Last Revised: February 20, 2016

Alpha Signals, Smart Betas and Factor Model Alignment

Terry Marsh and Paul Pfleiderer

Terry Marsh is Emeritus Professor of Finance at U.C. Berkeley and C.E.O. Quantal International Inc. in San Francisco, CA.
Email: terry.marsh@quantal.com

Paul Pfleiderer is C.O.G. Miller Distinguished Professor of Finance, Stanford University, Stanford, CA.
Email: pfleider@stanford.edu

ABSTRACT

The authors consider the case for augmenting risk models to be used in portfolio construction to reflect information embedded in the portfolio manager's alphas. They consider both "smart beta" models and cases where alpha signals are partly factor-driven but incorrectly perceived to be stock-specific. In smart beta cases, the authors argue that mechanically augmenting the risk model can cause losses by distorting an otherwise-correct factor structure. For those cases where asset specific alpha signals might unexpectedly be related to hidden systematic factors, errors of omission due to missing these hidden factors are shown to generally result in larger expected losses in portfolio efficiency than do errors of commission when nonexistent "phantom" factors are unintentionally included. When the alpha signals are very noisy, the practice of mechanically augmenting the risk model with a custom risk factor to offset that noise can improve portfolio efficiency. However, in those cases the custom risk factor has nothing to do with underlying sources of true risk that all investors face, but instead serves as a penalty that in a back door way tends to adjust for weak quality of the manager's alphas.

Key words: Factor Alignment Problem, Custom Risk Model, Model Error, Smart Beta, Alternative Beta, ETFs, Segmentation, Risk Factors, Macro Factors, Missing Factors, Phantom Factors, Risk Augmentation, Alphas, Eat Alpha, Mean-Variance Optimization

Recently a considerable amount of discussion concerning the practice of portfolio management has been focused on issues related to the "alignment" of alphas with the risk models used in portfolio construction. Related issues include the notion of "factor investing," custom risk models, and sources of "smart" beta returns that arise when it is realized "that a lot of what we thought was alpha is actually an alternative form of beta".[1] These discussions have produced various recommendations for augmenting and adjusting risk models given forecasted returns, which in a sense reverse the approaches suggested in Sharpe [1974] and Black and Litterman [1992] where forecasted returns would be revised given (among other things) changes in the risk model. In this article we develop some simple calibrated examples to illustrate the challenges one faces when attempting to align a risk model with information about expected returns. Our main message is that portfolio construction can in some circumstances be improved by risk model adjustments, but making mechanical adjustments without a good understanding of the sources of forecasted returns can make things worse. For instance, in a number of "smart beta" cases it is unambiguously a mistake to revise the risk model given those smart betas.

When the source of alphas for different assets is proprietary information or analysis, and the information being uncovered relates to sources that are independent across assets, little or no adjustment to a risk factor model is warranted. However, when the sources of alpha information are not necessarily independent across assets but may reflect in some part common factors that are missing from the risk model, then augmenting the risk model to align it with the latent factor influence on alphas obviously can improve the way in which the available information is used to

[1] Quote attributed to Yazann Romahi in "Threats to hedge fund managers' 'secret sauce'" by Robin Wigglesworth, *Financial Times*, February 18, 2016.

construct the portfolio. It is entirely plausible that such information about missing factor risk could lie hidden in perceived alphas.

While omitting sources of risk is clearly problematic, so is augmenting a risk model with "phantom" factors inferred from alphas. It is clear that there is a trade-off between the possible errors of omission and errors of commission. The calibrated examples we develop are designed to give an intuition for the potential magnitudes of the two types of error and the nature of the tradeoffs involved. For the examples we consider we find that, as a rule of thumb, errors of omission tend to be the most troublesome. This is because of the well-known tendency for any portfolio optimization approach, unless it is fairly tightly constrained, to take relatively extreme positions to exploit what appear to be "near-arbitrage" positions in any situation where perceived alpha spreads are believed to be associated with little or no systematic risk.

We also consider what one can infer from an improvement in portfolio performance brought about by adjusting the risk model to align it with alphas. If *ex post* portfolio performance is increased when an alpha-based "risk factor" is added to the risk model (e.g., a higher Sharpe ratio is achieved), can one conclude that this added risk factor was a hidden risk revealed by the alphas? We show that this is not necessarily the proper conclusion. If the alphas contain much more noise than signal, then adding a risk factor with asset exposures (loadings) that are aligned with the alphas can improve performance, but it does so not by capturing a fundamental source of risk in asset returns but rather by dampening what would otherwise be an inappropriate portfolio response to noisy alphas. In these cases augmentation of the risk model effectively imposes a penalty that tends to offset input errors. The augmentation must be customized for the nature of a given manager's alpha errors, and it is misleading to interpret this as a user-customized model of market risk. Indeed, when properly measured, market risk is by definition common to all investors.

Importance of Understanding the Source of Returns

Before getting into details, it is useful to define the general setting in which the above issues arise. For the most part we will frame the discussion within the mean/variance optimization paradigm, which continues to be the overarching workhorse in portfolio construction.[2] There one is concerned with the first two moments of the joint distribution of future asset returns. We will let (E,V) denote the joint distribution as perceived by investors having no special information or privileged insights, with E being the vector of expected returns and V a (positive definite) variance/covariance matrix. Of course, the relation between E and V is not an arbitrary one. The expected returns E will include the necessary compensation for taking the risks embedded in V needed to equate supply with demand. More specifically, in a simple equilibrium setting these expected returns can be explained as those needed for investors in the aggregate to be willing to hold the available assets in the supply given by the market where the asset holdings expose investors to the risks given by V.[3]

Now consider a portfolio manager who has some special information about asset returns that is not available to the public or a manager who has some special insights into how asset

[2]Of course, we recognize that more involved approaches that account for such things as differences in asset liquidity, the dynamics of changing investment opportunities and non-normal distributions are certainly employed in portfolio management, but we believe that many of the alignment issues can be addressed in the mean variance setting. Moreover, in a recent review of the literature, Markowitz [2014] suggests that the mean variance approach works reasonably well in practice in approximating a good range of expected utility solutions.

[3] In addition to market clearing, we also know that one set of these expected returns E must tautologically align cross-sectionally with corporate fundamentals like market-to-book, profitability, and capital investment if market values are "plugged" into the valuation identity in order to calculate E. The valuation formula is central in Miller and Modigliani [1961] and can be traced back to at least the earlier dividend discount formula in John Burr Williams [1938].

prices are set. That manager will generally have return forecasts that differ from E. Denote these forecasts as E^I where the superscript I indicates that these expected returns are based on special information. The difference between E^I and E gives us the vector of that manager's perceived alphas: $\alpha = E^I - E$. The big question that then arises is whether the existence of these alphas has any implications for asset risk: does the variance/covariance matrix describing returns continue to be V or should it be updated in some fashion to a new risk matrix V^I? One completely misleading answer is that V need not be adjusted at all since E^I can always be "explained" in a beta pricing sense based on the original V. Roll [1977][1980] showed that, for any given V it is always possible to linearly align a set of return premiums with risk measured as a V-implied beta with respect to any one of an infinite number of efficient portfolios. In other words, alpha and expected returns can always be mechanically aligned with risk irrespective of whether those returns are *ex post* or *ex ante*, but this alignment is nothing more than a mathematical tautology.[4]

Concerns that alphas may be accompanied by missing factors and a resulting misspecification of V have led some to consider augmenting V with additional factors, often by introducing a set of orthogonal "factor risk" portfolios that align with given alphas. As an example, consider the general approach taken by Fama and French[5] in addressing the "anomalously high" average returns that had been historically documented for stocks with low M/B and small market capitalization. Fama and French posit an underlying return generating process that, in addition to the traditional market factor, includes factors designed to align with

[4] MacKinlay [1995], MacKinlay and Pastor [2000], and Asgharian and Hansson [2005] extended the Roll analysis by couching the alignment exercise in terms of a unique optimal orthogonal portfolio that "augments" the factor model for the existing risk forecast in order to form the Roll tangency portfolio given E^I and V. But the exercise itself is again based on a mathematical tautology.
[5] For example, see Fama and French [2014].

these anomalies. These aligning factors are captured by the returns of portfolios formed from cross-sectional rankings of stocks based on characteristics like M/B and size.[6]

