MIT MANAGEMENT SLOAN SCHOOL

MIT Sloan School of Management

MIT Sloan School Working Paper 6113-20

Portfolio Choice with Path-Dependent Preferences

Mark Kritzman, Ding Li, Grace (Tiantian) Qiu and David Turkington

May 21, 2020

Portfolio Choice with Path-Dependent Preferences

Mark Kritzman
kritzman@mit.edu

Ding Li
liding@gic.com.sg

Grace (Tiantian) Qiu
graceqiu@gic.com.sg

David Turkington
dturkington@statestreet.com

THIS VERSION: May 21, 2020

Abstract

Many sophisticated investors rely on scenario analysis to select a portfolio. These investors define prospective economic scenarios, assign probabilities to them, translate the scenarios into expected asset class returns, and select the portfolio with the highest expected return or expected utility, given all these inputs. With this approach, the investor only considers single period outcomes. The authors propose a new approach to scenario analysis that enables investors to consider sequential outcomes. They define prospective scenarios, not as average values of economic variables, but as paths for these variables. And they measure the likelihood that these paths will prevail in the future based on their statistical similarity to the historical sequences of these variables. The authors also employ a novel forecasting technique called partial sample regression to map economic outcomes onto asset class returns. This process allows investors to evaluate portfolios based on the likelihood they will produce a certain pattern of returns over a specified investment horizon.

Portfolio Choice with Path-Dependent Preferences

Many investors select portfolios using a technique called scenario analysis. This approach requires them to define prospective economic scenarios, assign probabilities to them, translate the scenarios into expected asset class returns, and select the portfolio with the highest expected return or expected utility, given all these inputs.

Investors typically define a prospective scenario as a set of single period average values for key economic variables. We propose that investors instead define scenarios as sequences of values for the economic variables. Defining scenarios as paths rather than single period averages has several advantages.

First, it allows investors to consider interim outcomes which could lead them to make different choices than they would otherwise. Perhaps John Maynard Keynes said it best.

> In the long run, we are all dead. Economists set themselves too easy, too useless a task, if in tempestuous seasons they can only tell us that when the storm is long past, the ocean will be flat.
>
> John Maynard Keynes
>
> *A Tract on Monetary Reform* (1923)

When investors focus on single period outcomes, such as expected return or expected utility, they look beyond the metaphorical storms to the end of their investment horizon, which gives only a limited view of a portfolio's exposure to loss.

Second, we are better equipped to evaluate the relative likelihood of alternative scenarios if we have information about how they unfolded. It may be the case, for example, that the average outcomes for two sets of variables representing alternative scenarios occurred commonly throughout history, but when specified as multi-period paths, one scenario was without precedent, whereas the other was more usual.

Third, by defining scenarios as paths, we are better able to determine the relevance of historical observations that we rely upon to map economic scenarios onto asset class returns. As we describe later, we employ a novel forecasting procedure which filters historical observations according to their statistical relevance.

And fourth, information about the path of economic scenarios and their associated returns enables investors to make better informed tactical shifts ranging from simple rebalancing to active tilts away from a steady state investment posture.

We proceed as follows. We first describe how we define scenarios and estimate their likelihood of occurrence. Then we describe how we map economic scenarios onto asset class performance. We next illustrate our approach with a case study in which we generate a rich set of results. The case study captures not only the probable average performance of the alternative portfolios but the pattern of their returns and the dispersion of their performance across scenarios and through time. We conclude with a summary.

Scenarios and their Likelihood of Occurrence

We define scenarios as a set of outcomes for chosen economic variables. But rather than define a prospective economic scenario as an average value for economic growth and an average value for inflation, for example, we specify multiple values for each economic variable representing the early, middle, and late stages of a pattern. The pattern might cover multiple months, quarters, or years.

In order to estimate the relative likelihood of the prospective scenarios, we compute the statistical similarity of the prospective scenarios to the most recent economic experience using a statistic called the Mahalanobis distance. Effectively, we are asking: given the recent economic experience, how unusual would it be for one scenario to prevail going forward versus an alternative scenario?

