Futures Markets, Bayesian Forecasting and Risk Modeling

José Mario Quintana*, Carlos M. Carvalho†, James Scott‡ and Thomas Costigliola*

February 10, 2009

Chapter ? of

The Handbook of Applied Bayesian Analysis

Eds: Tony O'Hagan & Mike West
Oxford University Press

Abstract

Applications of the Bayesian approach to risk modeling regarding speculating trading strategies in Futures Markets is discussed in the context of the corresponding concepts of betting and investing, prices and expectations, and coherence and arbitrage-free pricing in the fields of Bayesian methodology and Finance.

*BEST, LLC, Hoboken, NJ 07030
†Booth School of Business, The University of Chicago
‡Department of Statistical Science, Duke University

1 Introduction

On the one hand, the media often refers (presumably in a pejorative sense) to certain trading activities, particularly those involving financial derivative instruments, as gambling. On the other hand, the origin of probability theory is traced back to a betting puzzle, on how to redistribute fairly the stakes of an interrupted game of chance, posed by the gambler Antoine Gombaud (a.k.a., Chevalier de Méré) to Blaise Pascal. Thus, it should not be surprising that a very strong bond exists between the fields of Subjective Probability and Derivative Finance. In this chapter, we explore overlapping concepts in these fields while focusing on the application of dynamic risk modeling *à propos* the hedging activity from the speculative perspective of a hedge fund manager.

Subjective expectations are motivated as fair prices of futures contracts in Section 2. The futures markets are presented, in Section 3, as a Bayesian market maker engine that dynamically reveals rational (coherent and proficient) expectations of random quantities as prices of futures contracts. A portfolio mean-variance efficiency generalization is motivated, in Section 4, as a sensible quantitative trading strategy for a hedge fund manager adopting the role of a Bayesian speculator (as opposed to the role of a Bayesian market maker) to highlight the critical role of hedging to ensue attractive risk-adjusted performance. Finally, general Bayesian dynamic models and specific Bayesian dynamic linear models are presented, in Section 5, to entertain a method, in Section 6, for assessing risk models in terms of their hedging effectiveness in the context of the risk-adjusted performance of trading strategies.

2 Subjective Expectations

Most children know how to make a sibling cut two pieces of a pie fairly: she would let her brother do it as long as he agrees that she can choose her piece first. To elicit from someone a current fair price of an item with a future settlement (payment and delivery), one can proceed in a similar way: the person is free to name any price as long as it has been agreed that one can decide afterwards how many items to buy from, or sell to, the individual. The transaction will happen in the future but at a specific place, date and price agreed at the present time (e.g., a gallon of regular gasoline in Hoboken, New Jersey on the last business day of the next month at a specified price). The individual should avoid naming a price that is too low (or too high) just in case he or she is forced to sell (or to buy) the item. Furthermore, a smart (money-seeking, risk-averse) person should name the expected future spot price of the item on the settlement date as the fair price because this value minimizes the maximum, and potentially huge, expected loss (assuming that the individual will buy, or sell, the items at the future spot price to accomplish the settlement).

The above betting scheme corresponds to de Finetti's (1931, 1974) operational foundation of Bayesian subjective probability and expectations in which the individual is forced to play the role of a liquid market maker. Interestingly, recent informal experiments performed by the first author confirm that personal fair prices (with future settlement dates), acquired by simultaneous elicitation from individuals familiar with the concept of coherence (the avoidance of becoming a sure loser in this setup), diverge. This is just a confirmation of the obvious: subjective personal fair prices (i.e., expectations) are subjective and personal; yet, this implies that a group of individually coherent people, operating independently, typically would act incoherently as an entity. It is even unfair to force a single person to act as a liquid market maker and one should never really quote fair prices because one would be potentially vulnerable unnecessarily. Only an immeasurably rich irrational individual or the ultimate genius would dare to do so for an extended period of time because, in the words of Barnard (1980), "fanatical insistence on freedom from 'incoherence' can lead to such complicatedly interrelated analyses of data as to go well beyond the capacity of our understanding." Yet, to a great extent, fair prices (with future settlement dates) of many standardized commodities began to be quoted almost three centuries ago, and they continue to be quoted in the present futures markets. This suggests the following questions: Are the futures markets coherent? Are their expectations worthwhile? Are they beatable? Furthermore, if they are, to what degree are they beatable?

3 Futures Markets

A futures contract represents the obligation to deliver a standardized commodity at a specified future maturity date and at a prearranged location and price. From an operational standpoint perfect liquid futures markets, as entities, are quoting fair prices (i.e., they are prepared to buy or sell at the quoted prices unlimited quantities of standardized commodities at future settlement dates). Perfect liquid markets, of course, do not exist; but, futures markets trade daily a variety of commodities (e.g., gallons of gasoline, S&P 500 Index units; Japanese Government Bonds, pounds of rice, shares of Google, barrels of WTI crude oil, euros vs. dollars, troy oz. of gold, etc.) with a total worth of trillions of dollars per day. Figure 3.1 shows, in random order, a snapshot of bid and ask prices of the US Treasury Bill Futures Contract and their corresponding stakes. It is apparent that the bid and ask prices are virtually indistinguishable. In addition, a $25,000,000 transaction would not have affected the bid and ask price quotes. Indeed, the spread between bid and ask prices of liquid futures markets are typically measured by a few basis points (a basis point is one percent of one percent) and the associated stakes are worth millions of dollars.

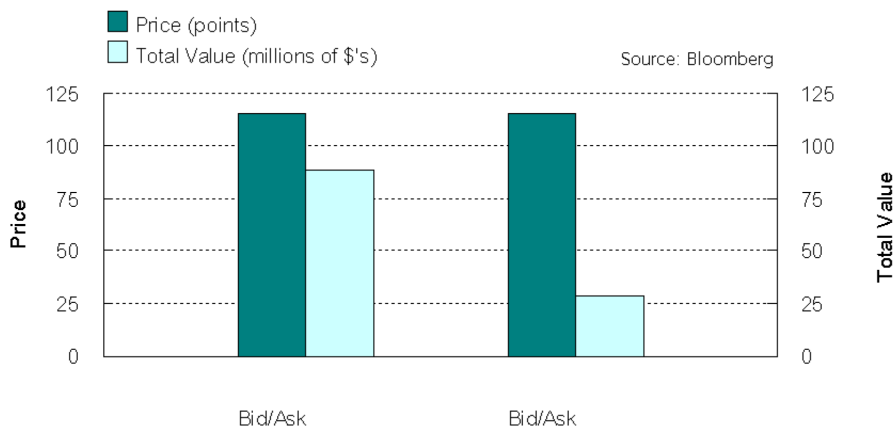

Figure 3.1: A Typical Liquid Market Snapshot

Are these liquid futures markets coherent over time? The straightforward answer is yes, and a simple financial economics explanation resembles the Anthropic Principle arguments: the fact that the question can be posed implies its answer (if the futures markets were incoherent they would have been an ideal fountain of wealth but only for a short while; afterwards they would have dried out and disappeared). More plausibly, as soon as traders would identify and exploit incoherencies in the markets, the trading activity would eventually restore coherent prices, since perfect liquid markets do not exist (prices are affected by supply and demand). Remarkably, although the futures markets participants might not be individually coherent, as an entity, they are acting in concert coherently by sharing between all of them two crucial pieces of information: the bid and ask prices. The ending value of a futures contract (i.e., the future spot price) can be tied to a general commodity (e.g., gallons of gasoline) but it also can be associated to an artificial item (e.g., S&P 500 index) and in that case is settled in cash. Further, in principle, the future spot price could be linked to any random variable yet to be revealed. For example, a well defined global average temperature measurement for the year 2025 could define a futures market where the global warmers and global coolers could settle their differences (actually, contracts of this kind but with near settlements are traded in the form of binary bets, which are described below).

