Digital Portfolios

Sanjiv R. Das

Sanjiv R. Das
is a professor at Santa Clara University, Santa Clara, CA.
srdas@scu.edu

Digital assets are investments with binary returns: the pay off is either very large or very small. Digital portfolios' return distributions are highly skewed and fat-tailed. A venture fund is a good example of such a portfolio. A Bernoulli distribution offers a simple representation of the payoff to a digital investment: a large payoff for a successful outcome and a very small (almost zero) payoff for a failed one. Digital investments typically offer a small probability of success around, 5% to 25% for new ventures (see Das, Jagannathan, and Sarin [2003]). In a recent paper, Fernandez, Stein, and Lo [2012] advocate the creation of a biomedical megafund of digital assets. The standard techniques used for mean–variance optimization aren't useful tools for optimizing digital assets. This article's analyses let us characterize digital assets' return distributions.

The intuitions obtained from the mean–variance setting may not carry over to portfolios of Bernoulli assets. As Bernoulli portfolios involve higher moments, the best way to diversify is by no means obvious. For instance, is diversification by increasing the number of assets in the digital portfolio always a good thing? Is it better to include assets with as little correlation as possible, or is there a sweet spot for optimal asset-correlation levels? Should all the investments be of even size, or is it preferable to take a few large bets and several small ones? Is a mixed portfolio of safe and risky assets better than one that offers a more uniform probability of success? These are all questions of interest to investors in digital type portfolios, including collateralized debt-obligation investors, venture capitalists, and venture fund investors.

We use a method based on standard recursion for modeling a Bernoulli portfolios' exact return distribution. Andersen, Sidenius, and Basu [2003] first developed this method for generating credit portfolios' loss distributions. We examine these portfolio's properties in a stochastic dominance framework to provide guidelines to digital investors. These guidelines are consistent with prescriptions from expected utility optimization:

1. Holding all else the same, it's better to have more digital investments. For example, a venture portfolio should seek to maximize market share.
2. As with mean–variance portfolios, lower asset correlation is better, unless the digital investor's payoff depends on the upper tail of returns.
3. A strategy of a few large bets and many small ones is inferior to one with bets of roughly the same size.
4. A mixed portfolio of assets with low and high probabilities of success is better than one in which all assets have the same probability of success.

MODELING DIGITAL PORTFOLIOS

Assume that an investor has a choice of n investments in digital assets—startup firms, for instance. The investments are indexed $i = 1, 2, \ldots, n$. Each investment has a probability of success, denoted q_i. If it succeeds, the payoff is S_i. With probability $(1 - q_i)$, the investment will not succeed, the start-up will fail, and the money will be entirely lost. Therefore, the payoff (cash flow) is

$$\text{Payoff} = C_i = \begin{cases} S_i & \text{with prob } q_i \\ 0 & \text{with prob } (1-q_i) \end{cases} \tag{1}$$

Calling this investment a Bernoulli trial reflects reality in digital portfolios. It mimics the venture capital business, for example. Consider two generalizations. First, we might extend the model and allow S_i to be random, i.e., drawn from a range of values. This will complicate the mathematics, but not enrich the model's results by much. Second, the failure payoff might be non-zero: an amount a_i. That gives us a pair of Bernoulli payoffs $\{S_i, a_i\}$. We can decompose these investment payoffs into a project with constant payoff a_i plus another project with payoffs $\{S_i - a_i, 0\}$, the latter being exactly the original setting with a failure payoff of zero. The version of the model we solve in this article, with zero failure payoffs, is without loss of generality.