Of course, we could choose a different period than the most recent economic experience to anchor our measure of the Mahalanobis distance. For example, if we believed that the recent experience was highly unusual and that conditions would revert to a more normal experience, we could anchor the Mahalanobis distance to values representing a more typical pattern.

The Mahalanobis distance was introduced in 1927 and modified in 1936 to analyze resemblances in human skulls among castes in India.[1] The measure is powerful and convenient because in a single number it characterizes the distance between two multivariate observations. In doing so, it accounts for the expected variation of each underlying variable from its average as well as the expected co-variation of each pair of underlying variables from their respective averages. Thus, a large Mahalanobis distance may result from greater dispersion in the values of one underlying variable from one observation to the next, or from pairs of deviations that are not particularly large, but which depart from the typical pattern of co-movement for the variables.

The usefulness of the Mahalanobis distance is evidenced by its application to a diverse set of challenges, including diagnosing liver disease (Su and Li, 2002), sleep apnea (Wang, Su, Chen and Chen, 2011) and breast cancer (Nasief, Rosado-Mendez, Zagzebshi and Hall, 2019), and detecting anomalies in self-driving vehicles (Khalastchi and Kaminka, 2010). Within the field of investing, it has been applied to measure financial turbulence (Chow, Jacquier, Kritzman and Lowry, 1999), estimate the likelihood of single-period economic scenarios (Czasonis, Kritzman, Pamir and Turkington, 2020), improve the forecast reliability of linear regression analysis (Czasonis, Kritzman and Turkington, 2020a) and forecast the correlation between stocks and bonds (Czasonis, Kritzman and Turkington, 2020b).

In this paper, we extend the work of Czasonis, Kritzman, Pamir, and Turkington (2020) by specifying scenarios as patterns rather than single period outcomes. In this context, the Mahalanobis distance d is computed as shown in Equation 1.[2]

$$d = (x - \gamma)'\Sigma^{-1}(x - \gamma) \qquad (1)$$

In Equation 1, d equals the Mahalanobis distance, x is a vector comprising the multi-stage values of a set of economic variables used to characterize a future scenario, γ reflects the recent multi-stage values of the economic variables, Σ is the historical covariance matrix of changes in values for those variables, and $'$ indicates a vector transpose. We express all vectors as column vectors.

As we stated earlier, we should have more confidence in the Mahalanobis distances between paths than the Mahalanobis distances between the average values of the paths because paths impose an additional condition and thereby contain more information.

The Mahalanobis distance is closely related to a scenario's probability of occurrence. Specifically, an observation with a high Mahalanobis distance will tend to occur less frequently than one with a low Mahalanobis distance. If we assume that the economic variables follow a multivariate normal distribution, we can measure the relative likelihood of scenarios precisely. The likelihood of an observation decays as the Mahalanobis distance increases. It decays according to an exponential function, which gives rise to the normal distribution. We measure the likelihood that we would observe a given scenario as shown in Equation 2.

$$likelihood \propto e^{-d/2} \tag{2}$$

In Equation 2, d equals the Mahalanobis distance, e is the base of the exponential function, and $\propto$ denotes a proportionality relationship.[3]

The likelihoods we compute are in comparable statistical units across scenarios, however they will not sum to one because we have only specified a subset of all possible outcomes. Therefore, we rescale the likelihoods to sum to one so that we may interpret them as probabilities. The next step is to map these economic scenarios onto expected returns for each of the asset classes we wish to consider.

Scenarios and Asset Class Returns

We apply a novel forecasting technique called partial sample regression (Czasonis, Kritzman and Turkington, 2020a) to convert our projected economic scenarios into estimates of expected returns for the asset classes we wish to consider.

Partial sample regression relies on a convenient but obscure mathematical equivalence. The prediction generated by a linear regression model may be written equivalently as a function of the weighted average of the past values of the dependent variable in which the weights are the relevance of the past observations for the independent variables. Relevance is equal to the sum of the statistical similarity of the past observations to the current values for the independent variables and the informativeness of the past observations. Both quantities are measured as Mahalanobis distances.