A futures contract where the future spot price can only take two values and one (and only one) potential value is zero (a.k.a., a binary bet) is of particular interest. Thus, a binary bet on an arbitrary event, the condition that defines a non zero value (e.g., the S&P500 Index will be up at the end of the day, Barack Obama will be the 2008 U.S. presidential election winner, etc.), is a futures contract that pays a specified amount (e.g., $100) if the event occurs and

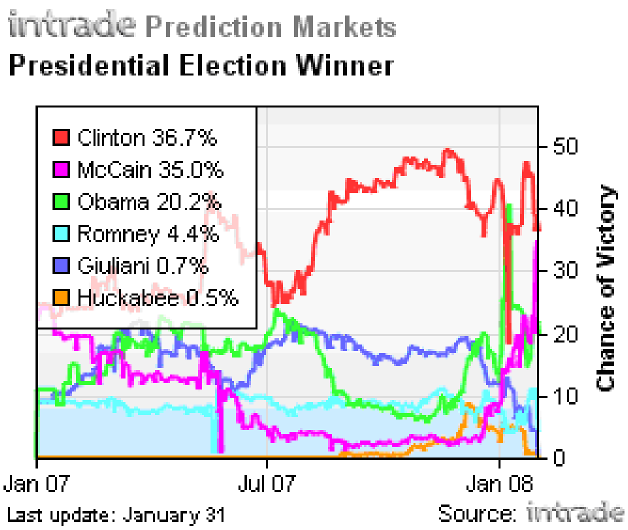

Figure 3.2: Market Implied Probabilities

zero otherwise. The concept of a binary bet, from a practical perspective, was introduced in the U.K. early in this millennium; however, from a theoretical perspective it can be traced back to de Finetti (1931), Ramsey (1926) and arguably to Bayes (1763). Indeed, his definition "The *probability of any event* is the ratio between the value at which an expectation depending on the happening of the event ought to be computed, and the chance of the thing expected upon it's happening" written in centuries old English might be hard to comprehend at first sight but it is easily illustrated by example. Bayes' insight behind his definition of probability of an event becomes apparent if we try to compute the expected payoff of a binary contract divided by the conditional reward ($100) according to a hypothetical market probability of its associated event,

$$\frac{E\left(\$\text{Payoff}\right)}{\$100} = \frac{\left(1 - P\left(\text{Event}\right)\right)\$0 + P\left(\text{Event}\right)\$100}{\$100} = P\left(\text{Event}\right)$$

Figure 3.2 depicts a partial time series of fair price transactions from the Irish Intrade exchange of several binary contracts for the U.S. presidential election winner (paying $100 if, and only if, the particular candidate wins). Thus, using the connection between fair prices and expectations, the time series of market fair prices represent the probabilities in percent (according to the market participants) of their associated events. These are time series of coherent market probabilities (they do not add up to 100% because other contracts of potential contenders, such as Al Gore, with small non-zero probabilities are not shown in the graph) assimilating relevant information as it is uncovered (e.g., the January primaries results). In fairness, the markets are as coherent as they are liquid, and this particular market is not as liquid (in terms of tight bid and ask spreads and deep associated stakes) as other markets. Yet, the implication is that since liquid markets ought to be coherent then liquid markets trading binary contracts must obey the rules of probability theory associated to coherence. That is, denoting generic events by A and B and a sure event by Ω, the market implied probabilities should comply with: $P(A)$ takes a unique value; $P(A) \geq 0$; $P(\Omega) = 1$; $P(A+B) = P(A) + P(B)$, where $A+B$ denotes the union of two inconsistent events, and $P(AB) = P(A)P(B|A)$ where $P(B|A)$ denotes a conditional binary bet on the event B that is called off if and only if A fails to happen. For completeness, a concise direct derivation of these coherent rules, following Quintana (2005), is presented in the Appendix A. Furthermore, as pointed out by de Finetti (1974), these probability rules can be alternatively derived indirectly from three basic rules regarding generic fair prices of payoffs connected to arbitrary random quantities (denoting the fair price of a contract tied to a random quantity X as $E(X)$ to emphasize that it is, or should be, its expectation): $E(X)$ takes a unique value; $E(X+Y) = E(X) + E(Y)$; and $X \leq a$ (meaning, a fixed quantity, a, is greater than or equal to any possible outcome of X) implies $E(X) \leq a$. Therefore, liquid futures markets ought to obey these rules and all their implications to remain coherent (e.g., for several random quantities, the associated fair prices' points must lie in the closed convex hull of the possible outcomes).

Are these implied market expectations worthwhile? Are the futures markets unbeatable? The supporters of the Efficient Markets Hypothesis (EMH) answer these questions affirmatively. The conjecture and its underlying argument can be traced to Bachelier's (1900) conclusion: "the mathematical expectation of the speculator is zero." The argument stretches convincingly the reasoning supporting the liquid markets' coherence as follows: if the futures markets were inefficient then the trading triggered by the speculators taking advantage of expected profitable opportunities would re-establish price levels that fully assimilate the relevant information and no more expected gains would remain, thus, effectively restoring the market efficiency. Empirical analyses (e.g., Fama 1970) support the EMH in a variety of setups, including binary bets (a.k.a., prediction) markets (e.g., Carvalho and Rickershauser (2009)) and the early period of the Dojima futures market (Hamori et al. 2001). Interestingly, the Dojima rice futures market, which can be regarded as the first Bayesian practitioner, started to quote fair prices in 1730; that is, 33 years before the work of the first Bayesian theorist was published posthumously.

The liquid futures markets' coherence and efficiency appears to be accomplishing an amazing feat. Yet, it can be explained because: first, key markets' speculators are well aware of the arbitrage-free concept (essentially the financial economics jargon for coherence) and eager to exploit any potential opportunities; second, the market consensus expectations (fair prices) integrate over a spectrum of presumably relevant pieces of information pondered by the conviction measured by the amount of money willing to be committed by each one of the participants; and third, a financial evolution process ensures that only the best fitted participants survive in the long run. Furthermore, the market coherence and efficiency ought to be dynamic; that is, coherent conditional market expectations (fair prices) are produced in real time assimilating available relevant information under all circumstances even while extreme events are happening. These liquid futures markets come across, as a whole, as an invincible juggernaut assimilating not only relevant information but also any applicable knowledge (e.g., financial economics, computing and statistical technologies among others) in its path.

4 Bayesian Speculation

Last section ending notwithstanding, one might sense a vulnerability on any market maker due to the willingness to buy or sell at the, virtually, same quoted price. Let us adopt henceforth the offensive role of a Bayesian speculator instead of the defensive attitude of a Bayesian market maker (just as one would rather choose a piece, as opposed to cutting two pieces fairly, of a pie). An encouraging indication, from the speculator's viewpoint of a hedge fund manager, is that eighty years after the aforementioned Bachelier's (1900) pronouncement, a disarming counter-argument was proposed by Grossman and Stiglitz (1980): If the markets were efficient there would not be any incentives to get, gather and process information at a certain non-zero cost, and incorporate it, by trading, into the markets; therefore, the market would not assimilate the relevant information and it would become inefficient. This leads to the EMH paradox: markets' efficiency implies markets' inefficiency, and markets' inefficiency implies markets' efficiency! As a smart (profit-seeking, risk-averse) speculator one could perceive markets incoherence and efficiency as black (when a sure profit is available) and white (when a positive expected gain is unattainable) with shades of gray in between. A general speculating strategy consists of two steps: first, form a (probabilistic) view of the markets; second, trade an advantageous book of bets based on that view.

