1. BAYESIAN NETWORK MODELS OF PORTFOLIO RISK AND RETURN

Catherine Shenoy and Prakash P. Shenoy

A Bayesian network is a tool for modeling large multivariate probability models and for making inferences from such models. A Bayesian network combines traditional quantitative analysis with expert judgement in an intuitive, graphical representation. In this paper, we show how to use Bayesian networks to model portfolio risk and return.

Traditional financial models emphasize the historical relationship between portfolio return and market return. In practice, to forecast portfolio return, financial analysts include expert subjective judgement about other factors that may affect the portfolio. These judgmental factors include special knowledge about the stocks in the portfolio that is not captured in the historical quantitative analysis.

We show how a Bayesian network can be used to represent a traditional financial model of portfolio return. Then we show how expert subjective judgement can be included in the Bayesian network model. The output of the model is the posterior marginal probability distribution of the portfolio return. This posterior return distribution can be used to obtain expected return, return variance, and value-at-risk.

The main goal of this paper is to show how Bayesian networks can be used to model portfolio risk and return. Bayesian networks have been used as a tool for modeling large multivariate probability models and for making inferences from such models (Pearl 1986, Lauritzen and Spiegelhalter 1988, Shenoy and Shafer 1990). A Bayesian network combines traditional quantitative analysis with an analyst's judgement in an intuitive, graphical representation. It allows an analyst to visualize the relationships among the variables in the model.

Finance models focus on the historical, quantitative relationships between economic variables. However, financial analysts usually combine historical data with qualitative information and judge how this information affects stock returns, market return, interest rates, or any other input to a portfolio model. For example, the anti-trust lawsuit against Microsoft affects the stock returns of many companies, but this type of information is difficult to incorporate in traditional return models. Bayesian networks are especially well suited for situations that combine quantitative and qualitative information. In this paper we provide an overview of how to combine traditional financial models with judgments about qualitative information in a Bayesian network framework.

Traditional portfolio return models are static. There is no systematic way to update results in the light of new information. A Bayesian network representation of portfolio return allows analysts to incorporate new information, to see the effect of that information on the return distributions for the whole network, and to visualize the distribution of returns, not just the summary statistics. In a Bayesian network an analyst first determines the qualitative structure of the model in an intuitive graphical way. A traditional portfolio model can easily be represented as a Bayesian network. From that basic structure, quantitative information is then added to the model. Any change to either the qualitative or quantitative structure of the model is immediately reflected as the model is updated. The output of the Bayesian network analysis is a distribution of portfolio returns based on the qualitative and quantitative structure of the model.

Most traditional financial models rely on strong implicit assumptions about the independence of various factors incorporated in the model. In a Bayesian network model, the analyst can explicitly model the dependence or independence of the factors. It is then possible to determine the sensitivity of the portfolio return to those simplifying assumptions by relaxing the assumptions.

Portfolio risk analysis is typically based on the assumption that the securities in the portfolio are well diversified. Portfolios that contain securities with several correlated risk factors do not meet the well-diversified criteria. Some portfolios by construction contain a predominant factor. Examples include sector or regional mutual funds. Other portfolios may be constrained in their ability to diversify. Examples include financial institutions' loan portfolios or an individual's personal portfolio. Using a Bayesian network model, we can examine the effect of risk concentration on portfolio risk. From the posterior return distribution, we calculate the portfolio variance and compare it to a non-diversifiable risk measure. Using this approach, we explore the dependence and independence assumptions used in traditional portfolio models.

Traditional financial analysis focuses on summary statistics — expected returns, beta, variance or standard deviation of returns. Recently value-at-risk analysis has emphasized consideration of the whole distribution of returns, or at least, the left-hand tail of the distribution. Since the output of the Bayesian network model is a posterior portfolio return distribution, we can also calculate the cutoff return for a value-at-risk calculation. As information is added to the network, the return distribution in the network reflects those changes and the cutoff value-at-risk is updated as well.

The rest of the paper proceeds as follows. In section two, we briefly discuss some traditional portfolio return models. In section three, we define and describe the semantics of a Bayesian network. In section four, we model a simple gold stock portfolio using a Bayesian network. Finally, in section five we discuss some modeling issues and limitations of Bayesian networks.

1.1. Traditional Portfolio Return Models

In a traditional portfolio analysis, the hypothesized relationship assumes that the rate of return on an asset is a linear function of the market rate of return and an asset specific factor, as follows:

$$R_i = a_i + b_i\, R_M + E_i \tag{1}$$

where R_i denotes return on asset i, a_i and b_i are constants, R_M denotes return on a market index, and E_i denotes an uncertain variable related to asset-specific factors.

Many studies have shown that some specific identifiable components of risk are not fully accounted for by just a market index. These studies have found that industry-specific risk, country-specific risk, and many other components account for correlation among individual securities. King (1966) finds that market factors explain 30% of return variation and industry factors explain an additional 10%. Goodman (1981) shows that in country-specific diversified portfolios, significant mis-measurement of risk occurs if the market proxy does not include global factors.

The arbitrage pricing theory (APT) (Ross 1976) and other multi-factor models (see Elton and Gruber 1997) extends the single factor model to account for these additional, identifiable factors. The multi-factor model can be represented as an expanded version of equation (1.1):

$$R_i = a_i + b_{i1}\, F_1 + b_{i2}\, F_2 + b_{ik}\, F_k + E_i \tag{2}$$

where $F_1, \ldots, F_k$ denote the k independent factors, and $b_{i1}, \ldots, b_{ik}$ are constants.