Unlike stock portfolios in which the asset choice set is assumed to be multivariate normal, digital asset investments have a joint Bernoulli distribution. Portfolio returns on these investments are unlikely to be Gaussian, and hence higher-order moments are likely to matter more. To generate a return distribution for a portfolio of digital assets, we must account for correlations across digital investments. We adopt the following simple correlation model. Define y_i as the performance proxy for the i-th asset. This proxy variable is simulated for comparison with a threshold performance level to determine whether the asset yielded a success or failure. It is defined by the following function, widely used in the correlated default-modeling literature. See Andersen, Sidenius, and Basu [2003] for an example.

$$y_i = \rho_i X + \sqrt{1-\rho_i^2}\, Z_i, \qquad i = 1 \ldots n \tag{2}$$

where $\rho_i \in [0, 1]$ is a coefficient that correlates threshold y_i with a normalized common factor $X \sim N(0, 1)$. The common factor drives the correlations between the portfolio's digital assets. We assume that $Z_i \sim N(0, 1)$ and Corr $(X, Z_i) = 0$, $\forall i$. The correlation between assets i and j is given by $\rho_i \times \rho_j$. The mean and variance of y_i are: $E(y_i) = 0$, $Var\,(y_i) = 1$, $\forall i$. Conditional on X, the values of y_i are all independent, as $\text{Corr}(Z_i, Z_j) = 0$.

We can formalize the probability model governing the digital investment's success or failure. We define a variable x_i, with distribution function $F(\cdot)$, such that $F(x_i) = q_i$, the digital investment's probability of success. Conditional on a fixed value of X, the probability of success of the i th investment is defined as

$$p_i^X = Pr[y_i < x_i \mid X] \tag{3}$$

Assuming F to be the normal distribution function, we have

$$\begin{aligned} p_i^X &= Pr\left[\rho_i X + \sqrt{1-\rho_i^2}\, Z_i < x_i \mid X\right] \\ &= Pr\left[Z_i < \frac{x_i - \rho_i X}{\sqrt{1-\rho_i^2}} \mid X\right] \\ &= \Phi\left[\frac{F^{-1}(q_i) - \rho_i X}{\sqrt{1-\rho_i}}\right] \end{aligned} \tag{4}$$

where $\Phi(.)$ is the cumulative normal density function. Given the level of the common factor X, asset correlation ρ, and the unconditional success probabilities q^i, we obtain the conditional success probability for each asset p_i^X. As X varies, so does p_i^X. For the numerical examples in this article we choose the function $F(x_i)$ to the cumulative normal probability function.

We use a fast technique for building up distributions for sums of Bernoulli random variables. In finance, Andersen, Sidenius, and Basu [2003] introduced this *recursion* technique in the credit portfolio-modeling literature.

We call a digital investment successful if it achieves its high payoff S_i. The portfolio's cash flow is a random variable $C = \sum_{i=1}^{n} C_i$. The portfolio's maximum possible cash flow is the sum of all digital asset cash flows, because each and every outcome is a success:

$$C_{\max} = \sum_{i=1}^{n} S_i \tag{5}$$

To keep matters simple, we assume that each S_i is an integer and we round off amounts to the nearest significant digit. If the smallest unit we care about is a million dollars, then each S_i will be in integer units of millions.

Recall that, conditional on a value of X, digital asset i's probability of success is given as p_i^X. The recursion technique will let us generate the portfolio cash flow probability distribution for each level of X. We then compose these conditional (on X) distributions using the marginal distribution for X, denoted $g(X)$, into the unconditional distribution for the entire portfolio. We define the probability of total cash flow from the portfolio, conditional on X, as $f(C|X)$. Then, the portfolio's unconditional cash flow distribution becomes

$$f(C) = \int_x f(C \mid X) \cdot g(X)\, dX \qquad (6)$$

Compute the distribution $f(C|X)$ numerically as follows:

We index the assets with $i = 1\ldots n$. Cash flow from all assets taken together will range from zero to C_{max}. Break this range into integer buckets, resulting in a total number of buckets N_B, each one containing an increasing level of total cash flow. We index these buckets by $j = 1\ldots N_B$, with the cash flow in each bucket equal to B_j. B_j represents the total cash flow from all assets (some pay off and some do not), and the buckets comprise the discrete support for the entire distribution of total portfolio cash flow. For example, suppose we had 10 assets, each with a payoff of $C_i = 3$. Then $C_{max} = 30$. A plausible set of buckets comprising the cash flow distribution's support would be: $\{0, 3, 6, 9, 12, 15, 18, 21, 24, 27, C_{max}\}$.