If the perceived alphas do indeed partly reflect hidden risk exposure, then using information about risk contained in the alphas to augment the risk model should improve portfolio construction, especially in extreme cases where alphas look like they imply “near arbitrage” opportunities when this hidden common risk is not taken into account. But a portfolio efficiency improvement due to the augmentation can’t be unambiguously interpreted as a remedy for missing risk. In particular, we show that in the situation where the alphas are simply pure noise, a mechanical augmentation of the variance-covariance matrix can in many cases improve portfolio efficiency by “cancelling out” that noise. Nevertheless, for risk control purposes like hedging, defining cVaR limits, or constructing a minimum volatility index, where V alone is important, we certainly would not want to use such a mechanically augmented V.[7]

A specific instance where it is very important to consider the source of the alphas is that of “smart beta.” The term “smart beta” is not precisely defined and, importantly, the economic factors that give rise to smart betas are often left mysterious. It may be that smart betas arise because of some underlying risk factor, but they could arise for reasons unrelated to risk. For

[6] It has become quite routine for academics to check whether observed historical excess returns can be explained by plausible risk augmentation in this way, though as Lewellen, Nagle and Shanken [2010, p. 176] point out: “…loadings on almost any proposed factor are likely to line up with expected returns – basically all that is required is for a factor to be (weakly) correlated with SMB or HML.” Another example can be found in Ceria, Saxena and Stubbs [2012] who advocate: “…augmenting the user risk model with an additional factor…[which aligns with alphas and]…which is orthogonal and uncorrelated with all the existing risk factors.” Lee and Stefek [2008] propose a custom risk model along the lines of “…building a new risk model that includes the alpha factors.” Augmenting a given existing model by adding additional factors that are possibly related to alphas is a distinct exercise from the augmentation approach presented in MacKinlay [1995] in which the variance-covariance matrix (V) is not changed.

[7] Jagannathan and Ma [2003] focus on the minimum variance portfolio and show that a covariance matrix modification that “…typically shrinks the larger elements of the covariance matrix towards zero” (p. 1676) can benefit minimum-variance portfolio construction even when it is equivalent to imposing a “wrong” constraint that portfolio weights be positive. Since augmenting the covariance matrix is more likely to increase than shrink larger elements, it is reasonable to conjecture that the augmentation would also have a negative effect in constructing a minimum volatility portfolio.

example, as is shown by Black [1972] and later in Frazzini and Pedersen [2014], when a significant portion of the market faces restrictions on holdings or other frictions, extra returns can be earned by unrestricted investors that are unrelated to an additional risk factor. Of course, it is always possible that the extra "smart beta" returns that an investor perceives do in fact arise because of some risk that the market is pricing but is not salient to the investor. In the next section we consider an example where these two possibilities exist and explore the losses that can occur when one case is mistaken for the other.

Even when a manager's alpha signals might appear to be mainly related to information, the possibility still exits that they contain some information about missing factor (systematic) risk. Since the true nature of the underlying information is not always explicitly known, there is model uncertainty. The costs of the model uncertainty involve both errors of omission when hidden factor risk is ignored and errors of commission that occur when the factor structure is "over-augmented" in ways that reduce full exploitation of expected return information.[8] We analyze the trade-off between these errors of omission and commission in the risk factor model within the context of a relatively general model for alpha signals.

Aligning Apples with Oranges:
Cross Sectional Smart Beta Premiums versus Factor Risk

In general one can think of alpha signals and smart betas as sources of information that can be used to construct portfolios to "beat" market performance. In some cases the reason why the alpha or smart beta creates perceived superior performance will be well understood. For example, someone who had looked very intelligently at the financial reporting of Enron before

[8]Given that Harvey, Liu, and Zhu [2016,] reports 314 factor candidates in the literature, there is no shortage of possible "phantom factors."

that company's fall could have generated information that was apparently not widely known by investors and not reflected in the market price.[9] In such a case the source of the alpha is the superior ability of the investor to use information to identify stocks that are mispriced. This superior information about fundamentals (e.g., future cash flows) can be used in a straightforward manner to generate superior performance as detailed in the next section. In other cases it is less clear what creates the ability to "beat" the benchmark. This is frequently the case with smart betas. Smart betas are often discovered empirically by finding that a tilt away from the benchmark toward stocks having certain attributes and away from stocks having other attributes has historically produced what appear to be superior returns on a risk adjusted basis. Unlike the Enron alpha case, the reason for this outperformance is often not obvious. In particular, since smart betas involve tilting toward certain types of stocks, there is always the possibility that at least part of the extra return gained arises from the presence of a hidden risk factor that the market is pricing. When a hidden risk factor is possible, one wants to make sure that the risk model is adjusted to account for the hidden factor and make it visible.

In this section we develop a simple illustrative example to highlight the challenges involved in making sure that the risk sources (if any) behind smart betas are accounted for. We look at a simple case where smart beta returns may arise in two different ways, one that does not involve hidden risk and one that does.

> **Case 1**: In this case the pricing anomalies behind the smart beta returns arise because of market segmentation and no hidden risk factors are involved. The pricing effects arising from segmentation that we analyze are based on the early insights and analysis found in Black [1972], which are subsequently developed further in Frazzini and Pedersen [2014]. In the basic segmentation model it is assumed that some investors are unable to take leveraged positions, i.e., unable to borrow to purchase risky assets. Both the leverage-constrained and leverage-unconstrained investors form optimal portfolios given their

[9] As is well known, James Chanos did just that.

respective constraints. For the market to clear, investors in the unconstrained group need to effectively "take the other side" of any deviations from holding the market portfolio that the constrained group make as they optimally adapt to their leverage constraint. The equilibrium pricing gives rise to a smart beta strategy that can be exploited by the unconstrained investors and (in our example) is realized by tilting toward value stocks and away from growth. Numerous "real world" impediments and frictions can lead to similar segmentation in investing and thus smart-betas.[10]

Case 2: In this case the extra "smart beta" return is in truth explained by a hidden catastrophic risk factor which means that a risk adjustment is necessary for proper portfolio construction. We calibrate this case so that the equilibrium expected returns that emerge are exactly the same as those in the segmentation Case 1, so all gains and losses involved in mistaking one case for the other hinge on having the right risk model.

If the source of the extra return produced by the smart beta exposure is absolutely clear so that one knows with certainty which case applies, then it is equally clear whether the risk model needs to be adjusted. However, if there is some uncertainty surrounding the source of the returns so that both Case 1 and Case 2 are possible, then both errors of commission (adjusting the risk model when no adjustment is needed) and errors of omission (not accounting for risk that is present) are possible. We look at the differential consequences of the two errors.

To compare errors of omission and commission, we first need to describe the equilibrium risk and return characteristics that each case produces. In the market segmentation case (Case 1) we assume that a significant fraction of the investing population that is tolerant of risk faces a leverage constraint. This will give rise to optimal investment weights that differ from the market-

[10] Other potential sources of segmentation include: home-country versus international investing; tax-loss selling "pressure" by substantial clienteles that is widely believed to induce non-risk-related calendar (January) effects in returns for subsets of assets; and potentially even unspecified "common animal spirits" that cause a significant set of investors to act "as if" they are constrained or myopic. In Merton [1987], a subset of investors doesn't hold small cap stocks about which they have no knowledge. In his case the exogenous constraint is expressed directly in terms of not holding a specified group of assets. In the case here, leverage is the exogenous constraint, and investors' choice of which stocks to hold is endogenous. Also, when investment decisions are delegated by investors to professional managers, Admati and Pfleiderer [1997] and Alankar, Blaustein and Scholes [2014] discuss how the use of benchmarks for performance assessment and fee determination can motivate managers to make portfolio choices that are sub-optimal for the investors, just as if the investors themselves were "constrained" to the suboptimal behavior (and conversely, proper alignment of returns and factor risk can be interpreted as a critical ingredient in better benchmark design). Investors' inability to trade human capital and self-imposed ESG/SRI investment constraints are other frequently encountered examples of shifts in the demand of subsets of investors away from simply holding equities at their market weights (the average investor holds investments at market weights given that the market clears).

index for both leveraged and unleveraged clienteles of investors. It also leads to "enhanced" risk adjusted returns for stocks with low-risk characteristics (value stocks). To see how the enhanced returns arise, we look specifically at the investment positions chosen when investors allocate across three assets that can be interpreted as ETFs or passive funds.[11] The first is a fund of low risk "value" stocks with an annualized volatility of 15%, the second is a fund composed of intermediate risk stocks that has a 20% volatility, and the third is a high risk "growth" stock fund with a 25% volatility. We simply assume that the cross-correlation between each pair of funds is 80%. This is shown in Exhibit 1.