Portfolio return, denoted by R_p, is defined as the weighted average of the individual returns that comprise the portfolio, $R_p = \sum_{i=1}^{n} w_i R_i$. Where w_i denotes the proportional amount invested in security I, and n denotes the number of securities in the portfolio. Portfolio variance, denoted by σ_p, is given by:

$$\sigma_P^2 = b_{1P}^2\sigma_{F_1}^2 + \ldots + b_{kP}^2\sigma_{F_1}^2 + w_1^2\sigma_{E_1}^2 + \ldots + w_1^2\sigma_{E_n}^2 + \sum_{j=1}^{n}\sum_{i=1}^{n} w_i w_j \operatorname{cov}\left(E_i, E_j\right) \tag{3}$$

where b_{kp} denotes the portfolio beta for the kth independent factor, $\sigma_{F_k}^2$ denotes the variance of the kth independent factor, and $\sigma_{E_i}^2$ denotes the residual asset-specific variance.

It is assumed that all asset-specific uncertain variables, $E_1, \ldots, E_k$, are mutually independent. So portfolio variance simplifies to:

$$\sigma_P^2 = b_{1P}^2\sigma_{F_1}^2 + \ldots + b_{kP}^2\sigma_{Fk}^2 + w_1^2\sigma_{E_1}^2 + \ldots + w_n^2\sigma_{E_n}^2 \tag{4}$$

Portfolio risk is divided into two components — diversifiable risk, $w_1^2\sigma_{E_1}^2 + \ldots + w_n^2\sigma_{E_n}^2$, and non-diversifiable risk, $b_{1P}^2\sigma_{F_1}^2 + \ldots + b_{kP}^2\sigma_{Fk}^2$. It is normally assumed that diversifiable risk is small since each w_i^2 is small. However, in study of bank loan portfolios, Chirinko and Guill (1990) find that assuming the covariance terms are zero leads to portfolio variances being under-estimated from 24.6% to 45.75%. For an equally weighted loan portfolio with 46 industries, the variance was underestimated by 36.36%.

1.2. Bayesian Networks

Bayesian networks have their roots in attempts to represent expert knowledge in domains where expert knowledge is uncertain, ambiguous, and/or incomplete. Bayesian networks are based on probability theory.

A Bayesian network model is represented at two levels, qualitative and quantitative. At the qualitative level, we have a directed acyclic graph in which nodes represent variables and directed arcs describe the conditional independence relations embedded in the model. Figure 1 shows a Bayesian network consisting of four discrete variables: Interest Rate (IR), Stock Market (SM), Oil Industry (OI), and Oil Company Stock Price (SP). At the quantitative level, we specify conditional probability distributions for each variable in the network. Each variable has a set of possible values called its *state space* that consists of mutually exclusive and exhaustive values of the variable. In Figure 1, e.g., Interest Rate has two states: ‘high’ and ‘low;’ Market has two states: ‘good’ and ‘bad;’ Oil Industry has two states: ‘good’ and ‘bad;’ and Oil Company Stock Price has two states: ‘high’ and ‘low.’ If there is an arc pointing from X to Y, we say X is a parent of Y. For each variable, we need to specify a table of conditional probability distributions, one for each configuration of states of its parents. Figure 1 shows these tables of conditional distributions—P(IR), P(SM | IR), P(OI), and P(SP | SM, OI).

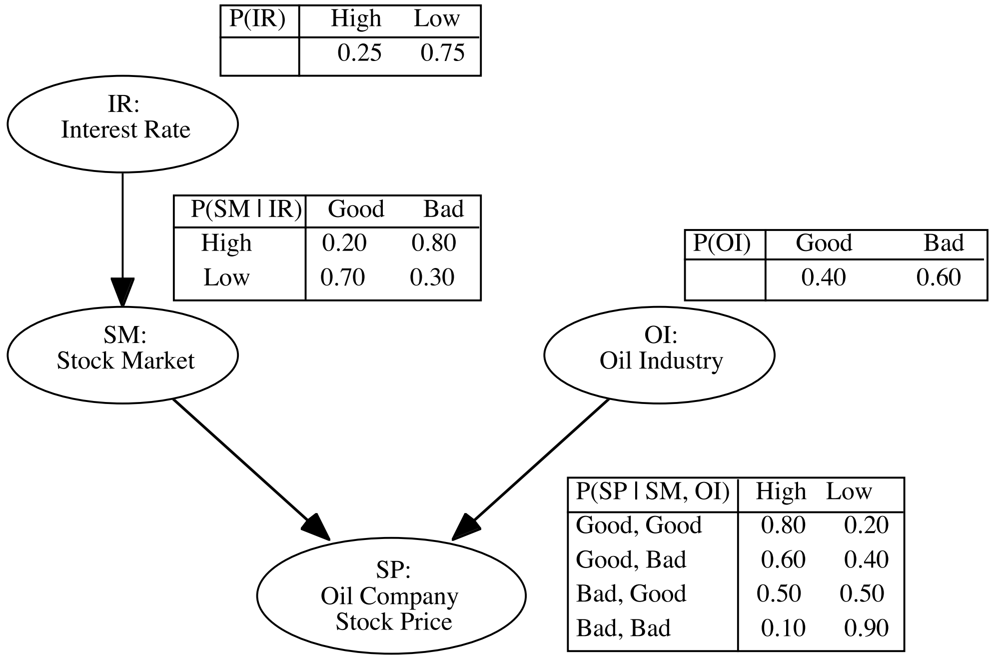
1.2.1. Semantics of Bayesian Networks

A fundamental assumption of a Bayesian network is that when we multiply the conditionals for each variable, we get the joint probability distribution for all variables in the network. In Figure 1, e.g., we are assuming that $P(IR, SM, OI, SP) = P(IR) \otimes P(SM \mid IR) \otimes P(OI) \otimes P(SP \mid SM, OI)$, where $\otimes$ denotes pointwise multiplication of tables. The rule of total probability tells us that $P(IR, SM, OI, SP) = P(IR) \otimes P(SM \mid IR) \otimes P(OI \mid IR, SM) \otimes P(SP \mid IR, SM, OI)$.