One way a Bayesian speculator can quantitatively implement the latter step is by maximizing the expected net payoff (a.k.a., excess return) of a book of bets $(\mathbf{w}_t'\mathbf{y}_t)$ at a certain time (t) for a certain holding period (e.g., a week) per unit of risk measured by its standard deviation (the so-called ex-ante information ratio); that is,

$$\max_{\mathbf{w}_t} \left(\frac{\underset{\mathbf{y}_t|\mathbf{x}_1,\mathbf{y}_1,\ldots,\mathbf{y}_{t-1},\mathbf{x}_t}{E} \mathbf{w}_t'\mathbf{y}_t}{\sqrt{\underset{\mathbf{y}_t|\mathbf{x}_1,\mathbf{y}_1,\ldots,\mathbf{y}_{t-1},\mathbf{x}_t}{V} \mathbf{w}_t'\mathbf{y}_t}} \right) \tag{4.1}$$

where $\mathbf{y}_t, \mathbf{x}_t, \mathbf{w}_t$ denote respectively the vector of net payoffs (the differences between buying prices and selling prices of the futures contracts), their associate vector of relevant explanatory variables (that presumably would contribute to its predictability) and the corresponding vector of trading (holding) amounts. Notice that the weights do not have to comply with the traditional investments constraints (i.e., they do not have to be non-negative, neither do they have to add up to one) since $\mathbf{w}_t$ represents a book of side bets as opposed to an allocation of investment capital (capital would still be required to support the book of bets to cover potential losses but it could be earning interest or be invested in any other liquid assets). The optimal information ratio is bounded by zero on the left hand side and by infinity on the right hand side; the former bound corresponds, from the speculator's point of view, to market efficiency whereas the latter implies an incoherent market condition (although market incoherence, in principle, could also occur in a situation with a finite information ratio). Thus, the information ratio provides quantitative means to evaluate to what degree the market expectations are beatable by measuring the speculator's expected gain per unit of risk taken in an optimal scenario.

This optimization step is equivalent to the following two-way Pareto optimization,

$$\max_{\mathbf{w}_t} \left(\left(\underset{\mathbf{y}_t|\mathbf{x}_1,\mathbf{y}_1,\ldots,\mathbf{y}_{t-1},\mathbf{x}_t}{E} \mathbf{y}_t \right) - \frac{1}{2}\lambda_t \left(\underset{\mathbf{y}_t|\mathbf{x}_1,\mathbf{y}_1,\ldots,\mathbf{y}_{t-1},\mathbf{x}_t}{V} \mathbf{y}_t \right) \right) \tag{4.2}$$

and optimal trades are given by,

$$\mathbf{w}_t^{\star} = \frac{1}{\lambda_t} \left(\underset{\mathbf{y}_t|\mathbf{x}_1,\mathbf{y}_1,\ldots,\mathbf{y}_{t-1},\mathbf{x}_t}{V} \mathbf{y}_t \right)^{-1} \underset{\mathbf{y}_t|\mathbf{x}_1,\mathbf{y}_1,\ldots,\mathbf{y}_{t-1},\mathbf{x}_t}{E} \mathbf{y}_t \tag{4.3}$$

where λ_t is an arbitrary inverse scaling factor representing the speculator's greed-fear tradeoff. The optimization (4.1) corresponds to the Modern Portfolio Theory (MPT) objective function but the feasible set for the weights is different: in this setup there are no constraints whereas in the MPT setup the weights are constrained to lie on the simplex and the solution formula (4.3) would not apply; instead, a quadratic programming technique would be required to obtain the solution. The online optimal book (portfolio) given by the formula (4.2) turns out to be also optimal in a multi-period utility maximization setup; this is discussed briefly in the Appendix A.

It is interesting to note at this point that once a liquid futures market has emerged the subjectivity (of probabilities and expectations elicited according to de Finetti's recipe) should disappear because a smart market maker should simply quote the liquid market price and hedge accordingly since this would not only make his expected loss null but the potential loss itself would be null (quoting a different price would not necessarily imply incoherence but a quasi-incoherence in the sense that he would be willing to buy at a higher price or to sell at a lower price than the market price giving a perfect arbitrage opportunity to another speculator to make sure money). However, if the expectations of the Bayesian speculator coincides with the futures market's expectations (i.e., prices), then the expected excess return of any book would necessarily be null. To avoid this pitfall the loophole should be closed and the elicitation process should, in principle, assume that the Bayesian speculator has no access to liquid futures markets during his hypothetical role as a market maker.

5 Bayesian Forecasting

The first step of the aforementioned trading strategy can be implemented quantitatively in the following manner: form the predictive density function of the payoffs for each holding period by a Bayesian updating process of an underlying Bayesian dynamic model. Coherency, as mentioned in the appendix, requires that conditional fair prices must correspond to the customary conditional expectations before the conditioning information is observed, but does not require that this correspondence must remain after the conditioning information is revealed. Although perceived by some as a weakness in the foundations of the Bayesian subjective approach, this feature is actually a blessing in disguise from the Bayesian speculator's perspective: on the one hand, the speculator could maintain consistency and rely routinely on the default Bayesian updating based on conditional probability and expectations since they ensure coherence for the predictive probabilistic view at each holding period, and this is an essential requirement for the strategy (otherwise, for example, one could dangerously conclude that there is a perfect arbitrage opportunity when in fact there is none); on the other hand, the Bayesian speculator is always free to consider and switch to an alternative Bayesian model at any time if a decisive new information, knowledge or technology becomes available (only a mad speculator would blindly commit perpetually to a fixed model).

A Bayesian dynamic (a.k.a., stochastic) model presumably would allow the trading strategy to adapt to regime changes and assimilate financial market shocks. Figure 5.1 represents the stochastic structure of a generic Bayesian dynamic model in a manner introduced by Quintana et al. (2003). This representation relies on the principle that a stochastic model is well defined if and only if its simulation is fully specified. Arrows denote inputs to a deterministic function at each node; double arrows denote generated random entity (vector, matrix, etc.) inputs whose distributions are fully specified (involving only known parameters) and they constitute the primary sources of randomness; and all the primary random entities are jointly independent. The Bayesian dynamic model consists of a Markovian evolution of the parameter time series $\mathbf{\Theta}_t$ coupled with contemporaneous random observations $\mathbf{y}_t$. The basic premise is that the model structure, described by the values of the system parameters $\mathbf{\Theta}_t$, is changing according to a stochastic process (as opposed to traditional rigid static formulations) but there is a degree of persistence making the sampling and filtering process worthwhile.

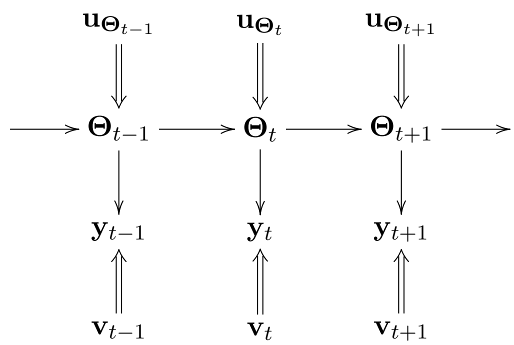

Figure 5.1: Functional network of a generic Bayesian dynamic model

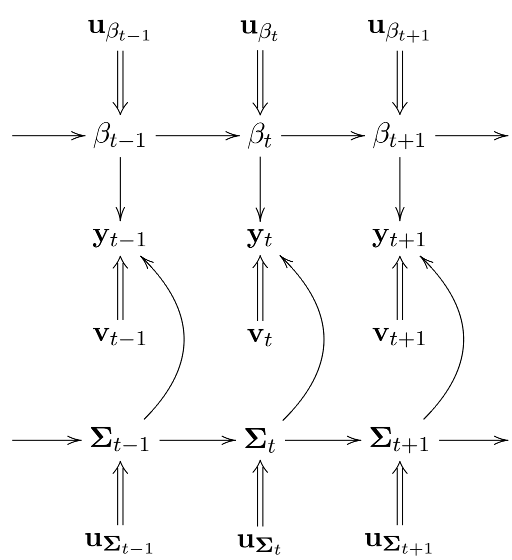

Figure 5.2: Functional network of a specific Bayesian DLM model

A Dynamic Linear Model (DLM) variant of the formulations entertained by West and Harrison (1997) is of particular interest when the speculator suspects that the markets are not assimilating efficiently all the available public (relevant) information; that is when the so-called semi-strong form of the EMH (Fama 1970) might be vulnerable. This model is a particular Bayesian dynamic model where the parameters are decomposed as follows: the observation equation, at each time period, takes the form of a multivariate multiple regression that relates the excess returns (or logarithmic excess returns that are more convenient from a modeling perspective) $\mathbf{y}_t$ and associated relevant available public information $\mathbf{x}_t$ via the regression coefficients β_t with a residual variance-covariance $\mathbf{\Sigma}_t$ representing the variance-covariance of the difference of the random observations and their corresponding expectations; and the random evolution of the parameters follows a form of random walk.

