

# MODELING FINANCIAL PORTFOLIOS USING BELIEF FUNCTIONS

**Catherine Shenoy • Prakash P. Shenoy**

School of Business, University of Kansas, 1300 Sunnyside Av, Lawrence, KS  
66045-7585

cshenoy@ku.edu • pshenoy@ku.edu

## 1. Introduction

The main goal of this paper is to demonstrate how the theory of belief functions [Dempster 1967, Shafer 1976, Kohlas and Monney 1995, Smets 1998] can be used to model financial portfolios. In particular, we are interested in modeling how a portfolio return distribution changes as we learn new information about the different factors that impact the portfolio.

In [Shenoy and Shenoy 2000], we describe how Bayes nets can be used to model financial portfolios. Here we build on the Bayes net framework, but we use belief functions instead to model the dependence of factors of a portfolio model. The theory of belief functions offers an alternative technique for modeling the dependence between the various factors that determine the return distribution of a portfolio. In particular, we use the semantics of ‘distinct evidence’ to model prior distribution and new evidence about factors that will change portfolio return and risk predictions.

We divide our modeling and analysis into two parts. The first part focuses on the historical, quantitative relationships between economic variables. In financial theory, this part is well described by the capital asset pricing model, the arbitrage pricing theory, and other asset pricing models that have been built on these foundations. The goal of these analyses is to predict the value of a variable as a function of others—an input-output analysis. There is no attempt to understand or explain the process by which the inputs are transformed to outputs. Our goal here is to attempt to model the process by which different factors are related to the variable of interest in a little more detail than those provided by financial theories. In [Shenoy and Shenoy 2000], we used the theory of Bayesian networks to describe the dependencies between a collection of variables. In this paper, we focus on the use of belief functions to model such dependencies.

The second part of a financial analysis is to study how changes in current economic events change historical relationships between the variables. We are

unaware of any formal models in financial theory that do this. Both steps are necessary for good predictions, but historically more attention has been devoted to the first step. The theory of belief functions provides a natural method for modeling evidence. And incorporating evidence into our prior model entails modeling evidence as a belief function and ensuring that there is no double-counting of uncertain evidence with our prior belief function so one can use Dempster's rule of combination to arrive at inferences based on the combined knowledge.

We use a portfolio of three gold stocks to illustrate our belief function model. Each stock is affected by three factors—the overall stock market, the gold market, and firm specific effects. Changes in any of these factors may affect the price, return, and/or risk of the stock. The output of the network is a marginal belief function for the portfolio and stock variables. Changes in the inputs affect the output distribution so that changes in risk and return can be assessed. A belief function can be summarized by a probability mass function and from this summary probability distribution, value-at-risk can be calculated.

This paper is organized as follows. In section two, we present a brief overview of belief functions. In section three, we construct the belief function portfolio model. This section includes the graphical model, a description of how belief functions can be generated, and how basic probability assignment functions are constructed. In section four, we illustrate how new evidence can be added to the model and the impact of the new evidence on the portfolio risk and return distribution. In section five, we conclude with a summary and some concluding remarks.

## 2. Basics of the Theory of Belief Functions

In this section, we sketch briefly the basics of the Dempster-Shafer theory of belief functions.

There are several equivalent ways to represent knowledge and/or evidence in the theory of belief functions. One can use either basic probability assignment (bpa) functions, or belief functions, or plausibility functions. We start by describing bpa functions. But first we set out our notation.

**Notation.** We use upper-case Roman alphabets to describe *variables*. Each variable  $X$  is associated with a finite set of its possible values denoted by  $\Omega_X$ . We call the elements of  $\Omega_X$  *states* of  $X$ , and denote them by bold-faced alphabets such as  $x_1, x_2$ , etc. We use lower-case alphabets to denote sets of variables. Suppose  $s$  is a subset of variables. The state space of  $s$ , denoted by  $\Omega_s$ , is the Cartesian product of the state space of  $X$  for  $X \in s$ , i.e.,  $\Omega_s = \times\{\Omega_X \mid X \in s\}$ .  $2^{\Omega_s}$  denotes the

set of all subsets of  $\Omega_s$ , and elements of  $2^{\Omega_s}$  (subsets of  $\Omega_s$ ) are denoted by lower-case bold-faced letter in Chicago font, **a**, **b**, **c**, etc.

**Basic probability assignment functions.** A bpa function for  $s$  is a function  $\mu: 2^{\Omega_s} \rightarrow [0, 1]$  such that  $\mu(\emptyset) = 0$ , and

$$\Sigma\{\mu(\mathbf{a}) \mid \mathbf{a} \in 2^{\Omega_s}\} = 1 \quad (1.1)$$

$\mu(\mathbf{a})$  can be interpreted as the proportion of belief that is assigned to proposition **a** exactly and nothing else. Propositions **a** such that  $\mu(\mathbf{a}) > 0$  are called *focal elements* of  $\mu$ . We will describe bpa functions by listing only the values of the bpa function for the focal elements.