Comparing the two, we notice that we are making the following assumptions: $P(OI \mid IR, SM) = P(OI)$, i.e., OI is independent of IR and SM; and $P(SP \mid IR, SM, OI) = P(SP \mid SM, OI)$, i.e., SP is conditionally independent of IR given SM and OI.

Notice that we can read these conditional independence assumptions directly from the graphical structure of the Bayesian network as follows. Suppose we pick a sequence of the variables in a Bayesian network such that for all directed arcs in the network, the variable at the tail of each arc precedes the variable at the head of the arc in the sequence. Since the directed graph is acyclic, there always exists one such sequence. In Figure 2 one such sequence is IR SM OI SP. The conditional independence assumptions in a Bayesian network can be stated as follows. For each variable in the sequence, we assume that it is conditionally independent of its predecessors in the sequence given its parents. The key point here is that missing arcs (from a node to its successors in the sequence) signify conditional independence assumptions. Thus the lack of an arc from IR to OI signifies that OI is independent of IR; the lack of an arc from SM to OI signifies that OI is independent of SM; and the lack of an arc from IR to SP signifies that SP is conditionally independent of IR given SM and OI.

Figure 1. A Bayesian Network with Conditional Probability Tables



P(IR)	High	Low
	0.25	0.75

P(SM | IR)	Good	Bad
High	0.20	0.80
Low	0.70	0.30

P(OI)	Good	Bad
	0.40	0.60

P(SP | SM, OI)	High	Low
Good, Good	0.80	0.20
Good, Bad	0.60	0.40
Bad, Good	0.50	0.50
Bad, Bad	0.10	0.90

In general, there may be several sequences consistent with the arcs in a Bayesian network. In such cases, the lists of conditional independence assumptions (associated with each sequence) are equivalent using the laws of conditional independence (Dawid 1979). There are other equivalent graphical methods for identifying conditional independence assumptions embedded in a Bayesian network graph (see Pearl (1988) and Lauritzen *et al*. (1990) for examples.).

1.1.2. Making Inferences in Bayesian Networks

Once a Bayesian network is constructed, it can be used to make inferences about the variables in the model. The conditionals given in Bayesian network representation specify the *prior* joint distribution of the variables. If we observe (or learn about) the values of some variables, then such observations can be represented by tables where we assign 1 for the observed values and 0 for the unobserved values. Then the product of all tables (conditionals and observations) gives the (unnormalized) *posterior* joint distribution of the variables. Thus the joint distribution of variables changes each time we learn new information about the variables.

In theory, the posterior marginal probability of a variable X, say P(X), can be computed from the joint probability by summing out all other variables except X one by one. In practice, such a naive approach is not computationally tractable when we have a large number of variables because the joint distribution has an exponential number of states and values. The key to efficient inference lies in the concept of *local computation* where we compute the marginal of the joint without actually computing the joint distribution. A key feature of a Bayesian network is that it describes a joint distribution from the local relationships—such as a node and its parents. Instead of tackling the whole collection of variables simultaneously, Bayesian networks use the concept of factorization. Factorization involves breaking down the joint probability distributions into subgroups called factors in such a way that the naive computations described above need only be performed within each subgroup. Since the state space of a subgroup is much smaller than that of the joint probability distribution, the calculations become manageable.

Bayesian networks can be used for two types of inference.[1] Often we are interested in the values of some target variables. In this case, we make inferences by computing the marginal of the posterior joint distribution for the variables of interest. Consider the situation described by the Bayesian network in Figure 1. Suppose we are interested in the true state of Oil Company Stock Price (SP). Given the prior model (as per the probability tables shown in Figure 1), the marginal distribution of SP is 0.502 for high and 0.498 for low. Now suppose we learn that Interest Rate is low. The posterior marginal distribution of SP changes to 0.554 for high and 0.446 for low. Suppose we further learn that the state of Oil Industry is good. Then the marginal distribution of SP changes to .71 for high and 0.29 for low. This type of inference is referred to as '*sum propagation*.'

Sometimes we are more interested in the configuration of all variables ("the big picture") rather than the values of individual variables. In this case, we can make inferences by computing the mode of the posterior joint distribution, i.e., a configuration of variables that has the maximum probability. Consider again the situation described by the Bayesian network in Figure 1. Given the prior model (as per the probability tables shown in Figure 1), the mode of the prior joint distribution is (low interest rate, good stock market, bad oil industry, low oil company stock price). Now suppose we learn that Interest Rate is high. The mode of the posterior joint distribution changes to (high interest rate, bad stock market, bad oil industry, high oil company stock price). This type of inference is referred to as '*max propagation*.'

The results of inference are more sensitive to the qualitative structure of the Bayesian network than the numerical probabilities (Darwiche and Goldszmidt 1994). For decision making, the inference results are robust with respect to the numerical probabilities (Henrion *et al*. 1994).

There are several commercial software tools such as Hugin (www.hugin.com) and Netica (www.norsys.com) that automate the process of inference. These tools allow the user to enter the Bayesian network structure graphically, enter the numerical details, enter any additional information, and then do inference of either type. The results of the inference are then shown graphically using bar charts.

1.3. A Bayesian Network Model of Multi-Factor Portfolio Return

In this section, we first describe a traditional security return model as a Bayesian network. Then we demonstrate how some of the independence assumptions in a traditional security return model can be relaxed using Bayesian networks.