This model's functional network can be decomposed as a marginal network (for $\mathbf{\Sigma}_t$) coupled with a conditional network (given $\mathbf{\Sigma}_t$), or vice versa in terms of β_t, in a hierarchical fashion as is depicted in Figure 5.2. A more detailed specification of this model, including specific functional forms and random generation procedure descriptions, is presented in the Appendix B.

To what degree are the market expectations beatable? A comprehensive discussion on this subject is beyond the scope of this chapter; however, the answer to the question depends, of course, on the respondent. According to simulated and, more importantly, actual trading of strategies based on the above described specific Bayesian DLM and a variant the information ratio optimization have generated attractive risk-adjusted returns (see Quintana 2005 and Quintana et al. 2003 for details). Nevertheless, the customary disclaimer applies: PAST PERFORMANCE IS NOT A GUARANTEE OF FUTURE RESULTS and the only way to beat the markets going forward is by taking some risk of losing money (unless market incoherencies are identified and exploited).

6 Risk Modeling

Risk reduction is more important than expected return enhancement in the sense that according to the ex-ante information ratio formula 4.1 a risk reduction of 50% is equivalent to an expected return enhancement of 100%. Someone might still try to argue pointing out that doubling the expected excess return at the same risk in some circumstances could be more appealing than halving the risk while maintaining the same expected return. However, the former book can be derived from the latter book using leverage (by doubling the bets; that is, the allocation weights). Thus, risk modeling is a crucial mission for a speculator seeking the best risk-adjusted performance (i.e., the best information ratio). In this section the focus is on the predictive ability, for hedging purposes, of the variance covariance of a book of bets ($\mathbf{w}_t'\mathbf{y}_t$); which is the variance appearing in the denominator of expression 4.1,

$$\underset{\mathbf{y}_t|\mathbf{x}_1,\mathbf{y}_1,\ldots,\mathbf{y}_{t-1},\mathbf{x}_t}{V} \mathbf{w}_t'\mathbf{y}_t = \mathbf{w}_t'\mathbf{V}_t\mathbf{w}_t \tag{6.1}$$

where V_t is short hand for the predictive variance of the book of bets $\underset{\mathbf{y}_t|\mathbf{x}_1,\mathbf{y}_1,\ldots,\mathbf{y}_{t-1},\mathbf{x}_t}{V} \mathbf{y}_t$.

Although risk and return are inherently connected, it is useful from a modeling perspective to analyze them as separately as possible. Because of the availability of leverage, a speculator trying to evaluate the expected return capability of a model cannot just set λ_t equal to zero in the optimization (4.2); one possible way to alleviate this difficulty is, for evaluation purposes, to force the weights to add up to one. However, because the book variance, given a direction, is directly proportional to the square of its norm, forcing the weights to add up to one would unfairly favor the minimum length book satisfying the constraint (i.e., the long equal weighted book) and punish other books as their length increase. This drawback can be avoided simply by constraining the potential weights to have a unitary norm instead. Thus, an ideal testing hedging task would be to find the minimum variance book lying on the hyper-circumference,

$$\min_{\mathbf{w}_t} \left(\mathbf{w}_t'\mathbf{V}_t\mathbf{w}_t\right), \ \textit{s.t.}\ \mathbf{w}_t'\mathbf{w}_t = 1 \tag{6.2}$$

or equivalently, to find the minimum $\mathbf{V}_t$'s Rayleigh quotient $((\mathbf{w}_t'\mathbf{V}_t\mathbf{w}_t)/(\mathbf{w}_t'\mathbf{w}_t))$. The solution to both problems can be derived directly but it is a known result in extremizing quadratic forms (e.g., Noble and Daniel 1987): the optimal weights correspond to the normalized eigenvector associated to the smallest eigenvalue (which itself is the value of the minimum variance book) of the spectral decomposition of $\mathbf{V}_t$; this is the counterpart of the other better known maximum variance result where the book (portfolio) corresponding to the largest eigenvalue exhibits the most variation (as discussed by Quintana and West 1987).

Although modeling assumptions about the payoff expectations would obviously influence risk predictions given the liquid markets' efficiency those expectations should not be far away from the markets' expectations. Thus, risk models evaluation results based on markets' payoff expectations should still apply and for this reason these markets' expectations are entertained henceforth. That is, it is assumed in this section that the expectation $\underset{\mathbf{y}_t|\mathbf{x}_1,\mathbf{y}_1,\ldots,\mathbf{y}_{t-1},\mathbf{x}_t}{E} \mathbf{y}_t$ corresponds to the market's expectation (i.e., it is null).

The merits of alternative risk models, since the expected excess returns are assumed to be null, can be assessed by comparing the cumulative sum of squares of the excess returns of these extreme books over time associated to each model, or equivalently, rolling (e.g., one year) ex-post (sample) standard deviations can be used to discriminate between models. Two alternative risk models embedded in the functional network structure (5.2), for $\beta_t = \mathbf{0}$, are considered in this study. The first, and the simplest, assumes that the Σ_t stochastic evolution follows a form of random walk that allows for a conjugate analysis with Wishart probability distributions; this model corresponds to the Stochastic Multiple-Factor Model discussed by Quintana et al. 2003 with the constraint $\beta_t = \mathbf{0}$ which was introduced by Quintana and West 1987. The second model, entertains a Graphical formulation for $\mathbf{\Sigma}_t$ that encompasses a conditional independence structure via a pattern of zeros in the corresponding precision matrix. As introduced by Carvalho and West (2007), the idea of conditioning the sequence of variance-covariance matrices on set of graphical constraints provides a parsimonious model for $\mathbf{\Sigma}$ potentially improving the empirical accuracy of book (portfolio) predictions and reducing the variation of realized returns. In this application, the set of constraints, i.e. the graph, was selected as the top model from a stochastic search procedure (Scott and Carvalho 2008) using observations prior to the period analyzed here. A comprehensive description of both models appears in the Appendix B.

The dataset used to evaluate the trading performance of the alternative dynamic variance-covariance models consists of weekly excess returns, from 5/25/1990 through 6/19/2008, derived from adjusted indexed futures prices provided by Bloomberg using the indexed price series. This futures dataset entails currency, government bonds and stock indices instruments typically traded by a Global Macro Hedge Fund and it is described at the end of this section. Figures 6.1–6.3 show the hedging performance of the full model versus the graphical model. The figures' left-hand sides depict the corresponding annualized one-year rolling ex-post standard deviations for each model. The right-hand sides show the ratio (full/graph) of these rolling standard deviations. The different pictures result from varying the discount factor as follows: $\delta = 0.97, 0.98, 0.99$. It is apparent from these figures that for a high discount factor the performances of the models are similar but as the values decrease the dominance of the graphical model over the full model becomes apparent. This result is consistent with the intuition that the regularization from the graphical structure is more relevant when information is scarcer (i.e, when the discount factor is smaller) and suggests that the hedging performance of the graphical model is more robust in relation to the choices of the discount factors making it very attractive.