Exhibit 1
Assumed Three Fund Environment

			Correlations		
	Market Weights	Volatility	Value	Neutral	Growth
Value	33.33%	15.00%	1.0	0.8	0.8
Neutral	33.33%	20.00%	0.8	1.0	0.8
Growth	33.33%	25.00%	0.8	0.8	1.0

We further assume that 40% of investable wealth is held by investors (Group "A") with relatively low risk tolerance (ρ= 0.4), while the rest of the wealth is held by investors more tolerant of risk (ρ= 0.9). Only 10% of the risk tolerant investors can use leverage to increase their risk exposures and obtain higher returns (Group "B"). The others (the constrained investors in "Group C") can only achieve higher returns by tilting their portfolios toward the growth sector and away from the value sector. In choosing their optimal portfolios, we assume that all investors use the correct risk model (that given in Exhibit 1) and in this sense they are *correctly aligning*

[11]Typical smart beta strategies in practice tend to group stocks cross-sectionally into funds like this. For example, Ang, Goetzmann and Schaefer [2009] define a "VALGRTH factor" as a zero-cost portfolio/fund that is long value stocks and short growth stocks, and then consider varying degrees of tilt toward that value portfolio and away from a benchmark. It is also virtually a computational necessity to group stocks when solving for a segmented equilibrium as here.

the risk model with their expected returns in whichever constraint situation they find themselves. Expected returns must be such that in the aggregate investors hold the sectors in their market proportions. We somewhat arbitrarily set the rate on cash at 2% and this will be the rate at which Group B investors can borrow from investors in Group A. Finally, we assume that cash is in zero net supply so whatever group B investors borrow, it must be equal to what group A investors wish to lend. Exhibit 2 shows the equilibrium holdings and expected returns.

Exhibit 2
Market-Clearing Equilibrium when a
Segment of Investors are Borrowing-Constrained

	Portfolio Holdings			Expected	CAPM	
	Group A	Group B	Group C	Return	Exp Ret	Alpha
Cash	20.69%	-137.93%	0.00%	2.00%	2.00%	0.00%
Value	57.14%	171.43%	0.35%	8.14%	7.53%	0.60%
Neutral	18.21%	54.64%	42.17%	9.53%	9.52%	0.01%
Growth	3.95%	11.86%	57.48%	10.97%	11.58%	-0.61%
Total	100.00%	100.00%	100.00%			

Investors in groups A and B hold stocks in identical proportions (around 72% of their stock holdings are in the value sector, 23% in the neutral sector and only 5% in the growth sector). The lending by investors in Group A allows the investors in Group B to take a highly leveraged position in the stock portfolio. The constrained investors in Group C, who are as risk tolerant as those in Group B, can only achieve higher returns by skewing their portfolios in the direction of growth. Ultimately this lowers the return of growth stocks, relative to what the CAPM would predict, and increases the returns of the value and neutral stocks, again relative to the CAPM prediction. As reported in Exhibit 2, the value-growth (VALGRTH) return spread with our calibration is 121 bps.

Our Case 1 specification includes the elements typically mentioned in practice as "defining" smart beta, *viz:* unconstrained investors wish to increase returns by following a passive investment policy which calls for tilting away from the market weights to capture a premium associated with a particular part of the investment universe (e.g., "value"). A first question for our calibration is whether an expected value-growth return spread of 121 bps is "close enough to reality" for the calibration to be interesting. We look to three empirical references that define value/low-risk stocks and portfolio strategies similarly to that here: (i) Lesham, Goldberg and Cummings [2015] show that portfolios that are constructed to tilt toward low book-to-price stocks while controlling for other sources of risk and real-world friction (which aren't in our framework either) had a 95 bps active geometric average return from January 1973 to December 2013; (2) The geometric average of annual returns on HML from Ken French's Web-site for the last 20 years (from 1995 to 2015) is 127 bps; and (3) the general industry consensus is that the *realized* value/growth return spread has been small or negative since the peak of the financial crisis in October 2008, i.e. inexpensive "value stock" valuations have instead been relatively expensive.[12] These recent realized spread returns are consistent with a low point estimate like 1.21% for the *ex ante* expected premium.

While the value-growth spread in the segmentation model appears broadly consistent with observed spreads, at least in recent times, we've found that it is generally difficult to calibrate "plain vanilla" borrowing-constrained segmentation models to produce return premiums greater than a couple of percent without implausibly extreme parameters. In particular, in their "betting against beta" paper cited earlier, Frazzini and Pedersen [2014] report empirical average annualized excess returns of 10.92% from 1926 to 2012 for U.S. equities, and excess

[12]Though even this statement is sensitive to the definition of value, e.g. in 2015 value stocks defined in terms of profitability/quality measures did *ex post* outperform.

returns for international equities in the period 1984-2012 that range from an annualized 1.32% for Australia to 11.76% for the Netherlands. In the simple type of segmentation model we use for our example here we are unable to calibrate anything close to observed double-digit returns.

Returning to risk alignment, note that the "excess returns" on value stocks and growth stocks are derived for Case 1 under the assumption that all investors are optimizing their portfolios using the *correct* variance-covariance matrix. Obviously there is no issue in Case 1 of *factor-risk* alignment nor is there any "factor investing" in a meaningful sense. The premiums that are obtained by the unconstrained investors are not due to their taking on some hidden factor risk but are instead the result of the demands of constrained investors that lead to value stocks being "underpriced" relative to what they would be in a world where no investors are constrained.

Proceeding now to Case 2 we switch to the assumption that the existence of a hidden risk factor is the source of the perceived smart beta extra returns where those extra returns turn out to be exactly the same as in Case 1: 0.60%, 0.01% and -0.61% for Value, Neutral and Growth. In Case 2 we are essentially assuming that the collective wisdom that ultimately prices assets in the market accounts for risk exposures that are not easily assessed and are potentially missed by some would-be smart-beta investors. There are, of course, many ways that a not easily detectable — but nevertheless priced-by-market-consensus — risk factor could explain observed alphas that are incorrectly attributed to a "smart beta" strategy. We consider a particularly simple possibility in which, in addition to the risks that are summarized in Exhibit 1, there is an additional risk of a catastrophic drawdown that affects all stocks and occurs with a very small probability π. Because it occurs with small probability it is not salient to many investors. In the event that the catastrophic outcome occurs, it causes a negative shock to the returns of stocks of type i equal to $-\gamma_i$. If no catastrophic shock occurs, then each stock has a very small positive return of $\pi\gamma_i/(1-\pi)$ on top of the returns produced by other risk exposures. This means that this additional catastrophic risk factor has an expectation equal to zero, as a factor representing a shock should. To proceed with a concrete example we assume that the shock occurs roughly every 40 years, so that $\pi = 0.025$ and the values for $(\gamma_{Value}, \gamma_{Neutral}, \gamma_{Growth})$ are $(0.598,\ 0.405,\ 0.203)$. Under these assumptions the (unconditional) variance-covariance matrix must be adjusted as shown in Exhibit 3. The equilibrium returns produced when there is market segmentation and the variance/covariance matrix is that given on the left in Exhibit 3 are identical to those produced when there is no market segmentation but the variance/covariance matrix is that given on the right. [13]

Exhibit 3
Comparison between the Covariance Matrices
in the Segmented Market Equilibrium and in the Catastrophe Factor Equilibrium

	Case 1 Variance Covariance Matrix			**Case 2 Variance Covariance Matrix (Includes Catastrophe Risk)**		
	Value	Neutral	Growth	Value	Neutral	Growth
Value	0.023	0.024	0.030	0.032	0.030	0.033
Neutral	0.024	0.040	0.040	0.030	0.044	0.042
Growth	0.030	0.040	0.063	0.033	0.042	0.064

An investor or manager who is optimally constructing his portfolio[14] in the Case 1 segmented world would make an error of commission if he or she were to use the augmented risk

[13] One technical adjustment must be made to make the equilibrium returns the same in the two models. In Case 1, the implied societal risk tolerance is approximately 0.46. It should be noted that this is less than the average risk tolerance in the economy (i.e., 0.66) and this difference arises because of the effects of the leverage constraints on a significant fraction of investors. In case 2, since there is additional risk due to the catastrophe factor, implied societal risk tolerance must be a bit higher for the market risk premium to remain the same. In the example we construct, implied societal risk tolerance is 0.52.

[14] We assume that the measurement of the premiums is noise-free. If they are not, cross-sectional firm characteristics like market-to-book and profitability that frequently define smart beta may well appear to explain the premiums simply because they serve as return proxies.

model shown on the right hand side in Exhibit 3. Similarly, an investor in the Case 2 world who does not include catastrophe risk in his risk calculations and instead uses the risk model on the left in Exhibit 3 makes an error of omission. How large are the respective errors that might be made by using the wrong risk model in these two settings? In Exhibit 4 we show one example of the losses that can occur by using the wrong risk model. We assume that the investor is unconstrained and has the high level of risk tolerance associated with Group C as described above (ρ= 0.9).