1.3.1. Description of a Gold Stock Portfolio Network

A security return model is a conditional expectations model and is usually estimated using least squares regression, so that

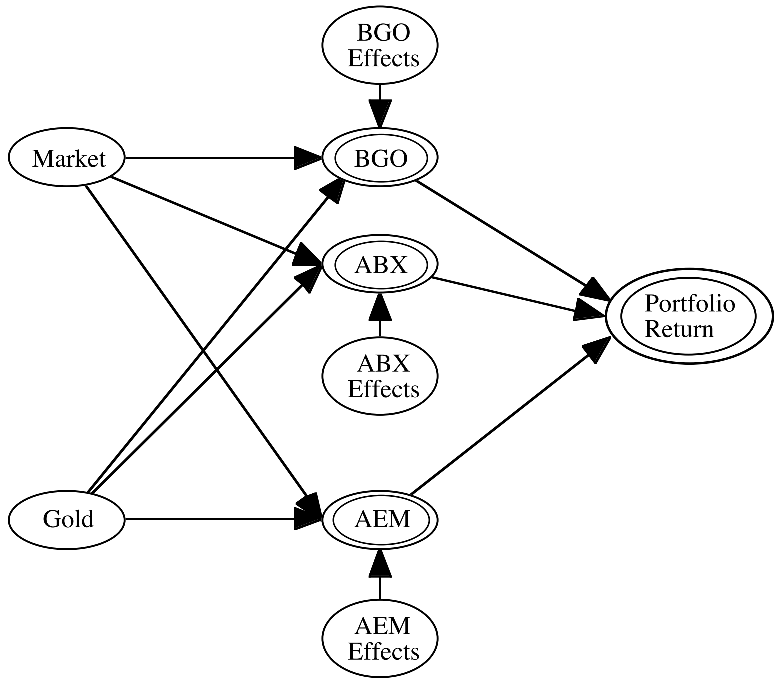
$$E\left(R_i \mid F_i, F_2, \ldots F_k\right) = a_i + b_{i1}F_1 + \ldots + b_{ik}F_k \tag{5}$$

We can easily regard this as a Bayesian network model where the factors F_1, …, F_k are regarded as mutually independent variables.

The following Bayesian network model considers the return on an equally weighted portfolio of three stocks from the gold mining industry. Their ticker symbols are BGO, ABX, and AEM. In the representation shown in Figure 2, market and gold returns are parents of each stock's return. In addition, each stock return has an idiosyncratic component. The portfolio return is a function of each stock return and the weight of each stock in the portfolio. In the graph, each random variable is shown as a node. A variable that is conditionally deterministic given the values of its parents is shown as a double-bordered node.

[1] Lauritzen and Speigelhalter (1988), Jensen *et al*. (1990) and Shenoy and Shafer (1990) have devised propagation algorithms to perform efficient probabilistic inference.

Figure 2. A Bayesian Network Model for a Portfolio



The graphical model is supplemented by numeric information about the conditional probability distributions. For each variable in the model, we define the conditional probability distributions given each combination of states in the parent nodes. In Figure 1, market, gold, BGO effects (e_{BGO}), ABX effects (e_{ABX}), and AEM effects (e_{AEM}) have no parents, so we specify a prior distribution for these nodes. BGO has market, gold and BGO effects as its parents. Since BGO is a conditionally deterministic node, we specify a functional relation for its value as a function of the states of its parents. We also define a similar relation for ABX and AEM. Finally, Portfolio Return has BGO, ABX, and AEM. Since Portfolio Return is also a conditionally deterministic variable, it has a unique deterministic state given by some functional relation.

1.3.2. Inputs for the Bayesian Network

The conditional relationship for each of the stocks can be specified as an equation, such as the multi-factor model specified in equation (1.2); as a discrete conditional probability table; or as a continuous conditional probability distribution. Any combination of historical data, forecasts, expert knowledge, or beliefs can be used to estimate the conditional relationships.

Initially for the primary inputs market, gold, and the individual stock effects we do not specify an explicit conditional relationship. We specify the a priori distribution for each of these as a normal distribution with parameters estimated over an arbitrary period from January 1996 through February 1998. We estimate weekly returns for these inputs. For example, using historical data over the estimation period, weekly market return has mean 0.55% and standard deviation 2.28%. These inputs are summarized in Table 1 below. For each stock effect, we assume a mean of zero and a standard deviation equal to the standard error of the regression equation.

For each of the stock returns, the distribution is conditioned on the market and gold returns and a stock-specific effect. The conditional distribution for each stock node is normal with mean based on equation (1.2) with factors of market and gold returns. The mean stock return is the estimated regression for each stock node. The standard error of the regression is the standard deviation of the distribution.

The regression estimates are based on weekly returns from January 1996 through February 1998. The result of the equation using the state values of the inputs determines the stock return node. For ABX the conditional relationship is the estimated regression equation:

$$\text{ABX return} = 0.17 + 0.366{*}\text{market} + 2.26{*}\text{gold} + e_{ABX}, \quad (6)$$

A summary of all of the inputs and regression coefficients used for each of the nodes is presented in Table 1.

Table 1. Parameter and Regression Estimates Used in Portfolio Network

	BGO	ABX	AEM	Market	Gold
Description	Bema Gold Corp.	Barrick Gold Corp.	Agnico Eagle Mines	S&P 500 Index	London PM Gold Fix
Average monthly return (%)	0.68	–0.11	–0.42	0.55	–0.23
Standard deviation (%)	12.55	5.00	6.58	2.28	1.43
Regression estimates:					
Intercept	1.48	0.17	0.05		
Market coefficient	0.26	0.37	0.27		
Gold coefficient	3.96	2.26	2.76		
standard error	11.27	3.63	5.27		

In order to generate the conditional return distributions for BGO, ABX, AEM, and the portfolio, we use Monte Carlo simulation. In the simulation, we specify the functional relationships between the nodes to generate the estimated return distributions. The portfolio return distribution is a simple average of the stock returns; that is, an equally weighted portfolio.

