investable and the RBSA output more intuitive, we include two constraints: The mimicking portfolio is fully invested, and it uses no shorting or explicit leverage. This allows for an intuitive interpretation of results, producing a clear signal of manager performance net of static factor tilts. The signal takes the form of the net alpha, for which a negative or statistically insignificant value suggests the manager did not produce a unique value add over the examination period.

2 See, for example, Mattu et al. (2016) and Brooks, Gould, and Richardson (2020).

3 See Stockton, Donaldson, and Chen (2019) for a discussion on the impact of high-yield bonds on portfolio performance.

4 HBSA is also referred to in the literature as portfolio- or characteristics-based style analysis.

5 Detailed comparison and analyses of either of these approaches is outside the scope of this paper. For a further breakdown of the benefits of RBSA for factor profile evaluation, see Zorina, Scholz, and Grim (2020). Bender, Hammond, and Mok (2014) note that RBSA is commonly used by industry practitioners. More sophisticated investors with access to the proper tools, data, and technical expertise can complement our method with an HBSA approach and a direct manager discussion for a more comprehensive performance assessment.

Constrained RBSA

The approach we chose to use could be described as constrained optimization to find the portfolio with the best historical fit:

$$R_t^M = \alpha + \sum_{j=1}^{J} w_j^F R_{j,t}^F + \varepsilon_t$$

where

R_t^M represents the manager's return for month t;

α represents alpha;

F represents selected factors;

w_j^F represents factor coefficients (that is, portfolio weights);

$R_{j,t}^F$ represents factor proxy return;

j represents factor proxy numbers from 1 to J; and

ε_t represents residual returns (through an error term).

We define *alpha* as the manager's excess return relative to the custom factor-mimicking portfolio. It could be driven by the manager's security selection and/or by good or bad luck.[6] Similar to Sharpe (1992), we impose two constraints:

$\sum_{j=1}^{J} w_j^F = 1$ Factor weights must sum to 1. This assumes the manager is fully invested.

$0 \le w_j^F \le 1$ All factor weights must be equal to or greater than 0 and less than or equal to 1. This assumption limits all factor exposures to long-only, disallowing any leveraged or short exposures.

These constraints ensure straightforward interpretation of the factor coefficients as the portfolio weight of each factor that would have led to the best fit between the returns of the factor-mimicking portfolio and those of the manager for a given period.[7] If the manager's net alpha is negative or not statistically significant, it's likely that the manager did not deliver unique value to end investors for the period studied.

RBSA: Part art, part science

RBSA done correctly can lead to valuable insights. However, the output can be misleading if a robust system of checks and balances is not in place. Although RBSA is a quantitative approach, professional judgment is still essential at different steps in the process. Improper model specifications can lead to incorrect interpretations of results and imprudent investment decisions. In this section, we address the most important methodology considerations to mitigate statistical biases and errors.

What factors to include?

There is no industry-wide consensus on which factors should be considered when attempting to mimic the performance of a manager. In theory, only those the manager is likely to be exhibiting should be included. However, that may not always be obvious, and at times the manager's factor profile may not align with expectations. A prudent starting point would be to consider factors that have a sound economic rationale for and extensive empirical evidence of driving fixed income active returns.

While a number of factors (such as credit, term, value, liquidity, momentum, carry, and low volatility) have been discussed in the fixed income literature, significant debate surrounds the efficacy of a few of them.[8] The main factors that drive fixed income returns are term and credit. More important from the practical perspective, low-cost, investable strategies are generally only available for products that have credit and/or term tilts.

Since the goal of this paper is to focus on assessing active fixed income managers versus investable, low-cost fixed income factors that investors could otherwise use, our analysis will not include the others. We include categories such as cash, international bonds, and specific subsets of U.S. bonds such as mortgage-backed securities (MBS) and Treasury inflation-protected securities (TIPS) because they represent additional exposures managers could use to express their investment views.[9]

6 Alpha could also result from model misspecification (such as a suboptimal choice of factor proxies or omitted factors).

7 In technical terms, it uses quadratic programming to determine the buy-and-hold weights of each factor strategy that minimize the squared error between the active manager and the long-only factor-mimicking portfolios' month-to-month returns, subject to the constraint that the mimicking portfolio is fully invested with no shorting or explicit leverage.

8 Bao, Pan, and Wang (2011) establish the liquidity factor in corporate bonds. Asness, Moskowitz, and Pedersen (2013) find evidence of consistent exposure to value and momentum in eight asset classes, including fixed income. Koijen et al. (2013) find strong predictive ability for a variety of fixed income assets arising from the carry factor.