Equation 3 defines the multivariate similarity between x_i and x_t, which is the opposite (negative) of the Mahalanobis distance between them.

$$similarity(x_i, x_t) = -(x_i - x_t)'\Omega^{-1}(x_i - x_t) \tag{3}$$

Here x_t is a vector of the current values of the independent variables, x_i is a vector of the prior values of the independent variables, the symbol $'$ indicates matrix transpose, and Ω^{-1} is the inverse covariance matrix of X where X comprises all the vectors of the independent variables. This measure takes into account not only how independently similar the components of the x_is are to those of the x_ts, but also the similarity of their co-occurrence to the co-occurrence of the x_ts. All else equal, prior observations for the independent variables that are more like the current observations are more relevant.

However, there is a second component to relevance. Observations that are more distant from their historical averages are more unusual and therefore more likely to be driven by events. These event-driven observations are potentially more informative.[4] Equation 4 defines the informativeness of a prior observation x_i as its multivariate distance from its average value, $\bar{x}$.

$$informativeness(x_i) = (x_i - \bar{x})'\Omega^{-1}(x_i - \bar{x}) \qquad (4)$$

The relevance of an observation x_i is equal to the sum of its multivariate similarity and its informativeness.

$$relevance(x_i) = similarity(x_i, x_t) + informativeness(x_i) \qquad (5)$$

In summary, prior periods that are like the current period but are different from the historical average are more relevant than those that are not.

Because linear regression is equivalent to a relevance-weighted average of the past values of the dependent variable, we specify Equations 6 and 7 which apply the same weights but only to a subset of the n most relevant observations.[5]

$$\hat{y}_t = \bar{y} + \frac{1}{2n}\sum_{i=1}^{n}\big(similarity(x_i, x_t) + informativeness(x_i)\big)(y_i - \bar{y}) \tag{6}$$

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \tag{7}$$

As we noted earlier, we improve the reliability of our forecasts by specifying the alternative scenarios as paths rather than as single period averages, because this refinement gives us more information upon which to assess the relevance of historical observations.

We next present a case study to illustrate our approach for assigning probabilities to multi-stage scenarios and for converting these scenarios into estimates of return sequences for the relevant asset classes.

Case Study

To illustrate our methodology, we consider six prospective economic scenarios, which we define as three-year patterns for economic growth and inflation. These scenarios pertain to the United States, and the variables are expressed in U.S. dollars. Most of our data comes from the Jordà-Schularick-Taylor (JST) Macrohistory Database (Release 4 from May 2019) which is publicly available at http://www.macrohistory.net/data.

While we acknowledge that the choice of scenarios is not one-size-fits-all, we believe that the following principles should apply.

1. The scenarios should span a comprehensive range of economic outcomes. Consider mean-variance analysis as a comparison. Mean-variance analysis implicitly accounts for all potential scenarios across a continuous distribution defined by an expected return and standard deviation. Scenario analysis, by construction, only considers a finite number of scenarios; however, these scenarios should be spread relatively evenly across an imaginary continuous distribution. It would not be helpful to specify scenarios that reside only on one side of this imaginary distribution, which could also lead to biased decision-making.
2. The scenarios should be relatively distinct from one another. If they are more redundant than distinct, the probabilities derived from their respective Mahalanobis distances will each be understated and unstable. Moreover, the probabilities assigned to the other scenarios may be understated. This principle is related to the first principle.
3. The chosen economic variables representing the scenarios should span the key fundamental drivers of future market behavior. The scenarios are representable by the economic variables chosen to define them. If the key features or narratives of a scenario are not captured by the variables, the inclusion of the scenario will not provide useful information. There is a large universe and multiple combinations of economic variables that investors may choose from. However, the choice of scenario variables should be holistic, parsimonious, and ideally orthogonal. This principle is also related to the first two principles.
4. The scenarios should be conceptually and empirically plausible. The scenarios used to guide asset allocation should capture the plausible set of future outcomes. If we define

a scenario whose combination of values is so statistically contrary to historical precedent and general market intuition, its likelihood of occurrence will be close to zero. While such extreme scenarios are useful for stress testing purposes, they would be less relevant for asset allocation, which should be based on a plausible range of outcomes.