Table 2 reports additional statistics for each of the conditional probability return distributions based on the simulation results. The average weekly portfolio return is 0.06% with a standard deviation of 6.15%.

Table 2. Conditional Probability Distributions for Bayesian Network

	Market	Gold	BGO	ABX	AEM	Portfolio
Mean	0.55	–0.23	0.72	–0.43	–0.15	0.06
standard deviation	2.28	1.43	12.73	6.60	4.95	6.15
Minimum	–6.94	–5.10	–48.86	–24.84	–17.46	–19.39
Maximum	8.69	4.48	49.73	20.15	18.64	18.52
5th percentile	–3.19	–2.57	–20.00	–11.34	–8.09	–10.32
50th percentile	0.55	–0.23	0.82	–0.54	–0.26	0.18
95th percentile	4.30	2.12	20.61	10.13	8.15	9.85

Figure 3. Portfolio Return Statistics from Bayesian Network Model

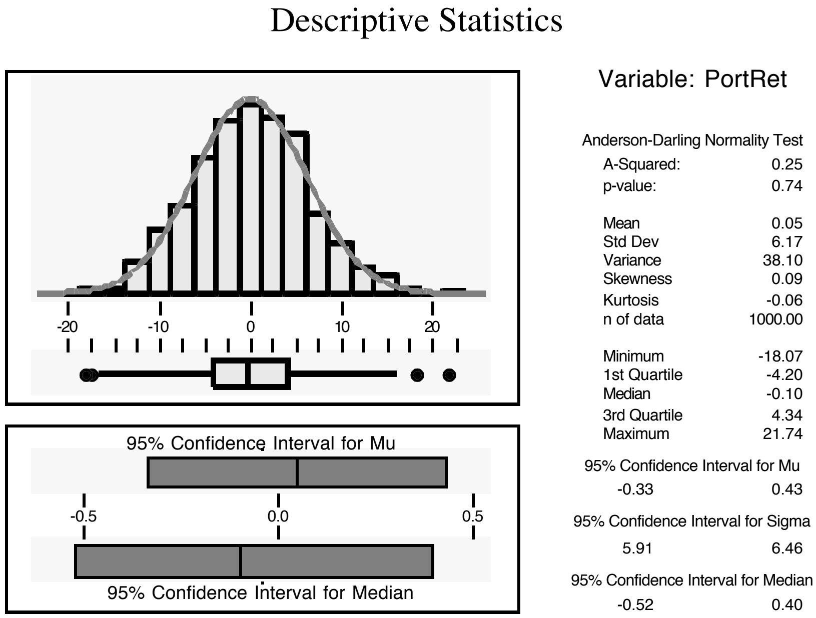


Figure 3 shows the simulated distribution of portfolio returns from the Bayesian network based on 10,000 iterations. Because we modeled the inputs to the network as normal distributions, the portfolio return distribution is also approximately normal. From the confidence interval for the mean and the median, we see that the mean return on this portfolio is not significantly different from zero. In Section 3.4 we compare this model and several other Bayesian network models to the actual portfolio return.

1.3.3. Bayesian Network as a Management Tool

A principal advantage of a Bayesian network representation of portfolio risk is in its flexibility as a management tool. In this section we show how new evidence can be entered into a network and how new information can be added to a network. Studies (see Henrion *et al*. (1994, 1996), and Pradhan *et al*. (1996)) have shown that the graphical representation of the conditional probabilities is the most important step in modeling. The exact numerical form is of secondary importance. Most decisions will be robust as long as the conditional independence relationships as encoded in the network are specified correctly. Managers usually have a good idea of the influences on a portfolio, but not their exact numerical form.

New information can be easily incorporated in the model. For example, suppose we learn that BGO will perform well in the next period if a new product is launched, but BGO will remain flat if the product is not launched on time. We also believe that it is very likely that the product launch will be on time. Specifically we model the evidence as a table of likelihoods as shown in Table 3 below where a return of 10 is four times more likely than a return of 0.

Table 3. Evidence for BGO

State for BGO	Likelihoods
Return of 0	0.20
Return of 10	0.80

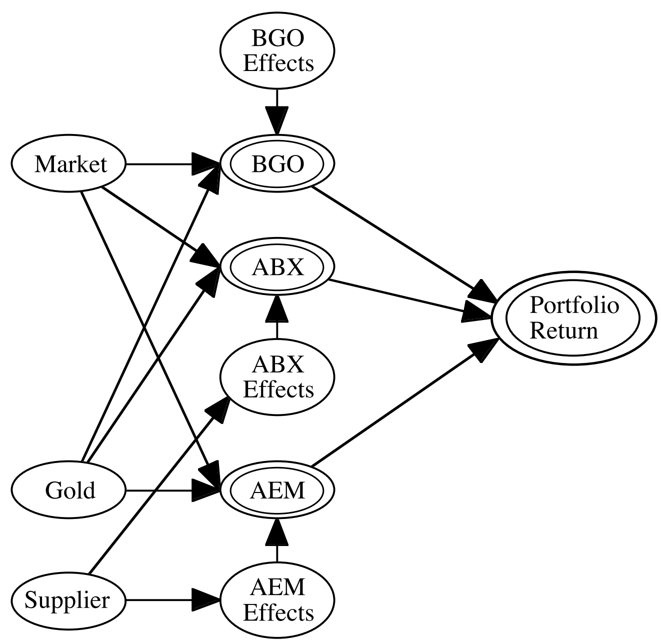
We incorporate this new evidence in the model, and recompute the marginals of the posterior distributions (as shown in Table 4) to reflect the new information.