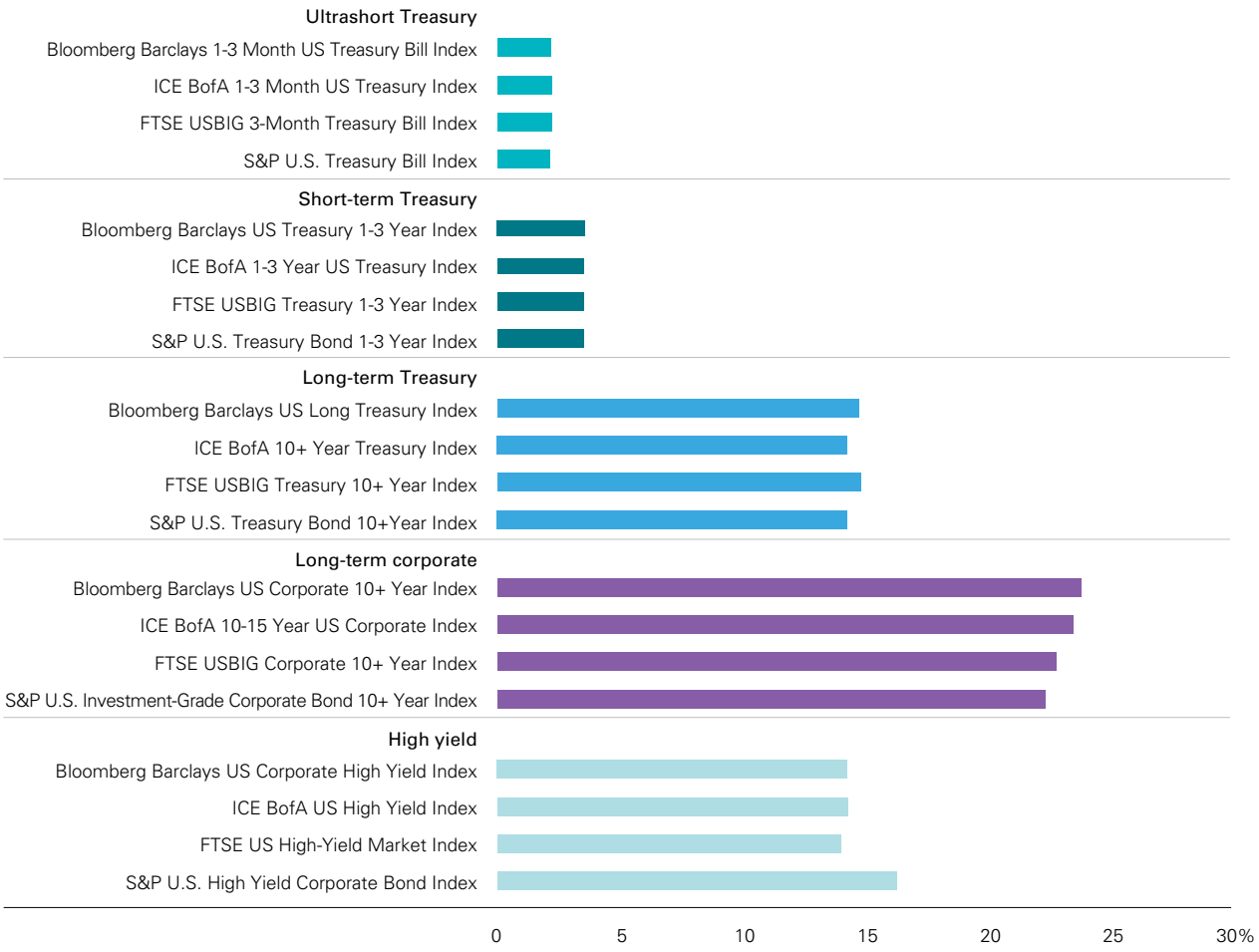
9 Developed and emerging-market bond fund returns are driven by the same fixed income factors but reflect different country exposures. The performance of the same factor in different countries can be materially different month to month, so including these additional proxies can potentially improve the ability to mimic the performance profile of a manager who invests a portion of the portfolio internationally.

What strategy should serve as the proxy for each factor?

Next, the investor must decide which investable strategy should represent the historical performance of each factor. In contrast to different potential proxies for equity factors, those that proxy for the same fixed income factors tend to perform in a very similar way year to year. This is because fixed income index categorization based on maturity ranges and credit quality are more universally agreed upon.[10] **Figure 2** shows that similar criteria led to similar returns in 2019.

Figure 2. Fixed income strategies that target the same factor exposures deliver very similar returns

Gross returns for 2019



Notes: This represents a sample, not an exhaustive list, of U.S. fixed income factor indexes. *Past performance is no guarantee of future returns. The performance of an index is not an exact representation of any particular investment, as you cannot invest directly in an index.* Benchmark comparative indexes represent unmanaged or average returns for various financial assets, which can be compared with funds' total returns for the purpose of measuring relative performance. Index performance does not reflect the deduction of fees and expenses.

Source: Vanguard calculations, using monthly gross return data from Thomson Reuters Datastream from January 1, 2019, to December 31, 2019.

10 This does not mean that their criteria are identical. For instance, maturity ranges may differ between different providers. While some providers define "intermediate" as 5 to 10 years, others can choose 3 to 7 or 1 to 10 years.

Managing multicollinearity

Using a set of long-only strategies to conduct RBSA comes with a potential statistical issue—multicollinearity—because their returns can be highly correlated. If two factors move in unison, it becomes difficult to distinguish whether one or both are influencing the manager's returns.[11] **Figure 3** shows that some fixed income factor strategies are highly correlated, while others (for example, ultrashort Treasury and high yield) exhibit low or even negative correlation with the majority of the term and credit factors.[12]

Ignoring the potential impact of multicollinearity can lead to a false sense of accuracy in factor weights. Standard errors, which are widely used to assess the accuracy of regression coefficients, can generate misleading estimates if they are not adjusted for multicollinearity.[13] Lobosco and DiBartolomeo (1997) developed an approach that helps take multicollinearity into account, adjusting standard errors for the amount of information that each factor adds to the RBSA.[14] When a factor's returns are very close to a linear combination of other factors used in the analysis, the added (or "unexplained") return and added (unexplained) volatility of that factor will be relatively small and its standard error will be relatively large.

In addition to making this adjustment, it's also considered a best practice to err on the side of including fewer factor strategies in the final factor-mimicking portfolio. This helps reduce the potential impact of multicollinearity. A factor that only marginally improves the "clonability" of a manager could thus be removed by lowering the number of factors.

It should be noted that when investors are building the factor portfolio in an attempt to replicate a manager, adding long-only fixed income factor strategies will reduce the weight of the broad investment-grade market index. In some cases, as we discuss later, this weight can be completely replaced by long-only fixed income factor strategies, as investment-grade market exposure is often a key driver of these strategies' month-to-month performance.

Figure 3. Correlations of fixed income factor strategy returns vary substantially

	Total U.S. market	Ultra-short Treasury	Short-term Treasury	Inter-mediate-term Treasury	Long-term Treasury	Short-term corporate	Inter-mediate-term corporate	Long-term corporate	High yield	Inter-national
Ultrashort Treasury	24%									
Short-term Treasury	70%	64%								
Intermediate-term Treasury	87%	39%	89%							
Long-term Treasury	85%	13%	56%	82%						
Short-term corporate	72%	12%	43%	47%	35%					
Intermediate-term corporate	84%	2%	38%	55%	55%	92%				
Long-term corporate	82%	-2%	27%	49%	66%	75%	92%			
High yield	16%	-19%	-25%	-23%	-16%	53%	53%	49%		
International	77%	21%	54%	71%	76%	49%	61%	64%	-1%	
Emerging market	53%	-3%	9%	21%	25%	63%	71%	68%	70%	31%

Low correlation (below 30%) Medium correlation (from 30% to 70%) High correlation (above 70%)