Suppose we have evidence, which if reliable, supports proposition **a**, and which if unreliable, supports  $\Omega_s$ . Also suppose that the probability of reliability of the evidence is  $p$ . In this case, we can model the evidence by a bpa function  $\mu_e$  such that  $\mu_e(\mathbf{a}) = p$ , and  $\mu_e(\Omega_s) = 1-p$ . A bpa function for  $s$  that represents complete ignorance is  $\mu_v(\Omega_s) = 1$ , it is referred to as the *vacuous* bpa function.

**Belief functions.** Related to each bpa function  $\mu$  are functions called belief  $\beta$  and plausibility  $\pi$ . A *belief function* for  $s$ ,  $\beta: 2^{\Omega_s} \rightarrow [0, 1]$  can be defined in terms of a bpa function  $\mu$  for  $s$  as follows:

$$\beta(\mathbf{a}) = \Sigma\{\mu(\mathbf{b}) \mid \mathbf{b} \subseteq \mathbf{a}\} \quad (1.2)$$

for each **a** in  $2^{\Omega_s}$ .  $\beta(\mathbf{a})$  can be interpreted as a conservative measure of belief in proposition **a**. Since proposition **b** logically entails all propositions **a** such that  $\mathbf{b} \subseteq \mathbf{a}$ , this explains the definition of  $\beta$  in terms of  $\mu$  in (1.2). Notice that as a consequence of (1.2) and the definition of  $\mu$  is that  $\beta(\emptyset) = 0$ , and  $\beta(\Omega_s) = 1$ . Also notice that unlike a probability function,  $\beta(\mathbf{a}) \geq \Sigma\{\beta(\{\mathbf{x}\}) \mid \mathbf{x} \in \mathbf{a}\}$ . If all focal elements of  $\mu$  are singleton subsets, then in this case, we do have  $\beta(\mathbf{a}) = \Sigma\{\beta(\{\mathbf{x}\}) \mid \mathbf{x} \in \mathbf{a}\}$  for all **a** in  $2^{\Omega_s}$ .

**Plausibility functions.** A *plausibility function*  $\pi$  for  $s$  corresponding to a bpa function  $\mu$  for  $s$  is a function  $\pi: 2^{\Omega_s} \rightarrow [0, 1]$  such that

$$\pi(\mathbf{a}) = \Sigma\{\mu(\mathbf{b}) \mid \mathbf{b} \cap \mathbf{a} \neq \emptyset\} \quad (1.3)$$

for each **a** in  $2^{\Omega_s}$ . It follows from (1.2) and (1.3) that  $\pi(\mathbf{a}) \geq \beta(\mathbf{a})$ , and  $\pi(\mathbf{a}) = 1 - \beta(\mathbf{a}^c)$ , where  $\mathbf{a}^c = \Omega_s - \mathbf{a}$  is the complement of **a** in  $\Omega_s$ . Thus, a plausibility function is not as conservative as the corresponding belief function. If all focal

elements of  $\mu$  are singleton subsets, then,  $\pi(\mathbf{a}) = \beta(\mathbf{a}) = \sum\{\beta(\{\mathbf{x}\}) \mid \mathbf{x} \in \mathbf{a}\}$  for all  $\mathbf{a} \in 2^{\Omega_s}$ .

**Discounting bpa functions.** Suppose  $\mu$  is a bpa function for  $s$ . Often we would like to discount bpa functions by some weight  $w$  where  $0 \leq w \leq 1$ . Let  $(1-w)\mu$  denote the discounted bpa function.  $(1-w)\mu$  is defined as follows:  $((1-w)\mu)(\mathbf{a}) = (1-w)\mu(\mathbf{a})$  for all  $\mathbf{a} \in 2^{\Omega_s - \{\Omega_s\}}$ , and  $((1-w)\mu)(\Omega_s) = (1-w)\mu(\Omega_s) + w$ . Intuitively, to discount a bpa function by  $w$  means to move a total of mass  $w$  from the focal elements of  $\mu$  to the state space of the domain of  $\mu$ . Thus, if  $w = 0$  (0% discounting), then  $1\mu = \mu$ , and if  $w = 1$  (100% discounting), then  $0\mu = \mu_v$ , where  $\mu_v$  is the vacuous bpa function for  $s$ .

**Dempster's rule of combination.** The rule for updating bpa functions in light of new evidence is called *Dempster's rule of combination* [Dempster 1967]. Suppose  $\mu_0$  is our prior bpa function for  $s$  and suppose we get evidence for  $t$  that is represented by bpa function  $\mu_1$  for  $t$ . The combination of  $\mu_0$  and  $\mu_1$ , denoted by  $\mu_0 \oplus \mu_1$ , is the bpa function for  $s \cup t$  defined as follows:  $(\mu_0 \oplus \mu_1)(\emptyset) = 0$ , and

$$(\mu_0 \oplus \mu_1)(\mathbf{c}) = K^{-1} \sum\{\mu_0(\mathbf{a})\mu_1(\mathbf{b}) \mid (\mathbf{a} \times \Omega_{t-s}) \cap (\mathbf{b} \times \Omega_{s-t}) = \mathbf{c}\} \quad (1.4)$$

for all non-empty  $\mathbf{c} \in 2^{\Omega_{s \cup t}}$ , where  $K = \sum\{\mu_0(\mathbf{a})\mu_1(\mathbf{b}) \mid (\mathbf{a} \times \Omega_{t-s}) \cap (\mathbf{b} \times \Omega_{s-t}) \neq \emptyset\}$  is the normalization constant. If the pieces of evidence on which  $\mu_0$  and  $\mu_1$  are based are "distinct," then  $\mu_0 \oplus \mu_1$  is the bpa representation of the aggregate of such evidence. In general,  $\mu \oplus \mu \neq \mu$ . Therefore, it is important to ensure that there is no double counting of evidence.