The discussion at the beginning of this section hints that the hedging performance dominance of the graphical models over the full models should be carried into risk-adjusted performance dominance. This is confirmed, in the case at hand, by the simulated risk-adjusted performance exhibited in Figures 6.4–6.6. The figures depict the one-year rolling ex-post information ratio attained when the risk models are coupled with the expected returns formulation of the Bayesian DLM model mentioned in Section 5 using proprietary explanatory variables that include measures of short, medium and long interest rates, liquidity and credit conditions, proxies for inflation, etc. These simulations begin after a buffer of three years used as a learning period for estimating the parameters associated to the payoffs expectations and take into account estimated transaction costs and employ a realistic production 4-way (mean-variance-size-turnover) Pareto optimization variant of the simpler theoretical 2-way (mean-variance) optimization discussed in Section 4 to promote the book's liquidity and control its transaction costs.

A more detailed description of the dataset follows to complete this section. The currencies

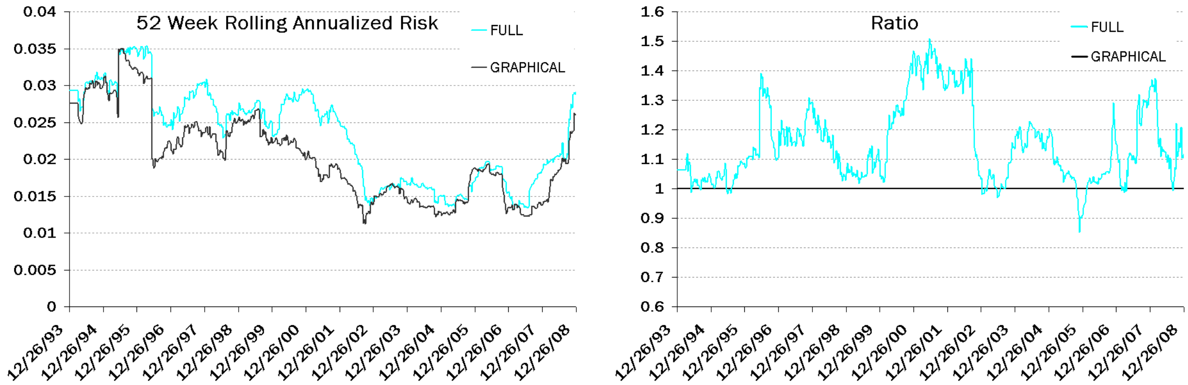

Figure 6.1: $\delta = 0.97$

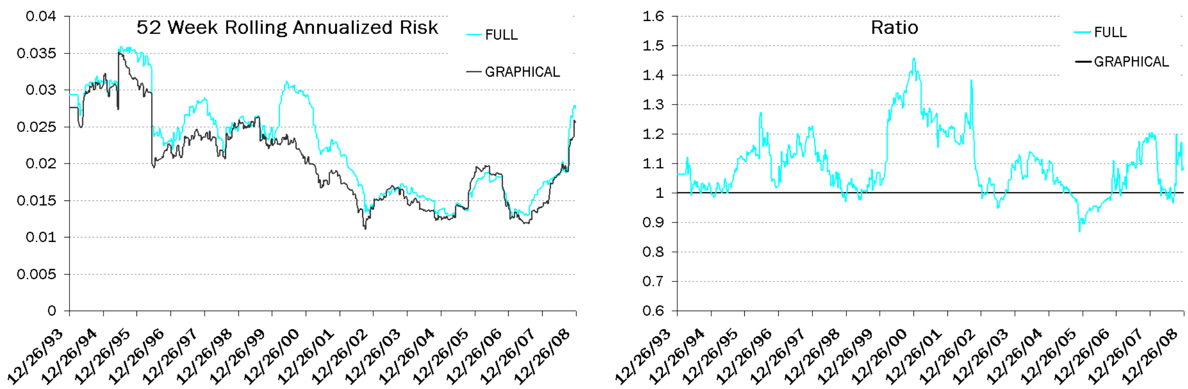

Figure 6.2: $\delta = 0.98$

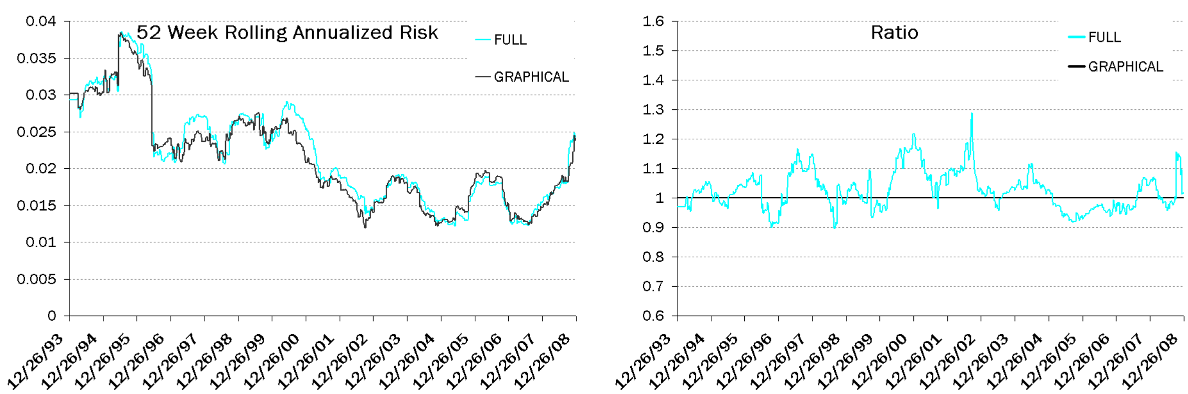

Figure 6.3: $\delta = 0.99$

weekly excess returns in the dataset were calculated as $y_t = \frac{C_t}{C_{t-1}} - 1$ where C_{t-1} and C_t are the specific indexed closing currency futures prices on consecutive Fridays (i.e., they are excess returns from a U.S. dollar speculator's perspective); the stock indices and the government bond weekly excess returns were calculated as $y_t = (\frac{P_t}{P_{t-1}} - 1)\frac{C_t}{C_{t-1}}$ where P_{t-1} and P_t are the indexed closing futures prices and C_{t-1} and C_t are the corresponding indexed closing currency futures prices (i.e., they are currency hedged excess returns from a U.S. dollar speculator's perspective).

The Bloomberg tickers of the liquid futures contracts considered are as follows,

Currencies: AD1 Crncy, CD1 Crncy, SF1 Crncy, BP1 Crncy, EC1 Crncy, JY1 Crncy.