Notes: Fixed income factor strategy benchmark returns are calculated using USD gross return indexes: Bloomberg Barclays US Aggregate Bond Index (USD), Bloomberg Barclays 1-3 Month U.S. Treasury Bill Index (USD), Bloomberg Barclays 1-3 Year US Treasury Bond Index (USD), Bloomberg Barclays US Intermediate Treasury Index (USD), Bloomberg Barclays US Long Treasury Index (USD), Bloomberg Barclays US Corporate Bond 1-5 Year Index (USD), Bloomberg Barclays US Credit Corporate 5-10 Year Index (USD), Bloomberg Barclays US Corporate 10+ Year Index (USD), Bloomberg Barclays US Corporate High Yield Index (USD), Bloomberg Barclays Global Aggregate ex-USD Bond Index USD-Hedged (USD), and FTSE Emerging Markets USD Government Bond Index Capped (USD). *The performance of an index is not an exact representation of any particular investment, as you cannot invest directly in an index.* Benchmark comparative indexes represent unmanaged or average returns on various financial assets, which can be compared with funds' total returns for the purpose of measuring relative performance. Index performance does not reflect the deduction of fees and expenses.

Source: Vanguard calculations, using monthly gross return data from Morningstar from January 1, 2000, to December 31, 2019.

11 A key goal of regression analysis is to isolate the relationship between each independent variable (factor return) and the dependent variable (manager return). A regression coefficient represents the average change in the dependent variable for a one-unit change in an independent variable, holding all of the other independent variables constant. However, when independent variables are correlated, changes in one may be associated with shifts in another. The stronger the correlation, the more difficult it is to estimate the actual effect of each variable.

12 These factor strategies do not represent separate factors but rather exhibit different levels of exposure to term and credit factors.

13 In this context, a larger standard error indicates a lower degree of confidence in the RBSA output's factor weight estimate.

14 Adjusting standard errors in this manner is not typically part of a statistical tool kit, but it can be done using standard programming software.

Costs count

A final important consideration of the factor-mimicking portfolio is that the comparison must take into account implementation frictions, including ongoing expenses (such as management fees) and those that are transaction-related (such as commissions and bid/ask spreads). If investors use indexes or a back-tested paper portfolio to represent all or part of the historical factor returns, they will need to estimate how real-world costs would affect those results. This allows for a fair comparison of results achievable in practice, because managers cannot avoid such costs.[15]

Assessing the clonability of a manager

To judge whether one or more factor strategies using RBSA could substitute for a manager, the investor must evaluate goodness of fit, confidence intervals for factor weights, and factor stability.

Goodness of fit and confidence intervals

Researchers use two key metrics to evaluate goodness of fit, or how similarly a custom factor-mimicking portfolio performs against a manager: adjusted R-squared and tracking error.[16] Although there is no consensus on the appropriate thresholds for these metrics, our analysis shows that an adjusted R-squared above 95% and a tracking error below 1.5% generally indicate a very good fit.

Adjusted standard errors help the investor evaluate confidence intervals for factor weights. This is critical in assessing the statistical significance of these weights to improve the robustness of the results.

Factor stability

An implicit assumption when conducting standard RBSA is that a manager's factor exposures are persistent. Sharpe (1992) and Lucas and Riepe (1996), among others, rightly point out that tilts exhibited in a prior period may have varied and may not represent the manager's future factor profile. Zorina, Scholz, and Grim (2020) discuss options available for an investor to determine whether the manager's factor exposure has been consistent and whether it is likely to persist. These options include:

Ask the manager.
Discuss why the manager has exhibited that factor profile in the past and whether or not it can be expected to persist.

Perform an "eyeball test."
Assess whether the factor weights appear fairly stable through different time periods using, for instance, 36-month rolling regression results.

Test for constant variance.
If the manager's factor exposure changes substantially over time, the regression residuals (errors) relative to the factor-mimicking portfolio should also experience substantial variability. Statistical tests (for example, the White test) can determine to what extent the regression residuals could be considered independent and identically distributed random noise.

Test for a structural break.
Divide the return history into subsamples and test whether the subsample regression results have statistically significant weight differences.

A tiered approach

Because of the many different fixed income factor strategies representing different exposures to term and credit factors and bond types such as international, TIPS, and MBS, we recommend approaching the RBSA in multiple rounds, evaluating whether changes in regression specifications in each round improve the fit of the factor-mimicking portfolio. This approach helps investors balance goodness of fit with managing multicollinearity.

Round 1: Credit profile. First, we attempt to identify the manager's credit quality preference by including factor categories that span the bond credit quality spectrum: total bond market, total market investment-grade corporate debt, and total market high-yield debt.[17] This allows us to identify credit quality preference without changing term factor risk too much, because the duration of the three total market categories is generally similar.

15 These costs vary by strategy, asset size, and manager. Using Morningstar data, Vanguard estimates that as of December 31, 2019, the asset-weighted U.S. fixed income factor strategies ETF expense ratio was 9 basis points (bps) and the equal-weighted expense ratio was 22 bps.

16 Tracking error indicates how much a manager's return differs from a specified benchmark (for the purposes of this analysis, the benchmark would be the factor-mimicking portfolio). The most common industry measure is the annualized standard deviation of excess return between manager and benchmark.

17 Some categories, such as below-investment-grade (high-yield) bonds, have limited availability of low-cost, investable index proxies. In these cases, the investor could consider a risk-controlled, low-cost, traditional active fund as a proxy.

By limiting the first round of RBSA to these strategies, we avoid a large number of variables in our initial RBSA and know where to focus our efforts next.

Round 2: Term profile. We then conduct a second round of RBSA to pinpoint the manager's historical duration profile preference. We only include factor strategies for different levels of term factor exposure for the credit-quality segment toward which the manager exhibits a statistically significant weight.[18]

Round 3: Manager's factor preferences for certain bond types. Last, we test whether bond-type tilts might improve the clonability of the factor-mimicking portfolio. Here we could include categories such as cash, broad international or emerging-market bond exposures, or specific subsets of bonds such as MBS and TIPS.

Manager cloning framework: A step-by-step guide

We now summarize our approach before showcasing examples using practical case studies.

Figure 4. A practical testing framework for assessing the extent to which a manager can be cloned

Step 1: Select an active manager to test

Choose an active fixed income fund in your portfolio or one being considered for inclusion, with a sufficient return history (preferably more than 36 months).