**Marginalization.** Finally, we will define marginalization for bpa functions. Suppose  $\mu$  is a bpa function for  $s$ , and suppose  $t \subset s$ . The marginal of  $\mu$  for  $t$ , denoted by  $\mu^{\downarrow t}$ , is the bpa function for  $t$  defined as follows:

$$\mu^{\downarrow t}(\mathbf{a}) = \sum\{\mu(\mathbf{b}) \mid \mathbf{b}^{\downarrow t} = \mathbf{a}\} \quad (1.5)$$

for each  $\mathbf{a} \in \Omega_t$ , where  $\mathbf{b}^{\downarrow t}$  denotes the subset of  $\Omega_t$  obtained by projecting each element of  $\mathbf{b}$  to  $t$ .

**Summarizing bpa functions by probability mass functions.** Often, we would like to summarize a bpa function  $\mu$  for  $s$  by a corresponding probability mass function for  $s$ ,  $\text{Pr}: \Omega_s \rightarrow [0, 1]$ . There are a couple of ways of doing this. One way is called the *pignistic transformation* [Smets 1990]. The pignistic transformation is defined as follows:

$$\text{Pr}(\mathbf{x}) = \sum\{\mu(\mathbf{a})/|\mathbf{a}| \mid \mathbf{a} \in 2^{\Omega_s} \text{ such that } \mathbf{x} \in \mathbf{a}\} \quad (1.6)$$

for each  $\mathbf{x} \in \Omega_s$ . It's easy to verify that  $\Pr$  as defined above is a probability mass function.

Another way to transform a bpa function to a probability mass function is to use plausibilities of singleton subsets suitably normalized. Suppose  $\mu$  is a bpa function for  $s$ . A *plausibility transformation* is a probability mass function  $\Pr: \Omega_s \rightarrow [0, 1]$  defined as follows:

$$\Pr(\mathbf{x}) = K^{-1} \pi(\{\mathbf{x}\}) \quad (1.7)$$

for each  $\mathbf{x} \in \Omega_s$ , where  $\pi$  is the plausibility function for  $s$  corresponding to bpa function  $\mu$ , and  $K = \sum \{\pi(\{\mathbf{x}\}) \mid \mathbf{x} \in \Omega_s\}$ . If  $\mu_v$  is the vacuous bpa function for  $s$ , then both the pignistic and the plausibility transformations of  $\mu_v$  result in the equally-likely probability mass function for  $s$ ,  $\Pr(\mathbf{x}) = 1/|\Omega_s|$  for each  $\mathbf{x} \in \Omega_s$ . However, in general, the two transformations lead to different probability mass functions. In this paper, we will use the plausibility transformation to summarize a bpa function.

### 3. A Simple Portfolio Model

We construct a simple portfolio model constituting of three stocks. Each of the stocks is a gold-mining stock. Each stock can be affected by changes in the stock market, the gold price, or other unspecified factors. This example is small enough to illustrate the computations, yet it includes the features of a large class of portfolios. It builds on the traditional framework of finance models such as the capital asset pricing model and multi-factor models. In a multi-factor model the return on a stock is:

$$r_i = \alpha + \beta_1 f_1 + \beta_2 f_2 + \dots + \beta_k f_k + \varepsilon_i \quad (2.1)$$

where  $r_i$  is the return on stock  $i$ ,  $f_k$  is the return on factor  $k$ ,  $\beta_k$  is the responsive of the stock to factor  $k$ , and  $\varepsilon_i$  is a random component of the return due to firm specific effects. Some of the assumptions of this model that can be relaxed in our more general belief function model are:

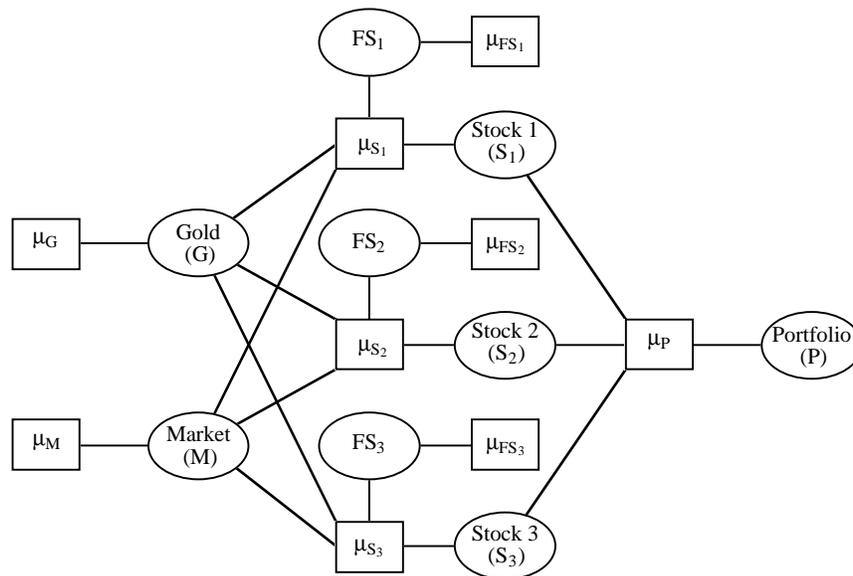
- the factors  $f_1$  to  $f_k$  are assumed to be uncorrelated;
- the expected value of the firm specific component,  $E(\varepsilon_i)$ , is equal to zero;
- there is no correlation between  $\varepsilon_i$  and any factor,  $f_1, \dots, f_k$ ; and
- the correlation between the firm specific effects for any firms  $i$  and  $j$  is equal to zero,  $\rho(\varepsilon_i, \varepsilon_j) = 0$ .

The model in (2.1) is a conditional probability model, where the distribution of the stock return is conditioned on the factor distributions and the firm specific effect distributions. In [Shenoy and Shenoy 2000], we show how an equivalent Bayes net models can be constructed. In this paper, however, we will describe a “knowledge-based” model in which we use knowledge to describe the relations between the variables. Also, we will use belief functions instead of probability distributions to model the uncertainty in the relationship among the variables.

### 3.1 The Graphical Structure of the Model

Our portfolio consists of three gold-mining stocks—Stock 1 ( $S_1$ ), Stock 2 ( $S_2$ ), and Stock 3 ( $S_3$ ). Each stock is influenced by three factors: the stock market (M), the price of gold (G), and other firm specific factors. We denote the firm specific factors for the three firms as  $FS_1$ ,  $FS_2$ , and  $FS_3$ , respectively. The firm specific factors are variables that include anything that will affect the firm that is not already included in the factors. Finally, we are interested in the return distribution of the portfolio (P).

**Figure 1:** The graphical structure of the portfolio model.



The graphical structure of the portfolio is shown in Figure 1 as a valuation network. There are two types of nodes in a valuation network. The elliptical nodes represent variables, and the rectangular nodes represent bpa functions. A valuation network is a bi-partite graph in which there are edges only between variable nodes

and bpa function nodes. The edges between a rectangular node and elliptical nodes denote the domain of the bpa function. Thus, for example, the domain of  $\mu_1$  is  $\{G, M, FS_1, S_1\}$ , and it represents how the return of Stock 1 is related to the values of gold, market, and  $FS_1$ . The bpa function  $\mu_P$  denotes how the portfolio return is related to the individual stock returns and is a deterministic function of the number of stocks of each firm in the portfolio. A fundamental assumption of a valuation network model is that all bpa functions in the model are based on distinct pieces of evidence and knowledge, i.e., there is no double counting of uncertain knowledge/evidence.

This portfolio structure is similar to traditional asset pricing models. For example, the capital asset pricing model describes individual stock returns as a conditional expectation of the market return and a residual term. Here we have the market level and the price of gold as factors affecting gold stocks. However, we will describe a knowledge-based model to describe the relationship between the variables.

Once we have specified the model completely by defining all bpa functions, we can evaluate the model by combining all bpa functions using Dempster's rule of combination to yield a joint bpa function. Then we compute the marginal of the joint bpa function for any variable of interest. In our portfolio model, we are interested in computing the marginal of the joint bpa function for the portfolio variable P and for each individual stock variables,  $S_1$ ,  $S_2$ , and  $S_3$ .

The output of the evaluation is bpa functions  $\mu^{\downarrow P}$  for P,  $\mu^{\downarrow S_1}$  for  $S_1$ ,  $\mu^{\downarrow S_2}$  for  $S_2$ , and  $\mu^{\downarrow S_3}$  for  $S_3$ , where  $\mu = \mu_G \oplus \mu_M \oplus \mu_{S_1} \oplus \mu_{S_2} \oplus \mu_{S_3} \oplus \mu_{FS_1} \oplus \mu_{FS_2} \oplus \mu_{FS_3} \oplus \mu_P$  is the *joint* bpa function. Since probability distributions are more familiar and easier to interpret than bpa functions, we summarize the output bpa functions  $\mu^{\downarrow P}$ ,  $\mu^{\downarrow S_1}$ ,  $\mu^{\downarrow S_2}$ , and  $\mu^{\downarrow S_3}$  by probability mass functions using the plausibility transformation.