Table 4. Revised Conditional Return Distributions

	BGO	Portfolio Return
Mean	4.30	1.25
standard deviation	4.95	3.69
Minimum	0.00	–12.06
Maximum	10.00	12.66
5th percentile	0.00	–4.73
50th percentile	0.00	1.09
95th percentile	10.00	7.27

A Bayesian network can also accommodate some types of information that is not easily incorporated in other types of models. Suppose an analyst learns that AEM and ABX have a common supplier whose favorable actions will affect both AEM and ABX. This new information can be added to the network using subjective probabilities. Figure 4 shows the addition of a new node that directly affects the nodes ABX effects (e_{ABX}) and AEM effects (e_{AEM}).

Figure 4. Revised Bayesian Network with Supplier Information



1.3.4. Additional Conditions in the Portfolio Return Model

The model we specify in Figure 2 is based on the traditional finance model that assumes the residual correlations are independent. We use the original return data to calculate the regression residuals and find the residual correlations among the stocks. The residual correlation is reported in Table 5. We find relatively large residual correlation between e_{ABX} and e_{AEM} and between e_{BGO} and e_{AEM}. We use this correlation data to specify three additional Bayesian network models that take into account these dependencies.

Table 5. Residual Correlations

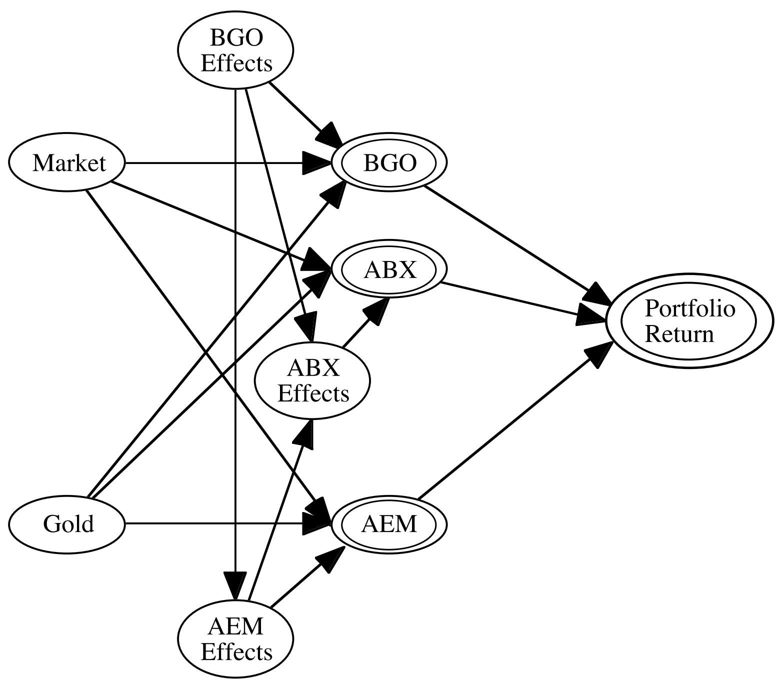
	BGO Effects	ABX Effects	AEM Effects
BGO Effects	1		
ABX Effects	0.1814	1	
AEM Effects	0.3367	0.4427	1

This high correlation indicates that there are still unmodeled factors which affect these stocks. If these factors are unknown, it is still possible to model the dependencies among the residuals. For any ordering of e_{BGO}, e_{AEM}, and e_{ABX}, using the multiplication rule, we have

$$f(e_{BGO}, e_{AEM}, e_{ABX}) = f(e_{BGO})\, f(e_{AEM} \mid e_{BGO})\, f(e_{ABX} \mid e_{BGO}, e_{AEM}). \tag{7}$$

Any other ordering of the residuals are equivalent to this ordering. This fact implies that a Bayesian network that captures all residual correlation can be specified with the following representation:

Figure 5. A Bayesian Network Model with Correlated Residuals



If X and Y are bivariate normal, the conditional distribution f(y | x) is given as:

$$f(y \mid x) \sim N\left(\frac{\rho\sigma_y}{\sigma_x}(x-\mu_x), (1-\rho^2)\sigma_y^2\right). \tag{8}$$

So if we assume that e_{BGO}, e_{ABX}, and e_{AEM} are multivariate normal, we specify three additional models with the following independence conditions: [2]

1. Model 2: $e_{BGO} \rightarrow e_{AEM}$, i.e., e_{AEM} and e_{BGO} are independent of e_{ABX};
2. Model 3: $e_{AEM} \rightarrow e_{ABX}$, i.e, e_{AEM} and e_{ABX} are independent of e_{BGO}; and
3. Model 4: $e_{BGO} \rightarrow e_{AEM} \rightarrow e_{ABX}$, i.e., e_{BGO} and e_{ABX} are conditionally independent given e_{AEM}.

To see the effect of adding these conditional probabilities we compare the actual returns over the three year period to each of the Bayesian network models and to the finance model. In Table 6, we find that each model slightly overestimates the mean and underestimates the standard deviation.

Table 6. Comparison of Actual Portfolio Return and Portfolio Return Models

Portfolio	Mean	Std(Risk)	% difference from actual	
			Mean	Std
Actual portfolio return	0.050	6.8		
Finance Model – Non-diversifiable risk	0.053	4.3	5.4	-35.9
1. Simple Bayes Net (BN)	0.049	6.2	-1.8	-8.8
2. BN with $e_{BGO} \rightarrow e_{AEM}$	0.053	6.6	5.1	-2.9
3. BN with $e_{AEM} \rightarrow e_{ABX}$	0.052	6.2	4.1	-8.8
4. BN with $e_{BGO} \rightarrow e_{AEM} \rightarrow e_{ABX}$	0.054	6.7	7.3	-0.4

Actual portfolio return is measured assuming weekly portfolio rebalancing. The mean of the actual portfolio return is the arithematic average of the weekly portfolio returns. For the finance model we measure non-diversifiable risk as $\sqrt{b_{p,m}^2 \sigma_m^2 + b_{p,g}^2 \sigma_g^2}$, where $b_{p,m}^2$ is the market portfolio beta squared, and $b_{p,g}^2$ is the gold portfolio beta squared. The portfolio beta is the weighted average of the appropriate gold or market return coefficients from the three regression equations. We see that modeling the effects with the largest residual correlations, $\rho_{BGO,AEM}$ and $\rho_{ABX,AEM}$ provides improvements in the risk and return estimation

Table 7 reports confidence intervals for the mean and standard deviation for the four Bayesian network models and the actual return. There are 115 weekly observations for the actual returns. The Bayesian network models are based on simulations of 10,000 iterations.