Step 2: Select factor candidates

Choose factor categories that cover credit quality (such as Treasury, investment-grade, or below-investment-grade) and term (short, medium, or long). Consider any other investable bond categories that may help mimic the manager's return profile (for example, international or inflation-linked).

Step 3: Select investable proxies for each factor category

Choose low-cost, long-only investable proxies with a sufficient return history from a similar investment universe as the manager.

Step 4: Conduct returns-based style analysis in rounds

Run a constrained regression of the manager's returns on the factor proxy returns in several rounds. First, assess the credit profile, then the term profile, and finally any other factors.

Step 5: Assess goodness of fit, factor stability, and alpha

Evaluate which factor categories and factor proxies can be used to substitute for the fund. Assess the goodness of fit measures, standard errors, and confidence intervals of the portfolio weights, factor stability, and net alpha of the fund to determine the extent to which returns can be cloned.

18 From the practical perspective, investable strategies representing term tilts are generally only available for Treasuries and investment-grade credit. Therefore, our analysis only includes the term tilts for these segments.

Software to use for conducting constrained RBSA

Morningstar Direct Presentation Studio has constrained regression functionality. The program can examine the basic criteria of both tracking error and adjusted R-squared and allows for eyeballing factor stability. This is a reasonable option for garnering basic insight into a manager's behavior.

Python is another powerful tool. It not only allows the user to run constrained regressions but can also provide more sophisticated statistical tests such as adjusted standard errors and structural break tests.

While some code has to be written, open-source packages such as scipy.optimize.minimize can be used by those with some basic coding background. We used Python for this paper. Other software packages that could undertake this analysis include RStudio and MATLAB.

Alternatively, some consulting firms and asset managers can also conduct constrained RBSA.

Implementation: Two practical case studies

In this section, we show how investors can apply our practical testing framework to determine to what extent a manager can be cloned. Our hypothetical case studies attempt to explain active U.S. fixed income manager returns using term, credit, and other factor categories. Our investable factor proxies consist of low-cost ETFs that provide exposure to the U.S. investment-grade bond market in its entirety, U.S. credit and term factors, and MBS, TIPS, and international bonds.[19]

As of December 31, 2019, the expense ratio for each of these investable strategies ranged from 4 to 23 bps per year.[20] Funds in each example are white-labeled (that is, we do not disclose the names of the active managers or factor ETFs used). We use the gross total return time series for the managers and factor proxies.[21] Each case study uses all available data; different data availability resulted in different time frames.

Although these case studies have been undertaken for U.S. fixed income managers and factors, the framework can be applied globally, depending on the availability of products.

Case study A: Cloning a manager

The first case study discusses a fund that can be cloned using our framework. Fund A is an actively managed, investment-grade fixed income fund with an expense ratio of 61 bps. We start our analysis by evaluating its credit exposure, then assess its term exposure, and finally determine whether the manager has exhibited tilts toward certain bond types, modifying the RBSA inputs each time until we get the final custom factor portfolio.

Fund A has outperformed a broad, investment-grade U.S. bond market index on a gross basis by 90 bps annually over the last 18 years. We follow the step-by-step approach detailed in Figure 4 using low-cost investable proxies for term, credit, and other factors.

By including factor proxies that span the bond credit quality spectrum in the RBSA (see **Figure 5**, Round 1), we find that much of the fund's historical outperformance relative to the market can be explained by a combination of the total U.S. bond market with a tilt toward some investment-grade and high-yield credit, with a tracking error of 0.97% and adjusted R-squared of 0.93.[22]

After including investment-grade credit factor proxies with different term factor exposure (Round 2 in Figure 5), we rerun the analysis and find that the fund seems to exhibit an intermediate-term profile. Fit is barely improved and net alpha is nearly unchanged, so the initial portfolio of total U.S. investment-grade market, total corporate, and high-yield credit may be sufficient. In Round 3, we experiment with adding other categories but find that these do not help improve fit and therefore do not include them in the final factor-mimicking portfolio.

We also see a decrease in the standard error of the intermediate-term corporate credit factor strategy that comes from removing the other corporate credit factor strategies. This illustrates the importance of limiting the number of factors in the RBSA.

19 Using investable factor proxies from the same index provider or asset manager is a useful starting point but not a requirement, because the framework is able to accommodate factor proxies from various providers.

20 For the purposes of our case studies, we assume that the factor-mimicking portfolios are rebalanced monthly with no transaction-related costs. Depending on the ETFs or mutual funds used and the size of the trades, the effect of the added costs could range from immaterial to significant.

21 The evaluator can consider using benchmark or simulated factor return history adjusted for estimated costs when the factor fund's live return time series is not long enough for the analysis and the alternative return history is derived from a strategy likely to be substantially similar to the factor fund's unique factor profile.

22 This example conceptually illustrates a regression method that is available as a built-in option in some statistical packages. Although it may only be available to sophisticated investors, it is a prudent approach to consider as it allows the user to control the number of independent regressors (factor strategies) while optimizing the selected fit criteria.

Figure 5. Constrained RBSA results for Fund A, 2002–2019

Factor-mimicking portfolio weights

	Alpha (annualized)	Total U.S. market	Treasury	Total corporate	Short-term corporate	Inter-mediate-term corporate	Long-term corporate	High yield	Tracking error (percentage)	Adjusted R^2
Round 1: Manager's credit exposure										
Factor strategies across credit quality spectrum	–1 bps	71%	0%	9%				21%	0.97	0.93
Standard errors	*(8 bps)*	*(8%)*	*(6%)*	*(3%)*				*(1%)*		
Round 2: Manager's term exposure										
Including all term and credit factor strategies	4 bps	65%		0%	7%	9%	0%	19%	0.94	0.94
Standard errors	*(8 bps)*	*(4%)*		*(40%)*	*(15%)*	*(14%)*	*(11%)*	*(1%)*		
Only non-zero term and credit factor strategies	4 bps	65%			7%	9%		19%	0.94	0.94
Standard errors	*(8 bps)*	*(5%)*			*(3%)*	*(2%)*		*(1%)*		
Only significant term and credit factor strategies	–4 bps	70%				10%		20%	0.95	0.94
Standard errors	*(7 bps)*	*(2%)*				*(2%)*		*(1%)*		

		Total U.S. market	Inter-mediate-term corporate	High yield	Cash	MBS	TIPS	Inter-national		
Round 3: Relevance of other factor strategies										
Including other factor strategies	–6 bps	43%	15%	19%	0%	14%	5%	5%	0.93	0.94
Standard errors	*(10 bps)*	*(9%)*	*(3%)*	*(1%)*	*(4%)*	*(7%)*	*(2%)*	*(4%)*		

Insignificant Significant at 1% Significant at 5%

Notes: Statistical significance is assessed using adjusted standard errors (Lobosco and DiBartolomeo, 1997). Standard error for the alpha term is approximate. Portfolio weights may not add up to 100% because of rounding. For the dynamic regression results performed after Round 2, see Figure 6.