### 3.2 Defining Basic Probability Assignment Functions

In this section, we describe the details of the bpa functions in the graphical model. First, we determine the state space of each of the variables. Next, we describe our knowledge and represent it by bpa functions. In our example, we define three states for each of the stocks, the market, gold, and the portfolio variables<sup>1</sup>: low (l), average (a), and high (h). To identify the states with the corresponding variables,

---

<sup>1</sup> It may be more natural to describe the state space as a continuous space. However, the computational theory of belief function is well developed only for the discrete case. Since without computational aids, a model becomes just an academic exercise, we will restrict our model to the case of discrete state spaces for all variables.

we will subscript the states with the variable name. Thus, for example,  $\Omega_G = \{l_G, a_G, h_G\}$ ,  $\Omega_M = \{l_M, a_M, h_M\}$ , etc. For the firm specific factors, we define two states: favorable (f) and unfavorable (u). Thus,  $\Omega_{FS_1} = \{f_{FS_1}, u_{FS_1}\}$ , etc.

The variables can be measured as returns or prices. For example, the ‘low’ state might be a “buy” price, the ‘high’ state might be a “sell” price, and ‘average’ might be a “hold” price. We can also define these states using historical returns over a specified period. For example, ‘low,’ may be defined as below the 12.5 percentile; high, above the 87.5 percentile; and ‘average’, between the 12.5 and 87.5 percentiles. The exact specification depends on the type of decision that the model will be used to support. A short-term day trading model may be more useful with prices, and a longer-term evaluation of the portfolio would likely use returns.

Returns and prices are typically described by continuous state spaces. But here we will restrict ourselves to the discrete case. However, the number of states can be expanded as needed. Smith [1993] describes procedures for discretizing continuous state spaces in order to use these variables in a decision problem.

First we define the bpa functions  $\mu_G$ ,  $\mu_M$ ,  $\mu_{FS_1}$ ,  $\mu_{FS_2}$ , and  $\mu_{FS_3}$ . Each of these bpa functions has a single variable in its domain. Their role is similar to the prior distributions in a Bayesian probability model. If we don’t have any information regarding these variables, we can always use the vacuous bpa function for this situation.

We illustrate the construction of a gold BPA with an example. One kind of information that frequently affects gold prices is news that a central bank is selling gold. Suppose that we have observed in the past that one of three circumstances occurs when central bank information is released. Many times the news will not change the gold price from its average state because the sale is relatively small. This state occurs 55% of the time. On occasion, because the sale is large or because the news is unexpected, the news will decrease the gold price. The amount of decrease is unknown. If the decrease is slight, the price remains in the average category. If the decrease is severe, the price dips to the low category. We assume that from past data, 15% of the time the price of gold has decreased. The remaining 30% of the time announcements of gold sale were made, there is no record of what happened to the price of gold. Therefore we assign this mass to the entire state space. Table 1 summarizes the Gold evidence from a central bank announcement.

**Table 1.** Bpa function  $\mu_G$  for gold

$\mathbf{a} \in 2^{\Omega_G}$	$\mu_G(\mathbf{a})$
$\{a_G\}$	0.55
$\{1_G, a_G\}$	0.15
$\Omega_G$	0.30

The bpa function  $\mu_M$  is shown in Table 2. It is derived from evidence similar to  $\mu_G$ .

**Table 2.** Bpa function for market

$\mathbf{a} \in 2^{\Omega_M}$	$\mu_M(\mathbf{a})$
$\{a_M\}$	0.45
$\{a_M, h_M\}$	0.40
$\Omega_M$	0.15

For now, we will assume that the bpa functions  $\mu_{FS_1}$ ,  $\mu_{FS_2}$ , and  $\mu_{FS_3}$  are vacuous. When we receive evidence about these firm specific factors, we can update these bpa functions.

Next we define the bpa functions  $\mu_{S_1}$ ,  $\mu_{S_2}$ , and  $\mu_{S_3}$ , which describe how each of the three gold stocks are related to the gold, market and the corresponding firm specific factor. Like the noisy-OR model in Bayes nets, we assume that the bpa function  $\mu_{S_1}$  decomposes into three separate distinct bpa functions—one that has  $\{M, S_1\}$  for its domain, one that has  $\{G, S_1\}$  in its domain, and one that has  $\{FS_1, S_1\}$  in its domain. This implicitly assumes that there are no interactions between the three causes of stock 1. Furthermore, we assume that  $\mu_{S_1}$  is a “weighted average” of these three bpa functions.

Table 3 shows the details of the bpa function  $\mu_{G-S_1}$  that describes the relationship between gold and stock 1 assuming gold is the only factor that influences stock 1. There are two focal elements. In the first, stock 1 tracks gold exactly. The probability assigned to this focal element is 0.7. In the second, we assume that the gold has a ‘lag’ effect on stock 1. The probability assigned to this element is 0.1. The remaining 0.2 probability mass is uncommitted.

**Table 3.** Bpa function  $\mu_{G-S_1}$ 

$\mathbf{a} \in 2^{\Omega_{\{G, S_1\}}}$	$\mu_{G-S_1}(\mathbf{a})$
$\{(l_G, l_{S_1}), (a_G, a_{S_1}), (h_G, h_{S_1})\}$	0.70
$\{(l_G, l_{S_1}), (a_G, a_{S_1}), (h_G, h_{S_1}), (h_G, a_{S_1}), (a_G, l_{S_1})\}$	0.10
$\Omega_{\{G, S_1\}}$	0.20

Tables 4 and 5 show a similar relationship between market and stock 1, and between  $FS_1$  and stock 1. For  $\mu_{G-S_1}$  and  $\mu_{M-S_1}$  we have uncommitted masses of 0.20. This means that our knowledge is based on fair amount of knowledge. If the uncommitted mass is high, then this means we have little evidence about the relationship between the variables.