Exhibit 4
Losses when the Wrong Risk Model is used by an Unconstrained and Very Risk Tolerant Investor

	Portfolio Weights				Segmented Model is Correct			Catastrophe Model is Correct		
	Cash	Value	Neutral	Growth	Exp Ret	Volatility	CE	Exp Ret	Volatility	CE
Portfolio for Segmented	-137.93%	171.43%	54.64%	11.86%	17.70%	37.59%	9.85%	17.70%	42.75%	7.55%
Portfolio for Catastrophe	-74.42%	58.14%	58.14%	58.14%	15.16%	32.53%	9.28%	15.16%	34.41%	8.58%
						Loss	0.57%		Loss	1.03%

Obviously Exhibit 4 displays a rather extreme case involving an aggressive risk tolerant investor, but as shown in Exhibit 5 the losses are still significant if risk tolerance is more moderate (ρ≈ 0.5).

Exhibit 5
Losses are Still Material when the Wrong Risk Model is used by a Moderately Risk Tolerant Unconstrained Investor

	Portfolio Weights				Segmented Model is Correct			Catastrophe Model is Correct		
	Cash	Value	Neutral	Growth	Exp Ret	Volatility	CE	Exp Ret	Volatility	CE
Portfolio for Segmented	-36.42%	98.29%	31.33%	6.80%	11.00%	21.55%	6.50%	11.00%	24.51%	5.18%
Portfolio for Catastrophe	0.00%	33.33%	33.33%	33.33%	9.54%	18.65%	6.17%	9.54%	19.73%	5.77%
						Loss	0.33%		Loss	0.59%

The important observation to make is that the error of omission (assuming there is no catastrophic risk factor when in truth there is) is almost twice the error of commission (falsely assuming a non-existent catastrophic risk factor).

This comparison of Cases 1 and 2 is not the whole story however. We considered only the one alternative, a catastrophe model Case 2 as an alternative to smart beta. But of course many other factor models beside catastrophe could also be consistent with the covariance matrix on the right-hand-side of Exhibit 3 – in practice it would be hard to differentiate models with say ten years of historical regression when catastrophes occur on average once every 40 years! Also in the Case 2 alternative we have assumed that the probability π of a catastrophic shock is constant, though it likely varies through time conditional on changes in the market environment that are important to investors. These changes manifest themselves as changes in aggregate market volatility — "volatility in volatility" — and "fat tails" in unconditional returns. There is also evidence that market risk aversion increases in bad states, and in recent years, there has also been significant market uncertainty associated with timing and consequences of "exit" from the long period of quantitative easing. In all these cases, the infrequent but significant downside results potentially complicate risk model alignment since they could plausibly result in a strong negative skew in returns.[15]

Nor is the indeterminacy in potential factor models that could be consistent with the observed risk premiums the only problem faced in comparing the pros and cons of factor alignment exercises. There are of course also many potential dimensions along which Case 1

[15] In particular the issue leads naturally a measure of a stock's downside tail risk – concepts like "lower tail dependence" (LTD) between a stock's downside returns and market conditions (Chabi-Yo, Ruenzi and Weigert [2015] and Weigert [2015]); "tail betas" and "systematic tail risk" (Kelly and Jiang [2014], van Oordt and Zhou [2013]); "extreme downside risk" (EDR) (Huang, Liu, Rhee, and Wu [2012]), "hybrid tail covariance risk" (H-TCR) (Bali, Cakici and Whitelaw [2014]); and separate risk-return relationships for bull and bear markets (Wu and Lee [2015]).

segmentation of asset demands could occur.[16] Even restricting ourselves to the segmentation environment as we've prescribed it here, there is still risk that the segmentation itself shifts in various ways in practice. For example, if previously constrained investors eventually work around their borrowing constraints, the smart-beta strategy could *ipso facto* become "crowded" and thereby erode the value premium (resulting in higher-than-anticipated short-term returns as the crowding occurs). In the other direction, in a crisis market situation, the investors who previously considered themselves unconstrained could find themselves facing unforeseen leverage-induced problems and, in unwinding positions in value stocks, potentially push prices down and subsequent premiums up. In short, in the presence of market-dependent risks like these, segmentation and factor risk are not necessarily mutually exclusive hypotheses to explain premiums.

Alpha Signals and Alignment with "Beta" Risk

In this section we continue to focus on the question of whether expected portfolio payoffs that an active manager hopes to gain by making deviations from a market benchmark are worth the risk they entail, and in particular whether the expectation and risk sides of the trade-off are well aligned. We assume now that the manager has an edge in the form of superior proprietary information and/or analytical skills that give rise to what we call alpha signals. In this situation, an alignment problem arises if the manager assumes that all of the alpha signals being used in

[16] For example, the cost of gathering information about investments plausibly translates into differences in demand for assets with domestic versus foreign and small versus large firm characteristics. Investor cash-flow horizons will make maturity and duration of cash payouts important, as has been long recognized in fixed income management (see also Lettau and Wachter [2007]). Empirically, Daniel and Titman [2011], Lewellen, Nagel, and Shanken [2010], and Harvey, Liu, and Zhu [2016] point to the large number of characteristics reported to align cross-sectionally with historical return premiums, just as we would expect if these other characteristics were even weakly cross-sectionally correlated with Fama-French HML or SMB. Bali, Brown, Murray and Tang [2015] directly hypothesize that a subset of investors have a "lottery demand" for stocks that cross-sectionally have a high probability of large price changes. Here, that lottery demand could follow from the leverage constraint.

portfolio construction are mainly about stock-specific returns when in reality the alpha signals partly contain a factor component that is pervasive across stocks but missing or otherwise mis-specified in the risk model being used. It is plausible that misalignments of this sort do occur and can be a problem in practice. For example, this is a potential risk when teams of buy-side portfolio analysts do traditional research on their own assigned sets of companies in a fund universe and analysts' output gets pooled at the fund level without fully accounting for common factors that might be present in the individual analyst assumptions and information sources.

To be concrete, we assume that returns can be described by the following familiar linear equity factor structure:

$$\tilde{R}_t^i = E_t\left(\tilde{R}_t^i\right) + \beta_{1,t}^i \tilde{F}_{1,t} + \beta_{2,t}^i \tilde{F}_{2,t} + \cdots + \beta_{K,t}^i \tilde{F}_{K,t} + \tilde{\varepsilon}_t^i \qquad (1)$$

where: $\tilde{R}_t^i$ is the return on asset *i* over *t*, a short time interval,[17] $E_t\left(\tilde{R}_t^i\right)$ is the market's expectation of the asset *i*'s return at time *t*; $\tilde{F}_{k,t}$, $k = 1,2,\ldots,K$ are the (mean-zero) factors whose influence is pervasive across the returns of a large number the assets in investors' universe; the coefficients $\beta_{k,t}^i$, $k = 1,\ldots,K$ are the exposures of the *i*th asset's return to the *k*th factor; and $\tilde{\varepsilon}_t^i$ is the component of asset *i*'s return due to asset-specific influences.

The standard expression for the *market equilibrium* expected return in period *t* for stock *i* in (1) can be written as:

$$E_t\left(\tilde{R}_t^i\right) = R_t^F + \lambda_{1,t}\beta_{1,t}^i + \lambda_{2,t}\beta_{2,t}^i + \cdots + \lambda_{K,t}\beta_{K,t}^i \qquad (2)$$

[17]We assume that the time interval is relatively short. Over a short interval linearity in (1) is reasonable, at least for stocks of companies with no leverage or reasonably low leverage. Of course, we don't want the interval length to be extremely short (e.g., a minute or a second), since for very short intervals microstructure effects can dominate.

In (2) the factor premiums $\lambda_{k,t},\ k=1,\ldots,K$ are reflected in the market price of each asset *i*, and R_t^F is the period *t* risk free rate. As it stands, the model itself says nothing about whether the market risk premium for any factor is positive, zero, or negative.

We assume that the manager has, at the beginning of period *t*, a set of signals S_t that allows him to predict $\tilde{\varepsilon}_t^i$ in (1) more accurately than the market's expectation that it is zero. Given the manager's signals, his alpha for the return on asset *i* is $E(\varepsilon_t^i \mid S_t)$, so his overall expected return for asset *i* becomes:

$$E_t\left(\tilde{R}_t^i \mid S_t\right) = R_t^F + E(\varepsilon_t^i \mid S_t) + \lambda_{1,t}\beta_{1,t}^i + \lambda_{2,t}\beta_{2,t}^i + \cdots + \lambda_{K,t}\beta_{K,t}^i \tag{3}$$

For example, the signals received by the manager could be the "true" values of the ε_t^i 's plus individual asset noise terms, i.e.:

$$s_t^i = \varepsilon_t^i + \eta_t^i, i = 1, \ldots, N \tag{4}$$

with the additional assumption that the random signal noise in (4) is independently distributed across different assets, i.e.: $E(\eta_t^i \eta_t^j) = 0$ for $i \neq j$.