We define our case study scenarios with two macro-economic variables, real GDP growth and inflation. Theoretically and empirically these two variables are the key drivers of future market behavior. They are also parsimonious and orthogonal, in the sense that they capture two distinct macroeconomic dimensions.

We construct six distinct and plausible scenarios by defining the three-year paths of these two macro variables. The baseline estimates are taken from the current Bloomberg consensus forecast for US GDP and inflation. The remaining scenarios are constructed by shocking GDP and inflation in pre-defined directions based on economic narratives for the scenarios. Two comments are useful to keep in mind.

The six scenarios are constructed following the COVID-19 pandemic, but they are by no means unique to current conditions. They could just as well represent alternative recovery paths following any economic or financial shock.

These scenarios are meant to be illustrative. Investors may wish to construct their own scenarios based on their own views or the opinions of experts. Our focus is not to propose specific scenarios, but rather to propose a comprehensive framework for conducting path-dependent scenario analysis.

Exhibit 1 shows the six scenarios we have chosen to consider.

Exhibit 1: Economic Scenarios

Scenarios		Current	Baseline (V)	Shallow V	U	W	Depression	Stagflation
Growth	Year 1	2.2%	-3.5%	-1.9%	-3.5%	-3.5%	-5.1%	-5.1%
Growth	Year 2	2.0%	3.8%	5.4%	0.0%	3.8%	-5.9%	-2.7%
Growth	Year 3	1.8%	2.3%	3.9%	3.9%	-4.2%	-7.4%	-0.9%
Inflation	Year 1	2.1%	1.0%	1.0%	1.0%	1.0%	-0.3%	2.3%
Inflation	Year 2	1.9%	1.7%	1.7%	0.4%	1.7%	-5.9%	4.2%
Inflation	Year 3	2.3%	2.0%	2.0%	0.7%	2.0%	-5.6%	5.8%
Probability (as of 2019)			21.5%	24.0%	30.1%	5.9%	2.4%	16.1%

The column labeled Current shows the three-year paths ended in 2019 for growth and inflation. These values serve as the anchor for computing each scenario's Mahalanobis distance. The columns to the right show the three-year paths of the prospective scenarios. The bottom row of Exhibit 1 gives the relative probabilities of the scenarios.

These probabilities are computed as follows.

1. We compute the time series of the yearly percentage change in real GDP per capita and the yearly percentage change in the consumer price index. We use the JST data set from 1927 to 2015, and data from the Federal Reserve of St Louis Economic Data Library (FRED) from 2016 to 2019.
2. We formulate time series to represent a three-year path of growth and inflation. For each year from 1929 to 2019, we include growth from two years ago, growth from one year ago, and growth from the current year, along with the inflation from two years ago, inflation from one year ago, and inflation from the current year.

3. We compute the covariance matrix from the changes in the values of the economic variables from one three-year period to the next three-year period (in other words, the differences in the paths of the variables).
4. We apply these inputs to Equation 1 to compute the scenarios' Mahalanobis distances, using the three-year paths for these variables ending in 2019 as the anchor
5. We use Equation 2 to convert the Mahalanobis distances into probabilities and rescale them to sum to 1.

Next we convert these scenarios into asset class returns. We consider three asset classes: U.S. stocks, U.S. bonds and U.S. cash, and we proceed as follows.

1. We obtain the time series of yearly stock total return, bond total return, and cash total return. We use the JST data set from 1927 to 2015, and we use the S&P 500, Bloomberg U.S. government bonds, and JP Morgan 3-month cash indices from 2016 to 2019.
2. We subtract annual inflation from each return to arrive at historical real returns, and we subtract the real returns of bills from those of stocks and bonds to convert the stock and bond returns to excess returns.
3. We compute the year-over-year change in the cash returns, which reflects a path of interest rate changes which we project from today's current interest rate levels.
4. For each year from 1929 to 2019, we formulate these three-year paths for each of the three asset classes in a vector (as we did for the economic variables) with nine elements.
5. We apply partial sample regression (with a subsample of the most relevant 25% of observations) to nine dependent variables and obtain the yearly changes of real cash

returns for three years, stock return premiums above cash for three years, and bond return premiums above cash for three years.