**Table 4.** Bpa function  $\mu_{M-S_1}$ 

$\mathbf{a} \in 2^{\Omega_{\{M, S_1\}}}$	$\mu_{M-S_1}(\mathbf{a})$
$\{(l_M, l_{S_1}), (a_M, a_{S_1}), (h_M, h_{S_1})\}$	0.40
$\{(l_M, l_{S_1}), (a_M, a_{S_1}), (h_M, h_{S_1}), (h_M, a_{S_1}), (a_M, l_{S_1})\}$	0.40
$\Omega_{\{M, S_1\}}$	0.20

**Table 5.** Bpa function  $\mu_{FS_1-S_1}$ 

$\mathbf{a} \in 2^{\Omega_{\{FS_1, S_1\}}}$	$\mu_{FS_1-S_1}(\mathbf{a})$
$\{(f_{FS_1}, h_{S_1}), (u_{FS_1}, l_{S_1})\}$	0.80
$\{(f_{FS_1}, h_{S_1}), (u_{FS_1}, l_{S_1}), (f_{FS_1}, a_{S_1})\}$	0.10
$\Omega_{\{FS_1, S_1\}}$	0.10

We assume that  $\mu_{S_1}$  is the “weighted average” of the three bpa functions  $\mu_{G-S_1}$ ,  $\mu_{M-S_1}$ , and  $\mu_{FS_1-S_1}$ . By weighted average, we mean we first discount<sup>2</sup> each bpa function by some weight  $1-w_i$  and then we combine the discounted bpa function using Dempster’s rule. If the weights used are  $w_G$ ,  $w_M$ , and  $w_{FS_1}$  where  $w_G + w_M + w_{FS_1} = 1$ , then  $\mu_{S_1} = w_G \mu_{G-S_1} \oplus w_M \mu_{M-S_1} \oplus w_{FS_1} \mu_{FS_1-S_1}$ . To determine the weights to be used, we could either assess these subjectively from an expert or we

<sup>2</sup> The process of discounting bpa functions is explained in Section 2.

could use an objective measure. We choose to do the latter, and we base the weights computed from a linear regression model using historical data as follows. The regression model of the stock  $i$  return ( $r_i$ ) given the return on the market ( $r_M$ ) and the return on gold ( $r_G$ ) is given by:  $r_i = \alpha + \beta_G r_G + \beta_M r_M + \epsilon_i$  where  $\epsilon_i$  is the residual. Let  $SS_G$ ,  $SS_M$ ,  $SS_{\text{error}}$ , and  $SS_{\text{total}}$  denote the sum of squares due to market, gold, error, and total, respectively. Then we define  $w_G = SS_G/SS_{\text{total}}$ ,  $w_M = SS_M/SS_{\text{total}}$ , and  $w_{FS_i} = SS_{\text{error}}/SS_{\text{total}}$ . Table 6 shows the weights that we used in our example to combine the basic probability assignments for each stock. We assume that  $\mu_{G-S_2}$  and  $\mu_{G-S_3}$  are similar to  $\mu_{G-S_1}$ ;  $\mu_{M-S_2}$  and  $\mu_{M-S_3}$  are similar to  $\mu_{M-S_1}$ ; and  $\mu_{FS_2-S_2}$  and  $\mu_{FS_3-S_3}$  are similar to  $\mu_{FS_1-S_1}$ . The differences in the three stocks are captured in our model by the weights shown in Table 6.

**Table 6.** Bpa combination weights for each stock

Stock 1	$w_i$	Stock 2	$w_i$	Stock 3	$w_i$
G-S <sub>1</sub>	0.60	G-S <sub>2</sub>	0.45	G-S <sub>3</sub>	0.50
M-S <sub>1</sub>	0.10	M-S <sub>2</sub>	0.15	M-S <sub>3</sub>	0.05
FS <sub>1</sub> -S <sub>1</sub>	0.30	FS <sub>2</sub> -S <sub>2</sub>	0.40	FS <sub>3</sub> -S <sub>3</sub>	0.45

Finally, we describe the bpa function  $\mu_p$  that describes the relationship between the three stocks and the portfolio. This relationship is a deterministic one depending on the number of stocks of each type that constitutes the portfolio. Assuming that we have equal number of funds in each stock, this deterministic relationship is defined by the bpa function shown in Table 7. This bpa function has only one focal element with value (that is necessarily) 1. Such bpa functions (with a single focal element) are called *deterministic*.

This completes our description of the belief function portfolio model. In the next section, we discuss making inferences from this model.