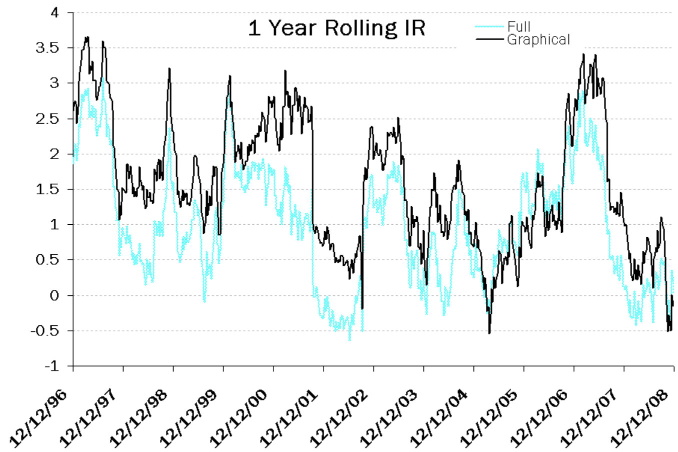

Figure 6.4: $\delta = 0.97$

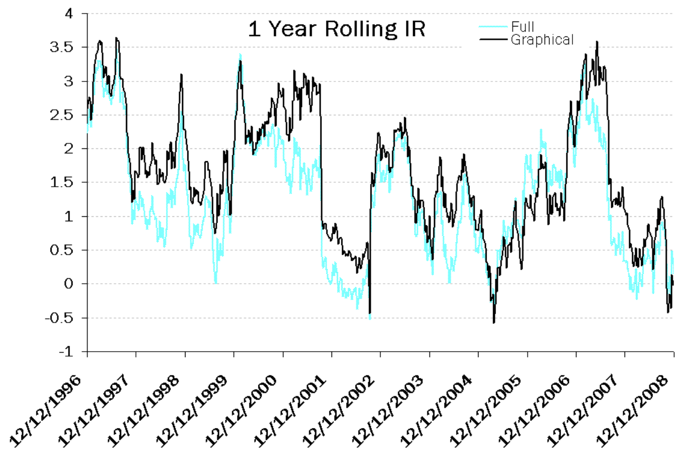

Figure 6.5: $\delta = 0.98$

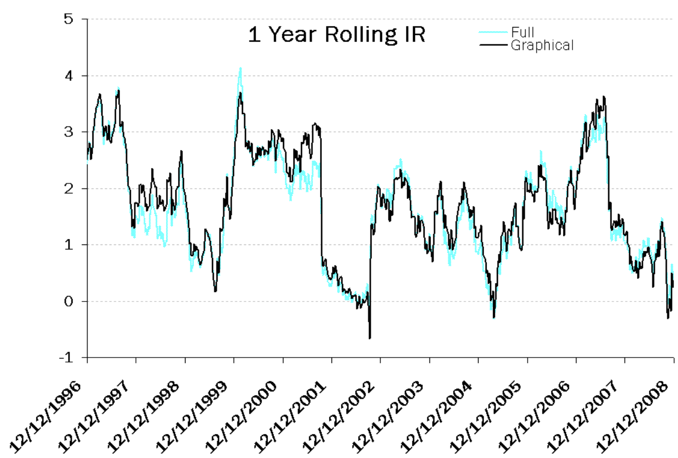

Figure 6.6: $\delta = 0.99$

Government Bonds: XM1 Comdty, CN1 Comdty, FB1 Comdty, RX1 Comdty, G 1 Comdty, JB1 Comdty, ZT1 Comdty, TY1 Comdty.

Stock Indices: XP1 Index, PT1 Index, SM1 Index, GX1 Index, IB1 Index, CF1 Index, Z 1 Index, ST1 Index, TP1 Index, EO1 Index, SP1 Index, ZI1 Index.

Indexing based on price ratio was used to smooth out jumps that occur when a contract rolls into a new month (i.e., the speculator is assumed to maintain the same exposure while rolling contracts). To adjust by ratio when a contract changes, the previous indexed prices were multiplied by the ratio of the price of the new contract on the roll date to the price of the previous contract just before the roll date. The rolling schedule used for the currency assets was 5 days prior to the contract expiration date. The equity indices roll schedule was 4 days prior to the contract expiration date with the exception of the French CAC40 Equity Index (CF1 Index) — which was 5 days prior to expiration. The bond assets were rolled 5 days prior to the contract expiration with the following exceptions: CN1 Comdty — 28 days prior to expiration, G 1 Comdty — 35 days prior to expiration, TY1 Comdty — 25 days prior to expiration, XM1 Comdty and ZT1 Comdty — on the expiration date. Adjustments to imply excess returns were performed for the latter due to the Australian ad hoc quoting methods (described in its Bloomberg description page). The Deutsche mark was considered instead of the Euro prior to 1999. Finally, where futures prices were not available for the entire history considered, implied fair future prices were calculated based on suitable proxies.

7 Conclusions

According to subjective formulations of probability theory, Bayesians are, in principle, playing the role of altruist bookmakers in a betting scheme. Correspondingly, liquid financial futures markets act virtually as Bayesian market makers quoting fair prices that, consequently, represent objective coherent consensus expectations of the underlying commodities (including currencies, government bonds, stock indices, etc.). This amazing feat is accomplished by market participants acting in concert, albeit driven by selfish motivation, and putting into practice the Japanese proverb “none of us is as smart as all of us.” Furthermore, the futures markets, as an entity, assimilates not only all the relevant information but also the applicable methodologies and technologies resulting in a formidable apparatus providing key economic ongoing expectations that, according to the Efficient Markets Hypothesis advocates, cannot be improved.

Although the futures markets appear to be coherent, a Bayesian speculator driven by dynamic models predictions, which are hypothetically elicited without access to the markets to allow back subjectivity, could be able to generate attractive returns, given the risk taken, by trading in different markets as a counterpart. Regarding risk-adjusted performance, risk modeling and risk predictions are at least as important as the expected payoffs’ formulations. Moreover, the hedging ability of different risk models can be assessed by comparing their success in producing normalized books (portfolios) with minimum risk; by this criterion, graphical models seem to be more robust having either better or similar hedging effectiveness in relation to the corresponding models without a conditional independence structure. Furthermore, their associated risk-adjusted performance exhibits, as expected, a similar dominance relationship.

Ultimately, resistance is futile, since the market Borg would dynamically assimilate any Bayesian speculating trading strategy that is implemented and would reward (or punish) the Bayesian speculator to the exact extent to which the futures markets’ expectations improve (or deteriorate) as a result of this particular trading activity.

References

Bachelier, L. (1900). *Théorie de la Spéculation*. Annales de l’École normale supérieure. Paris. Translated from French in *The Random Character of Stock Market Prices*. (P.H. Cootner, ed.). Cambridge, MA: The MIT Press, 1964, 17-78.

Barnard, G. A. (1980). Discussion of H. Akaike ‘Likelihood and the Bayes procedure’ and A. P. Dawid ‘A Bayesian Look at Nuisance Parameters’. In J. M. Bernardo, M. H. DeGroot, D. V. Lindley, and A. F. M. Smith (Eds.), *Bayesian Statistics*, pp. 185–189. Valencia: Univeristy Press.

Bayes, T. (1763). An Essay Towards Solving a Problem in the Doctorine of Chances. *Phil. Trans. Roy. Soc. London 53*, 370–418. Published Posthumously. Electronic edition: `http://www.stat.ucla.edu/history/essay.pdf`.

Carvalho, C. and M. West (2007). Dynamic matrix-variate graphical models. *Bayesian Analysis 2*, 69–96.

Carvalho, C. M. and J. Rickershauser (2009). Volatility in Prediction Markets: A Measure of Information Flow in Political Campaigns. In A. O'Hagan and M. West (Eds.), *The Hanbook of Applied Bayesian Analysis*. Oxford University Press.

Dawid, A. P. and S. L. Lauritzen (1993). Hyper-Markov laws in the statistical analysis of decomposable graphical models. *The Annals of Statistics 3*, 1272–1317.

de Finetti, B. (1931). Sul Significato Soggettivo della Probabilità. *Fundamenta Mathematicae 17*, 298–329. Translated into English as "On the Subjective Meaning of Probability," in Paola Monari and Daniela Cocchi (Eds., 1993) "Probabilit e Induzione," Clueb, Bologna, 291-321.

de Finetti, B. (1974). *Theory of Probability, a Critical Introductory Treatment*. Wiley.

DiBartolomeo, D. (2003). Portfolio Optimization: The Robust Solution. In *Prudential Securities Quantitative Conference*. Electronic edition: `http://www.northinfo.com/documents/45.pdf`.

Fama, E. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance 25*, 383–417.

Goldstein, M. (2005). Subjective Bayesian Analysis: Principles and Practice. In *Case Studies in Bayesian Statistics, Workshop 8–2005*. Electronic edition `http://lib.stat.cmu.edu/bayesworkshop/2005/goldstein.pdf`.