6. We cumulate the changes in cash returns to get the total cash returns for each year, and we add these to the premiums of stocks and bonds to get the total real returns for stocks and bonds.

Exhibit 2 shows the expected real returns of stocks, bonds, and cash associated with each of the economic scenarios.

Exhibit 2: Expected Asset Class Real Returns

Real Returns		Baseline (V)	Shallow V	U	W	Depression	Stagflation
Stocks	Year 1	16.3%	19.7%	-3.1%	19.5%	-22.2%	1.7%
Stocks	Year 2	5.8%	7.0%	13.4%	-7.2%	-22.0%	-1.9%
Stocks	Year 3	-0.5%	1.1%	7.2%	-3.2%	8.7%	-6.0%
Bonds	Year 1	4.8%	4.9%	6.4%	-1.1%	3.9%	-0.8%
Bonds	Year 2	-1.4%	-0.7%	2.5%	-0.2%	13.8%	-10.6%
Bonds	Year 3	4.9%	2.8%	5.1%	10.0%	22.4%	-10.0%
Cash	Year 1	1.4%	1.4%	1.3%	1.9%	3.2%	0.8%
Cash	Year 2	0.5%	0.5%	1.5%	1.4%	8.8%	-0.5%
Cash	Year 3	0.1%	0.1%	1.1%	0.7%	8.1%	-2.1%

The asset class returns shown in Exhibit 2 seem remarkably consistent with the scenario descriptions, which should give us confidence in our choice of scenarios and in our process for mapping the scenarios onto asset class returns.

Next, we specify the portfolios we wish to consider, which are shown in Exhibit 3.

Exhibit 3: Alternative Portfolio

	Conservative	Moderate	Aggressive
Stocks	40%	60%	80%
Bonds	50%	35%	20%
Cash	10%	5%	0%

By applying these portfolio weights to the asset class returns, we derive the return paths of these portfolios for each of the economic scenarios, as shown in Exhibit 4.

Exhibit 4: Portfolio Return Paths

Portfolios		Baseline (V)	Shallow V	U	W	Depression	Stagflation
Conservative	Year 1	9.1%	10.5%	2.1%	7.4%	-6.6%	0.3%
Conservative	Year 2	1.7%	2.5%	6.8%	-2.9%	-1.0%	-6.1%
Conservative	Year 3	2.3%	1.8%	5.6%	3.8%	15.5%	-7.6%
Moderate	Year 1	11.5%	13.6%	0.5%	11.4%	-11.8%	0.8%
Moderate	Year 2	3.0%	4.0%	9.0%	-4.3%	-7.9%	-4.9%
Moderate	Year 3	1.4%	1.6%	6.2%	1.6%	13.5%	-7.2%
Aggressive	Year 1	14.0%	16.8%	-1.2%	15.4%	-17.0%	1.2%
Aggressive	Year 2	4.4%	5.5%	11.2%	-5.8%	-14.8%	-3.7%
Aggressive	Year 3	0.6%	1.4%	6.8%	-0.6%	11.5%	-6.8%

The question now arises as to how we should evaluate these alternative portfolios. With conventional scenario analysis, in which case we would have just a single average return for each scenario, we would compute the weighted average return across the scenarios and select the portfolio with the highest return. Or we could specify a utility function and select the portfolio with the highest expected utility. However, because we specified our scenarios as paths, we have a richer set of data by which to evaluate the alternative portfolios. Exhibit 5 presents a variety of metrics by which to evaluate the portfolios.