Define $P(k, B_j)$ as the probability of bucket j's cash flow level B_j if we account for the first k assets. For example, if we had just three assets, with payoffs of value 1, 3, and 2 respectively, then we would have seven buckets, i.e., $B_j = \{0, 1, 2, 3, 4, 5, 6\}$. After accounting for the first asset, the only possible buckets with positive probability would be $B_j = 0, 1$, and after the first two assets, the buckets with positive probability would be $B_j = 0, 1, 3, 4$. We begin with the first asset, then the second and so on in order, and compute the probability of seeing the returns in each bucket. Each probability is given by the following recursion:

$$P(k+1, B_j) = P(k, B_j)\,[1 - p_{k+1}^X] + P(k, B_j - S_{k+1})\, p_{k+1}^X, \quad k = 1, \ldots, n-1 \qquad (7)$$

The probability of a total cash flow of B_j after considering the first $(k + 1)$ firms is equal to the sum of two probability terms. First, the probability of the same cash flow B_j from the first k firms, given that firm $(k + 1)$ did not succeed. Second, the probability of a cash flow of $B_j - S_{k+1}$ from the first k firms and the $(k + 1)$-st firm does succeed.

We start this recursion from the first asset, after which the N_B buckets are all of probability zero, except for the bucket with zero cash flow (the first bucket) and the one with S_1 cash flow:

$$P(1, 0) = 1 - p_1^X \qquad (8)$$

$$P(1, S_1) = p_1^X \qquad (9)$$

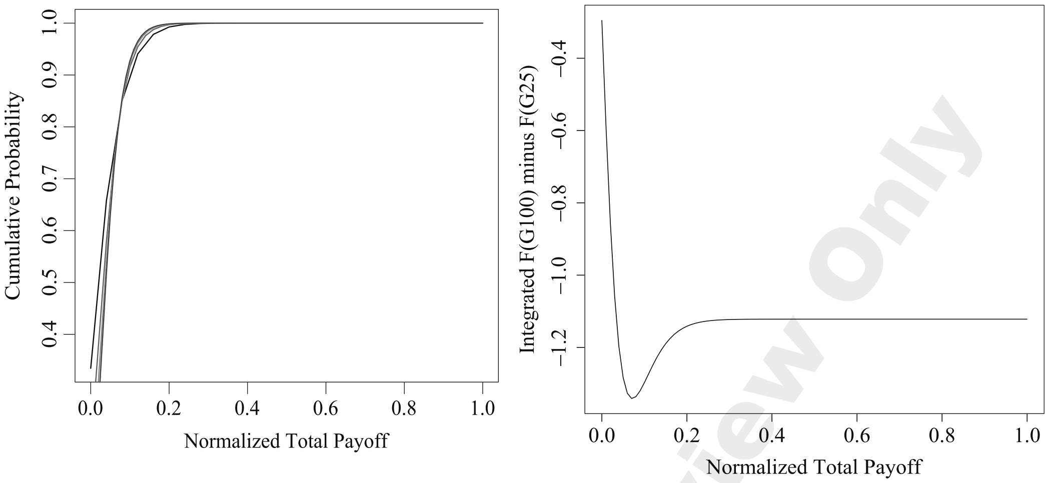
All the other buckets will have probability zero $P(1, B_j \neq \{0, S_1\} = 0$. With these starting values, we can run the system from the first asset to the n-th one by repeatedly applying Equation (7). The entire distribution $P(n, B_j)$, is conditional on a given value of X. We compose all the distributions that are conditional on X into a single cash-flow distribution using Equation (6), numerically integrating over all values of X.