Assuming for simplicity that the signals and noise terms are jointly normal in distribution, the alphas are:

$$\begin{gathered} E_t\left(\tilde{R}_t^i \mid Signals_t\right) \equiv E(R_t^i) + Alpha_{i,t} = E(R_t^i) + w_{\varepsilon,i} s_t^i, \text{ where} \\ w_{\varepsilon,i} = \frac{\sigma_{\varepsilon,i}^2}{\sigma_{\varepsilon,i}^2 + \sigma_{\eta,i}^2} \end{gathered} \tag{5}$$

Equation (5) formalizes the common intuition that the higher the noise $\sigma^2_{\eta,i}$ in the signal, the smaller should be the update in expectations after seeing the signal: Poor quality signals don't tell a manager much about returns, so if $\sigma^2_{\eta,i}$ is high relative to $\sigma^2_{\varepsilon,i}$, then $w_{\varepsilon,i}$ is low. It is also possible that the signal doesn't contain genuine information because the signal is already discounted into stock prices, i.e. the innovation in stock *i*'s return has already occurred; the effects of this case can effectively be captured by assuming that $\sigma^2_{\eta,i}$ is very high relative to $\sigma^2_{\varepsilon,i}$. Note that in (5) we are positing no "alpha factors" and reserve the term "factor" for its traditional meaning of a pervasive influence across a vector of time-series variables, here stock returns.

Finally, we extend the basic model (5) to encompass a situation where the information signals have implications for the valuation of more than just a single stock, i.e. they contain macro content. In this case the alpha signals s^i_t *(i=1,...,N)* for the *i*th stock in (4) are assumed to be actually generated as follows:

$$s^i_t = \xi^i_t + \theta^i_{K+1,t} F_{K+1,t} + \eta^i_t \tag{6}$$

where: ξ^i_t is the "true" stock-specific component of the signal s^i_t, and the ε^i_t in (4) becomes: $\varepsilon^i_t = \xi^i_t + \theta^i_{K+1,t} F_{K+1,t}$, where we assume that $F_{K+1,t}$ is independent of all other factors. (Note that if $\theta^i_{K+1,t} = 0$ we are back to the earlier instance of stock-specific signals). The idea is that there is a missing factor $F_{K+1,t}$ that has a latent influence on alpha signals across stocks. Without loss of generality, we can also continue to assume that $E(\eta_i \eta_j) = 0$ for $i \neq j$.

To illustrate the impact of various errors that can arise in portfolio solutions given the manager's alpha signals and the factor model presented in (1) − (6), we consider the following four scenarios that might define the true underlying "state of the world."

- *Base Case – No Alpha, No Additional Factors*: For the base case, we assume an S&P 500 universe and use the Quantal equity factor structure as it was forecasted on August 9, 2014, along with market weights for that day. On that day the forecast S&P 500 volatility was approximately 20% annualized, and we assume that at that time the market aggregate risk aversion was consistent with an *ex ante* expected market risk premium of 5%. Letting *V* denote the S&P forecasted variance-covariance matrix, the vector of expected excess returns $E^{(Base)}$ is assumed to be: $E^{(Base)} = \rho V^{(Base)} w$, where *w* is the vector of market weights and *ρ* is the measure of societal risk tolerance.

- *Case 1 – Pure Alphas, No Additional Factors*: To keep our alpha specification for (5) and (6) generic and transparent, we do the following: (i) Sort the S&P stocks by ascending market weights and number them 1 to 500; (ii) Define a 500x1 vector *h* with $h_i = 1$ for stock *i* if *i* is even, and $h_i = -1$ if *i* is odd; (iii) Define the vector of alphas to be $\pi h / 10$ so that expected returns (including the alpha) are $E^{(1)} = E^{(base)} + \pi h / 10$. The parameter π reflects the strength of the manager's alpha and in our illustrations below we consider values for π ranging from 0% to 2%. At the high end of 2%, π translates into roughly 0.17 in terms of the familiar information coefficient. Note that since this is the pure alpha case and there are no additional factors, $V^{(1)} = V^{(Base)}$.

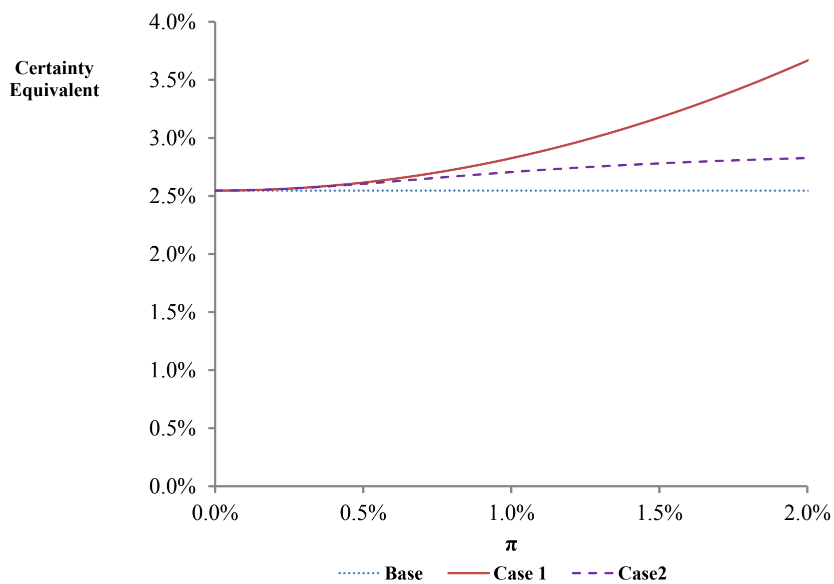
- *Case 2 – Alphas accompanied by an Additional Factor*: In this case, expected returns are the same as in Case 1, i.e. $E^{(2)} = E^{(1)} = E^{(Base)} + \pi h / 10$, but we now assume that underlying these alphas is an additional source of systematic risk, which we take to be an independent factor with loadings equal to *h* and volatility equal to *π*. This means that the true variance-covariance matrix is $V^{(2)} = V^{(Base)} + \pi^2 hh'$.

- *Case 3 – An Additional Factor with no Alphas*: Finally, for the sake of completeness we consider the case where the additional factor described in Case 2 is present, but there are no alphas, which means that $E^{(3)} = E^{(Base)}$ and $V^{(3)} = V^{(2)} = V^{(Base)} + \pi^2 hh'$.

It should be noted that the additional factor assumed to be present in Cases 2 and 3 adds essentially no risk to the market portfolio. This is because *h* is constructed so that $w'h \approx 0$. Thus the market portfolio is the optimal portfolio in both the Base Case and Case 3.

In Exhibit 6 we show for each of the cases the certainty equivalent return that can be achieved by an investor with average risk tolerance. (We do not include Case 3, since as just discussed, that scenario and the base scenario are essentially equivalent as the market portfolio is optimal in each and the additional factor in scenario 3 adds no risk to the market portfolio.)

Exhibit 6
Certainty Equivalents for the Base Case and Cases 1 and 2 as a function of Alpha Signal Strength π



Note that, not surprisingly, the certainty equivalent that can be achieved in Case 1 exceeds that achievable in Case 2 because exploiting the alphas in Case 2 involves taking on additional systematic risk over what is entailed in exploiting the alphas in Case 1.

Now consider the various errors that can be made by incorrectly assessing the true state of the world as depicted by the above cases. We begin by assuming that the true state of the world is defined by the Base Case (there are no alphas) but a manager believes there are alphas and also believes that the appropriate scenario is either Case 1 or Case 2.