Exhibit 5: Portfolio Metrics

	Baseline (V)	Shallow V	U	W	Depression	Stagflation	Probability Weighted	Worst
Probabilities	21.5%	24.0%	30.1%	5.9%	2.4%	16.1%		
Cumulative return								
Conservative	13.4%	15.3%	15.1%	8.3%	6.8%	-13.0%	9.6%	-13.0%
Moderate	16.5%	20.1%	16.3%	8.3%	-7.8%	-11.1%	11.8%	-11.1%
Aggressive	19.7%	24.9%	17.4%	8.1%	-21.2%	-9.2%	13.9%	-21.2%
Maximum drawdown								
Conservative	0.0%	0.0%	0.0%	-2.9%	-7.6%	-13.3%	-2.5%	-13.3%
Moderate	0.0%	0.0%	0.0%	-4.3%	-18.8%	-11.8%	-2.6%	-18.8%
Aggressive	0.0%	0.0%	-1.2%	-6.3%	-29.3%	-10.2%	-3.1%	-29.3%
Within-horizon loss								
Conservative	0.0%	0.0%	0.0%	0.0%	-7.6%	-13.0%	-2.3%	-13.0%
Moderate	0.0%	0.0%	0.0%	0.0%	-18.8%	-11.1%	-2.2%	-18.8%
Aggressive	0.0%	0.0%	-1.2%	0.0%	-29.3%	-9.2%	-2.5%	-29.3%
Worst loss								
Conservative	0.0%	0.0%	0.0%	-2.9%	-6.6%	-7.6%	-1.6%	-7.6%
Moderate	0.0%	0.0%	0.0%	-4.3%	-11.8%	-7.2%	-1.7%	-11.8%
Aggressive	0.0%	0.0%	-1.2%	-5.8%	-17.0%	-6.8%	-2.2%	-17.0%
Number of negative years								
Conservative	0	0	0	1	2	2	0.43	2
Moderate	0	0	0	1	2	2	0.43	2
Aggressive	0	0	1	2	2	2	0.79	2

The first panel shows the annualized cumulative return for each of the portfolios. This information is the same as we would generate had we used single period average values to define the scenarios. The remaining panels provide information that would not be known had we not defined the scenarios as paths. The second panel shows the maximum drawdown assuming we only observe the portfolios annually. The third panel shows within-horizon loss, which measures how far the portfolios might decline below their initial values during the three-year horizon. It differs from maximum drawdown, because it does consider losses that occurred from a higher value than the portfolios' initial values. The final two panels show the worst annual lose and the number of annual losses for each portfolio during the three-year horizon. All this information is presented for each scenario as well as a weighted average across the scenarios and a worst-case outcome.

If we only cared about cumulative return, we would select the aggressive portfolio. We might believe that this metric reflects risk because the probability-weighted cumulative returns consider a wide range of both positive and negative outcomes. But cumulative return reveals nothing about the extremes in performance that occur within the horizons. Because we specified the scenarios as paths, we are better able to observe each portfolio's within-horizon exposure to loss, which might incline us more toward the moderate or conservative portfolio.

There is another advantage to defining scenarios as paths. We might value consistency. The returns displayed in Exhibit 4 enable us to measure consistency. One approach would simply be to compute the spread between the highest and lowest return across the scenarios and through time for each portfolio. Although somewhat informative, these spreads do not consider the relative likelihood of the scenarios. A large spread between two unlikely scenarios might not suggest the same level of inconsistency as would a tighter spread between two more likely scenarios. We can address this issue by computing the standard deviation of returns across the scenarios and across the years for each portfolio in Exhibit 4, considering the scenarios' relative probabilities. These spreads and standard deviations are shown in Exhibit 6.

Exhibit 6: Consistency

Portfolio	Spread	Standard Deviation
Conservative	23.1%	4.9%
Moderate	25.4%	5.9%
Aggressive	33.8%	7.2%

Not surprisingly, the conservative portfolio offers the greatest degree of consistency across scenarios and through time.

Summary

We introduced a variation to scenario analysis in which prospective economic scenarios are defined as paths for economic variables rather than as single horizon averages, and we discussed several advantages to this approach. We then showed how to modify the methodology first introduced by Czasonis, Kritzman, Pamir, and Turkington (2020) to assign probabilities to these multi-stage scenarios. Next, we applied a novel forecasting technique called partial sample regression to map these multi-stage economic scenarios onto return paths for asset classes. We illustrated this new approach with a case study. We produced a variety of path-dependent metrics by which to evaluate alternative portfolios, which revealed that preferences based on path-dependent outcomes could lead investors to choose a different portfolio than they would otherwise choose in the absence of this richer set of results.