Source: Vanguard calculations, using monthly gross return data from Morningstar from January 1, 2002, to December 31, 2019.

The factor analysis of Fund A shows that this investment-grade fixed income fund has been employing a long-term tilt to high-yield credit to achieve outperformance.[23]

We assess the factor exposure stability of Fund A using the total U.S. investment-grade market, intermediate-term corporate and high-yield credit, and two approaches: full sample regression (to test for constant variance) and rolling-window regressions (the eyeball test).

The best way to assess factor stability using the full sample regression is to test if the variance of the error term remains constant over the entire sample. If the manager's factor exposure changes significantly over time, the regression residuals (errors) relative to the factor-mimicking portfolio should also experience substantial variability. We run the White test for the entire sample, which results in a p-value of 0.25. The null hypothesis of constant variance is not rejected at the 10% significance level, and we conclude that the error variance is constant—that is, heteroscedasticity is not present. (Heteroscedasticity occurs when the standard deviations of a predicted variable, monitored over time periods or different values of an independent variable, are nonconstant. It is a violation of the assumptions for the ordinary least squares linear regression analysis, and it can affect standard errors and invalidate inferences obtained from this analysis.)

We also assess factor stability by reviewing a plot of dynamic rolling-window regressions, estimating the 36-month rolling portfolio weights that most closely mimic the manager's returns over each subperiod (the eyeball test). **Figure 6** shows that the exposures were relatively stable over time.

Figure 6. Dynamic regression results (36-month rolling window) for Fund A, 2002–2019

Notes: The White test for the entire sample results in a p-value of 0.25, and the null hypothesis (homoscedasticity) is not rejected at the 10% significance level. We conclude that the error variance is constant (heteroscedasticity is not present). The dynamic regression is run from 2002; the chart starts at 2005 because the regression window length is 36 months.

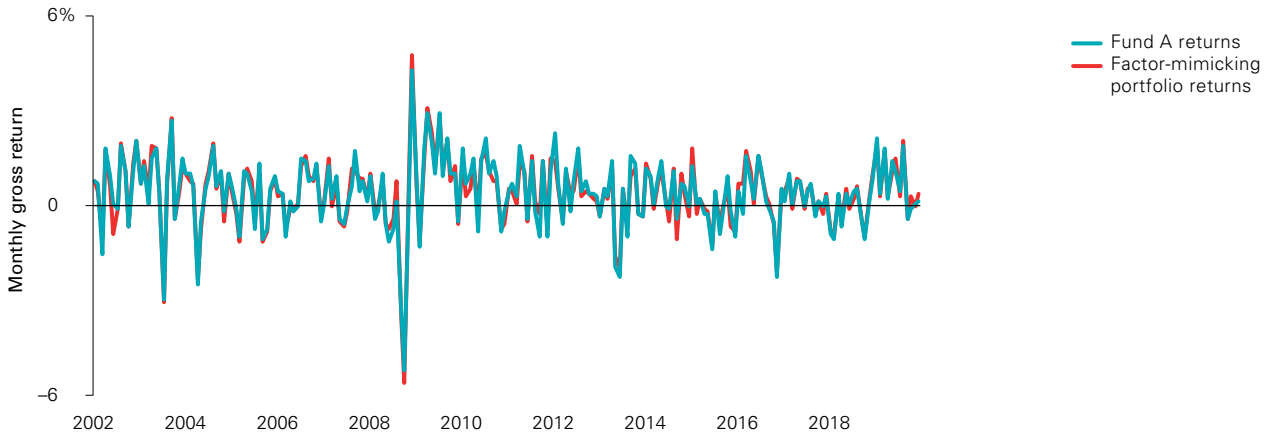
Source: Vanguard calculations, using monthly gross return data from Morningstar from January 1, 2002, to December 31, 2019.

23 High-yield credit outperformed investment-grade credit over the time period observed.

This portfolio tracks the actual manager's returns very closely, with a 1.0% tracking error and 0.94 adjusted R-squared. The performance difference (gross alpha) between the portfolio and the manager is not statistically significant at 10%, showing that the majority of excess returns have been delivered through stable factor tilts.

Stable factor exposure was confirmed by both the eyeball test and a formal test for heteroscedasticity. We conclude that Fund A can be cloned using cheaper (approximately 8 bps[24] versus 61 bps) and more transparent off-the-shelf factor products. **Figure 7** shows that the final portfolio's returns closely track those of the manager.

Figure 7. Fund A's historical returns and those of the final factor-mimicking portfolio show very similar behavior