PORTFOLIO CHARACTERISTICS

Armed with this established machinery, there are several questions a digital-portfolio investor (a venture capitalist, for instance) may pose. First, is there an optimal number of assets, i.e., ceteris paribus, are more assets better than fewer assets, assuming no span of control issues? Second, are Bernoulli portfolios different from mean-variance portfolios, in that is it always better to have less asset correlation than more asset correlation? Third, is it better to have an even investment weighting across assets, or might it be better to take a few large bets among many smaller ones? Fourth, is a high dispersion of probability of success better than a low dispersion? These questions are very different from the ones facing investors in traditional mean-variance portfolios. We shall examine each of these questions in turn.

EXHIBIT 1

Distribution Functions for Returns from Bernoulli Investments as the Number of Investments (*n*) Increases



Notes: Using the recursion technique we compute the probability distribution of the portfolio payoff for four values of n = {25,50,75,100}. The distribution function is plotted in the left panel. There are four plots, one for each n, and if we look at the bottom left of the plot, the leftmost line is for n = 100. The next line to the right is for n = 75, and so on. The right panel plots the value of $\int_0^u [G_{100}(x) - G_{25}(x)]dx$ *for all u ∈ (0, 1), and confirms that it is always negative. The correlation parameter is* $\rho = 0.25$

EXHIBIT 2

Expected Utility for Bernoulli Portfolios as the Number of Investments (*n*) Increases

η	E(C)	Pr[C > 0.03]	Pr[C > 0.07]	Pr[C > 0.10]	Pr[C > 0.15]	E[U(C)]
25	0.05	0.665	0.342	0.150	0.059	–29.259
50	0.05	0.633	0.259	0.084	0.024	–26.755
75	0.05	0.620	0.223	0.096	0.015	–25.876
100	0.05	0.612	0.202	0.073	0.011	–25.433

Notes: The exhibit reports the portfolio statistics for n = {25,50,75,100}. Expected utility is given in the last column. The correlation parameter is $\rho = 0.25$*. The utility function is* $U(C) = (0.1+C)^{1-\gamma}/(1-\gamma)$*,* $\gamma = 3$

How Many Assets?

With mean-variance portfolios, in which the portfolio's mean return is fixed it is better to have more securities in the portfolio, because diversification reduces portfolio variance. In mean-variance portfolios, higher-order moments do not matter. But with portfolios of Bernoulli assets, increasing the number of assets might exacerbate higher-order moments, even though it will reduce variance. Beyond a point, it may not be worthwhile to increase the number of assets (n).

To assess this issue we conduct an experiment. We invest in n assets each with payoff of $1/n$. If all assets succeed, the total (normalized) payoff is 1. This normalization is only to make the results comparable across different n, and is without loss of generality. We assume that the correlation parameter is $\rho_i = 0.25$, for all i. To make results easy to interpret, we assumed that assets are identical, with a success probability of $q_i = 0.05$ for all i. Using the recursion technique, we computed the probability distribution of the portfolio payoff for four values of $n = \{25, 50, 75, 100\}$. The distribution function is plotted in Exhibit 1, left panel. There are four plots, one for each n. At the bottom left of the plot, the leftmost line is for $n = 100$. The next line to the right is for $n = 75$, and so on.

One way to determine whether greater n is better for a digital portfolio is to investigate whether a portfolio of n assets stochastically dominates one with fewer than n assets. On examining the shapes of the distribution functions for different n, we see that it is likely that as n increases, we obtain portfolios that exhibit second-order stochastic dominance (SSD) over portfolios with smaller n. The return distribution when $n = 100$ (denoted G_{100}) dominates that for $n = 25$ (denoted G_{25}) in the SSD sense, if $\int_x x\, dG_{100}(x) = \int_x x\, dG_{25}(x)$, and $\int_0^u [G_{100}(x) - G_{25}(x)]\, dx \leq 0$ for all $u \in (0,1)$. That is, G_{25} has a mean-preserving spread over G_{100}, or G_{100} has the same mean as G_{25} but lower variance, implying superior mean-variance efficiency. To show this we plotted the integral $\int_0^u [G_{100}(x) - G_{25}(x)]\, dx$ and checked the SSD condition. We found that this condition is satisfied (see Exhibit 1). As others have verified, SSD implies mean-variance efficiency as well.