Exhibit 7
Loss When True State is the Base Case,
but the Manager Incorrectly believes Alphas Exist

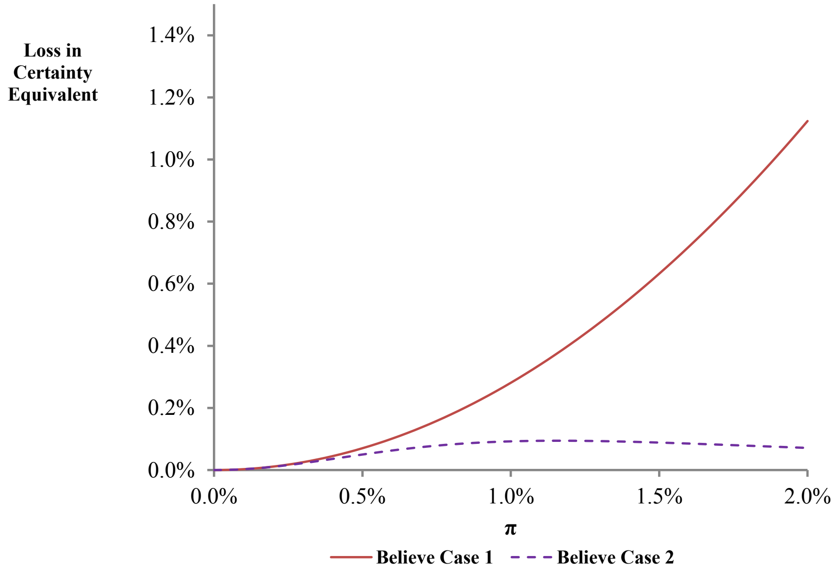
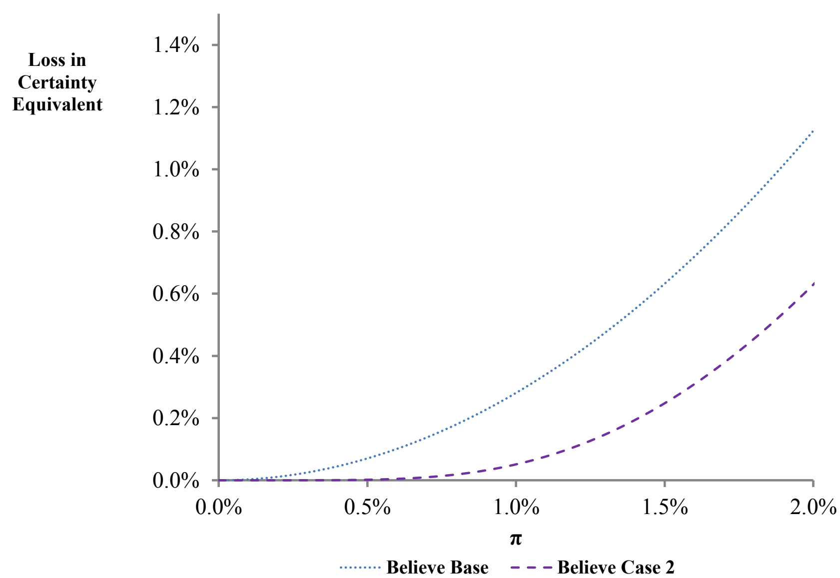


Exhibit 7 shows the anticipated degradation in portfolio performance when the portfolio is tilted in the direction of alphas that don't exist. As can be seen, that loss is substantially reduced when the manager mistakenly believes that Case 2 holds instead of mistakenly believing that Case 1 is correct. The additional risk factor acts as a penalty on extreme positions that would otherwise be taken in light of the mistaken alphas. In this respect an incorrect risk model helps improve portfolio performance: one mistake offsets another. Of course it is the fortuitously applied risk penalty that is responsible for the improvement, not correctly aligning the risk model *per se*. Indeed, the appendix to this article shows that if we know the alphas to be "junk," as in Exhibit 7, we can always add a penalty to the variance-covariance matrix that completely eliminates the effects of the alpha errors on the portfolio weights.[18] The penalty could be labeled a "customization" of the risk model that improves performance, but it might better be called a forecast error correction technique that is delivered by tinkering with the risk model.

[18] Of course, if the manger knew the alphas to be junk, then he or she could (and should) ignore them in the first place.

Now consider what happens when Case 1 is the true state of the world, but the manager either misses alphas (an error of omission) or assumes falsely that factor risk is present (an error of commission). The resulting losses in this scenario are plotted in Exhibit 8.

Exhibit 8
Loss When True State is Case 1
but Manager Incorrectly believes either the Base Case or Case 2



From Exhibit 8 we can see that at least over the range of π plotted, the loss incurred when assuming the Base Case is true when in fact Case 1 applies is roughly equivalent to the loss incurred when the manager falsely assumes Case 1 holds but the Base Case is actually true (in Exhibit 7). However, this is only true for small alphas. It can be shown that as π becomes very large in the case depicted in Exhibit 7, the losses become extreme. For large π, assuming nonexistent alphas is more costly than missing true alphas. We saw in Exhibit 7 that falsely assuming a factor risk is beneficial (it reduces losses) when there are in fact no true alphas. In the case depicted in Exhibit 8 where alphas do in fact exist, falsely assuming factor risk (believing

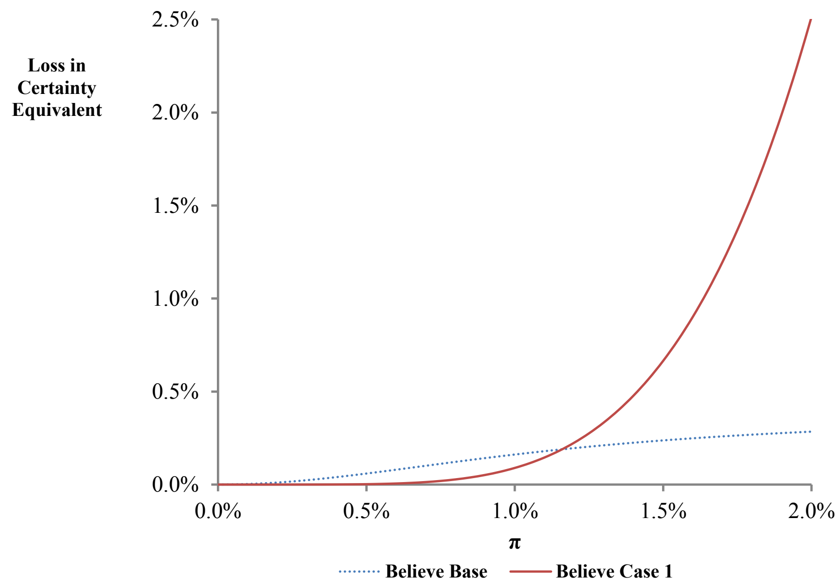
Case 2 applies) is now costly since it causes the manager to hold back on fully exploiting the alphas.

Next we consider the situation where alphas exist along with an underlying factor that partly drives them. In this case there are two errors that can be made. First, the manager who falsely believes that Case 1 applies correctly assumes the alphas exist, but mistakenly believes there is no underlying factor risk. In this case the manager will "over exploit" the alphas, earning extra return but taking on hidden risk of which he is unaware. Second, the manager who mistakenly believes that the Base Case applies avoids the consequences of extra factor risk but misses out on capturing any of the extra return available from the alphas that does in fact exist. Exhibit 9 shows that for low levels of π the loss is greater when the alphas are mistakenly ignored. This is because when π is small, the manager's response to associated alphas will be small, even when the underlying factor risk is ignored.[19] For small π the gains in expected return achieved by exploiting the alphas are first order, while the increases in the underlying and "hidden" factor risk are of second order. This is related to the observation that even a very risk averse person will benefit by taking on a small favorable bet if the outcome of that bet is independent of all other risks to which he is exposed. Since the underlying hidden factor risk is unrelated to any of the other risks in the manager's portfolio, this risk is not consequential for small bets and the manager who ignores it and captures alpha is better off than the manager who ignores the alphas altogether. However, as π increases, the manager who falsely believes that Case 1 applies takes on larger and larger positions and eventually the hidden factor risk comes to dominate and the over exploitation of alphas becomes more costly than ignoring the alphas entirely.

---

[19] The response to alphas is limited by the individual asset idiosyncratic risk.