[2] Model 1 is the Bayesian network shown in Figure 2 where we are assuming that the three effects are mutually independent.

Table 7. 95% Confidence Intervals for μ and σ

	Conf. Interval for μ			Conf. Interval for σ		
	$\bar{x}$	Lower	Upper	s	Lower	Upper
Actual	0.050	–1.200	1.300	6.8	6.1	7.6
Model 1	0.049	–0.334	0.432	6.2	6.0	6.4
Model 2	0.053	–0.355	0.460	6.6	6.3	6.8
Model 3	0.052	–0.331	0.435	6.2	6.0	6.4
Model 4	0.054	–0.365	0.472	6.7	6.1	7.6

1.4. Modeling Issues and Limitations

Bayesian networks are able to incorporate different types of information. In this section we address the issues of how to find inputs to the network, discrete vs. continuous probability distributions, and some model limitations.

1.4.1. Inputs to the network

Two types of inputs are needed for the network. First, factors that affect each asset return in the portfolio have to be identified. Then the conditional probability distributions for the asset returns have to be specified. Both types of inputs can be any combination of empirical data, expectations, judgment, or forecasts.

Many empirical studies have attempted to identify the factors that cause variation in security returns. Roll and Ross (1980), Dhrymes *et al.* (1984), Chen *et al.* (1986), Elton and Gruber (1997) identify factors such changes in inflation, industrial production, and yield spreads. Other portfolio specific factors may also be important. For example, a geographically limited portfolio would have a factor relating regional economic conditions to the stock returns. Common production inputs, customers, and other special circumstances can also be included.

Empirical analysis tools such as linear regression, factor analysis, time series analysis, neural nets and data mining techniques can all be used to generate the conditional probabilities for the dependent nodes. These tools examine the historical relationship among the nodes. Using these analytical tools may be equivalent to using current financial models, if the independence assumptions are the same. For example, Model 1 is equivalent to a traditional multi-factor model because each specific stock effect is assumed to be mutually independent, and all inputs are based on historical data.

Judgment and forecasts can be added to the model by revising the conditional probability tables for the nodes, by revising the priors for nodes with no parents, or by adding new nodes. Studies (see Henrion *et al.* 1994, 1996), and Pradhan *et al.* (1996)) have shown that the graphical representation of the probability model is the most important step in modeling. The numerical details of the probability model are of secondary importance. Most decisions will be robust as long as the conditional independence relationships as encoded in the network are specified correctly. Managers usually have a good idea of the influences on a portfolio, but not their exact numerical form. Discrete conditional distributions can be used as approximations of the exact form of the distribution.

The Bayesian network representation forces the modeler to make explicit judgments of the causal structure of the model. Traditional statistical models have an implicit causal structure that is not always appropriate. The decision-maker can examine the effect of assuming independent residuals and other factors in the model. For example, a model may have a geographic factor and a market factor. In a multi-factor model it is usually assumed the factors are independent; however, the market and geographic factor may not be independent. In a Bayesian network the dependence between the two factors can also be modeled.

1.4.2. Value-at-Risk

The value at risk (VAR) for a portfolio is the expected maximum loss over a target horizon within a given confidence interval. Recent, large corporate losses in Orange County, Barings Bank, Daiwa and others have received media and regulatory attention. Academicians, regulators, and financial managers have asserted to the need for a better method of summarizing risk. The value at risk metric is one way to quantify portfolio risk. This measure is easily incorporated into a Bayesian network model. The VAR measure is a linear function of the portfolio return defined as:

$$VAR = W_0 R^*, \tag{9}$$

where W_0 is initial investment and R^* is the cutoff portfolio return for the ith percentile. Using Model 1 at a cutoff percentile of 5, R^* and is equal to –10.32 percent, and $VAR = W_0(-0.1032)$ for a period of one week. An important consideration for VAR measures is the time over which risk is measured.

1.1.3. Limitations of Bayesian Network Models

A major limitation in using Bayesian networks to model portfolio returns is determining the graphical structure of the Bayesian network model. A graphical structure can either be obtained subjectively from an expert or one can be induced from data. The latter technique is the subject of current research in the uncertain reasoning literature (see (Heckerman 1997) for a recent survey). Once a graphical structure is obtained, determining the numerical parameters of the model is straightforward when securities are publicly traded and when data is often readily available.

If all variables in a Bayesian network are discrete, then the marginal distribution of any variable can be computed exactly using the local computational algorithms proposed, e.g., by Pearl (1988), Lauritzen and Spiegelhalter (1988), and Shenoy and Shafer (1990). These algorithms are encoded in commercial software such as Netica (www.norsys.com) and Hugin (www.hugin.com). Since security returns are usually modeled as continuous variables, exact computation of the Bayesian network is not always possible. We can either discretize the distributions or use simulation methods. We can compute the posterior marginals approximately using Monte Carlo methods (see, e.g., Henrion 1988). As the number of variables grows, even Monte Carlo methods require an inordinate amount of sampling for a decent approximation. In such cases, Markov Chain Monte Carlo methods have been proposed for faster convergence (see, e.g., Gilks *et al*. 1996).