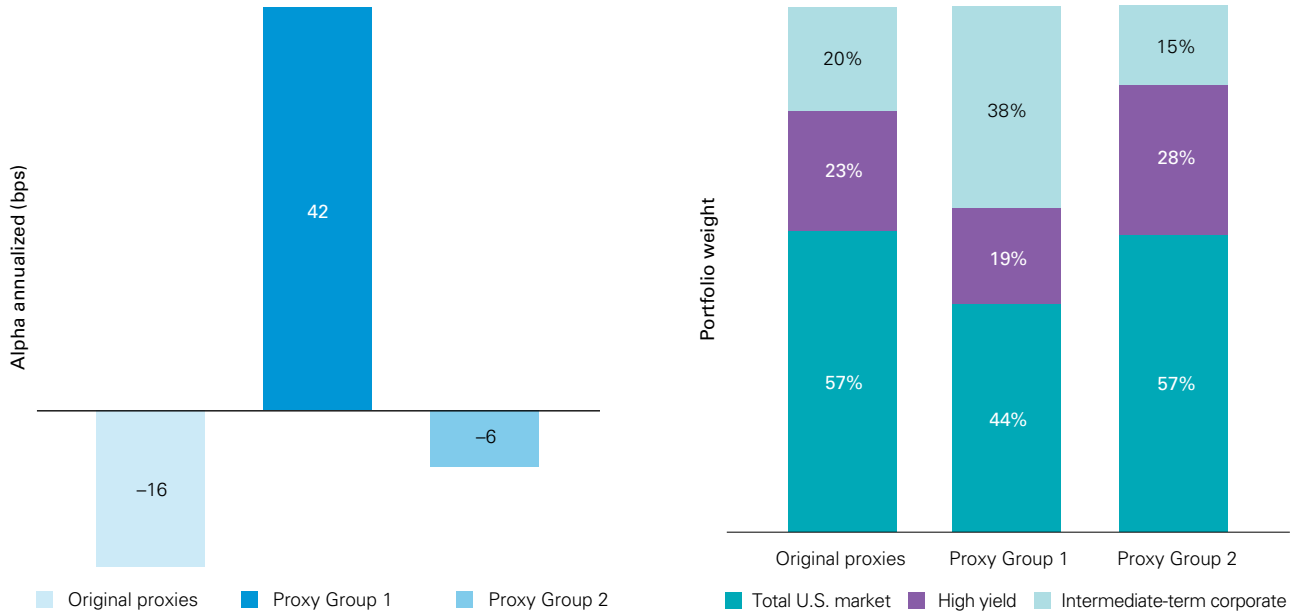
Source: Vanguard calculations, using monthly gross return data from Morningstar from January 1, 2002, to December 31, 2019.

24 This represents the weighted expense ratio of the factor proxy products used in the final factor-mimicking portfolio.

What if we had selected different factor proxies for the analysis?

Different factor products could proxy each factor strategy. They could have different weighting methodologies, be active (if highly risk controlled) or passive, or be run by different asset management firms. To illustrate the impact a different set of factor proxies can have on the RBSA results, we test the same factor strategies with two different asset managers' sets of proxies in **Figure 8**.

Figure 8. Proxy selection mainly affects Fund A's alpha and portfolio weights



	Adjusted R-squared	Tracking error
Original proxies	0.91	0.93%
Proxy Group 1	0.90	0.97%
Proxy Group 2	0.90	0.97%

Note: The return history used starts in 2010 because of limited historical data for proxy groups.

Source: Vanguard calculations, using monthly gross return data from Morningstar from January 1, 2010, to December 31, 2019.

Interestingly, the goodness of fit measures are very close for all three sets of proxies, but the portfolio weights and alphas are different.[25] Because it is also possible that the factor stability is different between these sets of proxies, it is important to properly assess the custom factor portfolio.

The clonability of a manager as assessed by its alpha and goodness of fit can depend on the factor proxies chosen. We encourage investors to try different combinations of factors and proxies until they are comfortable with the result.

Case Study B: Can every manager be cloned?

Not every manager generates returns solely through stable tilts toward factors. Managers often try to add value for end investors through security selection. This is particularly true for those who maintain relatively few holdings and tend to have high levels of idiosyncratic risk that factor strategies cannot mimic. Another source of potential manager alpha is factor timing, which results in factor exposures varying over time and can be challenging to clone.

In the following case study, we analyze a fund that our framework determines is not clonable. Fund B represents a multisector U.S. bond fund with an expense ratio of 91 bps. We start with the broad U.S. bond market and then add U.S. fixed income factor strategies, assessing alpha, goodness of fit, and the statistical significance of the coefficients (see **Figure 9**).

Fund B has outperformed the broad market on a gross basis by a staggering 240 bps per year. Its tracking error is also very high at 6.5%. This means that its return profile has been very different.

Once more, we follow the step-by-step approach detailed in Figure 4, using low-cost investable proxies for term, credit, and other categories. By including factor proxies that span the bond credit quality spectrum in the RBSA (see Figure 9, Round 1), we find that Fund B's credit quality profile is best explained by a combination of investment-grade and high-yield credit, although outperformance and poor fit persist.

After including investment-grade credit factor proxies with different term factor exposure (Round 2 in Figure 9), we rerun the analysis and find that the fund seems to exhibit a tilt toward short-term and intermediate-term corporate credit. Fit is barely improved, and the alpha of the manager increases. In Round 3, we experiment with adding other factors but find that these do not help improve fit and therefore do not include them in the final factor-mimicking portfolio.

We can also see the decrease in the standard error of the short-term and intermediate-term corporate credit that comes from removing the other factor strategies.

25 A potential source of this difference is the term structure of intermediate investment-grade credit. While some proxies target bonds with five to ten years left to maturity, others choose three to seven or one to ten years.

Figure 9. Constrained RBSA results for Fund B, 2000–2019

Factor-mimicking portfolio weights

	Alpha (annualized)	Total U.S. market	Treasury	Total corporate	High yield	Tracking error (percentage)	Adjusted R^2
Round 1: Manager's credit exposure							
Factor strategies across credit quality spectrum	91 bps	0%	0%	67%	33%	4.7	0.56
Standard errors	*(35 bps)*	*(38%)*	*(30%)*	*(13%)*	*(6%)*		

	Alpha (annualized)	Total U.S. market	Short-term Treasury	Intermediate-term Treasury	Long-term Treasury	Short-term corporate	Intermediate-term corporate	Long-term corporate	High yield	Tracking error (percentage)	Adjusted R^2
Round 2: Manager's term exposure											
Including all term and credit factor strategies	130 bps	0%	0%	0%	0%	32%	37%	0%	31%	4.5	0.58
Standard errors	*(38 bps)*	*(62%)*	*(61%)*	*(55%)*	*(12%)*	*(47%)*	*(47%)*	*(15%)*	*(6%)*		
Only non-zero term and credit factor strategies	130 bps					32%	37%		31%	4.5	0.59
Standard errors	*(33 bps)*					*(10%)*	*(10%)*		*(4%)*		

Insignificant Significant at 1% Significant at 5%

Notes: Statistical significance is assessed using adjusted standard error (Lobosco and DiBartolomeo, 1997). Standard error for the alpha term is approximate. Portfolio weights may not add up to 100% because of rounding. For the dynamic regression results performed after Round 2, see Figure 10.

Source: Vanguard calculations, using monthly gross return data from Morningstar from January 1, 2000, to December 31, 2019.