Exhibit 9
Loss When True State is Case 2
but Manager Incorrectly believes either Base Case or Case 1

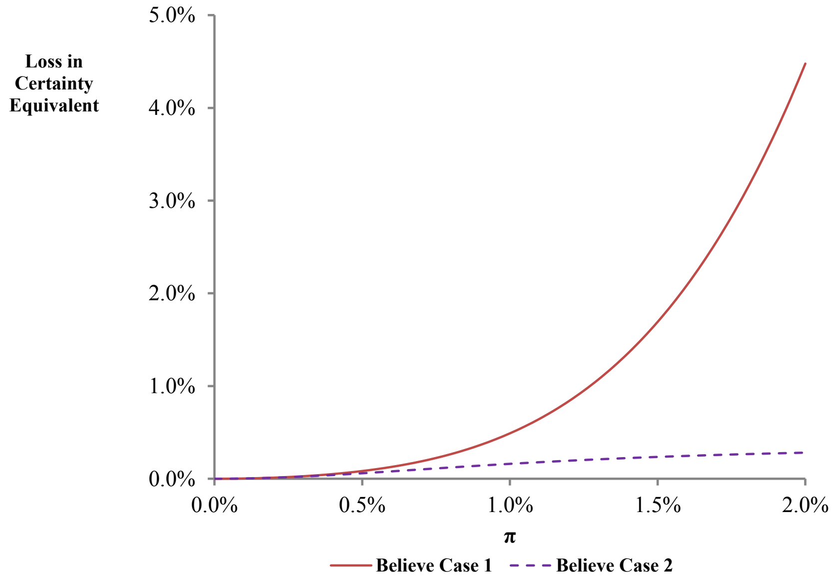


Finally, for the sake of completeness we consider in Exhibit 10 the case where there are no true alphas but the risk factors need to be augmented by the systematic "π" risk we have been positing might exist. In this instance, the losses are very large for a manager who believes that to the contrary he has alphas and there is no additional systematic risk associated with them. The manager loses nearly 4.5% in portfolio certainty equivalent as the misperceived alphas and missing risk are amplified in the mean-variance optimization.

Exhibit 10

Loss When True State is Case 3

but Manager Incorrectly believes either Case 1 or Case 2



In summary the consequences of errors of omission and errors of commission in portfolio construction are generally quite different. A manager who uses alphas that are related to a factor exposure, but omits this factor from the risk model, commits an error of omission. As seen in our example, the losses associated with this error of omission can be as much as four times the losses that would occur if this factor did not exist but the manager mistakenly included it. Including the phantom factor would be an error of commission. Of course the relative losses depend upon specific assumptions about alphas and factor structure, but we have found that the basic result is quite general over a wide range of calibrated variations of the basic exercise we described.

Summary and Discussion

We have examined problems that can arise when the risk model used in portfolio construction is not consistent with the sources of alphas that cause the manager to over or underweight assets. For example, in practice a misalignment problem can arise if there is a common factor that underlies alpha signals, but this factor is not fully reflected in the risk model used in portfolio construction. As a remedy, Ceria, Saxena and Stubbs [2012] suggest:[20] "…augmenting the user risk model with an additional factor…[which aligns with alphas and]…which is orthogonal and uncorrelated with all the existing risk factors." Lee and Stefek [2008] propose: "...building a new risk model that includes the alpha factors." Here we have discussed cases where adjustments made in this spirit can lead to improvements in portfolio construction, but we have also shown that there are other cases where these adjustments are not warranted and lead to suboptimal portfolios.

An obvious case in which one can be led astray by making these adjustments occurs when the alphas are truly "smart betas," i.e., premiums that are *not* associated with extra risk taking. These "free lunch" smart beta premiums are generally pursued by tilting away from the market portfolio based on cross sectional characteristics like value/growth, but the sources of the gains are often left mysterious. To illustrate with a specific example the challenges in getting alphas aligned with the risk model we modeled a case where "free lunch" smart beta premiums arise, one that is based on real-world borrowing constraints as first considered in Black [1972] and later in Frazzini and Pedersen [2014]. The beta premiums arise when investors who are borrowing-constrained demand riskier assets (growth stocks), which leads to pricing that induces

[20]See also Saxena and Stubbs [2015].

borrowing-unconstrained investors to overweight less risky assets (value stocks). In equilibrium the unconstrained investors earn a "smart beta" premium for overweighting the value stocks and "supplying" (*via* underweighting) the growth stocks to the borrowing-constrained segment of investors. In such a segmentation/smart beta world, the so-called smart beta premium does not stem from a specific risk exposure associated with value stocks and any attempt to augment a factor model to "align" it with observed premiums is misdirected and results in an opportunity loss.

When the economic foundations for alphas and "smart betas" are well understood, it will generally be clear whether the risk model needs to be adjusted and what the nature of any adjustments should be. In practice, however, there is often considerable uncertainty about the underlying reasons for the existence of various perceived alphas and smart betas. This is because the alphas and smart betas are often discovered by searching for anomalous statistical regularities in past returns, and these statistical searches don't necessarily reveal the reason for anomalous return differentials.[21] A "smart beta" tilt could be a tilt toward higher returns created by the type of frictions in the market that produce favorable pricing for unconstrained investors along the lines just described. On the other hand, the tilt could be exploiting higher returns that are in fact risk premiums associated with a risk that is not well understood, such as a catastrophe risk that is not salient to the manager but nevertheless is priced in the market, perhaps to a limited extent.

[21]In practice time-variation in return premiums and concomitant risk model parameters complicate the estimation and thus the alignment problem. Cochrane [2011] summarizes the evidence that premiums are time varying and Bali [2008] is able to fit time-varying risk premiums and conditional "betas" that are functions of cross-sectional industry and SMB and HML characteristics. In practice, managers' most insightful alphas often involve anticipating shifts in the financial environment ahead of, or better than, the market, while risk models are generally data-based and necessarily backward looking.

Our examples suggest that the portfolio construction losses associated with mistakenly believing that one is in an innocuous "smart beta" world when in fact the premiums reflect higher catastrophe risk exposure are plausibly twice the losses from making the opposite error. Nevertheless, the opposite error results in a loss that one would also like to avoid if possible.

We also considered cases where the source of a manager's alphas is an informational or analytical edge over the market involving stock specific information. As a point of departure in this case, it is straightforward to see that if the alphas are "true and independent alphas" accurately reflecting proprietary stock-specific analyst signals and the user risk model is well specified, then no factor model adjustment is needed and any adjustment would lead to suboptimal results. On the other hand, when the individual stock alpha signals are not cross sectionally independent but include some common factor influences that are missing from the risk model, augmenting the risk model to include the missing factor risk is clearly required. The errors of omission associated with not including these common factors are accentuated by the standard mean-variance optimizer's well-known tendency to take extreme solutions when alphas "look attractive" relative to the understated risks. As in the smart beta case, correcting for potential errors associated with the omission of factor risk needs to be balanced against the cost of potential errors of commission in adding nonexistent or phantom factors. We also observe that when the alphas are mostly or wholly noise, better portfolio construction outcomes can be achieved by a "custom augmentation" of the risk model that accounts for the weakness of the alpha signals, but these adjustments merely add a penalty in the optimization to prevent inappropriate responses to weak or nonexistent alphas. These penalties, by construction, need have nothing to do with missing risk factors. Indeed, in some applications the Bayesian and resampling procedures that have been proposed to deal with alpha noise typically serve the same operational purpose as a penalty, e.g. Brown [1976], Michaud [1989], and Scherer [2002]. The portfolio-improving "customization" should be interpreted as a correction to a given manager's alphas, not to the model of market risks that are faced in common by all investors in the same market.

To focus on the alignment issues in as stark a way as possible we did not include portfolio constraints in our analysis. But as just noted, a risk penalty in the form of a factor augmentation to offset alpha errors obviously operates on portfolio solutions in much the same way as do various constraints or trading penalties and costs.[22] Saxena, Martin and Stubbs [2012] and Saxena and Stubbs [2015] point to long-only or active weight constraints, turnover constraints and trading costs as examples, all of which are common in practice. Of course the constraints can also often be interpreted as a robustness stand-in or "band aid" for incompletely specified aspects of the portfolio problem.

Our bottom line is that attention might be more profitably paid to understanding the economic foundations for smart betas or alpha rather than to *ad hoc* imposition of constraints or "risk" penalties in portfolio construction. Reassuringly, we often encounter parallel reasoning in practice in fund manager screening. In particular, just as here, asset owners try to understand both the drivers of a manger's returns and the concomitant risk factors that they likely entail. Also, in that consideration of both drivers and risk factors, there are potential errors of omission to be traded off against potential errors of commission, in alignment with our focus here!

[22]Indeed "soft" constraints with quadratic penalties or market impact trading costs modeled as a quadratic function of trade size act essentially the same as a portfolio variance-risk penalty.

REFERENCES

Admati, A.R., and P. Pfleiderer. "Does It All Add Up? Benchmarks and the Compensation of Active Portfolio Managers." *The Journal of Business*, Vol. 70 No. 3 (1997), pp. 323-350

Alankar, A., P. Blaustein and M.S. Scholes. "The Cost of Constraints: Risk Management, Agency Theory and Asset Prices," Working Paper, February 2, 2014.

Ang, A., W. N. Goetzmann, and S.M. Schaefer. "Evaluation of Active Management of the Norwegian Government Pension Fund – Global," Report, December 14, 2009.