**Table 7.** Bpa function  $\mu_P$

$\mathbf{a} \in 2^{\Omega_{\{S_1, S_2, S_3, P\}}}$	$\mu_P(\mathbf{a})$
$\{(l_{S_1}, l_{S_2}, l_{S_3}, l_P), (l_{S_1}, l_{S_2}, a_{S_3}, l_P), (l_{S_1}, l_{S_2}, h_{S_3}, a_P),$ $(l_{S_1}, a_{S_2}, l_{S_3}, l_P), (l_{S_1}, a_{S_2}, a_{S_3}, a_P), (l_{S_1}, a_{S_2}, h_{S_3}, a_P),$ $(l_{S_1}, h_{S_2}, l_{S_3}, a_P), (l_{S_1}, h_{S_2}, a_{S_3}, a_P), (l_{S_1}, h_{S_2}, h_{S_3}, a_P),$ $(a_{S_1}, l_{S_2}, l_{S_3}, l_P), (a_{S_1}, l_{S_2}, a_{S_3}, a_P), (a_{S_1}, l_{S_2}, h_{S_3}, a_P),$ $(a_{S_1}, a_{S_2}, l_{S_3}, a_P), (a_{S_1}, a_{S_2}, a_{S_3}, a_P), (a_{S_1}, a_{S_2}, h_{S_3}, a_P),$ $(a_{S_1}, h_{S_2}, l_{S_3}, a_P), (a_{S_1}, h_{S_2}, a_{S_3}, a_P), (a_{S_1}, h_{S_2}, h_{S_3}, h_P),$ $(h_{S_1}, l_{S_2}, l_{S_3}, a_P), (h_{S_1}, l_{S_2}, a_{S_3}, a_P), (h_{S_1}, l_{S_2}, h_{S_3}, a_P),$ $(h_{S_1}, a_{S_2}, l_{S_3}, a_P), (h_{S_1}, a_{S_2}, a_{S_3}, a_P), (h_{S_1}, a_{S_2}, h_{S_3}, h_P),$ $(h_{S_1}, h_{S_2}, l_{S_3}, a_P), (h_{S_1}, h_{S_2}, a_{S_3}, h_P), (h_{S_1}, h_{S_2}, h_{S_3}, h_P)\}$	1

### 3.3 Inferences from the Portfolio Model

The belief function model is evaluated by first combining all of the basic probability assignments using Dempster's rule, and then computing the marginal of the joint for the variables of interest. When there are a large number of variables, it may be computationally intractable to compute the joint bpa function. Nevertheless, one can compute the marginal of the joint bpa function without explicitly computing the joint bpa function using so-called local computation [Shenoy and Shafer 1990, Shenoy 1994]. We use a software implementation of belief functions called MacEvidence [Hsia and Shenoy 1989, Shafer *et al.* 1988] to compute the marginals of the joint bpa function.

Table 8 describes the marginal bpa function for the portfolio variable. As we said earlier, bpa functions are difficult to interpret. We can compute the corresponding belief function  $\beta^{\downarrow P}$  or corresponding plausibility function  $\pi^{\downarrow P}$ . These are shown in Table 8. Alternatively, we can summarize the bpa function by a probability mass function using the plausibility transformation. The first step is to find plausibilities for each state. For example  $\pi^{\downarrow P}(l_P) = 0.009 + 0.014 + 0.143 + 0.714 = 0.880$ . Table 9 summarizes the plausibilities of each of the states of P and the corresponding probability mass function. We emphasize that when we summarize a bpa function by a probability mass function, some information is necessarily lost.

**Table 8.** Marginal bpa function  $\mu^{\downarrow P}$ , belief function  $\beta^{\downarrow P}$ , and plausibility function  $\pi^{\downarrow P}$

$\mathbf{a} \in 2^{\Omega_P}$	$\mu^{\downarrow P}(\mathbf{a})$	$\beta^{\downarrow P}(\mathbf{a})$	$\pi^{\downarrow P}(\mathbf{a})$
$\{l_P\}$	0.009	0.009	0.880
$\{a_P\}$	0.102	0.102	0.976
$\{h_P\}$	0.001	0.001	0.746
$\{l_P, h_P\}$	0.014	0.024	0.898
$\{l_P, a_P\}$	0.143	0.254	0.999
$\{a_P, h_P\}$	0.017	0.120	0.991
$\Omega_P$	0.714	1.000	1.000

**Table 9.** Plausibilities of singleton subsets of P

$\Omega_P$	$\pi^{\downarrow P}$	$\Pr(P)$
$l_P$	0.880	0.338
$a_P$	0.976	0.375
$h_P$	0.746	0.287
Total	2.601	1.000

In a similar manner, we first compute the marginal bpa functions for  $S_1$ ,  $S_2$ , and  $S_3$ , and then the summary probability mass functions using the plausibility transformations. These are shown in Table 10.

**Table 10.** Probabilities mass functions for the three stock variables

$\Omega$	$\Pr(S_1)$	$\Pr(S_2)$	$\Pr(S_3)$
low	0.356	0.368	0.377
average	0.338	0.302	0.289
high	0.306	0.331	0.335

## 4. Entering Evidence

Belief functions can be used to represent conflicting evidence from different sources. For example, consider a report that Asian jewelry makers are entering the gold market to stockpile inventory. There is a small (10%) chance that this report is just idle speculation on the part of some analyst. If the report is genuine, it is not certain from the report just how much inventory the Asian jewelry makers will stockpile. Based on previous knowledge, there is a 15% chance that the report is genuine and that they will stockpile a significant amount leading to high gold returns. There is a 75% chance that the report is genuine and that they will stockpile enough gold to result in either average or high gold returns. This second piece of evidence can be represented by a bpa function  $\mu_{G_2}$  as shown in Table 11.