Figure 10 shows that Fund B's factor exposures were very unstable over time. It is possible that the fund employed security selection or factor timing and thus delivered excess returns that could not be imitated using static tilts to off-the-shelf factor products.

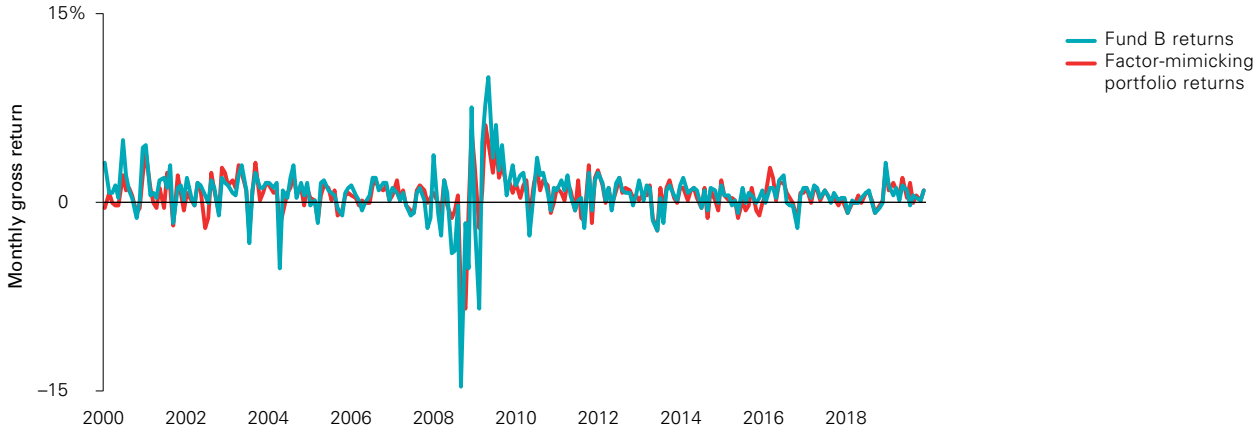
While it may be possible to find some other, more suitable off-the-shelf investable fixed income factor proxies to replicate the return profile of Fund B, because the gross alpha is so high, the fit is so poor, and tweaking the fixed income factor proxies does not tend to lead to substantial difference in fit, we conclude that this fund is likely not clonable using the factor strategies. **Figure 11** shows that the final factor-mimicking portfolio does not clone Fund B's historical returns very well.

Figure 10. Dynamic regression results (36-month rolling window) for Fund B, 2000–2019

Notes: The White test for the entire sample results in a p-value of 0.00; the null hypothesis (homoscedasticity) is rejected at the 1% significance level. We conclude that the error variance is not constant (that is, heteroscedasticity is present). The dynamic regression is run from 2000; the chart starts at 2003 because the regression window length is 36 months.

Source: Vanguard calculations, using monthly gross return data from Morningstar from January 1, 2000, to December 31, 2019.

Figure 11. Fund B's historical returns and those of the final factor-mimicking portfolio do not move in lockstep



Source: Vanguard calculations, using monthly gross return data from Morningstar from January 1, 2000, to December 31, 2019.

What proportion of U.S. active fixed income managers may be clonable using our framework?

To answer this question, we began with all actively managed U.S. fixed income funds in the Morningstar Direct database, including conventional mutual funds and ETFs that were still available as of December 31, 2019. We then removed any that had less than five years of return data or were missing gross return or expense ratio information. We also eliminated all funds that invested in U.S. bank loans, U.S. high-yield bonds, and U.S. nontraditional bond categories and were left with 625 managers.

We applied the fixed income credit quality and term factor strategies used in the case studies to find the tracking-error-minimizing portfolio for each manager for the five years ended December 31, 2019. We began with the following nine U.S. fixed income factor strategies: total U.S. market; high yield; and short, intermediate, and long term factor proxies for the Treasury and investment-grade segments, as well as a cash term using ultrashort Treasuries.

We ran a constrained RBSA for each fund in our sample to determine the three strategies from our choice set that best explained performance over the five years ended December 31, 2019. We found that much of the funds' historical performance could be explained by a combination of total U.S. market, high yield, and ultrashort Treasuries.

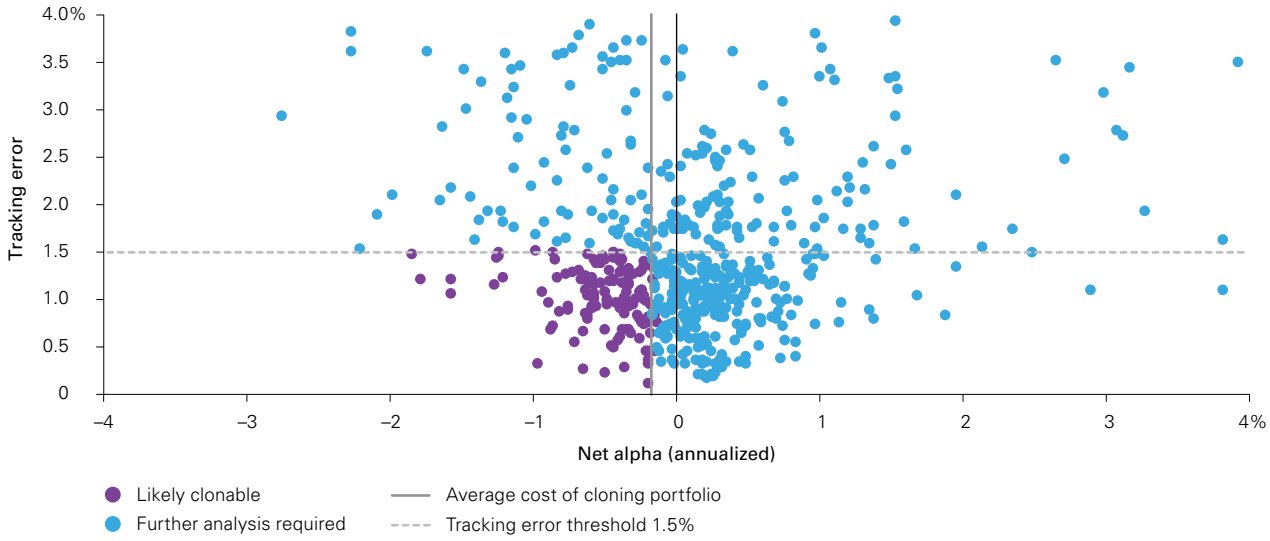
These factor strategies not only generated the highest proportion of statistically significant factor-mimicking portfolio weights and the highest average weights in our sample, they also exhibited low correlation with each other (see Figure 3), which helped control multicollinearity. Using our framework and these factor strategies, we built a factor-mimicking portfolio and estimated the annualized net alpha and tracking error for each fund in our sample.[26]

26 If the net alpha for the manager is negative or not statistically significant, it's likely that the manager did not deliver unique value to end investors for the period studied.