Asgharian, H., and B. Hansson. "Evaluating the Importance of Missing Factors using the Optimal Orthogonal Portfolio Approach." *Journal of Empirical Finance* Vol. 12, (2005), pp. 556-575.

Bali, T. G. "The Intertemporal Relation Between Expected Returns and Risk." *Journal of Financial Economics*, Vol. 87, (2008), pp. 101-31.

Bali, T.G., N. Cakici, and R.F. Whitelaw. "Hybrid Tail Risk and Expected Stock Returns: When Does the Tail Wag the Dog?" *Review of Financial Studies*, Advance Access publication, September 15, 2014.

Bali, T.G., S.J. Brown, S. Murray, and Y. Tang. "Betting against Beta or Demand for Lottery," *Presentation to Pacific Basin Finance Economics Accounting and Management 23rd Annual Conference*, Ho Chi Minh City, July 16, 2015.

Black, F. "Capital Market Equilibrium with Restricted Borrowing." *Journal of Business*, Vol. 45, No. 3, (1972), pp. 444-455.

Black, F., and R. Litterman. "Global Portfolio Optimization." *Financial Analysts Journal*, September-October, Vol. 48, No. 5, (1992), pp. 28-43.

Brown, S. "Optimal Portfolio Choice Under Uncertainty: A Bayesian Approach," Ph.D. Dissertation, University of Chicago, 1976.

Ceria, S., A. Saxena, and R.A. Stubbs. "Factor Alignment Problems and Quantitative Portfolio Management." *The Journal of Portfolio Management,* Vol. 38, No. 2, Winter (2012), pp. 29-43.

Chabi-Yo, F., S. Ruenzi, and F. Weigert. "Crash Sensitivity and the Cross-Section of Expected Stock Returns," Working Paper, Ohio State University, March, 2015. http://ssrn.com/abstract=2011746

Cochrane, J. H. "Presidential Address: Discount Rates." *The Journal of Finance*, Vol LXVI, No. 4, (2011), pp. 1047-1108.

Daniel, K., and S. Titman. "Testing Factor-Model Explanations of Market Anomalies," Working Paper, Columbia University, 2011.

Fama, E.F., and K.R. French. "A Five-Factor Asset Pricing Model," Working Paper, March, 2014.

Frazzini, A., and L.H. Pedersen. "Betting against beta." *Journal of Financial Economics*, Vol. 111 (2014), pp. 1-25.

Harvey, C.R., Y. Liu, and H. Zhu. "…and the Cross-Section of Expected Returns." *The Review of Financial Studies*, Vol. 29 (2016), pp. 5-68.

Huang, W., Q. Liu, S., G. Rhee, and F. Wu. "Extreme downside risk and expected stock returns." *Journal of Banking and Finance*, Vol. 36 (2012), pp. 1492-1502.

Jagannathan, R., and T. Ma. "Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps." *The Journal of Finance*, Vol. 63, No. 4 (2003), pp. 1651-1683.

Kelly, B., and H. Jiang. "Tail Risk and Asset Prices." *Review of Financial Studies*, Advance Access, July 28, 2014.

Lee, J-H., and D. Stefek. "Do Risk Factors Eat Alphas?" *Journal of Portfolio Management*, Vol. 34, No. 4 (2008), pp. 12-24.

Lesham, R., L.R. Goldberg, and A. Cummings. "Optimizing Value," Manuscript, Aperio Group. Forthcoming in *Journal of Portfolio Management*.

Lettau, M., and J. A. Wachter. "Why is Long-Horizon Equity Less Risky? A Duration-Based Explanation of the Value Premium." *The Journal of Finance,* Vol. 62, No. 1 (2007), pp. 55-92.

Lewellen, J., S. Nagel, and J. Shanken. "A Skeptical Appraisal of Asset Pricing Tests." *Journal of Financial Economics*, Vol. 96 (2010), pp. 175-194.

MacKinlay, A.C. "Multifactor Models Do Not Explain Deviations from the CAPM." *Journal of Financial Economics*, Vol. 38 (1995), pp. 3-28.

MacKinlay, A. C., and L. Pastor. "Asset Pricing Models: Implications for Expected Returns and Portfolio Selection." *The Review of Financial Studies*, Vol. 13, No. 4, Winter (2000), pp. 883-916.

Markowitz, H. "Mean–variance Approximations to Expected Utility." *European Journal of Operational Research,* Vol. 234 (2014), pp. 346–355.

Merton, R.C. "A Simple Model of Capital Market Equilibrium with Incomplete Information." *The Journal of Finance* Vol. 42, No. 3 (1987), pp. 483-510.

Michaud, R.O. "The Markowitz Optimization Enigma: Is 'Optimized' Optimal?" *Financial Analysts Journal*, January-February (1989), pp. 31-42.

Miller, M.H., and F. Modigliani. "Dividend Policy, Growth, and the Valuation of Shares." *The Journal of Business*, Vol. 34, No. 4 (1961), pp. 411-433.

Roll, R. "A Critique of the Asset Pricing Theory's Tests: Part I." *Journal of Financial Economics*, Vol. 4 (1977), pp. 129-176.

Roll, R. "Orthogonal Portfolios." *Journal of Financial and Quantitative Analysis*, Vol. XV, No. 5 (1980), pp. 1005-1023.

Saxena, A., C. Martin, and R.A. Stubbs. "Aligning Alpha and Risk Factors, a Panacea to Factor Alignment Problems," *Axioma Research Paper No. 38*, February 22, 2012.

Saxena, A., and R. A. Stubbs. "Augmented Risk Models to Mitigate Factor Alignment Problems," *Journal Of Investment Management*, Vol. 13, No. 3 (2015), pp. 57–79.

Scherer, B. “Portfolio Resampling: Review and Critique.” *Financial Analysts Journal*, Vol. 58, No. 6 (2002), pp. 98-109.

Sharpe, W.F. “Imputing Expected Security Returns from Portfolio Composition.” *Journal of Financial and Quantitative Analysis,* Vol. 3, June (1974), pp. 463-472.

Van Oordt, M., and C. Zhou. “Systematic Tail Risk,” DNB Working Paper No. 400, November 2013. http://ssrn.com/abstract=2367225

Weigert, F. “Crash Aversion and the Cross-Section of Expected Stock Returns Worldwide,” University of St. Gallen Working Paper on Finance No. 2013/25, June 2015. http://ssrn.com/abstract=2309538

Williams, J. B., *The Theory of Investment Value*, Amsterdam: North-Holland Publishing Company, 1938. Reprinted: Fraser Publishing, Burlington Vermont, 1997.

Wu, S-J., and W-M. Lee. “Intertemporal Risk-return Relationships in Bull and Bear Markets.” *International Review of Economics and Finance,* Vol. 38 (2015), pp. 308–325.

Appendix A
Applying a "Factor Penalty" to correct for False Alphas in Expected Returns

Define:

The $N \times N$ Variance-Covariance matrix is: V

The true $N \times 1$ Expectation of Excess Returns Vector is: μ

The simple mean-variance optimization with no constraints is:

$$\underset{w}{Max}\left(\mathrm{w}'\mu - \frac{1}{2\rho}\mathrm{w}'Vw \right)$$

The well-known solution is: $w^* = \rho V^{-1}\mu$. Assume that the manager erroneously believes that the expected return vector is $\mu + \hat{\alpha}$, so that $\hat{\alpha}$ is a vector of "false" alphas.

If $\hat{\alpha}'V^{-1}\mu > 0$, then the variance covariance matrix can be augmented by a risk factor γ so that the portfolio constructed using the expected returns with the false alphas and this augmented risk model is the optimal portfolio. For this to be true, the γ must satisfy:

$$(V + \gamma\gamma')^{-1}(\mu + \hat{\alpha}) = V^{-1}\mu.$$

Pre-multiplying this by $V + \gamma\gamma'$ gives:

$$\begin{aligned} \mu + \hat{\alpha} &= (V + \gamma\gamma')V^{-1}\mu \\ &= \mu + \gamma\gamma'V^{-1}\mu, \end{aligned}$$

which means that $\hat{\alpha} = \gamma\gamma'V^{-1}\mu$, which further implies that:

$$\gamma = \frac{\hat{\alpha}}{\sqrt{\hat{\alpha}'V^{-1}\mu}}.$$

Augmenting V by the "risk factor" γ creates a quadratic penalty that fully eliminates the responses to the false alphas.

Note that if $\hat{\alpha}'V^{-1}\mu \le 0$, it will generally not be possible to fully eliminate the response to the false alpha by augmenting the risk model, but it will almost always be possible to add a "penalty" factor to significantly improve portfolio construction.