Grossman, S. and J. Stiglitz (1980). On the Impossibility of Informationally Efficient Markets. *American Economic Review 70*, 393–408.

Hamori, S., N. Hamori, and D. A. Anderson (2001). An Empirical Analysis of the Efficiency of the Osaka Rice Market During Japan's Tokugawa Era. *Journal of Futures Markets 21*, 861–874.

Jones, B., C. Carvalho, A. Dobra, C. Hans, C. Carter, and M. West (2005). Experiments in stochastic computation for high-dimensional graphical models. *Statistical Science 20*, 388–400.

Lauritzen, S. L. (1996). *Graphical Models*. Clarendon Press, Oxford.

Markowitz, H. M. (1959). *Portfolio Selection: Efficient Diversification of Investments* (2nd ed. ed.). Oxford: Blackwell.

Noble, B. and J. Daniel (1987). *Applied Linear Algebra* (Second ed.). Englewood Cliffs: Prentice Hall.

Quintana, J. M. (1992). Optimal Portfolios of Forward Currency Contracts. In J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith (Eds.), *Bayesian Statistics 4*, pp. 753–762. Oxford University Press.

Quintana, J. M. (2005). Bayesian Efficient Trading. In N. Kolev and P. Morettin (Eds.), *Proceedings of the Second Brazilian Conference on Statistical Modelling in Insurance and Finance*. Institute of Mathematics and Statistics at the University of Sao Paulo.

Quintana, J. M., V. Lourdes, O. Aguilar, and J. Liu (2003). Global Gambling. In J. M. Bernardo, M. J. Bayarri, J. O. Berger, A. P. Dawid, D. Heckerman, A. F. M. Smith, and M.West (Eds.), *Bayesian Statistics 7*, pp. 349–367. Oxford University Press. (with discussion).

Quintana, J. M. and M. West (1987). Multivariate time series analysis: New techniques applied to international exchange rate data. *The Statistician 36*, 275–281.

Ramsey, F. P. (1926). "Truth and Probability". In: Ramsey, 1931, *The Foundations of Mathematics and other Logical Essays*, Ch. VII, p.156-198, edited by R.B. Braithwaite, London: Kegan, Paul, Trench, Trubner & Co., New York: Harcourt, Brace and Company.

Scott, J. G. and C. M. Carvalho (2008). Feature-inclusion stochastic search for gaussian graphical models. *Journal of Computational and Graphical Statistics 17*, 790–808.

West, M. and P. J. Harrison (1997). *Bayesian Forecasting and Dynamic Models* (Second ed.). New York: Springer-Verlag.

Zelner, A. and V. K. Chetty (1965). Prediction and Decision Problems in Regression Models from the Bayesian Point of View. *J. Amer. Statist. Assoc. 60*, 608–616.

Appendices

A Broader Context and Background

A.1 Foundations: Probabilty Axioms Derivation

An implication of the probability axioms (showing the steps to get sure money from an otherwise incoherent market maker) is concisely presented in Quintana (2005) as follows:

1. $P(A)$ is unique (otherwise, buy it low and sell it high).
2. $P(A) \geq 0$ (otherwise, buy it if $P(A) < 0$).
3. $P(\Omega) = 1$ (otherwise, buy it if $P(\Omega) < 1$, or sell it if $P(\Omega) > 1$; recall, Ω denotes a sure event).
4. $P(A + B) = P(A) + P(B)$ (a synthetic binary bet for $A + B$ is a binary bet for A plus a binary bet for B; $A + B$ denotes the union of two inconsistent events).
5. $P(AB) = P(A)P(B \mid A)$ (a synthetic binary bet for AB is $P(B \mid A)$ units of a binary bet on A plus a contingent binary bet on $B \mid A$ if A happens, using the contingent payoff to buy it; recall, a binary bet on $B \mid A$ is called off if A does not happen).

However, there are two caveats worth mentioning: First, Axiom 4 only implies finite additivity, and its implied generality is somewhat controversial. Second, the price $P(B \mid A)$ is quoted beforehand; but the market maker could quote a different one when and if A happens, even if this is the only extra information revealed (then again, as pointed out by Goldstein (2005), typically extra relevant information would also become simultaneously available). There is an alleviating provision, the market maker cannot tell in advance, without becoming incoherent, that her future quote will always be higher (or lower) afterwards.

A.2 Foundations: Bayesian Portfolio Optimization

The optimization (4.2) corresponds to the Markowitz (1959) mean-variance approach to finding optimal allocation weights in a traditional investment setup but without any of the traditional investment constraints for the weights to lie on the simplex. Although this approach, where only the first two predictive moments of the excess returns are in play, is not a panacea it seems to work well in practice. This direct predictive formulation is equivalent to solving first the optimization problem conditional on the model's unknown parameters and taking the utility expectation afterwards, as prescribed by the early work of Zelner and Chetty (1965), and naturally deals with the uncertainty inherent in the unknown modeling parameters. In contrast, there is an industry (e.g., DiBartolomeo 2003) to compensate for naive non-Bayesian alternatives that rely on plug-in estimates.

Moreover, it can be justified by the Bayesian paradigm of maximizing expected utility and it is the solution to a multiple-period decision problem even though it can be implemented myopically. This can be induced by defining the multiple-period utility as a weighted additive (where the components are the single-period utilities). In this framework, the associated stochastic dynamic programming problem is broken down into several isolated single-period optimizations and each optimal portfolio can be determined online requiring only single-step conditional predictive distributions for the payoffs. In addition, the single-period optimization (4.2) can be derived from the principle of utility maximization when the single-period utility is set as the difference of (positive) target payoff minus a square error penalty for deviating from that target.

This approach and its motivations are discussed in (Quintana 1992; Quintana et al. 2003) and references therein. It is worth noting that, despite common Wall Street wisdom to the contrary, the speculator does not necessarily have to assume normality, nor miss potentially unlimited opportunities (due to the presence of market incoherence) as it would be the case when using a naive quadratic loss formulation.

B Models and Computations

B.1 Models

Generic Bayesian Dynamic Model

The generic statistical model depicted by the functional network in Figure 5.1 and described by the corresponding specific functional forms for the observations and parameters random evolution,

$$\begin{aligned} \mathbf{y}_t &= f\left(\mathbf{\Theta}_t, \mathbf{v}_t\right) \\ \mathbf{\Theta}_t &= g\left(\mathbf{\Theta}_{t-1}, \mathbf{u}_{\mathbf{\Theta}_t}\right) \end{aligned} \qquad \text{A.1}$$

coupled with independent random generation procedures for $\mathbf{v}_t$ and $\mathbf{u}_{\mathbf{\Theta}_t}$ has a conditional probabilistic structure; this generic model specification can be defined in a traditional way by the set of stochastic Markovian parametric probability density functions $p(\mathbf{\Theta}_t \mid \mathbf{\Theta}_{t-1})$ and the set of observational probability density (i.e., likelihood) functions $p(\mathbf{y}_t \mid \mathbf{\Theta}_t)$. For notational convenience we are omitting all the given available information (e.g., $\mathbf{x}_1, \mathbf{y}_1, \ldots, \mathbf{y}_{t-1}, \mathbf{x}_t$). The initial parameter, without loss of generality, can be set $\mathbf{\Theta}_0 = 0$. The updating recurrences for the evolution, prediction and filtering are respectively:

$$\begin{aligned} p\left(\mathbf{\Theta}_t\right) &= \int_{\mathbf{\Theta}_{t-1}} p\left(\mathbf{\Theta}_t \mid \mathbf{\Theta}_{t-1}\right) p\left(\mathbf{\Theta}_{t-1}\right) d\mathbf{\Theta}_{t-1} \\ p\left(\mathbf{y}_t\right) &= \int_{\mathbf{\Theta}_t} p\left(\mathbf{y} \mid \mathbf{\Theta}_t\right) p\left(\ \mathbf{\Theta}_t\right) d\mathbf{\Theta}_t \\ p\left(\mathbf{\Theta}_t \mid \mathbf{y}_t\right) &= \frac{p\left(\mathbf{y}_t \mid \mathbf{\Theta}_t\right) p\left(\mathbf{\Theta}_t\right)}{p\left(\mathbf{y}_t\right)} \end{aligned}$$

that is, the steps involve the computation of two marginal densities and one conditional density.