Notes

This material is for informational purposes only. The views expressed in this material are the views of the authors, are provided "as-is" at the time of first publication, are not intended for distribution to any person or entity in any jurisdiction where such distribution or use would be contrary to applicable law, and are not an offer or solicitation to buy or sell securities or any product. The views expressed do not necessarily represent the views of State Street Global Markets®, State Street Corporation® and its affiliates, Windham Capital Management, and/or GIC Private Limited and its affiliates (collectively, Singapore's sovereign wealth fund and referred to as "GIC"). While State Street, Windham Capital Management and GIC have collaborated for purposes of conducting research and developing this paper, State Street, Windham Capital Management and GIC are not engaged in any joint venture, affiliated in any way, or collectively providing or offering any services or products.

References

Chow, G., E. Jacquier, M. Kritzman, and K. Lowry. 1999. "Optimal Portfolios in Good Times and Bad." *Financial Analysts Journal*, 55 (3): 65-71 (May/June).

Czasonis, M., M. Kritzman, and D. Turkington. 2020a. "Addition by Subtraction: A Better Way to Forecast Factor Returns (And Everything Else)." Forthcoming in *The Journal of Portfolio Management*, Fall.

Czasonis, M., M. Kritzman, and D. Turkington. 2020b. "The Stock-Bond Correlation." *MIT Sloan Working Paper 6108-20.*

Czasonis, M., M. Kritzman, B. Pamir and D. Turkington. 2020. "Enhanced Scenario Analysis." *The Journal of Portfolio Management*, 46 (4): 69-79 (March).

Mahalanobis, P. C. 1927. "Analysis of Race-Mixture in Bengal." *Journal of the Asiatic Society of Bengal*, vol. 23: 301-333.

Mahalanobis, P. C. 1936. "On the Generalised Distance in Statistics." *Proceedings of the National Institute of Sciences of India*, 2 (1): 49-55.

Lin, R., E. Khalastchi, and G. Kaminka. 2010. "Detecting Anomalies in Unmanned Vehicles Using the Mahalanobis Distance." The Maverick Group, Computer Science Department, Bar-Ilan University, Ramat-Gan, Israel 52900.

Nasief, H., I. Rosado-Mendez, J. Zagzebski, and T. Hall. 2019. "A Quantitative Ultrasound-Based Multi-Parameter Classifier for Breast Masses," *Ultrasound in Medicine & Biology*, 45 (7): 1603-1616 (July).

Su, C. and S. Li. "A MD Based Classifier for Diagnosing Diseases. 2002 "*Journal of Chinese Institute of Industrial Engineers*, 19 (5): 41-47. September.

Wang, P., C. Su, K. Chen, and N. Chen. 2011. "The Application of Rough Set and Mahalanobis Distance to Enhance the Quality of OSA Diagnosis." *Expert Systems with Applications*, 38 (6): 7828-7836 (June).

---

[1] See Mahalanobis (1927) and Mahalanobis (1936).

[2] The Mahalanobis distance is often multiplied by $\frac{1}{N}$ so that the average distance score across the data set equals 1. This is just a scaling factor that we exclude for purposes of our analysis. It is sometimes shown as the square root of this quantity, which is another form of scaling.

[3] The probability density function for the multivariate normal distribution has a similar form, but includes a constant term that ensures the cumulative probability of all possible outcomes equals one. The scaling is irrelevant to our analysis because we are interested in the relative probabilities of a discrete set of scenarios which we rescale to sum to one.

[4] For further discussion of noise-driven versus event-driven observations and their relationship to estimating risk, see Chow, G., E. Jacquier, M. Kritzman, and K. Lowry (1999).

[5] See Czasonis, Kritzman, and Turkington (2020) for a thorough discussion of this technique.