In our example, we assumed an equally weighted portfolio. If a manager is evaluating a currently held portfolio, the weights will change as the stock prices of the component returns change. Therefore assuming constant weights implies constant rebalancing of the portfolio (to maintain the constant weights) as the stock prices change. Of course, it is not possible to rebalance a portfolio without incurring transaction costs, so the actual return from this type of a portfolio would be lower. It is possible to construct a Bayesian network that calculates portfolio return based on share prices and constant number of shares held. However, such a model is quite different from a traditional finance model, and is the subject of future research.

1.5. Conclusions

The main goal of this paper is to propose Bayesian networks as a tool for modeling portfolio returns. Bayesian networks allow us to explicitly model the dependence between the various factors that affect portfolio return. Also, recent advances in the uncertain reasoning literature allow one to compute the marginal posterior distribution of the portfolio return even when we have a multivariate probability model with many variables. The marginal distribution of portfolio return can be dynamically updated (using Bayes rule) as we observe the values of some of the variables.

References

Chen, N., R. Roll and S. A. Ross (1986), "Economic Forces and the Stock Market," *Journal of Business*, **59**, 386–403.

Chirinko, R. S. and G. D. Guill (1991), "A framework for assessing credit risk in depository institutions: Toward regulatory reform," *Journal of Banking and Finance*, **15**, 785–804.

Darwiche, A. and M. Goldszmidt (1994), "On the relation between kappa calculus and probabilistic reasoning," in R. L. Mantaras and D. Poole (eds.), *Uncertainty in Artificial Intelligence: Proceedings of the Tenth Conference*, 145–153, Morgan Kaufmann, San Francisco, CA.

Dawid, A. P. (1979), "Conditional independence in statistical theory (with discussion)," *Journal of the Royal Statistical Society, Series B*, **41**(1), 1–31.

Dhrymes, P. J., I. Friend, and N. B. Gultekin (1984), "A Critical Reexamination of the Empirical Evidence on the Arbitrage Pricing Theory," *Journal of Finance*, **39**(2), 323–346.

Elton, E. J., M. J. Gruber and C. R. Blake (1997), "Common factors in mutual fund returns," Working Paper S-97-42, The NYU Salomon Center for Research in Financial Institutions and Markets, New York University, NY.

Elton, E. J., and M. J. Gruber (1997), *Modern Portfolio Theory and Investment Analysis*, 4th Ed., John Wiley & Sons, New York.

Gilks, W. R., S. Richardson, and D. J. Spiegelhalter (1996), *Markov Chain Monte Carlo in Practice*, Chapman & Hall, London.

Goodman, L. S. (1981), "Bank Lending to Non-OPEC LDCs: Are Risks Diversifiable?," *FRBNY Quarterly Review*, 10–20.

Heckerman, D. (1997), "Bayesian networks for data mining," *Data Mining and Knowledge Discovery*, **1**, 79–119.

Henrion, M. (1988), "Propagating uncertainty in Bayesian networks by probabilistic logic sampling," in J. F. Lemmer and L. Kanal, N. (eds.), *Uncertainty in Artificial Intelligence*, **2**, 149–164, North-Holland, Amsterdam.

Henrion, M., A. Darwiche, M. Goldszmidt, G. Provan, and B. Del Favero (1994), "An experimental comparison of infinitesimal and numerical probabilities for diagnostic reasoning," *Proceedings of the Fifth Workshop on the Principles of Diagnosis*, 131–139.

Henrion, M., M. Pradhan, B. Del Favero, K. Huang, G. Provan and P. O'Rorke (1996), "Why is diagnosis using belief networks insensitive to imprecision in probabilities," in E. Horvitz and F. Jensen (eds.), *Uncertainty in Artificial Intelligence: Proceedings of the Twelfth Conference*, 307–314, Morgan Kaufmann, San Francisco.

Jensen, F. V., S. L. Lauritzen and K. G. Olesen (1990), "Bayesian updating in causal probabilistic networks by local computation," *Computational Statistics Quarterly*, **4**, 269–282.

Jorion, P. (1997), *Value at Risk: The New Benchmark for Controlling Market Risk.*, McGraw Hill, NY.

Kao, D. (1993), "Illiquid securities: Pricing and performance measurement," *Financial Analysts Journal*, 28–35.

Keefer, D. L. (1994), "Certainty equivalents for three-point discrete-distribution approximations," *Management Science*, **40**, 760–772.

Lauritzen, S. L. and D. J. Spiegelhalter (1988), “Local computations with probabilities on graphical structures and their application to expert systems (with discussion),” *Journal of Royal Statistical Society, Series B*, **50**(2), 157–224.

Lauritzen, S. L., A. P. Dawid, B. N. Larsen and H.-G. Leimer (1990), “Independence properties of directed Markov fields,” *Networks*, **20**(5), 491–505.

Pearl, J. (1986), “Fusion, propagation and structuring in belief networks,” *Artificial Intelligence*, **29**, 241–288.

Pearl, J. (1988), *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann, San Mateo, CA.

Pradhan, M., M. Henrion, G. Provan, B. Del Favero, and K. Huang. (1996), The sensitivity of belief networks to imprecise probabilities: An experimental investigation, *Artificial Intelligence*, **85**(1–2), 363–397.

Roll, R., and Ross, S. A. (1980), “An Empirical Investigation of the Arbitrage Pricing Theory,” *Journal of Finance*, **35**(5), 1073-1103.

Ross, S. A. (1976), “The Arbitrage Theory of Capital Asset Pricing,” *Journal of economic Theory*, **13**, 341– 360.

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.

Smith, J. E. (1993), “Moment methods for decision analysis,” *Management Science*, **39**, 340–358.