Specific Bayesian DLM

The observational functional equations and the associated random generation procedures (given $\mathbf{\Sigma}_t$) of the generic Bayesian DLM, depicted in Figure 5.2, correspond to a multi-variate multiple regression model with a global scaling factor for the residual variance covariance and are defined as follows, $\mathbf{y}_t = \mathbf{X}_t \beta_t + C\left(\mathbf{\Sigma}_t\right) \mathbf{v}_t$, where $\mathbf{v}_t \sim N\left(0, \mathbf{I}\right)$ and $C\left(\mathbf{\Sigma}_t\right)$ denotes the Cholesky decomposition of $\mathbf{\Sigma}_t$ (i.e., $(\mathbf{y}_t \mid \beta_t) \sim N\left(\mathbf{X}_t\beta_t, \mathbf{\Sigma}_t\right)$.)

The aim of the evolution functional equations and the (independent) random generation procedures corresponding to the network depicted in Figure 5.2 (given $\mathbf{\Sigma}_t$) is to provide a bridge from a posterior conjugate distribution at time $t-1$ to a prior conjugate distribution at time t. That is, to evolve from $\beta_{t-1} \sim N\left(\mathbf{m}_{t-1}, \mathbf{C}_{t-1}\right)$ (i.e., $\beta_{t-1} \sim N\left(\mathbf{m}_{t-1}, \mathbf{C}_{t-1}\right)$) to $\beta_t \sim N\left(\mathbf{m}_t^\star, \mathbf{C}_t^\star\right)$ where

$$\begin{aligned} \mathbf{m}_t^\star &= \mathbf{G}\mathbf{m}_{t-1} && \text{A.2a} \\ \mathbf{C}_t^\star &= \mathbf{C}_{t-1} + \mathbf{G}_t\mathbf{U}_t\mathbf{G}_t' && \text{A.2b} \end{aligned} \qquad \text{(B.1)}$$

This is accomplished by the following functional equations and random generation procedures, $\beta_t = \mathbf{G}_t\beta_{t-1} + U_t\mathbf{u}_{\beta_t}$ where $\mathbf{u}_{\beta_\mathbf{t}} \sim N\left(0, I\right)$; U_t denotes a variance-covariance matrix associated to the evolution of β_t. Conversely, the evolutional functional equations and the (independent) random generation procedures for the evolution of $\mathbf{\Sigma}_t$ is chosen aiming to produce a dynamic conjugate updating procedure given β_t . The details of this evolution are found in Quintana et al. (2003) and can be summarized as follows: First, pre-multiply and post-multiply $\mathbf{\Sigma}_t$ by the inverse of the Cholesky decomposition of the scale parameter of its inverse Wishart distribution (this new random matrix has also a Wishart distribution but with an identity matrix as its scale parameter). Second, find the Cholesky decomposition of this new random matrix. Third, multiply the diagonal elements of the latter by suitable transformations of beta independent random variables to discount their degrees of freedom. Finally, reverse the transformations of the first and second steps. The entire process is similar to taking a computer apart, tweaking one component, and putting it back together. The end result of choosing this form of evolution is that (given β_t) if the posterior distribution of $\mathbf{\Sigma}_{t-1} \sim IW\left(d_{t-1}, \mathbf{S}_{t-1}\right)$ is an inverse Wishart then the prior distribution of $\mathbf{\Sigma}_t \sim IW\left(d_t^\star, \mathbf{S}_t^\star\right)$ where

$$\mathbf{S}_t^\star = \delta_t \mathbf{S}_{t-1} \quad \text{A.3a}$$

$$d_t^\star = \delta_t d_{t-1} \quad \text{A.3b}$$

and $0 < \delta_t \leq 1$ is interpreted as an information discount factor.

This provides a framework where the sequence of estimates of $\mathbf{\Sigma}_t$ keep adapting to new information while further discounting past observations.

Variance-Covariance Graphical Model

A Graphical model is a probability model that characterizes the conditional independence structure of a set of random variable by a graph (Lauritzen 1996; Jones, Carvalho, Dobra, Hans, Carter, and West 2005). Graphs provide a way to decompose the sample space into subsets of variables generating efficient ways to model conditional and marginal distributions locally. In high-dimensional problems, graphical model structuring is a key approach to parameter dimension reduction and, hence, to scientific parsimony and statistical efficiency when appropriate graphical structures can be identified.

In the context of a multivariate normal distribution, conditional independence restrictions are simply expressed through zeros in the off-diagonal elements of the precision matrix (inverse of the covariance matrix), establishing a parsimonious way to model covariance structures. Let a p-vector $\mathbf{y} \sim N(0, \mathbf{\Sigma})$ and $\mathbf{\Omega} = \mathbf{\Sigma}^{-1}$ with elements ω_{ij}. Write $G = (V, E)$ for the undirected graph whose vertex set V corresponds to the set of p random variables in $\mathbf{y}$, and E contains the elements (i, j) for which ω_{ij} are not equal to 0. The canonical parameter $\mathbf{\Omega}$ belongs to $M(G)$, the set of all positive-definite symmetric matrices with elements equal to zero for all $(i, j) \notin E$.

In working with decomposable Gaussian graphical models, Dawid and Lauritzen (1993) defined a family of conjugate Markov probability distributions called *hyper-inverse Wishart*. If $\mathbf{\Omega} \in M(G)$, the hyper-inverse Wishart

$$\mathbf{\Sigma} \sim HIW_G(d, \mathbf{S})$$

has a degree-of-freedom parameter b and location matrix $\mathbf{S} \in M(G)$ implying that each clique $C \in \mathcal{C}$, $\mathbf{\Sigma}_C \sim IW(d, \mathbf{S}_C)$ where $\mathbf{S}_C$ is the diagonal block of $\mathbf{S}$ corresponding to the vertices in C.

As shown by Carvalho and West (2007), graphical structuring can be incorporated in matrix normal DLMs to provide parsimonious models for the innovation variance-covariance matrix $\mathbf{\Sigma}$. For a given decomposable graph G, take the hyper-inverse Wishart as a conjugate prior for $\mathbf{\Sigma}$; it turns out that the closed-form, sequential updating theory of DLMs can be generalized to this richer model class.

This is also true for the case described in the previous section where $\mathbf{\Sigma}$ is time-varying. In detail, if the posterior at time $t-1$ is

$$\mathbf{\Sigma}_{t-1} \sim HIW_G(d_{t-1}, \mathbf{S}_{t-1})$$

the prior at time t take the form

$$\mathbf{\Sigma}_t \sim HIW_G(\delta d_t b_{t-1}, \delta_t \mathbf{S}_{t-1}).$$

This development is based on the same stochastic model of independent beta shocks applied to the diagonal elements of the cholesky decomposition of $\mathbf{\Sigma}_{t-1}$ presented above.

B.2 Computations

Specific Bayesian DLM Computations

The specific DLM described above does not allow for tractable exact recursive updating formulas; however, as a result of the choice for the specific Markovian parametric evolution, given the time series $\mathbf{\Sigma}_t$, the conventional updating recursions (West and Harrison 1997) apply for updating the hyper parameters of the normal distribution corresponding to the parameter β_t; and conversely, given the time series β_t, the conventional updating recursions (West and Harrison 1997) apply for updating the hyper parameters of the inverse Wishart distribution corresponding to the parameters $\mathbf{\Sigma}_t$. This suggests a Gibbs sampling scheme for the implementation of the model.