We also examine whether higher n portfolios are better for a power utility investor with utility function, $U(C) = \frac{(0.1+C)^{1-\gamma}}{1-\gamma}$, where C is the Bernoulli portfolio's normalized total payoff. Expected utility is given by $\Sigma_C U(C)\, f(C)$. We set the risk aversion coefficient to $\gamma = 3$ which is in the standard range in the asset-pricing literature. Exhibit 2 reports the results. The expected utility increases monotonically with n. For a power utility investor, having more assets is better than having fewer, keeping the portfolio's mean return constant. Economically, and specifically for venture capitalists, this highlights the goal of capturing a larger share of the number of available ventures. The results from the SSD analysis are consistent with those of expected power utility.

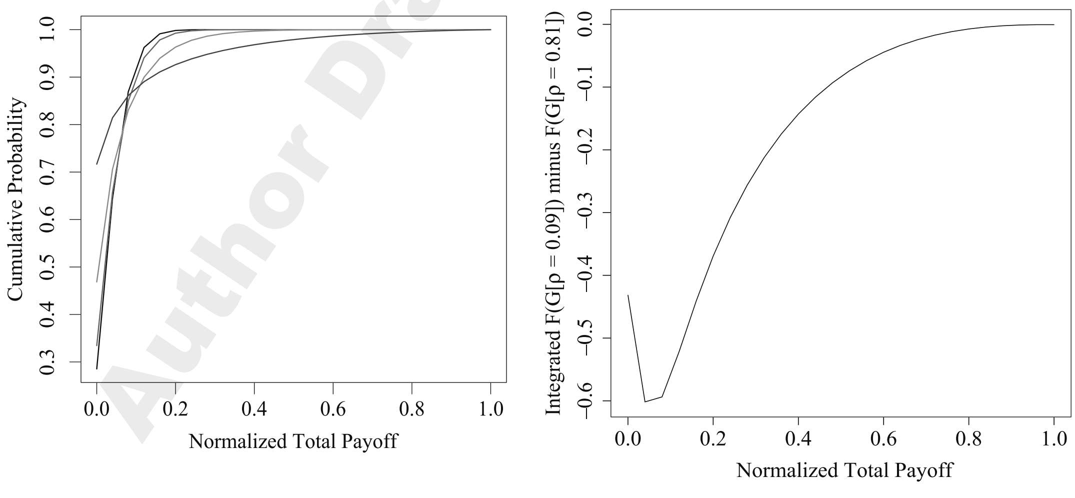
We have abstracted away from issues around the span of investor management. Investors actively play a role in their invested digital portfolios assets, and increasing n beyond a point may become costly, as modeled in Kanniainen and Keuschnigg [2003].

The Impact of Correlation

As with mean-variance portfolios, we expect that increases in payoff correlation for Bernoulli assets will adversely affect portfolios. In order to verify this intuition we analyze portfolios, keeping all other variables the same, but changing correlation. In the previous subsection, we set the correlation parameter to be $\rho = 0.25$. Here, we examine four levels of the correlation parameter: $\rho = \{0.09, 0.25, 0.49, 0.81\}$. For each correlation level, we computed the normalized total payoff distribution. The number of assets is fixed at $n = 25$ and

Exhibit 3

Distribution Functions for Returns from Bernoulli Investments as the Correlation Parameter (ρ^2) Increases



Notes: Using the recursion technique we computed the distribution of the portfolio payoff for four values of $\rho = \{0.09, 0.25, 0.49, 0.81\}$ ordered from top to bottom on the left plot, respectively. The distribution function is plotted in the left panel. The right panel plots the value of $\int_0^u [G_{\rho=0.09}(x) - G_{\rho=0.81}(x)]dx$ for all $u \in (0,1)$, and confirms that it is always negative