To avoid double counting of uncertain knowledge,  $\mu_{G_2}$  is based on the evidence alone (disregarding all other evidence).

**Table 11.** Additional evidence assessment for gold

$\mathbf{a} \in 2^{\Omega_G}$	$\mu_{G_2}(\mathbf{a})$
$\{h_G\}$	0.15
$\{a_G, h_G\}$	0.75
$\Omega_G$	0.10

If we include this piece of evidence, and re-compute the marginals for the three stock variables and the portfolio variable, the results summarized as probability distributions are shown in Table 12 below.

**Table 12:** Revised probability mass functions for stock and portfolio variables

$\Omega$	$\Pr(S_1)$	$\Pr(S_2)$	$\Pr(S_3)$	$\Pr(P)$
low	0.292	0.323	0.326	0.279
average	0.364	0.318	0.307	0.398
high	0.344	0.359	0.367	0.323

As expected, comparing with the prior distributions (as seen in Table 10), the probability of the portfolio being in the low state decreases, and the probability of being in the average or high state increases. The same is also true for each of the stock variables.

## 5. Summary and Conclusion

In this paper, we demonstrate how the theory of belief functions may be used to model financial portfolios. In particular, we show how to model portfolio changes as we learn new information about different factors that influence individual stocks or a portfolio. Financial theory focuses primarily on describing historical relationship between financial variables, and they do this using regression models. Regression models do not attempt to either understand or explain how the inputs are converted to outputs—they simply attempt to predict the outputs from the inputs using data. Our goal in this paper is to describe a knowledge-based model that describes the dependencies between all variables based on knowledge of relationships between variables. We use the theory of belief functions to describe the relationships and make inferences from the model. We also show how new information or evidence can be modeled as independent belief functions and then be used to update the predictions.

## Acknowledgements

This paper has benefited from comments and discussions with Philippe Smets, Jürg Kohlas, Rajendra Srivastava, and one anonymous reviewer.

## References

- Dempster, A. P. (1967), "Upper and lower probabilities induced by a multivalued mapping." *Annals of Mathematical Statistics*, **38**, 325–339.
- Hsia, Y. T. and P. P. Shenoy (1989), "An evidential language for expert systems," in Z. W. Ras (ed.), *Methodologies for Intelligent Systems*, **4**, 9–16, North-Holland, Amsterdam.
- Kohlas, J. and P.-A. Monney (1995), *A Mathematical Theory of Hints: An Approach to Dempster-Shafer Theory of Evidence*, Lecture Notes in Economics and Mathematical Systems, **425**, Springer-Verlag, Berlin.
- Shafer, G. (1976), *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, NJ.
- Shafer, G., P. P. Shenoy, and R. P. Srivastava (1988), "AUDITOR'S ASSISTANT: A knowledge engineering tool for audit decisions," in R. P. Srivastava and J. E. Rebele (eds.), *Auditing Symposium IX*, 61–84, School of Business, University of Kansas, Lawrence, KS.
- Shenoy, P. P. (1994), "Using Dempster-Shafer's belief-function theory in expert systems," in R. R. Yager, M. Fedrizzi and J. Kacprzyk (eds.), *Advances in the Dempster-Shafer Theory of Evidence*, 395–414, Wiley, New York.

- Shenoy, P. P. and G. Shafer (1990), "Axioms for probability and belief-function propagation," in Shachter, R. D., T. S. Levitt, J. F. Lemmer and L. N. Kanal (eds.), *Uncertainty in Artificial Intelligence*, **4**, 169–198, North-Holland, Amsterdam. Reprinted in: Shafer, G. and J. Pearl (eds.), *Readings in Uncertain Reasoning*, 575–610, 1990, Morgan Kaufmann, San Mateo, CA.
- Shenoy, C. and P. P. Shenoy (2000), "Bayesian Network Models of Portfolio Risk and Return," in Y. S. Abu-Mostafa, B. LeBaron, A. W. Lo, and A. S. Weigend (eds.), *Computational Finance 1999*, 87–106, MIT Press.
- Smets, P. (1990), "Constructing the pignistic probability function in a context of uncertainty," in Henrion, M., R. D. Shachter, L. N. Kanal, and J. F. Lemmer (eds.), *Uncertainty in Artificial Intelligence*, **5**, 29–40, North-Holland, Amsterdam.
- Smets, P. (1998), "The transferable belief model for quantified belief representation," in P. Smets (ed.), *Handbook of Defeasible Reasoning and Uncertainty Management Systems: Quantified Representation of Uncertainty and Imprecision*, **1**, 267–301, Kluwer, Netherlands.
- Smith, J. E. (1993), "Moment methods for decision analysis," *Management Science*, **39**, 340–358.