The results indicated that a surprisingly high 22% of the funds were in the "likely clonable" category, as defined by a tracking error of less than 1.5% and net annualized alpha below the cost of their respective factor-mimicking portfolios (see **Figure 12**).[27] Because the three factor strategies may not necessarily be the most appropriate for cloning every manager, the analysis provided a rough estimate of the extent to which U.S. active fixed income managers could be cloned using factor funds.

To obtain a more precise estimate of the number of active funds that can be cloned, our full framework must be applied on an individual fund level. This allows investors to select the factor strategies and their investable proxies most suitable for replicating a particular active manager while managing multicollinearity and accounting for the precise fund costs.

Figure 12. A large percentage of U.S. active fixed income managers could potentially be replaced



Notes: The chart illustrates individual U.S.-domiciled active funds investing in U.S. fixed income. The tracking error and alpha values are calculated versus the custom, investable factor-mimicking portfolio for each manager. The factors used in the analysis include the total U.S. market, total below-investment-grade (high-yield) credit, and ultrashort Treasuries acting as a cash term. The average cost of cloning is estimated to be 20 bps per year, which represents the weighted expense ratio of the factor proxy products of an average factor-mimicking portfolio.

Source: Vanguard calculations, using monthly gross return data from Morningstar from January 1, 2015, to December 31, 2019.

27 The number of funds that were within a 1.5% tracking error and had statistically insignificant net annualized alpha above the cost of their respective factor-mimicking portfolios was zero.

Conclusion

In this paper, we presented an accessible method to test to what extent a traditional active manager has generated a return profile that could be imitated using stable tilts toward well-known factors. Augmenting previous returns-based analysis research, we describe both the full process needed for a robust test and the judgment necessary when addressing important statistical features to ensure that investors do not draw incorrect inferences from the results.

Our analysis employs long-only investable factor strategies to create a fair alternative investment option to a long-only active fixed income manager or a multimanager portfolio subject to real-world costs. The framework builds a bridge between pure academic research on both factor investing and econometric testing and the practical implementation of the insights for investors through increasingly available off-the-shelf fixed income factor products.

Our approach has a number of practical applications, including enhancing factor-oriented due diligence conducted on managers. If the active manager is found to have generated returns that could have been more or less mimicked by factor strategies, investors may be able to produce a similar factor profile with greater transparency, more risk control, and lower costs, thereby raising the bar for some active managers. This type of analysis helps investors both evaluate prospective managers and ensure that managers in the portfolio are adding unique value and their style of investing is true-to-label.

References

Asness, Cliff, Tobias J. Moskowitz, and Lasse H. Pedersen, 2013. Value and Momentum Everywhere. *The Journal of Finance* 68(3): 929–985.

Bao, Jack, Jun Pan, and Jiang Wang, 2011. The Illiquidity of Corporate Bonds. *The Journal of Finance* 66(3): 911–946.

Bender, Jennifer, P. Brett Hammond, and William Mok, 2014. Can Alpha Be Captured by Risk Premia? *The Journal of Portfolio Management* 40(2): 18–29.

Blake, Christopher R., Edwin J. Elton, and Martin J. Gruber, 1993. The Performance of Bond Mutual Funds. *The Journal of Business* 66(3): 371–403.

Brooks, Jordan, Tony Gould, and Scott Richardson, 2020. Active Fixed Income Illusions. *The Journal of Fixed Income* Spring 2020, 29(4): 5–19.

Carhart, Mark M., 1997. On Persistence in Mutual Fund Performance. *The Journal of Finance* 52(1): 57–82.

Christopherson, Jon A., and Frank C. Sabin, 1999. How Effective Is Effective Mix? *Russell Research Commentary*: 1–18.

Fama, Eugene F., and Kenneth R. French, 1993. Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics* 33(1): 3–56.

Israel, Ronen, Diogo Palhares, and Scott Richardson, 2018. Common Factors in Corporate Bond Returns. *Journal of Investment Management* 16(2): 17–46.

Koijen, Ralph S.J., Tobias J. Moskowitz, Lasse Heje Pedersen, and Evert B. Vrugt, 2013. Carry. *Journal of Financial Economics* 127(2): 197–225.

Lobosco, Angelo, and Dan DiBartolomeo, 1997. Approximating the Confidence Intervals for Sharpe Style Weights. *Financial Analysts Journal* 53(4): 80–85.

Lucas, Lori, and Mark W. Riepe, 1996. *The Role of Returns-Based Style Analysis: Understanding, Implementing, and Interpreting the Technique*. Chicago: Ibbotson Associates, Inc.

Mattu, Ravi K., Mukundan Devarajan, Steve Sapra, and Dzmitry Nikalaichyk, 2016. *Fixed Income Manager Selection: Beware of Biases*. PIMCO Quantitative Research.

Roberts, Daren, Thomas Paradise, and Chris Tidmore, 2019. *Global Active Bond Fund Returns: A Factor Decomposition*. Valley Forge, Pa.: The Vanguard Group.

Sharpe, William F., 1992. Asset Allocation: Management Style and Performance Measurement. *The Journal of Portfolio Management* 18(2): 7–19.

Stockton, Kimberly, Scott Donaldson, and Simone Chen, 2019. *Junk or Jewel? Assessing the Role of High-Yield Bonds in a Diversified Portfolio*. Valley Forge, Pa.: The Vanguard Group.

White, Halbert, 1980. A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. Econometrica: *Journal of the Econometric Society* 48(4): 817–838.

Zorina, Inna, Reimar Scholz, and Douglas M. Grim, 2020. *Can Active Equity Managers Be Cloned Using Factor Funds? A Practical Testing Method*. Valley Forge, Pa.: The Vanguard Group.

Connect with Vanguard® > vanguard.com

CFA® is a registered trademark owned by CFA Institute.

ISGFIX 022021