Exhibit 4

Expected Utility for Bernoulli Portfolios as the Correlation (ρ) Increases

ρ	E(C)	Pr[C > 0.03]	Pr[C > 0.07]	Pr[C > 0.10]	Pr[C > 0.15]	E[U(C)]
0.3^2	0.05	0.715	0.356	0.131	0.038	–28.112
0.5^2	0.05	0.665	0.342	0.150	0.059	–29.259
0.7^2	0.05	0.531	0.294	0.170	0.100	–32.668
0.9^2	0.05	0.283	0.186	0.139	0.110	–39.758

Notes: The exhibit reports the portfolio statistics for ρ = {0.09, 0.25, 0.49, 0.81}. The last column shows expected utility. The utility function is $U(C) = (0.1 + C)^{1-\gamma}/(1-\gamma)$, $\gamma = 3$

the probability of success of each digital asset is 0.05 as before.

Exhibit 3 shows the probability distribution function of payoffs for all four correlation levels. The SSD condition is met: lower-correlation portfolios stochastically dominate (in the SSD sense) higher-correlation portfolios. We also examine changing correlation in the context of a power utility investor with the same utility function as in the previous subsection. Exhibit 4 shows the results, which confirm that as with mean-variance portfolios, Bernoulli portfolios also improve if their assets have low correlation. Digital investors should optimally attempt to diversify their portfolios. Insurance companies are a good example: they diversify risk across geographical and other demographic divisions.

Uneven Bets?

Digital-asset investors are often faced with the question of whether to bet even amounts across digital investments, or to invest with different weights. We explore this question by considering two types of Bernoulli portfolios. Both have $n = 25$ assets, each with a success probability of $q_i = 0.05$. The first has equal payoffs: 1/25 each. The second portfolio has payoffs that monotonically increase, i.e., the payoffs are equal to $j/325$, $j = 1, 2, \ldots, 25$. The sum of the payoffs in both

Exhibit 5

Expected Utility for Bernoulli Portfolios when the Portfolio Comprises Balanced Investing in Assets vs. Imbalanced Weights

		Probability that C > *x*							
Wts	*E*(*C*)	*x* = 0.01	*x* = 0.02	*x* = 0.03	*x* = 0.07	*x* = 0.10	*x* = 0.15	*x* = 0.25	*E*[*U*(*C*)]
Balanced	0.05	0.490	0.490	0.490	0.278	0.169	0.107	0.031	–33.782
Imbalanced	0.05	0.464	0.437	0.408	0.257	0.176	0.103	0.037	–34.494

Notes: Both the balanced and imbalanced portfolio have $n = 25$ *assets, each with a success probability of* $q_i = 0.05$. *The first has equal payoffs, 1/25 each. The second portfolio has payoffs that monotonically increase, i.e., the payoffs are equal to* $j/325$, $j = 1, 2, \ldots, 25$. *The sum of the payoffs in both cases is 1. The correlation parameter is* $\rho = 0.55$. *The utility function is* $U(C) = (0.1 + \mathbf{C})^{1-\gamma}/(1-\gamma)$, $\gamma = 3$.

Exhibit 6

Expected Utility for Bernoulli Portfolios when the Portfolio Comprises Balanced Investing in Assets with Identical Success Probabilities vs. Investing in Assets with Mixed Success Probabilities

		Probability that C > *x*							
Probs	*E(C)*	*x* = 0.01	*x* = 0.02	*x* = 0.03	*x* = 0.07	*x* = 0.10	*x* = 0.15	*x* = 0.25	*E*[*U(C)*]
Uniform	0.10	0.701	0.701	0.701	0.502	0.366	0.270	0.111	–24.625
Mixed	0.10	0.721	0.721	0.721	0.519	0.376	0.273	0.106	–23.945

Notes: Both the uniform and mixed portfolios have $n = 26$ *assets. In the first portfolio, all the assets have a success probability equal to* $q_i = 0.10$. *In the second portfolio, half the firms have a success probability of 0.05 and the other half have a probability of 0.15. All investments have a payoff of 1/26. The correlation parameter is* $\rho = 0.55$. *The utility function is* $U(C) = (0.1 + C)^{1-\gamma}/(1-\gamma)$, $\gamma = 3$.

cases is 1. Exhibit 5 shows the investor utility where the utility function is the same as in the previous sections. The utility for the balanced portfolio is higher than that for the imbalanced one. The balanced portfolio evidences SSD over the imbalanced portfolio, but the return distribution has fatter tails when portfolio investments are imbalanced. Investors seeking to distinguish themselves by taking greater risk in their early careers may be better off with imbalanced portfolios.

Mixing Safe and Risky Assets

Is it better to have assets with a wide variation in probability of success or with similar probabilities of success? To examine this, we consider two portfolios of $n = 26$ assets. In the first portfolio, all the assets have a success probability equal to $q_i = 0.10$. In the second portfolio, half the firms have a success probability of 0.05 and the other half have a probability of 0.15. All investments have a payoff of 1/26. The probability distribution of payoffs and the expected utility for the same power utility investor (with $\gamma = 3$) are given in Exhibit 6. Mixing the portfolio between investments with both high and low probabilities of success results in higher expected utility than does choosing investments with similar probabilities of success. Imbalanced success probability portfolios also evidence SSD over portfolios with similar success rate investments. This result does not have a natural analog in the mean-variance world with non-digital assets. For empirical evidence on the efficacy of various diversification approaches, see Lossen [2006].

CONCLUSIONS

Digital asset portfolios are different from mean-variance ones because their asset returns are Bernoulli, with small success probabilities. We used a recursion technique borrowed from the credit-portfolio literature to construct the payoff distributions for Bernoulli portfolios. We find that many intuitions for these portfolios are similar to those for mean-variance portfolios: diversification by adding assets is useful, low correlations between investments is good. However, we also find that investors prefer uniform bet sizes over a variety of small and large bets. Instead of constructing portfolios with assets having uniform success probabilities, it is better to have a mix of assets some with low success probabilities and others with high success probabilities Venture funds are just one possible application. These insights augment the standard understanding obtained from mean-variance portfolio optimization.

Using this approach is simple. Investors need the expected payoffs of the assets C_i, success probabilities q_i, and the average correlation between assets, given by a parameter ρ. Broad statistics on these inputs are available, for venture investments from papers such as Das, Jagannathan, and Sarin [2003]. Fernandez, Stein and Lo's [2012] biotechnology fund, on the other hand, requires scientists to help determine inputs.

With that data, it is easy to optimize a digital asset fund's portfolio. Investors can easily extend the technical approach to features that include the cost of investor effort as the number of projects grows (Kanniainen and Keuschnigg [2003]), syndication, and so on. The marketplace shows a growing number of portfolios with digital assets, and the results of this article provide useful insights and intuition for asset managers.

The modeling section shows just one way in which to model joint success probabilities using a common factor. There are other ways, too, such as modeling joint probabilities directly, making sure that they are consistent with each other, which may be mathematically tricky. The results may differ from the ones developed here for some different system of joint success probabilities. The system we adopt here with a single common factor X may also be extended to more than one common factor, an approach often taken in the default literature. We leave these interesting extensions for future work.

ENDNOTE

I am grateful to George Chacko, Bob Hendershott and Hoje Jo for helpful comments on the paper.

REFERENCES

Andersen, L., J. Sidenius, and S. Basu. "All Your Hedges in One Basket." *Risk* (November 2003), pp. 67-72.

Das, S., M. Jagannathan, and A. Sarin. "Private Equity Returns: An Empirical Examination of the Exit of Asset-Backed Companies." *Journal of Investment Management*, Vol. 1, No. 1 (2003), pp. 152-177.

Fernandez, J.M., R. Stein, and A. Lo. "Commercializing Biomedical Research Through Securitization Techniques." *Nature Biotechnology* (2012), online.

Kanniainen, V., and C. Keuschnigg. "The Optimal Portfolio of Start-Up Firms In Asset Capital Finance." *Journal of Corporate Finance*, Vol. 9, No. 5 (2003), pp. 521-534.

Lossen, U. "The Performance of Private Equity Funds: Does Diversification Matter?" Discussion Papers 192, SFB/TR 15, University of Munich, 2006.